ISSN 2986-3929

Volume 1, Nomor 4, Agustus 2023

UNIVERSITAS UDAYANA

JNATIA

Jurnal Nasional Teknologi Informasi dan Aplikasinya

Program Studi Informatika
Fakultas Matematika dan Ilmu Pengetahuan Alam
Universitas Udayana

Susunan Dewan Redaksi
Jurnal Nasional Teknologi Informasi dan Aplikasinya (JNATIA)
Volume 1, Nomor 4, Agustus 2023

JNATIA Volume 1, Nomor 4, Agustus 2023 p-ISSN: 2986-3929
Jurnal Nasional Teknologi Informasi dan Aplikasinya

Daftar Isi

JNATIA Volume 1, Nomor 4, Agustus 2023 p-ISSN: 2986-3929
Jurnal Nasional Teknologi Informasi dan Aplikasinya

Prototype Aplikasi Edukasi Cyberbullying Berbasis Mobile Android

Fiki Nur Rahman[a1], Shendy Aditayahya Wardana[a2], Pradityo Utomo[b3]

[a]Program Studi Manajemen Informatika, Universitas Merdeka Madiun, Indonesia
[1]fiki.nr13@gmail.com
[2]shendywardana22@gmail.com
[3] pradityo@unmer-madiun.ac.id

Abstract

As technology develops rapidly, it has not only positive but also negative impacts; one of the negative impacts is the case of cyberbullying. Cyberbullying is an unpleasant action towards other people on the internet. One way to reduce cyberbullying is to create educational applications about cyberbullying. This application will provide an explanation of what cyberbullying is and how to prevent cyberbullying by providing an explanation that is light and easy to understand. This application is based on Android mobile, so it can more easily reach more users, using Java programming and display design using figma.

Keywords: *Cyberbullying, Education Application, Mobile Android*

1. Pendahuluan

Perkembangan teknologi informasi dan komunikasi (TIK) semakin pesat di era globalisasi dan telah merambah ke segala bidang kehidupan masyarakat. Perkembangan pesat ini telah menyebabkan banyak perubahan perilaku manusia di seluruh dunia. Perubahan perilaku ini disebabkan oleh transformasi interaksi masyarakat yang awalnya bersifat visual, audible, dan tactile menjadi interaksi virtual[1]. Salah satu kasus akibat dari perkembangan teknologi ini adalah Cyberbullying. Cyberbullying adalah suatu tindakan mengolok-olok, mempermalukan dan mengintimidasi seseorang menggunakan media internet atau digital. Dalam hal ini, Cyberbullying termasuk di dalamnya komentar sentimental negatif yang mengandung unsur Cyberbullying dalam pesan pribadi tertentu yang tidak baik di media sosial. Dampak negatif adanya Cyberbullying diantarannya dapat memengaruhi pola pikir korban dan menyebabkan depresi sedang hingga berat, gangguan emosi, dan rasa tidak aman, yang dapat menyebabkan depresi, penyerangan, dan bahkan bunuh diri[2]. Peristiwa Cyberbullying ini tidak bisa dianggap remeh mengingat dampaknya yang besar.

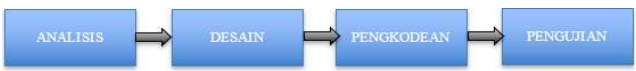
Berdasarkan data nasional yang diambil dari Komisi Perlindungan Anak Indonesia(KPAI) dalam kurun waktu 2011 sampai dengan 2019 tercatat ada pengaduan kasus Cyberbullying yang angkanya mencapai 2473 laporan dan kasus tersebut selalu terus mengalami peningkatan di setiap tahunnya[3]. Dikarenakan hal tersebut diperlukan suatu solusi untuk mengurangi kasus tersebut agar tidak mengalami kenaikan. Suatu cara yang bisa dilakukan adalah dengan membuat sebuah media pemahaman salah satunya adalah media edukasi. Media edukasi yang dipilih harus sesuai dengan perkembangan dunia digital layaknya sebuah smartphone yang memberikan berbagai macam kemudahan[4]. Berdasarkan hal tersebut maka aplikasi edukasi bisa menjadi suatu alternatif media pemahaman. Media aplikasi yang dibuat harus menunjukan pembelajaran yang baik[5].

Android adalah sistem operasi yang dapat digunakan pada ponsel dan tablet berbasis Linux dengan layar sentuh. Basis sistem operasi Android adalah kernel Linux open source, sehingga pengembang dapat menggunakan sistem operasi Android untuk membuat aplikasi Android sendiri kapan saja, yang dapat digunakan di berbagai perangkat mobile[6]. Pemilihan aplikasi berbasis mobile android karena aplikasi berbasis mobile android merupakan teknologi yang

berkembang dan bersifat bisa digunakan dimana saja sehingga aplikasi berbasis mobile android ini sangat cocok untuk penerapan aplikasi edukasi[7]. Berdasarkan hal tersebut dibuatlah sebuah prototype aplikasi edukasi Cyberbullying berbasis mobile android. Prototype aplikasi ini diharapkan dapat digunakan untuk menyelesaikan masalah perundungan yang terjadi di sosial media atau yang lebih dikenal dengan Cyberbullying[8]. Hasil aplikasi ini kemudian diverifikasi menggunakan metode black box testing, karena menawarkan banyak keuntungan dibanding metode verifikasi lainnya[9]. Aplikasi ini memiliki banyak materi pembelajaran tentang Cyberbullying dimana pengguna akan diberikan materi yang mudah dipahami dan dipelajari.

2. Metode Penelitian

Metode penelitian yang digunakan adalah metode penelitian model pengembangan sekuensial linier atau model waterfall. Metode waterfall sering disebut metode konvensional atau classic life cycle[10]. Metode ini merupakan metode yang sering di gunakan dalam tahap pengembangan atau prototype. Model waterfall memberikan pendekatan berurutan atau bertahap, dimulai dengan tahap analisis, desain, pengkodean, pengujian dan support[11]. Alur waterfall dapat digambarkan sebagai berikut:



Gambar 1. Alur Metode Waterfall

Tahapan dalam Model waterfall sebagai berikut:

a. Tahapan Analisis

Pada tahapan ini dilakukan analisa pengumpulan data apa yang dibutuhkan untuk pengembangan aplikasi, untuk pengembangan berikutnya diperlukan hardware berupa laptop atau pc dan software berupa android studio dan figma untuk desain tampilan. Selain mendefinisikan Hardware dan Software yang akan digunakan pada pengembangan system juga dilakukan pendefinisian kebutuhan data dan informasi terkait Cyberbullying yang diperlukan dalam pengembangan aplikasi.

b. Tahapan Desain

Setelah melakukan analisis kebutuhan tahap berikutnya melakukan design tampilan aplikasi yang akan dibuat, pada aplikasi ini diperlukan beberapa tampilan dan fasilitas yang diperlukan diantarannya tampilan awal aplikasi, tampilan menu utama, tampilan tiap menu dan tampilan menu quiz. Dan aplikasi yang akan dikembangkan telah ditetapkan ada sejumlah 5 tampilan dasar yaitu : splash screen, start page, Home page, menu penjelasan, dan tampilan quiz.

c. Tahapan Pengkodean

Setelah melakukan desain tampilan tahap berikutnya melakukan pengkodean, dari desain dan prototype yang telah disusun sebelumnya kemudian di transformasikan ke dalam bentuk Bahasa pemrograman. Untuk bahasa pemrograman yang di gunakan adalah bahasa pemrograman java karena mudah untuk dikembangkan dalam berbagai jenis aplikasi[12].

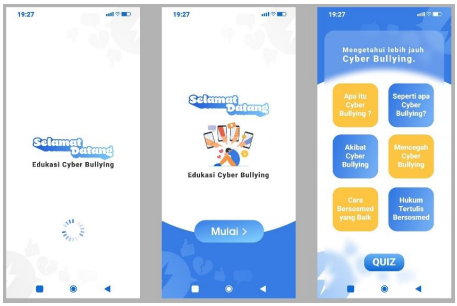
d. Tahapan Pengujian Dan Support

Setelah penyusunan kode program selesai tahap selanjutnya melakukan pengujian terhadap aplikasi yang telah dibuat, apakah aplikasi tersebut dapat berjalan dengan baik sesui dengan yang diharapkan. Untuk pengujian menggunakan metode black box testing

karena berfokus terhadap spesifikasi fungsional program sehingga pengujian hanya berfokus pada kesesuaian output aplikasi yang telah dibangun[13].

3. Hasil dan Pembahasan

Berdasarkan hasil prototype pada pembahasan sebelumnya, diperoleh hasil pengembangan aplikasi sebagai berikut :



Gambar 2. Tampilan Splash Screen, Start Page, Home Page

3.1. Splash Screen

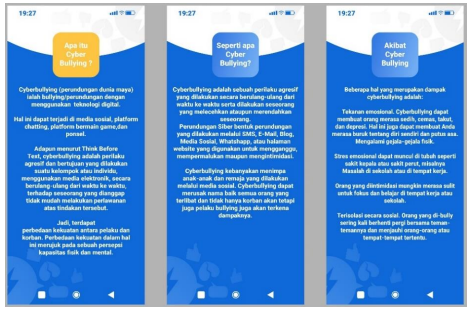
Tampilan awal aplikasi ketika berhasil di pasang disertai voice pada saat membuka aplikasi sebelum masuk ke tampilan menu awal aplikasi.

3.2. Start Page

Pada menu ini pengguna diharuskan menekan tombol “mulai” untuk masuk ke menu utama aplikasi dan kemudian pengguna dihadapkan pada beberapa pilihan menu aplikasi yang dapat diakses sesuai kebutuhan dari pengguna.

3.3. Home Page

Pada tampilan ini berisi beberapa button penjelasan materi edukasi yang dapat dipilih dengan menekan salah satu button maka, Aplikasi akan berpindah ke layar yang telah di pilih. Beberapa button diantarannya button Apa itu Cyberbullying, button Seperti Apa Cyberbullying, button Akibat Cyberbullying, button Cara Bersosmed Yang Baik, button Hukum Tertulis Bersosmed, button Mencegah Cyberbullying dan button Quiz.



Gambar 3. Tampilan Menu Penjelasan

3.4. Menu Penjelasan

Pada bagian ini berisi tentang materi edukasi yang sebelumnya telah dipilih dibagian home page disertai dengan penjelasan yang mudah untuk dipahami, seperti materi tentang apa itu cyberbullying, seperti apa cyberbullying, akibat cyberbullying, cara bersosmed yang baik, hukum tertulis bersosmed, dan mencegah cyberbullying.

Gambar 4. Tampilan Menu Quiz

3.5. Quiz

Tampilan ini berisi menu Quiz, Terdapat soal dan jawaban yang dapat dipilih kemudian klik tombol "Lanjut" untuk melanjutkan ke soal selanjutnya dan tombol "Kembali" untuk kembali ke soal sebelumnnya.

4. Kesimpulan

Berdasarkan pembahasan dan uraian di atas dapat disimpulkan bahwa prototype aplikasi ini bisa dikatakan berhasil. Dengan adanya aplikasi ini diharapkan dapat menjadi tempat edukasi atau pembelajaran mengenai tindakan Cyberbullying dan juga dapat mengurangi aksi Cyberbullying yang selama ini terjadi.

Daftar Pustaka

[1] N. K. P. Pratiwi, A. A. S. L. Dewi, and I. M. M. Widyantara, "Penegakan Hukum Terhadap Pelaku Cyber Bullying di Kepolisian Daerah Bali pada Masa Pandemi Covid-19," J. Prefer. Huk., vol. 3, no. 1, pp. 190–195, 2022, doi: 10.22225/jph.3.1.4682.190-195.
[2] A. S. Hutagalung, A. B. P. Negara, and E. E. Pratama, "Aplikasi Pendeteksi Cyberbullying Terhadap Komentar Postingan Media Sosial Instagram dengan Metode Naïve Bayes Classifier Berbasis Website," J. Sist. dan Teknol. Inf., vol. 9, no. 3, p. 364, 2021, doi: 10.26418/justin.v9i3.44843.
[3] Carolus Borromeus Mulyatno, "Pengembangan Game Athena Versi 1.0 Sebagai Edukasi Cyberbullying Pada Remaja," J. Pendidik. dan Konseling, vol. 4, pp. 1349–1358, 2022.
[4] N. Juliansyah, Herlinda, and B. D. Theodora, "Perancangan Aplikasi Edukasi Menghitung Luas Dan Keliling Bangun Datar Berbasis Android," Semnas Ristek (Seminar Nas. Ris. dan Inov. Teknol., vol. 5, no. 1, pp. 397–403, 2021, [Online]. Available: http://proceeding.unindra.ac.id/index.php/semnasristek/article/view/4927/835
[5] N. Supriono and F. Rozi, "Pengembangan Media Pembelajaran Bentuk Molekul Kimia Menggunakan Augmented Reality Berbasis Android," JIPI (Jurnal Ilm. Penelit. dan Pembelajaran Inform., vol. 3, no. 1, pp. 53–61, 2018, doi: 10.29100/jipi.v3i1.652.
[6] S. Sulistyowati, E. Gunawan, and L. Rusdiana, "Aplikasi Game Edukasi Matematika Tingkat Dasar Berbasis Android," J. Teknoinfo, vol. 16, no. 1, p. 107, 2022, doi: 10.33365/jti.v16i1.806.
[7] R. Aditya, V. H. Pranatawijaya, and & P. B. A. A. Putra, "Rancang Bangun Aplikasi Monitoring Kegiatan Menggunakan Metode Prototype," J. Inf. Technol. Comput. Sci., vol. 1, no. 1, pp. 47–57, 2021.
[8] G. Gibeon, I. G. Ngurah, and A. Cahyadi, "Perancangan UI dan UX pada Aplikasi Pencari Indekos di Sekitar Kampus Universitas Udayana," vol. 1, pp. 789–796, 2023.
[9] A. Budiman, P. Utomo, and S. Rahayu, "Pengembangan Aplikasi Deteksi Dini Serangan Hama Padi Berbasis Android," J. Terap. Abdimas, vol. 4, no. 1, p. 33, 2019, doi: 10.25273/jta.v4i1.3805.
[10] F. Supandi, W. Desta P, Y. Ambar S, and M. Sudir, "Analisis Resiko Pada Pengembangan Perangkat Lunak Yang Menggunakan Metode Waterfall Dan Prototyping," Pros. Semin. Nas. Din. Inform. 2018 (SENADI 2018), vol. 2, no. 1, pp. 83–86, 2019, [Online]. Available: http://prosiding.senadi.upy.ac.id/index.php/senadi/article/view/86
[11] A. Oktaviani, D. Sarkawi, and A. Priadi, "Perancangan Aplikasi Penjualan Dengan Metode Waterfall Pada Koperasi Karyawan Rsud Pasar Rebo," Petir, vol. 11, no. 1, pp. 9–24, 2018, doi: 10.33322/petir.v11i1.3.
[12] A. P. C. Udaksana and W. R. Kusaeri, "Rancang Bangun Aplikasi Digital School Dengan Java NetBeans IDE 8.1," Irons, pp. 332–336, 2018, [Online]. Available: https://jurnal.polban.ac.id/proceeding/article/view/1118/918
[13] D. S. Purnia, A. Rifai, and S. Rahmatullah, "Penerapan Metode Waterfall dalam Perancangan Sistem Informasi Aplikasi Bantuan Sosial Berbasis Android," Semin. Nas. Sains dan Teknol. 2019, pp. 1–7, 2019.

Halaman ini sengaja dibiarkan kosong

JNATIA Volume 1, Nomor 4, Agustus 2023 p-ISSN: 2986-3929
Jurnal Nasional Teknologi Informasi dan Aplikasinya

Low-level Images Extraction Features pada Algoritma SVM untuk Mengidentifikasi Kematangan Jeruk

I Made Agus Rama Wijaya[a1], I Ketut Gede Suhartana[a2]

[a]Program Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam,
Universitas Udayana
Jalan Raya Kampus Udayana, Bukit Jimbaran, Kuta Selatan, Badung, Bali Indonesia
[1]ramawijaya1415@gmail.com
[2]ikg.suhartana@unud.ac.id

Abstract

Terkadang jika melakukan identifikasi secara manual oleh petani buah masih memunculkan peluang terjadinya human error saat panen. Untuk itu, penelitian ini dilakukan bertujuan untuk pelatihan klasifikasi terhadap buah jeruk guna mengurangi error rate dalam pemanenan. Kematangan buah dipisah ke dalam dua kelas yaitu matang dan belum matang. Dimana dataset yang digunakan merupakan data sekunder citra buah jeruk dengan total sebanyak 400 buah, yaitu 200 citra matang dan 200 citra belum matang. Metode yang digunakan adalah Support Vector Machine kernel linear dengan penggabungan Low-level Image Extraction Features, yaitu warna dengan color histogram, tekstur dengan metode Gray Level Co-occurance Matrix (GLCM), dan bentuk dengan kalkulasi kontur. Pembagian klasifikasi dilakukan dengan membagi dataset menjadi 20% data uji dan 80% data latih. Hasil klasifikasi pada penelitian mendapatkan nilai akurasi sebesar 96,34%.

Keywords: *SVM, Klasifikasi, Low-level Extraction Features*

1. Pendahuluan

Sektor pertanian merupakan salah satu penopang terbesar pendapatan masyarakat Indonesia yang juga menjadi target pembangunan nasional [1]. Dimana salah satu sub sektornya yang memegang peran dalam hal tersebut adalah sektor perkebunan. Dunia perkebunan memiliki berbagai jenis tanaman yang dapat dibudidayakan baik berupa tanaman holtikultura, rempah-rempah, buah-buahan, dan kacang-kacangan. Pada tahun 2022 kemarin, kelompok buah-buahan memiliki hasil produksi yang tergolong sangat banyak, yaitu sekitar 28.368.404 ton [2]. Untuk mendapatkan hasil panen yang berlimpah dan berkualitas, diperlukan ketelitian dalam memanen dalam waktu yang lebih singkat. Petani buah juga merupakan seorang manusia yang tetap memiliki peluang terjadinya human error. Melalui jumlah panen yang sangat banyak, diperlukan sebuah alat bantu untuk melakukan klasifikasi kematangan buah agar mendapatkan hasil panen yang lebih berkualitas dengan error rate yang lebih rendah.

Pada perkembangan zaman seperti sekarang ini, hampir seluruh bidang kehidupan menerapkan konsep otomatisasi dalam menyelesaikan permasalahan. Konsep tersebut tidak jauh dari penggunaan Artificial Intelligence. Dimana Artificial Intelligence atau kecerdasan buatan diartikan sebagai sebuah mesin atau sistem yang didesain agar dapat menyelesaikan pekerjaan-pekerjaan dengan kemampuan penyelesaian masalah seperti kecerdasan manusia [3]. Umumnya beberapa parameter yang dijadikan acuan bagi para petani buah untuk mengidentifikasi kematangannya adalah warna, tekstur, dan bentuk. Dengan demikian, konsep kecerdasan buatan akan dapat membantu para petani buah melakukan panen dengan mengenali kematangan buah menggunakan pemrosesan citra dengan tiga parameter acuan tersebut.

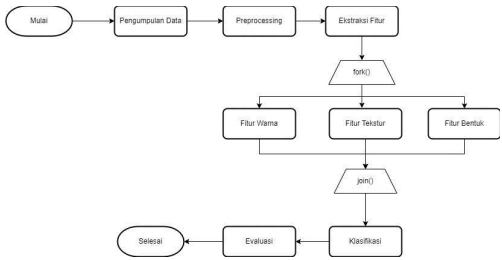
Sebelumnya terdapat beberapa artikel penelitian terkait klasifikasi kematangan buah dengan objek serta metode pengolahan yang berbeda-beda, seperti yang ditulis oleh Juprianus dkk, Ismail dkk, dan Rizal dkk. Ketiga artikel menggunakan objek penelitian buah di antaranya adalah buah melon dan buah kopi. Berdasarkan penggunaan algoritma Support Vector Machine,

penerapan klasifikasi dengan metode ekstraksi fitur warna HSV menghasilkan tingkat akurasi yang sama pada buah melon dan kopi yaitu 86,67%. Sedangkan, penggunaan metode ekstraksi fitur tekstur dengan Gray Level Co-occurence Matrix (GLCM) menghasilkan tingkat akurasi yang lebih kecil dibandingkan dengan ekstraksi fitur warna yaitu 80% [4] [5] [6].

Melalui studi literatur mengenai artikel terkait dapat disimpulkan bahwa algoritma Support Vector Machine (SVM) cocok dan dapat digunakan untuk melakukan klasifikasi berdasarkan ekstraksi fitur input yang dipilih. Namun, setiap ekstraksi fitur yang dipilih tentu memiliki hasil yang berbeda-beda. Oleh karena itu, artikel ini bertujuan untuk menjelaskan penelitian yang menerapkan Low-Level Image Extraction Feature yang meliputi warna, tekstur, dan bentuk dari buah jeruk sehingga dapat mengklasifikasi kematangannya secara otomatis dan konsisten serta memiliiki performa yang lebih baik.

2. Metode Penelitian

Dalam mencapai hasil terkait klasifikasi kematangan buah dengan Low-level Image Extraction Feature berikut adalah alur penelitian yang digunakan.



Gambar 1. Alur Penelitian

2.1. Pengumpulan Data

Metode pengumpulan data bersifat sekunder yang mengambil dataset berupa citra jeruk. Data yang diambil sebanyak 406 citra dibagi menjadi dua kelas, yaitu 203 citra jeruk matang dan 203 citra jeruk belum matang. Pengumpulan data dilakukan melalui sebuah website bernama roboflow universe.

2.2. Preprocessing

Preprocessing merupakan tahap awal dalam melakukan pemrosesan citra yang bertujuan menyetarakan dan membersihkan data agar siap untuk diolah lebih lanjut. Beberapa tahap preprocessing yang dilakukan antara lain:

a. Cropping

Proses ini merupakan tahap untuk memotong gambar sesuai dengan objek yang akan di analisis, pada kasus ini yaitu jeruk. Sehingga, latar belakang dari gambar tidak akan dianalisis. Selain itu, tujuan dilakukan cropping adalah menyesuaikan komposisi citra agar citra input menjadi satu ukuran rasio yang sama.

b. Image Denoising

Noise pada data citra sangat rawan muncul ketika data tersebut diambil. Pada dataset yang digunakan terdapat noise yang mengganggu karena menghilangkan beberapa detail penting dari citra. Oleh karena itu, perlu dilakukan tahap ini guna memperbaiki kualitas gambar.

c. Konversi Ruang Warna

Umumnya citra direpresentasikan dengan 3 channel yaitu R (red), G (green), dan B (blue). Namun, dalam pemrosesan data citra khususnya untuk mengekstraksi fitur tekstur dan bentuk memerlukan data citra yang sudah dikonversi menuju 1 channel yang merepresentasikan intensitas dari tiap pikselnya. Konversi dapat dilakukan dengan proses grayscaling, dimana dapat dikalkulasi dengan rumus berikut [7].

$$\text{Grayscale} = 0.299 * \text{R} + 0.587 * \text{G} + 0.114 * \text{B} \quad (1)$$

2.3. Ekstraksi Fitur

Ekstraksi fitur adalah bagian penting yang menentukan bagaimana sebuah model dapat melakukan klasifikasi. Fitur-fitur input yang sudah diekstraksi akan menjadi parameter seleksi. Dalam penelitian ini, fitur yang digunakan adalah Low-level Image Extraction Feature yang meliputi:

a. Fitur Warna

Persebaran nilai piksel dari tiap-tiap channel warna dapat dilihat melalui sebuah histogram yang disebut dengan color histogram. Fitur warna yang diekstraksi yaitu rata-rata (mean), variasi (variance), dan median dari persebaran tiap-tiap pikselnya.

b. Fitur Tekstur

Tekstur merupakan salah satu parameter uji yang digunakan untuk mengidentifikasikan kematangan buah. Penelitian ini menggunakan metode Gray Level Co-occurence Matrix yang akan menghasilkan beberapa informasi mengenai teksur yang dihitung melalui persamaan berikut.

$$\text{Contrast} = \sum_{i,j=0}^{l-1} P_{i,j}(i-j)^2 \quad (2)$$

$$\text{Dissimilarity} = \sum_{i,j=0}^{l-1} P_{i,j}|i-j| \quad (3)$$

$$\text{Homogeneity} = \sum_{i,j=0}^{l-1} \frac{P_{i,j}}{1+(i-j)^2} \quad (4)$$

$$\text{ASM} = \sum_{i,j=0}^{l-1} P_{i,j}^2 \quad (5)$$

$$\text{Energy} = \sqrt{\text{ASM}} \quad (6)$$

$$\text{Correlation} = \sum_{i,j=0}^{l-1} P_{i,j}\left[\frac{(i-\mu_i)(j-\mu_j)}{\sqrt{(\sigma_i^2)(\sigma_j^2)}}\right] \quad (7)$$

c. Fitur Bentuk

Ciri bentuk yang akan digunakan dalam penelitian ini berdasarkan nilai kontur sebuah objek dalam suatu citra. Dimana kontur tersebut merupakan rangkaian piksel yang saling terhubung mengikuti tepi objek luar yang akan dianalisis.

2.4. Klasifikasi

Klasifikasi adalah proses pengelompokkan suatu objek ke dalam kategori atau kelas yang sudah ditentukan. Model klasifikasi akan mempelajari dan mengenali pola dari tiap-tiap data yang ada. Pola tersebut kemudian dijadikan sebuah acuan untuk mengklasifikasikan sebuah data baru yang masuk. Dalam model yang dirancang, keseluruhan dataset akan dipecah menjadi 2 bagian, yaitu 80% sebagai data latih dan 20% sebagai data uji yang sudah disusun berdasarkan fitur-fitur ekstraksi.

2.5. Evaluasi

Evaluasi sebuah model penting dilakukan guna mengetahui performa yang dihasilkan. Salah satu metode evaluasi umum digunakan adalah Confusion Matrix. Confusion Matrix adalah tabel visualisasi performa yang dihasilkan sebuah model berdasarkan nilai aktual dan nilai prediksi. Melalui metode tersebut dihasilkan akurasi, precision, recall, dan F1-score berdasarkan perhitungan berikut.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{8}$$

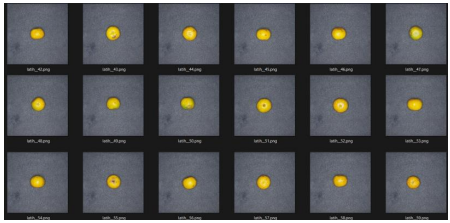
$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{9}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{10}$$

$$\text{F1} - \text{score} = \frac{2 * \text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \tag{11}$$

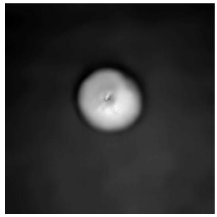
3. Hasil dan Pembahasan

Pengujian sistem klasifikasi yang dirancang terlebih dahulu mempersiapkan dataset berupa citra jeruk sebanyak 406 buah citra baik matang maupun belum matang seperti terlihat pada gambar 2.



Gambar 2. Data Citra Jeruk

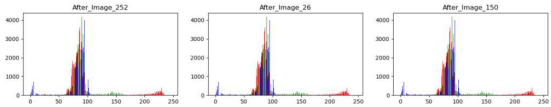
Proses selanjutnya adalah melakukan preprocessing dengan melakukan cropping, denoising, dan konversi ruang warna menuju bentuk grayscale.



Gambar 3. Sampel Gambar setelah Preprocessing

Setelah preprocessing data, barulah mencari nilai ekstraksi fitur yang disusun ke dalam bentuk statistikal

3.1. Ekstraksi Fitur Warna



Gambar 4. Histogram Persebaran Warna

Dalam pengambilan fitur warna, ruang warna citra akan tetap di pertahankan dalam bentuk RGB dikarenakan informasi warna pada buah jeruk merupakan faktor penting dalam proses klasifikasi kematangan. Berikut adalah beberapa sampel hasil ekstraksi fitur warna dengan mengambil nilai mean, median, dan variance.

Tabel 1. Ekstraksi Fitur Warna

Gambar	Mean	Variance	Median
matang_1	47.159930	530.106332	44.0
matang_2	45.815177	356.916839	44.0
matang_3	46.597610	400.679478	44.0
matang_4	45.437872	360.951889	43.0
mentah_1	113.679408	768.227978	114.0
mentah_2	115.027623	872.323936	115.0
mentah_3	114.621233	886.929013	114.0
mentah_4	115.571023	880.800658	115.0

3.2. Ekstraksi Fitur Tekstur

Melalui Gray Level Co-occurrence Matrix sebagai metode yang digunakan untuk melakukan ekstraksi fitur berdasarkan pola dan distribusi intensitas piksel yang direpresentasikan dengan enam parameter properti yaitu contrast, dissimilarity, homogeneity, ASM, energy, dan correlation. Namun, perlu diingat karena metode ini menggunakan distribusi intensitas piksel, maka ruang warna citra harus dikonversi ke dalam bentuk grayscale. Dalam penelitian ini, metode GLCM

menggunakan empat arah yaitu 0º, 45º, 90º, dan 135º. Keempat arah tentunya menghasilkan nilai ekstraksi yang berbeda-beda. Dengan demikian, ekstraksi fitur keseluruhan arah akan dirata-ratakan sehingga menghasilkan satu nilai ekstraksi fitur disetiap parameternya. Beberapa sampel hasil ekstraksi fitur tekstur yaitu sebagai berikut terlihat pada tabel 2.

Tabel 2. Ekstraksi Fitur Tekstur

Gambar	Contrast	Dissimilarity	Homogeneity	ASM	Energy	Correlation
matang_1	38.479755	2.060968	0.642851	0.013979	0.118207	0.957197
matang_2	26.704400	1.696761	0.655467	0.014899	0.122035	0.954070
matang_3	31.164091	1.849961	0.645623	0.014901	0.122042	0.952856
matang_4	25.984982	1.710584	0.649876	0.013708	0.117047	0.955123
mentah_1	70.709757	4.223336	0.327098	0.003562	0.059658	0.892224
mentah_2	83.465761	4.558104	0.305166	0.003590	0.059904	0.892677
mentah_3	80.439571	4.419954	0.313390	0.003441	0.058642	0.898330
mentah_4	81.672679	4.594225	0.301253	0.003678	0.060626	0.893906

3.3. Ekstraksi Fitur Bentuk

Pengambilan fitur bentuk diekstraksi melalui atribut-atribut yang merepresentasikan bentuk atau kontur sebuah jeruk dari keseluruhan bagian dalam citra. Informasi yang dapat diekstraksi berdasarkan nilai kontur adalah luas (area) dan keliling (perimeter). Berikut adalah beberapa hasil ekstraksi fitur bentuk yang diperoleh.

Tabel 3. Ekstraksi Fitur Bentuk

Gambar	**Area**	**Perimeter**
matang_1	2999.0	207.823374
matang_2	2465.5	191.438598
matang_3	2942.0	203.480230
matang_4	2482.0	191.195958
mentah_1	2856.5	208.066014
mentah_2	3341.5	231.722869
mentah_3	3665.5	228.066015
mentah_4	3582.0	225.137082

3.4. Hasil Evaluasi

Tabel 4. Confusion Matrix

		Prediksi	
		1	0
Aktual	1	46	1
	0	2	33

Berdasarkan tabel tersebut, total jumlah data uji yaitu sebanyak 82 buah dimana 1 merepresentasikan kondisi matang dan 0 merepresentasikan kondisi belum matang. Dengan nilai

yang terdapat pada tabel dapat dihasilkan beberapa nilai parameter uji seperti akurasi, precision, recall, dan F1-score sebagai berikut.

$$\text{Accuracy} = \frac{46+33}{82} = 96{,}34\% \quad (12)$$

$$\text{Precision} = \frac{46}{46+2} = 95{,}83\% \quad (13)$$

$$\text{Recall} = \frac{46}{46+1} = 97{,}87\% \quad (14)$$

$$\text{F1} - \text{score} = \frac{2 * 0{,}9583 * 0{,}9787}{0{,}9583 + 0{,}9787} = 96{,}83\% \quad (15)$$

4. Kesimpulan

Sistem klasifikasi kematangan buah dengan pemrosesan citra yang dikembangkan dapat melakukan identifikasi dengan hasil yang cukup baik. Sebelum melalui tahap pemrosesan, data citra harus dibersihkan terlebih dahulu dengan tahap preprocessing. Preprocessing yang dilakukan antara lain cropping, image denoising, dan konversi ruang warna dengan tujuan analisis pada tekstur dan kontur. Ekstraksi fitur dilakukan dengan menggabungkan Low-level Image Extraction Features yang meliputi histogram warna, GLCM, dan analisis kontur. Sehingga, hasil yang didapatkan setelah melalui uji Confusion Matrix memilki performa yang sangat baik yaitu 96,34% akurasi, 95,83% precision, 97,87% recall, dan 96,83% F1-score. Hal ini menunjukkan kombinasi dari ketiga fitur pada Low-level Image Extraction Features memberikan dampak yang cukup besar terhadap performa algoritma Support Vector Machine.

Daftar Pustaka

[1] U. Isbah e R. Y. Iyan, “Analisis Peran Sektor Pertanian Dalam Perekonomian Dan Kesempatan Kerja Di Provinsi Riau,”Jurnal Sosial Ekonomi Pembangunan, vol. 7, no 19, pp. 45-54, 2016.

[2] BPS, Statistik Indonesia, Jakarta: Badan Pusat Statistik, 2023.

[3] M. S. Y. Lubis, “Implementasi Artificial Intelligence Pada System Manufaktur Terpadu,” Seminar Nasional Teknik (SEMNASTEK) UISU, vol. 4, no 1, pp. 1-7, 2021.

[4] J. Rusman e N. Pasae, “Prototype Sistem Penyortir Buah Kopi Arabika Berdasarkan Tingkat Kematangan Menggunakan Metode Support Vector Machine,” TEKNIKA, vol. 12, no 1, pp. 65-72, 2023.

[5] Ismail, N. Arifin e Prihastinur, “Klasifikasi Kematangan Buah Naga Berdasarkan Fitur Warna Menggunakan Algoritma Multi-Class Support Vector Machine,” JINTEKS (Jurnal Informatika Teknologi dan Sains), vol. 5, no 1, p. 121–126, 2023.

[6] R. A. Saputra, D. Puspitasari e T. Baidawi, “Deteksi Kematangan Buah Melon dengan Algoritma Support Vector Machine Berbasis Ekstraksi Fitur GLCM,” Jurnal Infortech, vol. 4, no 2, pp. 200-206, 2022.

[7] T. Akbar, M. F. B, M. A. Amir, A. A. N. Risal, N. A. A. Safanah e M. M. Fakhri, “Sulsel Typical Batik Motif Classification Using Neural Network Method With GLCM Feature Extraction,” DECODING, vol. 1, no 1, pp. 24-33, 2023.

Halaman ini sengaja dibiarkan kosong

JNATIA Volume 1, Nomor 4, Agustus 2023 p-ISSN: 2986-3929
Jurnal Nasional Teknologi Informasi dan Aplikasinya

Klasifikasi Teks Spam dengan Algoritma Support Vector Machine dan Chi – Square

Getzbie Alfredo Tpoy[a1], Agus Muliantara[a2]

[a]Program Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Udayana
Jalan Raya Kampus Udayana, Bukit Jimbaran, Kuta Selatan, Badung, Bali Indonesia
[1]getzbiealfredo123@gmail.com
[2]muliantara@unud.ac.id

Abstract

Spam messages are messages that contain false information, commonly regarding events, banking, insurance, bills, advertisements, and viruses. To address the issue of spam, classification can be performed on the received messages. Classification can be done by separating texts that contain spam messages from texts that contain legitimate (ham) messages. In this study, spam text classification was conducted using the Support Vector Machine algorithm, feature selection using Chi-Square. The Chi-Square feature selection method was performed using percentages of 20%, 40%, 60%, and 80%, with accuracy, precision, recall, and F1-Score as the measured values. The result of study obtained was an accuracy of 98.82% with an F1-Score of 93.05% at a feature selection percentage of 60%, using the RBF kernel. Feature selection with percentages of 20%, 40%, and 80% resulted in accuracies of 97.93%, 98.29%, and 98.02%, respectively. These accuracies were better compared to the accuracy without feature selection, which was 97.57%.

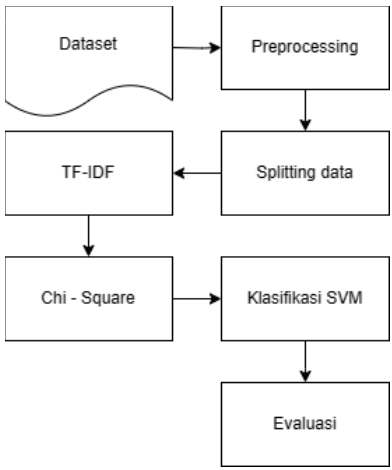
Keywords: *Chi - Square, spam, support vector machine*

1. Pendahuluan

Majunya teknologi memberikan banyak manfaat bagi banyak orang. Manfaat yang diterima berbanding lurus dengan permasalahan yang ditimbulkan, salah satu permasalahan yang muncul adalah maraknya email Spam. Pesan spam merupakan sebuah pesan yang berisi informasi palsu, umumnya mengenai event, perbankan, asuransi, tagihan, iklan, dan virus[1]. Spam umumnya disebarkan secara terus – menerus, sehingga beberapa pengguna dapat menerima banyak pesan spam dalam satu waktu, hal ini memberikan rasa resah dan cukup mengganggu. Cara untuk mengatasi permasalahan spam dengan melakukan klasifikasi terhadap pesan yang diterima. Klasifikasi dapat dilakukan dengan memisahkan teks yang berisikan pesan palsu dengan teks yang berisikan pesan tidak palsu(ham). Memisahkan pesan asli dan palsu memang dapat dilakukan secara langsung oleh manusia, namun tentunya akan menyulitkan jika teks yang dipisahkan berjumlah sangat banyak, sehingga diperlukan bantuan komputasi dengan menggunakan algoritma klasifikasi. Beberapa penelitian sebelumnya yang berkaitan dengan klasifikasi pesan spam, penelitian Syam [2] klasifikasi komentar spam pada instagram menggunakan metode support vector machine memperoleh nilai precision 97.33%, nilai recall 97.33%, dan akurasi 97.33%. Penelitian Ghani [1] Email Spam Filtering dengan Algoritma Random Forest dengan evaluasi menggunakan confusion matrix diperoleh hasil akurasi 94,22% dan AUC 0,98. Penelitian Syafii [3]Klasifikasi SMS Spam Dengan Komparasi Metode SVM Dan Naïve Bayes diperoleh hasil akurasi sebesar 0.94 pada metode Naïve Bayes dan 0.93 pada metode Support Vector Machine. Berdasarkan beberapa metode klasifikasi dan studi kasus yang dilakukan sebelumnya, dalam penelitian ini akan dilakukan klasifikasi teks spam menggunakan algoritma Support Vector Machine dengan seleksi fitur Chi-Square. Diharapkan dengan digunakannya seleksi fitur, performa klasifikasi spam pada email dengan algortitma Support Vector Machine dapat ditingkatkan.

2. Metode Penelitian

Alur pada penelitian sesuai pada gambar 1, pertama akan dilakukan pengambilan dataset yang akan digunakan dalam penelitian. Preprocessing dilakukan untuk membersihkan data sebelum pemrosesan. Kemudian dilakukan pemisahan data training dan testing. Pembobotan menggunakan TF-IDF. Menyeleksi fitur dengan menggunakan metode CHI-SQUARE. Melakukan klasifikasi dengan Support Vector Machine. Terakhir melakukan evaluasi.



Gambar 1. Alur Penelitian

2.1. Dataset

Dataset yang digunakan dalam penelitian ini adalah dataset sekunder yang contohnya dapat dilihat pada tabel 1. Data didapat dari Kaggle.com berupa email yang berisikan teks email spam dan ham. Data berjumlah 5169 dengan rincian 87% data spam dan 13% data ham.

Tabel 1. Contoh Dataset

Teks	Label
URGENT! Your Mobile No. was awarded à£2000 Bonus Caller Prize on 5/9/03 This is our final try to contact U! Call from Landline 09064019788 BOX42WR29C, 150PPM	Spam
I know you are thinking about malaria. But relax, children can't handle malaria. She would have been worse, and it was gastroenteritis. If she takes enough time to replace her loss her temp will reduce. And if you give her malaria meds now, she will just vomit. It's a self-limiting	Ham

Teks	Label
illness she has which means in a few days it will completely stop	

2.2. Text Preprocessing

Text preprocessing dilakukan untuk membersihkan data yang akan diproses di tahap selanjurtnya. Tahapan preprocessing yang dilakukan adalah case folding, cleaning, tokenizing, filtering/stopwords removal, stemming.

a. Case Folding

Hasil case folding pada tabel 2, merupakan tahapan untuk mengubah semua huruf yang ada pada teks menjadi huruf kecil.

Tabel 2. Case Folding

Teks	Case Folding
I know you are thinking malaria. But relax, children can't handle malaria. She would have been worse and its gastroenteritis. If she takes enough to replace her loss her temp will reduce. And if you give her malaria meds now, she will just vomit. It's a self-limiting illness she has which means in a few days it will completely stop	i know you are thinking malaria. but relax, children can't handle malaria. she would have been worse and its gastroenteritis. if she takes enough to replace her loss her temp will reduce. and if you give her malaria meds now, she will just vomit. it's a self-limiting illness she has which means in a few days it will completely stop

b. Cleansing

Cleansing merupakan tahapan untuk membersihkan teks dari karakter yang tidak perlu seperti tanda baca, link, serta emoticon. Contoh hasil cleansing dapat dilihat pada tabel 3.

Tabel 3. Cleansing

Teks	Cleansing
I know you are thinking malaria. But relax, children can't handle malaria. She would have been worse and its gastroenteritis. If she takes enough to replace her loss her temp will reduce. And if you give her malaria meds now, she will just vomit. It's a self-limiting illness she has which means in a few days it will completely stop	i know you are thinking malaria but relax children cant handle malaria she would have been worse and its gastroenteritis if she takes enough to replace her loss her temp will reduce and if you give her malaria meds now she will just vomit its a self-limiting illness she has which means in a few days it will completely stop

c. Tokenizing

Tokenizing merupakan tahapan untuk memecah kalimat pada teks menjadi term atau kata. Kata – kata yang terdapat pada dokumen akan dipecah menjadi kata tunggal yang hasilnya dapat dilihat pada tabel 4.

Tabel 4. Tokenizing

Teks	Tokenizing
i know you are thinking malaria. but relax, children can't handle malaria. she would have been worse and its gastroenteritis. if she takes enough to replace her loss her temp will reduce. and if you give her malaria meds now, she will just vomit. it's a self-limiting illness she has which means in a few days it will completely stop	['i', 'know', 'you', 'are', 'thinking', 'malaria', 'but', 'relax', 'children', 'cant', 'handle', 'malaria', 'she', 'would', 'have', 'been', 'worse', 'and', 'its', 'gastroenteritis', 'if', 'she', 'takes', 'enough', 'to', 'replace', 'her', 'loss', 'her', 'temp', 'will', 'reduce', 'and', 'if', 'you', 'give', 'her', 'malaria', 'meds', 'now', 'she', 'will', 'just', 'vomit', 'its', 'a', 'self-limiting', 'illness', 'she', 'has', 'which', 'means', 'in', 'a', 'few', 'days', 'it', 'will', 'completely', 'stop']

d. Stopword Removal

Langkah ini merupakan tahapan dalam menghapus kata yang dirasa tidak perlu dalam teks. Kata tersebut berupa kata umum dan dianggap tidak relevan. Hasilnya ditunjukan pada tabel 5.

Tabel 5. Stopword

Teks	Stopword
['i', 'know', 'you', 'are', 'thinking', 'malaria', 'but', 'relax', 'children', 'cant', 'handle', 'malaria', 'she', 'would', 'have', 'been', 'worse', 'and', 'its', 'gastroenteritis', 'if', 'she', 'takes', 'enough', 'to', 'replace', 'her', 'loss', 'her', 'temp', 'will', 'reduce', 'and', 'if', 'you', 'give', 'her', 'malaria', 'meds', 'now', 'she', 'will', 'just', 'vomit', 'its', 'a', 'self-limiting', 'illness', 'she', 'has', 'which', 'means', 'in', 'a', 'few', 'days', 'it', 'will', 'completely', 'stop']	['know', 'thinking', 'malaria', 'relax', 'children', 'cant', 'handle', 'malaria', 'would', 'worse', 'gastroenteritis', 'takes', 'enough', 'replace', 'loss', 'temp', 'reduce', 'give', 'malaria', 'meds', 'vomit', 'self', 'limiting', 'illness', 'means', 'few', 'days', 'completely', 'stop']

e. Stemming

Stemming merupakan tahapan untuk mengubah sebuah kata yang ada Kembali ke kata aslinya, atau menjadi kata dasar yang hasilnya ditunjukan pada tabel 6.

Tabel 6. Stemming

Teks	Stemming
['know', 'thinking', 'malaria', 'relax', 'children', 'cant', 'handle', 'malaria', 'would', 'worse', 'gastroenteritis', 'takes', 'enough', 'replace', 'loss', 'temp', 'reduce', 'give', 'malaria', 'meds', 'vomit', 'self', 'limiting', 'illness', 'means', 'days', 'completely', 'stop']	['know', 'thinking', 'malaria', 'relax', 'children', 'cant', 'handl', 'malaria', 'would', 'wor', 'gastroent', 'take', 'enough', 'replac', 'loss', 'temp', 'reduc', 'give', 'malaria', 'med', 'vomit', 'self', 'limit', 'ill', 'mean', 'day', 'complet', 'stop']

2.3. Splitting Data

Data akan dipisah menjadi data latih dan data uji. Rincian pemisahan data adalah 80% data untuk pelatihan, 20% data untuk pengujian.

2.4. TF-IDF

Pemobotan dilakukan dalam menentukan angka atau nilai pada frekuensi sebuah kata sebagai bobot yang dapat digunakan untuk pemrosesan selanjutnya [4]Metode yang digunakan dalam pembobotan adalah TF-IDF. TF-IDF memberikan bobot yang berbeda berdasarkan frekuensi term di dokumen, dan frekuensi term di seluruh dokumen. Tahapannyanya adalah menghitung TF (1), menghitung inverse DF (2), dan terakhir menghitung TF-IDF (3).

$$tft = 1 + \log(tft) \quad (1)$$

Keterangan;

tft: jumlah kemunculan term t

$$idft = log(\frac{\mathrm{D}}{\mathrm{df_t}}) \quad (2)$$

Keterangan;
$\mathrm{idf_t}$: inverse frekuensi dokumen
D : banyaknya dokumen
$\mathrm{df_t}$: jumlah dokumen yang mengandung term t

$$\mathrm{W_{t,d}} = \mathrm{tf_t} \times \mathrm{idf_t} \quad (3)$$

Keterangan;

W : bobot term (t) dalam dokumen (d)
$\mathrm{tf_t}$: jumlah kemunculan term t
$\mathrm{idf_t}$: invers frekuensi dokumen yang mengandung term t

2.5. Chi – Square

Penyeleksian dilakukan untuk menghilangkan fitur yang tidak relevan dalam proses klasifikasi [5]. Penerapan seleksi fitur yang cocok dapat meningkatkan hasil evaluasi yang didapatkan. Pada penelitian ini digunakan metode chi – square untuk meyeleksi fitur. Chi – Square menggunakan ilmu statistika untuk menguji independensi sebuah term pada kategorinya. Yang menjadi peristiwa dalam fitur seleksi ini adalah kemunculan fitur dan kemunculan kategori. Perhitungan chi – square ditunjukan dalam [6] (4)

$$\mathrm{x^2(t,c)} = \frac{\mathrm{N(AD-CB)^2}}{\mathrm{(A+C)(B+D)(A+B)(C+D)}} \quad (4)$$

Keterangan:

$\mathrm{x^2(t,c)}$: nilai chi – square term t disetiap kategori c
t : kata (fitur)
c : kategori
N : banyak dokumen latih
A : jumlah dokumen kategori c yang terdapat term t
B : jumlah dokumen di kategori bukan c yang terdapat term t
C : jumlah dokumen kategori c yang tidak terdapat term t
D : jumlah dokumen di kategori bukan c yang tidak terdapat term t

2.6. Support Vector Machine

Support Vector Machine adalah metode komputasi dalam melakukan prediksi baik permasalahan pengklasifikasian maupun regresi [7] SVM akan mencari hyperplane yang

optimal dengan margin maksimal untuk memisahkan kelas. Beberapa perhitungan SVM, decision function (5).

$$f(x) = \text{sign}\,(w \cdot x + b) \tag{5}$$

Dengan · merupakan sekalar sehingga (6)

$$w \cdot x = w^{T}x \tag{6}$$

Perhitungan margin terbesar dengan persamaan (7).

$$\frac{1}{\|\vec{w}\|} \tag{7}$$

quadratic problem untuk mencari titik minimal ditunjukan dalam persamaan (8) constrait atau kendala persamaannya (9).

$$\min_{\vec{w}} \tau(w) = \frac{1}{2}\|\vec{w}\| \tag{8}$$

$$yi(w \cdot x_i + b) \geq -1, \forall i \tag{9}$$

Persamaan (8)(9) dapat direduksi dengan menggunakan fungsi langrange, persamaan ditunjukan (10)

$$L(w, b) = \frac{1}{2}(w.w) - \sum_{i=1}^{m} a_i\,(y_i(w.x_i + b) - 1 \tag{10}$$

Dimana a_i merupakan langrange multipliers dan nilai $a_i \geq 0$.

Terdapat pula kernel, yaitu ruang berdimensi tinggi untuk SVM memetakan data.

2.7. Evaluasi

Evaluasi dilakukan dengan menggunakan dengan menghitung nilai akurasi(11), recall(12), precision(13), dan F1-Score(14). Kemudian membandingkan hasil yang diperoleh berdasarkan presentase seleksi fitur, pada kernel linear dan kernel gaussian atau Radial Basis Function (RBF).

$$\text{Akurasi} = \frac{TP+TN}{TP+FP+TN+FN} \times 100\% \tag{11}$$

$$\text{Precision} = \frac{TP}{TP+FP} \tag{12}$$

$$\text{Recall} = \frac{TP}{FP+FN} \tag{13}$$

$$F1 - \text{Score} = \frac{2(\text{Precision}+\text{Recall})}{(\text{Precision}+\text{Recall})} \tag{14}$$

3. Hasil dan Pembahasan

Studi ini dilakukan dengan tujuan mencari tahu pengaruh seleksi fitur terhadap kinerja pengkalsifikasian spam pada algoritma klasifikasi. Seleksi dengan metode Chi-Square dilakukan dengan menggunakan presentasi 20%, 40%, 60%, 80% nilai evaluasi yang dipakai adalah akurasi, presisi, nilai recall, dan F1 – Score dengan algoritma klasifikasi Support Vector Machine. Data dipetakan pada linear dan RBF.

3.1. Pengujian Dengan Menggunakan Seleksi Fitur

Tabel 7. Hasil Pengujian Kernel Linear Dengan Seleksi Fitur

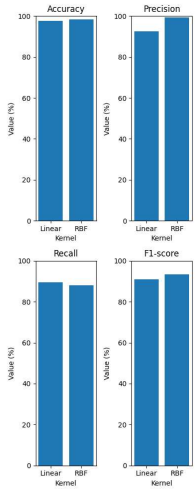
Presentase seleksi fitur	Kernel			
	Linear			
	Akurasi	**Precision**	**Recall**	**F1-Score**
20%	97.57%	94.28%	87.41%	90.72%
40%	97.57%	92.46%	89.40%	90.90%
60%	97.57%	92.46%	89.40%	90.90%
80%	97.75%	93.75%	89.40%	91.52%

Hasil yang diperoleh dari pengujian dengan menggunakan seleksi fitur dapat dilihat pada table diatas. Akurasi pada kernel linear diperoleh hasil konsisten dari seleksi fitur dengan presentase 20%, 40%, dan 60% yaitu akurasi sebesar 97.57%. Akurasi pada kernel linear meningkat menjadi 97.75% pada presentase seleksi fitur 80%. Kemudian nilai F1- Score mengalami peningkatan pada setiap presentase seleksi fitur, dengan nilai F1 – Score sebesar 91.52%.

Tabel 8. Hasil Pengujian Kernel RBF Dengan Seleksi Fitur

Presentase seleksi fitur	Kernel			
	RBF			
	Akurasi	**Precision**	**Recall**	**F1-Score**
20%	97.93%	99.23%	85.43%	91.81%
40%	98.29%	99.25%	88.07%	93%
60%	**98.82%**	97.81%	88.74%	**93.05%**
80%	98.02%	97%	87.41%	92.30%

Akurasi pada kernel RBF memperoleh Nilai hasil yang lebih baik dibandingkan dengan nilai yang diperoleh pada kernel linear. Akurasi terendah pada kernel RBF adalah 97.93% pada seleksi fitur dengan presentase 20%. Akurasi tertinggi pada kernel RBF adalah 98.82% pada seleksi fitur dengan presentase 60%. Kemudian nilai F1-Score tertinggi pada kernel RBF adalah 93.05% pada seleksi fitur dengan presentase 60%. Perbandingan hasil yang diperoleh dapat dilihat pada gambar 2.



Gambar 2. Rata Rata Hasil Pengujian Dengan Seleksi Fitur

Dari hasil pengujian dua kernel diatas, kernel RBF dengan presentase seleksi fitur sebesar 60% menghasilkan nilai akurasi yang lebih baik dibandingkan kernel Linear dalam penggunaan seleksi fitur.

3.2. Pengujian Tanpa Menggunakan Seleksi Fitur

Tabel 9. Hasil Pengujian Tanpa Menerapkan Seleksi Fitur

Kernel							
Linear				RBF			
Akurasi	**Precision**	**Recall**	**F1-Score**	**Akurasi**	**Precision**	**Recall**	**F1-Score**
97.57%	91.89%	90.41%	90.96%	98.02%	97%	87.41%	92.30%

Hasil yang diperoleh dari pengujian tanpa menggunakan seleksi fitur adalah nilai akurasi pada kernel linear sebesar 97.57% dengan F1 – Score sebesar 90.06%, nilai akurasi pada kernel RBF sebesar 98.02% dengan F1 – Score sebesar 92.30%. Dari hasil table diatas, kernel RBF menghasilkan nilai akurasi yang lebih baik dibandikan kernel Linear tanpa seleksi fitur.

Berdasarkan dua pengujian yang dilakukan dapat dilihat peningkatan performa dengan melakukan seleksi fitur pada presentase tertentu, dibandingkan tanpa menggunakan seleksi fitur. Hasil pada kernel RBF dengan semua presentase seleksi fitur yang digunakan, menghasilkan akurasi yang lebih baik jika dibandingkan dengan hasil yang diperoleh tanpa menerapkan seleksi fitur. Sedangkan pada kernel linear dengan menggunakan seleksi fitur 80% memperoleh nilai akurasi yang lebih baik dibandingkan hasil pengujian pada kernel yang sama, tanpa menggunakan seleksi fitur.

4. Kesimpulan

Dari hasil pengujian yang diterapkan dalam beberapa pengujian, kesimpulan yang diperoleh penggunaan seleksi fitur Chi – Square dalam pengklasifikasian teks spam dengan menggunakan metode Support Vector Machine dapat meningkatkan performa klasifikasi. Niali terbaik yang diperoleh adalah nilai akurasi 98.82% dengan F1 – Score 93.05% dengan presentase seleksi fitur 60%, pada kernel RBF. Hasil akurasi presentase lainnya pada kernel RBF juga menunjukan hasil akurasi yang lebih tinggi jika dibandingkan tanpa menerapkan seleksi fitur yaitu presentase 20%, 40%, dan 80% menghasilkan nilai akurasi berturut – turut 97.93%, 98.29%, dan 98.02%, lebih baik daripada hasil akurasi tanpa seleksi fitur, yaitu akurasi sebesar 97.57%.

Daftar Pustaka

[1] M. A. Ghani dan A. Subekti, “Email Spam Filtering Dengan Algoritma Random Forest,” IJCIT (Indonesian Journal on Computer and Information Technology, vol. 3, no. 2, hlm. 216–221, 2018.

[2] A. T. Syam dkk., “Klasifikasi Komentar Spam Pada Instagram Menggunakan Metode Support Vector Machine,” vol. 6, 2020, [Daring]. Tersedia pada: https://journal.uniku.ac.id/index.php/buffer

[3] M. Hajat Syafii, J. Margonda Raya No, dan J. Barat, “Klasifikasi Sms Spam Dengan Komparasi Metode Svm Dan Naïve Bayes,” Jurnal METHODIKA, doi: 10.1007/s00500.

[4] F. D. Ananda dan Y. Pristyanto, “Analisis Sentimen Pengguna Twitter Terhadap Layanan Internet Provider Menggunakan Algoritma Support Vector Machine,” MATRIK: Jurnal Manajemen, Teknik Informatika dan Rekayasa Komputer, vol. 20, no. 2, hlm. 407–416, Mei 2021, doi: 10.30812/matrik. v20i2.1130.

[5] A. Z. Amrullah, A. Sofyan Anas, M. Adrian, dan J. Hidayat, “Analisis Sentimen Movie Review Menggunakan Naive Bayes Classifier Dengan Seleksi Fitur Chi Square,” Jurnal, vol. 2, no. 1, 2020, doi: 10.30812/bite. v2i1.804.

[6] M. Imron Maulana dan A. Andy Soebroto, “Klasifikasi Tingkat Stres Berdasarkan Tweet pada Akun Twitter menggunakan Metode Improved k-Nearest Neighbor dan Seleksi Fitur Chi-square,” 2019. [Daring]. Tersedia pada: http://j-ptiik.ub.ac.id

[7] M. N. Muttaqin dan I. Kharisudin, “Analisis Sentimen Pada Ulasan Aplikasi Gojek Menggunakan Metode Support Vector Machine dan K Nearest Neighbor,” UNNES Journal of Mathematics, vol. 10, no. 2, hlm. 22–27, 2021, [Daring]. Tersedia pada: http://journal.unnes.ac.id/sju/index.php/ujm

Halaman ini sengaja dibiarkan kosong

JNATIA Volume 1, Nomor 4, Agustus 2023 p-ISSN: 2986-3929
Jurnal Nasional Teknologi Informasi dan Aplikasinya

Evaluasi Desain UI pada Prototype Aplikasi Influencer Marketing "Endorsfy" dengan Metode SEQ

Anak Agung Ngurah Mahadana Apta Gotra[a1], Anak Agung Istri Ngurah Eka Karyawati[a2]

[a]Program Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Udayana
Jalan Raya Kampus Udayana, Bukit Jimbaran, Kuta Selatan, Badung, Bali Indonesia
[1]ngurahmahadanaapta@student.unud.ac.id
[2]eka.karyawati@unud.ac.id

Abstract

The success of a digital product relies heavily on the inseparable relationship between User Interface (UI) and User Experience (UX). Although UI is a fundamental aspect of UX, it's imperative to acknowledge that the UI's design, functionality, and ease of use are pivotal factors that can make or break a digital product's success. A top-notch UI design can significantly enhance the UX, while a deficient and convoluted UI can lead to a less-than-stellar user experience. In this study, we undertake the task of evaluating the UI quality of "Endorsfy" – an Influencer Marketing Mobile Application. We employ the Single Ease Question (SEQ) method and the Maze tool to assess the UI quality and determine if it meets the standards for optimal usability. Additionally, this evaluation will consider the usability aspect of the application to ensure that it delivers a seamless user experience. Our ultimate objective is to guarantee that "Endorsfy" achieves the highest standards for UI quality and usability, providing users with unparalleled experience when using the application.

Keywords: *User Interface, Single Ease Question, Influencer Marketing, Mobile, Usability*

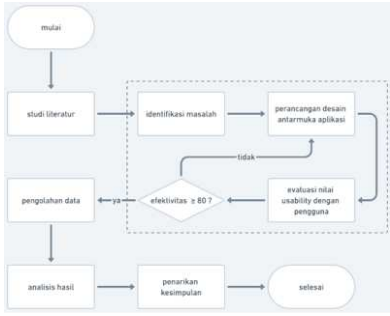
1. Pendahuluan

Usaha mikro, kecil dan menengah (UMKM) berkembang pesat pada sejumlah daerah di Indonesia. Berdasarkan data yang dirilis oleh Kementerian Koperasi dan Usaha Kecil dan Menengah (Kemenkop UKM) menunjukkan bahwa tingginya jumlah total usaha UMKM di Indonesia mencapai 8,71 juta pada tahun 2022 [1]. Meskipun sektor UMKM memiliki potensi yang besar untuk menggerakkan ekonomi dan menciptakan lapangan kerja, namun faktanya banyak UMKM yang sulit berkembang. Beberapa faktor yang menjadi kendala dalam pengembangan UMKM antara lain kurangnya akses pasar, keterbatasan modal, serta minimnya pengetahuan tentang teknologi [2]. Di sisi lain, kecanggihan teknologi membuat para pelanggan menginginkan kemudahan untuk berbelanja kapan saja dan di mana saja tanpa harus pergi dari rumah. Hal tersebut muncul akibat pengaruh dari pandemi yang telah melanda akhir-akhir ini. Menurut laporan "*Navigating Indonesia's E-Commerce: Omnichannel as the Future of Retail*", sebelum pandemi hanya 11% konsumen yang berbelanja secara eksklusif online, namun pada awal tahun 2021 persentase tersebut meningkat drastis menjadi 25,5% [3]. Peningkatan tren belanja online ini juga tidak lepas dari pengaruh *influencer* di media sosial. Sebuah studi menunjukkan sekitar 61% konsumen di Indonesia membeli produk yang direkomendasikan oleh *influencer* [4]. Melihat peningkatan tren belanja online yang dipengaruhi oleh *influencer* ini, tak heran jika *influencer marketing* menjadi salah satu strategi pemasaran yang efektif dan efisien dalam memasarkan produk UMKM melalui platform digital [5]. Namun, dengan munculnya ide tersebut, muncul pula tantangan baru, yaitu bagaimana merancang aplikasi apik yang mudah digunakan agar para pelaku UMKM bisa mendapatkan *influencer* yang sesuai dengan usaha mereka dan para *influencer* bisa mendapatkan UMKM yang sesuai dengan konten mereka. Inilah yang menjadi latar belakang dilakukannya sebuah penelitian untuk merancang sebuah antarmuka aplikasi *mobile* yang bernama "Endorsfy". Aplikasi "Endorsfy" akan memungkinkan pengguna untuk mencari *influencer* atau UMKM yang sesuai dengan kebutuhan mereka. Pengguna akan dimanjakan dengan fitur-fitur canggih seperti pencarian *influencer* berdasarkan kategori, lokasi,

dan jumlah *followers*, serta pencarian UMKM berdasarkan produk atau jenis usaha. Selain itu, pengguna juga dapat berinteraksi langsung dengan *influencer* atau UMKM melalui fitur *chat* atau *direct message*. Untuk memastikan bahwa desain aplikasi memenuhi standar kualitas UI dan UX, maka *usability testing* menjadi hal yang sangat penting untuk dilakukan. *Usability testing* diperlukan agar dapat mengamati langsung bagaimana pengguna berinteraksi dengan desain yang telah dirancang [6]. Adapun metode yang digunakan untuk melakukan *usability testing* dalam penelitian ini adalah SEQ (*Single Ease Question*). SEQ adalah sebuah metode kuesioner yang diberikan setelah responden melakukan sebuah tugas dalam suatu pengujian. SEQ terdiri dari satu pertanyaan dengan skala Likert 1-7 mulai dari sangat sulit hingga sangat mudah [6]. Metode SEQ ini juga cukup sering digunakan pada penelitian-penelitian sebelumnya, seperti pada penelitian "Analisis User Experience pada Website Waste4change Menggunakan Metode Single Ease Question" oleh Dinar, dkk [6], "Usability User Interfacedan User Experience Media Pembelajaran Kamus Kolok Bengkala Berbasis Android" oleh Joko Santoso [7], ataupun "Pengujian UI/UX dengan System Usability Scale dan Single Ease Question pada Aplikasi Pantau untuk Monitoring Perkembangan Penanaman Tanaman di Lahan Hijau" oleh Novita dkk [8]. Melalui penelitian ini, penulis berharap dapat memastikan bahwa antarmuka aplikasi "Endorsfy" memenuhi standar kualitas UI dan UX, sehingga dapat memberikan pengalaman pengguna yang optimal saat menggunakan aplikasi ini.

2. Metode Penelitian

Penelitian ini mengadopsi metode SEQ sebagai alat untuk mengukur tingkat kemudahan/kesulitan pengguna dalam menggunakan desain yang dirancang. Metode ini, dapat mengumpulkan data kuantitatif yang berharga dalam mengevaluasi usabilitas aplikasi. Proses penelitian ini melibatkan serangkaian tahapan yang harus dilakukan secara sistematis, yang dapat dilihat pada Gambar 1.



Gambar 1. Alur Penelitian

Tahap identifikasi masalah dilakukan dengan metode *User Journey Mapping*. *User Journey Mapping* dipilih karena memiliki kemampuan untuk menggali masalah lebih dalam yang terdapat pada suatu kegiatan atau prosedur yang melibatkan pengguna bahkan para pemangku kepentingan terkait [9]. Setelah melakukan identifikasi masalah dan memahami pengalaman pengguna, tahap selanjutnya adalah membuat desain antarmuka aplikasi. Tahap ini dimulai dengan membuat *wireframe*, yang nantinya *wireframe* tersebut akan digunakan sebagai landasan dalam membuat desain yang lebih rinci dan lengkap. Setelah *wireframe* maka dilanjutkan dengan membuat desain yang lebih rinci (h*igh fidelity*), tahap ini sudah memperlihatkan elemen-elemen desain seperti *layout*, *palette*, *icon*, dan *font*. Untuk membuat

desain-desain tersebut, penulis menggunakan *tool* desain Figma. Tahap terakhir dalam penelitian ini adalah pengujian usabilitas. Pengujian ini dilakukan untuk menentukan apakah desain aplikasi yang dibangun sudah memenuhi standar atau tidak. Untuk itu, penulis akan menggunakan metode SEQ atau *Single Ease Question* dengan jumlah responden sebanyak 20 orang. Jumlah ini paling efisien untuk pengujian usability [10]. Penulis juga akan menggunakan *tool* pengujian usabilitas bernama Maze untuk mempermudah pengujian. Dengan mengumpulkan data dari responden, maka akan dapatkan nilai efektivitas penggunaan aplikasi berdasarkan rumus berikut.

$$Effectiveness = \frac{Number\ of\ tasks\ completed\ successfully}{Total\ number\ of\ tasks\ undertaken} \times 100\% \quad (1)$$

3. Hasil dan Diskusi

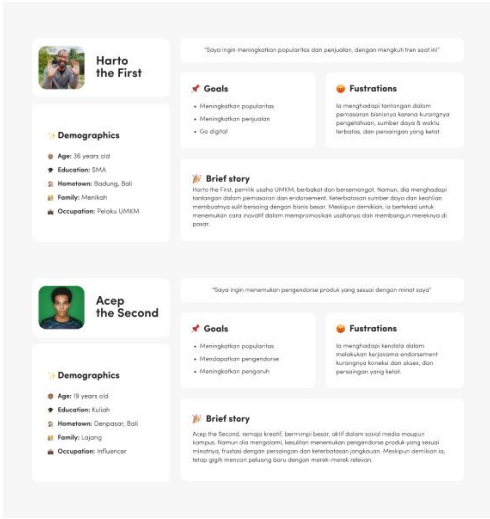
Hasil dan diskusi dari penelitian ini mencakup tahap perancangan desain hingga hasil evaluasi dengan metode SEQ. Selama pengujian, responden akan diberikan sebuah tugas untuk menjalankan *prototype* aplikasi yang telah dirancang dan memberikan penilaian berdasarkan skala Likert 1-7. Penilaian dari responden akan menjadi penentu apakah desain antarmuka aplikasi tersebut layak sesuai standar kualitas UI dan UX atau tidak.

3.1. User Journey Map

Gambar 2. User Journey Map

User journey map di atas menggambarkan perjalanan pengguna mulai dari sebelum mereka mengenal produk hingga setelah mereka menggunakannya. Melalui *user journey map*, kita dapat memperoleh informasi tentang tingkat kenyamanan pengguna, pengalaman pengguna secara detail, serta peluang untuk meningkatkan kualitas produk. Dengan memahami perjalanan pengguna secara menyeluruh, kita dapat meningkatkan kualitas produk sehingga pengguna merasa lebih puas dan terkesan dengan produk yang kita tawarkan.

3.2. User Persona

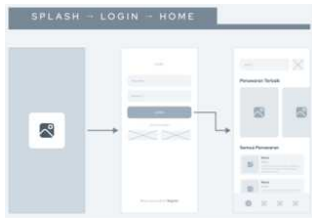


Gambar 3. User Persona

User persona di atas merupakan representasi fiksi dari pengguna ideal yang terbentuk dari hasil evaluasi *user journey map*. *User persona* juga menggambarkan karakteristik, kebutuhan, dan tujuan pengguna. Dengan menggunakan *user persona*, maka desain yang dihasilkan dapat menjawab keluhan pengguna.

3.3. Wireframe dan Wireflow

a. Alur dari Splash Screen Menuju Halaman Utama



Gambar 4. Login

Gambar 3 menampilkan rancangan desain dari *splash screen* yang menampilkan logo aplikasi. Kemudian, pengguna akan langsung diarahkan ke halaman *login/register*, di mana mereka dapat memasukkan informasi yang diperlukan untuk membuat akun baru atau masuk ke akun yang sudah ada. Setelah proses *login/register* selesai, pengguna akan diarahkan ke halaman utama yang menyajikan berbagai fitur dan konten menarik.

b. Alur dari Halaman Utama Menuju Mengajukan Kolaborasi dan Mengecek Statusnya

Gambar 5. Mengajukan Kolaborasi

Gambar 4 menampilkan rancangan desain dari halaman utama hingga halaman untuk mengajukan kolaborasi. Ketika pengguna membuka halaman utama, mereka akan disambut dengan beragam *influencer*. Pengguna dapat memilih *influencer* yang ingin diajak berkolaborasi dengan masuk ke halaman informasi detail *influencer* tersebut. Setelah mempertimbangkan dengan seksama, pengguna dapat mengirimkan pengajuan kolaborasi mereka kepada *influencer* terkait. Setelah pengajuan dikirim, pengguna dapat dengan mudah mengecek status pengajuannya atau kembali ke halaman utama untuk mengeksplorasi lebih banyak lagi.

c. Alur dari Status Pengajuan Menuju Pesan Diskusi dan Membayar Uang Muka

Gambar 6. Diskusi dan Membayar Uang Muka

Gambar 5 menampilkan rancangan desain dari halaman status pengajuan hingga halaman pembayaran uang muka. Setelah mengajukan kolaborasi, pengguna dapat melihat status pengajuannya di halaman status pengajuan. Di sana, mereka memiliki dua opsi: berdiskusi terlebih dahulu melalui fitur *chat* atau langsung melakukan pembayaran uang muka. Jika pengguna ingin berkomunikasi lebih lanjut sebelum memastikan kolaborasinya, mereka dapat menggunakan fitur *chat* yang tersedia. Namun, jika pengguna sudah siap untuk melanjutkan dengan pembayaran, mereka dapat langsung menuju halaman pembayaran uang muka. Setelah berhasil membayar, pengguna akan diarahkan kembali ke halaman utama.

d. Alur dari Halaman Utama Menuju Profile dan Logout

Gambar 7. Logout

Gambar 6 menampilkan rancangan desain dari halaman utama, lalu pengguna dapat mengakses halaman profil mereka untuk melihat informasi pribadi dan pengaturan akun. Di halaman profil, pengguna juga akan menemukan opsi untuk melakukan *logout*. Jika pengguna ingin *logout* dari akun mereka, mereka dapat mengklik tombol keluar untuk mengakhiri sesi mereka. Setelah *logout*, pengguna akan diarahkan kembali ke halaman *login*.

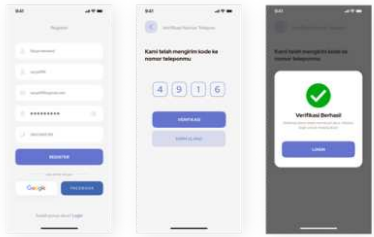
3.4. High Fidelity

a. Tampilan Splash Screen dan Onboarding Screen

Gambar 8. Splash dan Onboarding Screen

Gambar 7 menampilkan desain dari *splash screen* dan *onboarding screen*. Pada *splash screen*, logo aplikasi tampil dengan jelas, memberikan kesan profesional dan menarik perhatian pengguna. Kemudian, pengguna akan dialihkan ke *onboarding screen*, di mana mereka akan melihat kalimat-kalimat yang memberikan sedikit penjelasan tentang aplikasi "Endorsfy". Tujuan dari *onboarding screen* ini adalah untuk memberikan gambaran awal kepada pengguna tentang apa yang dapat mereka harapkan dari aplikasi ini. Dengan kombinasi antara logo yang mencolok dan kalimat-kalimat yang informatif, pengguna akan merasa tertarik dan termotivasi untuk melanjutkan penggunaan aplikasi "Endorsfy" dan mengeksplorasi lebih lanjut fitur-fitur yang ditawarkan.

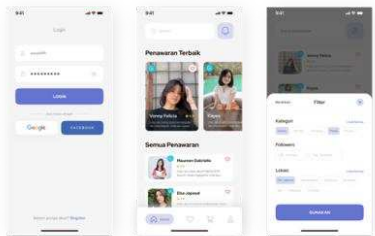
b. Tampilan Register



Gambar 9. Halaman Register

Gambar 8 menampilkan desain dari halaman *register* hingga selesai. Pada halaman *register*, pengguna diberikan dua opsi yang memudahkan mereka dalam mengakses aplikasi, yaitu dengan menggunakan *shortcut login* jika sudah memiliki akun atau membuat akun baru dengan mengisi formulir data diri. Pengguna dapat dengan mudah mengisi data diri yang diperlukan untuk proses registrasi. Setelah pengguna mengisi data diri, mereka akan menerima kode verifikasi yang harus dimasukkan untuk menyelesaikan proses registrasi dengan sukses. Hal ini memastikan keamanan dan keotentikan informasi pengguna.

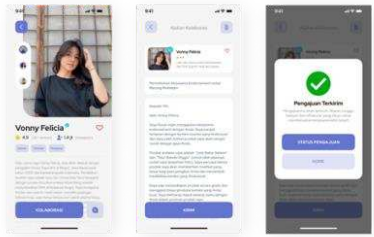
c. Tampilan Login, Halaman Utama, dan Filter



Gambar 10. Halaman Login, Home, dan Filter

Gambar 9 menampilkan desain dari halaman utama, *login*, dan filter. Pada halaman *login*, pengguna diberikan kemudahan untuk memasukkan *username* dan *password* mereka agar dapat masuk ke dalam aplikasi dengan cepat dan aman. Setelah login, pengguna akan langsung disambut dengan tampilan halaman utama yang menampilkan *influencer-influencer* ternama. Selain itu, pengguna juga memiliki kebebasan untuk mengatur kriteria pencarian mereka melalui fitur filter. Mereka dapat menyesuaikan preferensi seperti lokasi dan jumlah *followers* untuk mendapatkan hasil pencarian yang lebih spesifik dan sesuai dengan kebutuhan mereka.

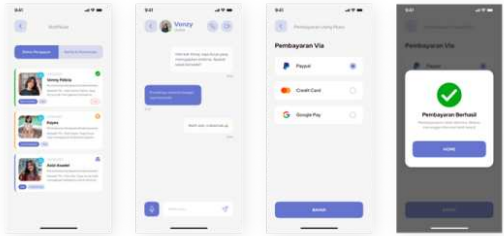
d. Tampilan Detail Influencer dan Pangajuan Kolaborasi



Gambar 11. Halaman Detail Influencer dan Pengajuan Kolaborasi

Gambar 10 menampilkan desain dari halaman detail *influencer* dan pengajuan kolaborasi. Pada halaman detail *influencer*, pengguna dapat menemukan informasi yang sangat berguna tentang *influencer* yang mereka pilih. Mereka dapat melihat kategori atau *niche* yang menjadi keahlian *influencer*, jumlah *followers'* yang dimiliki, serta rating atau penilaian dari pengguna lane. Di sisi lain, halaman pengajuan kolaborasi memberikan pengguna akses untuk mengirimkan pengajuan kolaborasi kepada *influencer* yang mereka pilih. Pengguna dapat mengisi pesan yang ingin disampaikan kepada *influencer* dan bahkan melampirkan file-file tertentu yang diperlukan dalam kolaborasi tersebut. Dengan fitur ini, pengguna dapat berkomunikasi secara langsung dengan *influencer* dan menyampaikan gagasan serta kebutuhan kolaborasi mereka secara detail.

e. Tampilan Status Pengajuan, Chat, dan Pembayaran



Gambar 12. Halaman Status Pengajuan, Chat, dan Pembayaran

Gambar 11 menampilkan desain dari halaman status pengajuan, *chat*, dan pembayaran. Pada halaman status pengajuan, pengguna dapat dengan mudah melihat perkembangan dari pengajuan mereka. Mereka dapat melihat apakah pengajuan telah diterima, ditolak, atau sedang berada dalam proses kolaborasi. Hal ini memberikan transparansi dan memudahkan pengguna untuk memantau status kolaborasi mereka dengan *influencer*. Selanjutnya, halaman *chat* memungkinkan pengguna untuk berkomunikasi langsung

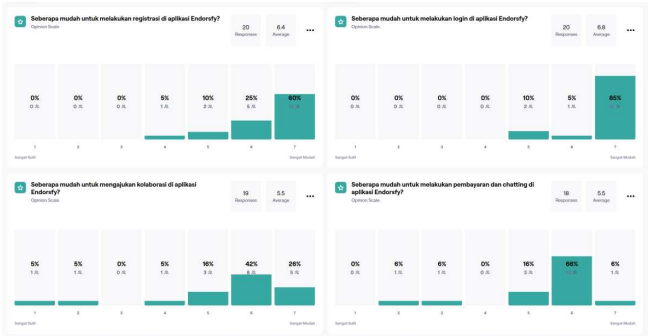
dengan *influencer* yang akan atau sedang berkolaborasi dengan mereka. Melalui fitur ini, pengguna dapat membahas dan menyelesaikan detail pekerjaan secara efektif. Komunikasi yang lancar dan terintegrasi dalam aplikasi membantu pengguna dan *influencer* untuk saling berinteraksi dengan lebih baik. Terakhir, halaman pembayaran menampilkan pilihan metode pembayaran yang tersedia untuk pengguna. Dengan tampilan yang jelas, pengguna dapat memilih metode pembayaran yang paling sesuai bagi mereka. Tombol yang disediakan memudahkan pengguna untuk memproses pembayaran dengan cepat dan mudah.

f. Tampilan Profile dan Logout

Gambar 13. Halaman Profile

Gambar 12 menampilkan desain dari halaman *profile*. Pada halaman profile, pengguna dapat mengakses informasi akun mereka dan melakukan pengaturan yang diperlukan. Dengan mengklik tombol keluar, pengguna akan *logout* dari akun mereka dan diarahkan kembali ke halaman *login*. Fitur *logout* ini membantu menjaga privasi dan keamanan pengguna, serta memungkinkan mereka untuk kembali masuk ke akun mereka jika diperlukan.

3.5. Hasil Pengujian



Gambar 14. Hasil Pengujian

Gambar 13 membuktikan keberhasilan *usability testing* aplikasi "Endorsfy" menggunakan metode *Single Ease Question* atau SEQ. Dapat dilihat bahwa sebanyak 18 dari 20 responden berhasil menyelesaikan semua *task* dengan rata-rata nilai yang cukup tinggi di setiap *task*-nya. Skor efektivitas yang diperoleh dari perhitungan rumus mencapai 96% ($\frac{77}{80} \times 100\%$), menunjukkan keberhasilan desain antarmuka yang luar biasa. Dengan kata lain, aplikasi "Endorsfy" berhasil memberikan pengalaman pengguna yang mudah dan nyaman ketika digunakan.

4. Kesimpulan

Hasil pengujian usabilitas aplikasi "Endorsfy" menggunakan metode *Single Ease Question* (SEQ) menghasilkan hasil yang mengagumkan. Dari 20 responden yang menguji desain tampilan aplikasi ini, hanya dua orang yang tidak berhasil menyelesaikan seluruh *task*. Rata-rata nilai pada setiap *task*-nya juga menunjukkan kepuasan yang tinggi, yaitu 6,4 pada *task* pertama, 6,8 pada *task* kedua, dan 5,5 pada *task* ketiga dan keempat. Selain itu, skor efektivitas yang dihasilkan juga fantastis, yaitu sebesar 96%, membuktikan bahwa desain tampilan aplikasi "Endorsfy" sangat efektif dan mudah digunakan oleh pengguna. Dengan hasil pengujian yang memuaskan seperti ini, aplikasi "Endorsfy" memiliki potensi besar untuk menjadi aplikasi yang sukses dan bernilai di masa depan.

Daftar Pustaka

[1] E. Santika, “Jumlah UMKM di Indonesia Sepanjang 2022, Provinsi Mana Terbanyak?” *databoks.katadata.co.id*, Feb. 02, 2023. https://databoks.katadata.co.id/datapublish/2023/02/02/jumlah-umkm-di-indonesia-sepanjang-2022-provinsi-mana-terbanyak (accessed Apr. 30, 2023).

[2] A. Afriyadi, “Bisnis UMKM Sulit Tumbuh, Ini Masalahnya,” *finance.detik.com*, Nov. 02, 2021. https://finance.detik.com/berita-ekonomi-bisnis/d-5793600/bisnis-umkm-sulit-tumbuh-ini-masalahnya (accessed Apr. 30, 2023).

[3] F. Ulya, “Riset: Masyarakat Lebih Banyak Belanja Online Dibanding Offline,” *kompas.com*, Oct. 22, 2021. https://money.kompas.com/read/2021/10/22/211000926/riset--masyarakat-lebih-banyak-belanja-online-dibanding-offline (accessed Apr. 30, 2023).

[4] “80% of social media users in Asia who follow influencers are likely to purchase products recommended by the influencers,” *nielsen.com*, Sep. 14, 2022.

https://www.nielsen.com/news-center/2022/80-of-social-media-users-in-asia-who-follow-influencers-are-likely-to-purchase-products-recommended-by-the-influencers/ (accessed Apr. 30, 2023).

[5] A. Bhaskoro, "Mengenal Metode Pemasaran 'Influencer Marketing,'" *dailysocial.id*, Jan. 04, 2022. https://dailysocial.id/post/mengenal-metode-pemasaran-influencer-marketing (accessed Apr. 30, 2023).

[6] D. Hariri, H. Hannie, and I. Purnamasari, "Analisis User Experience pada Website Waste4change Menggunakan Metode Single Ease Question," *JIWP*, vol. 8, no. 13, pp. 95–108, 2022.

[7] J. Santoso, "Usability User Interface dan User Experience Media Pembelajaran Kamus Kolok Bengkala Berbasis Android," *JSI*, vol. 12, no. 2, pp. 174–181, 2018.

[8] N. Ningrum, I. Mulyono, and Z. Umami, "Pengujian UI/UX dengan System Usability Scale dan Single Ease Question pada Aplikasi Pantau untuk Monitoring Perkembangan Penanaman Tanaman di Lahan Hijau," *Sens 7*, vol. 7, no. 1, 2022.

[9] A. Nurfitri, I. Aknuranda, and H. Az-Zahra, "Pemetaan User Journey untuk Sistem Informasi Praktik Kerja Lapangan Fakultas Ilmu Komputer Universitas Brawijaya," *J-PTIIK*, vol. 3, no. 8, pp. 7542–7548, 2019.

[10] J. Nielsen, "How Many Test Users in a Usability Study?" *nngroup.com*, Jun. 03, 2012. https://www.nngroup.com/articles/how-many-test-users/ (accessed Apr. 30, 2023).

Halaman ini sengaja dibiarkan kosong

JNATIA Volume 1, Nomor 4, Agustus 2023 p-ISSN: 2986-3929
Jurnal Nasional Teknologi Informasi dan Aplikasinya

Perancangan UI/UX Aplikasi Tanda Bahaya untuk Perlindungan Anak "SafeKid" Berbasis Mobile

Ratri Desy Christirahma[a1], Anak Agung Istri Ngurah Eka Karyawati[a2]

[a]Program Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam,
Universitas Udayana
Jalan Raya Kampus Udayana, Bukit Jimbaran, Kuta Selatan, Badung, Bali Indonesia
[1]ratridesy161202@gmail.com
[2]eka.karyawati@unud.ac.id

Abstract

During the last 36 years, the National Center for Missing and Exploited Children has received more than 5 million calls about cases of missing children. In 2019 alone, the NCMEC assisted law enforcement and families with more than 29,000 cases of missing children. In Asia itself, it is estimated that there are around 20,000 to 200,000 missing children. More precisely in China and Southeast Asia. To help reduce the rate of loss of children in Indonesia, the authors created a design solution using the prototyping and UML methods. By using this prototype, users will be able to track the child's location and provide assistance in the form of a help button if something happens to the child. With this solution, it is hoped that it can help children who are in an emergency and parents who have lost their children so as to suppress the growth of cases of missing children.

Keywords: *Mobile, User Interface, User Experience, Child Protection, Location Tracking*

1. Pendahuluan

Perkembangan kasus penculikan anak semakin hari semakin bertambah. Hal ini sudah menjasi isu yang krusian dan memerlukan perhatian serius. Di Indonesia sendiri sudah terjadi 28 kasus penculikan anak sepanjang 2022. Angka tersebut terbilang meningkat dari tahun 2021 yang masih 15 kasus penculikan anak. Data tersebut diambil dari laporan Komisi Perlindungan Anak Indonesia. [1] Kasus penculikan anak bisa terjadi dikarenakan oleh banyak motif dan metode yang digunakan oleh penculik. Beberapa motif penculikan yang sering digunakan yaitu untuk mecari keuntungan finansial, eksploitasi seksual, adopsi illegal, pengambilalihan anak, Pendidikan, hinga perdagangan anak. Metode yang digunakannya pun juga cukup beragam. Dari memberikannya bantuan, hingga membawanya secara paksa akan dilakukan penculik untuk melancarkan aksinya. [2]. Oleh karena itu orang tua sekarang harus lebih waspada untuk menjaga anaknya. Di lain hal, saat ini teknologi juga semakin berkembang pesat. Dari semua kalangan pasti menggunakan teknologi. Mulai dari anak anak hingga orang tua pasti sudah tidak asing lagi dengan penggunaan teknologi di kehidupan sehari – harinya. Cara untuk menangani kasus anak hilang juga bisa dengan memanfaatkan sebuah teknologi. Salah satu contohnya yaitu dengan sebuah aplikasi yang dapat melacak anak dan anak dapat mengirimkan sinyal bantuan jika berada dalam keadaan darurat. Aplikasi ini akan memberikan sarana bagi orang tua untuk meningkatkan pengawasan dan keamanan anak. Dengan adanya permasalahan tersebut, terbentuklah ide yang melatarbelakangi dilakukannya penelitian ini dengan tujuan untuk mendesain sebuah aplikasi berbasis mobile bernama "SafeKid" yang diharapkan dapat membantu para orang tua dalam melindungi anaknya agar tidak terjadi hal – hal yang tidak dinginkan. Aplikasi "SafeKid" adalah semacam aplikasi SOS yang digunakan untuk mengirim sinyal bahaya kepada orang yang terhubung. Tidak hanya itu, orang yang terhubung juga dapat melihat lokasi dimana si pengirim sinyal berada. Jika dirasa berada di keadaan darurat, pengguna juga dapat langsung menghubungi pihak kepolisian atau 911. Dikarenakan aplikasi ini ditujukan untuk anak anak, maka aplikasi harus dibuat dengan desain yang sederhana tetapi juga mempermudah bagi pengguna. Melalui penelitian ini diharapkan dapat menghasilkan rancangan

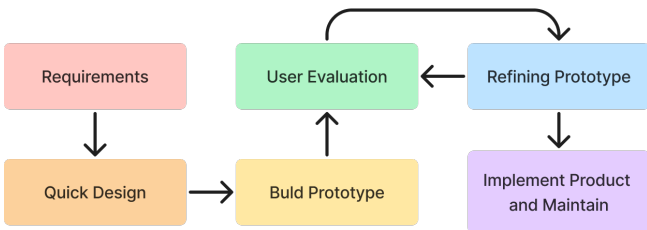
design antarmuka aplikasi “SafeKid” yang sesuai dengan kebutuhan orang tua dan dapat memudahkan anak dari segi UI maupun UX.

2. Metode Penelitian

Pada penelitian ini digunakan metode penelitian Prototype. Metode ini digunakan karena dapat membantu dalam perancangan aplikasi “SafeKid” berbasis mobile.

2.1 Metode Prototype

Metode prototype adalah sebuah teknik dalam melakukan pengembangan sistem perangkat lunak dengan menggunakan prototype untuk menggambarkan sistem agar pemilik sistem mempunyai gambaran jelas mengenai sistem yang akan dibangun oleh tim pengembang. Berikut adalah tahapan dari metode prototype.



Gambar 1. Model Prototype

a. Requirements

Pada tahap ini adalah tahapan pertama prototype dimulai. Tahapan awal yang dilakukan yaitu melakukan pengumpulan data dan analisis terhadap kebutuhan sistem. Data yang dikumpulkan berdasarkan dari kebutuhan pengguna yang akan dikembangkan dan diidentifikasi untuk dipahami lebih dalam. Pada tahap ini akan akan dilakukan diskusi antara klien dan tim developer untuk mendiskusikan detail sistem seperti apa yang diinginkan oleh user.

b. Quick Design

Tahap selanjutnya yaitu quick design. Pada tahap ini akan dilakukan pembuatan desain sederhana sebuah aplikasi yang akan dikembangkan dengan tujuan untuk memberikan gambaran singkat tentang sistem yang akan dibuat berdasalkan hasil diskusi dari tahapan pertama. Karena pada tahapan ini hanya sebuah desain sederhana, maka desain yang dihasilkan hanya inti dari aplikasi saja dan tidak mencakup keseluruhan sistem.

c. Build Prototype

Tahap ketiga yaitu build prototype. Pada tahap ini dilakukan pembuatan prototype berdasarkan hasil dari quick design sebelumnya. Di tahap ini lebih berfokus pada

mengimplementasikan fitur fitur yang akan menggambarkan konsep dan fungsionalitas utama dari aplikasi tersebut.

d. User Evaluation

Pada tahap user evolution akan dilakukan evaluasi terhadap pengguna awal. Di tahap ini, prototype yang telah dibuat berdasarkan kebutuhan pengguna akan dipresentasikan di depan klien supaya segera dilakukan evaluasi dan penilaian. Setelah dilakukan presentasi prototype tersebut, klien dapat memberi evaluasi berupa saran dan komentar terkait aplikasi yang sudah dibuat.

e. Refining Prototype

Setelah dilakukan tahap user evaluation, selanjutnya yaitu refining prototype atau memperbaiki prototype dimana prototype sebelumnya yang sudah di evaluasi akan diperbaiki berdasarkan hasil evaluasi sebelumnya. Pada tahap ini, pengembang akan terus melakukan perbaikan pada prototype hingga klien menyetujui prototype yang dibuat dan akhirnya dapat lanjut ke tahap selanjutnya.

f. Implement Product and Maintain

Pada tahap terakhir yaitu implementasi produk dan pemeliharaan akan dilakukan. Produk sistem tersebut akan dikembangkan lagi oleh developer dari prototype hingga menjadi sebuah aplikasi yang dapat berjalan dan digunakan semsetinya. Setelah produk jadi, produk akan dilakukan pemeliharaan agar produk tetap berjalan dengan baik tanpa ada kendala apapun.

3. Hasil dan Diskusi

3.1 Analisis Kebutuhan

Aplikasi ini dikembangkan dengan tujuan untuk membantu para orang tua dalam mengawasi anaknya dan membantu anak – anak yang sedang berada di keadaan darurat.

a. Kebutuhan Fungsional

1. Aplikasi menampilkan Splash Screen.
2. Aplikasi dapat melakukan login dan registrasi.
3. Aplikasi dapat mengirimkan sinyal darurat.
4. Aplikasi bisa melacak lokasi pengirim sinyal.
5. Aplikasi bisa meminta pertolongan kepada pihak berwajib saat ada sinyal darurat.
6. Aplikasi bisa memberikan arah ke titik lokasi sinyal diberikan.
7. Aplikasi bisa memberikan notifikasi jika terdapat sinyal darurat yang masuk.

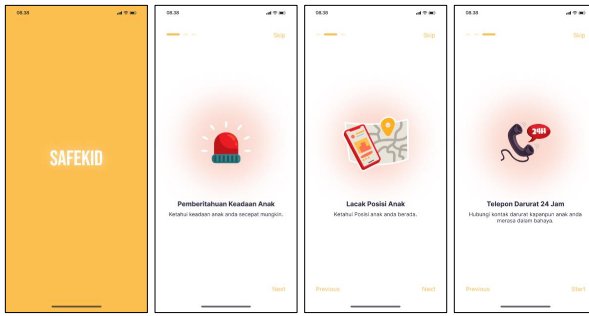
b. Kebutuhan Non Fungsional

1. Perangkat lunak yang digunakan dalam perancangan aplikasi ini adalah MacOS Monterey, Figma, Adobe Illustrator, Visual Studio Code.
2. Perangkat keras yang digunakan dalam perancangan aplikasi ini adalah Macbook Pro intel core i5.

3.2 Implementasi Sistem

a. Tampilan Splash Screen dan Onboarding Screen

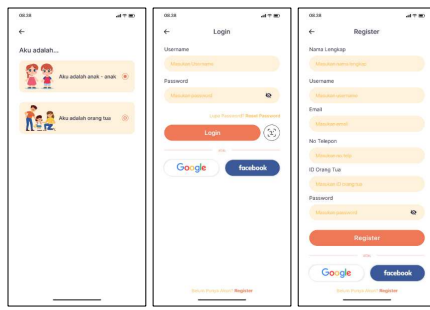
Gambar 2 adalah desain tampilan antarmuka splash screen dan onboarding screen dimana splash screen hanya berisi logo dari aplikasi “SafeKid”. Setelah splash screen muncul, selanjutnya halaman onboarding screen akan muncul berupa perkenalan singkat dari aplikasi “SafeKid”.



Gambar 2. Splash dan Onboarding Screen

b. Tampilan Login Page dan Register Page

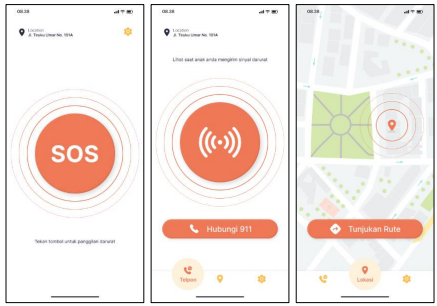
Gambar 3 adalah tampilan antarmuka dari login dan register. Tetapi sebelum ke halaman login, aplikasi akan memunculkan halaman dimana user akan ditanya apakah dia anak anak atau orang tua karena fitur yang didapatkan nantinya akan berbeda. Setelah user memilih, user akan login seperti biasa atau register terlebih dahulu jika belum mempunyai akun.



Gambar 3. Login dan Register

c. Tampilan Home Page

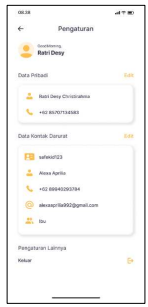
Gambar 4 adalah desain tampilan antarmuka homepage dari aplikasi “SafeKid”. Jika user login sebagai anak anak, maka aplikasi hanya akan memunculkan tampilan berupa button SOS. Hal ini dilakukan agar mempermudah anak anak sebagai penggunanya untuk menggunakan aplikasi “SafeKid”. Jika user login sebagai orang tua, aplikasi akan memunculkan halaman homepage yang berisi sinyal. Sinyal tersebut akan memberikan notifikasi apakah anak mereka sedang dalam bahaya atau tidak. Jika di keadaan darurat, user bisa langsung menghubungi 911 dengan menekan button yang tersedia. Tidak hanya itu, user juga bisa melacak lokasi anak dibagian page lokasi. Di halaman tersebut user bisa langsung mengeluarkan rute untuk menuju kesana melalui Google Maps.



Gambar 4. Homepage

d. Tampilan Pengaturan

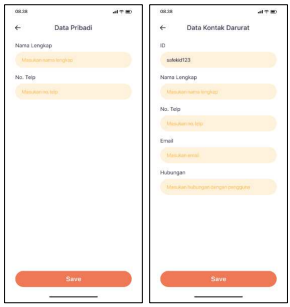
Gambar 5 adalah desain tampilan antarmuka dari halaman pengaturan. Di halaman ini menampilkan informasi data pribadi dan kontak darurat. User dapat mengedit data tersebut dan nantinya sistem akan mengarahkan ke halaman edit.



Gambar 5. Pengaturan

e. Tampilan Edit Pengaturan

Gambar 6 adalah tampilan dari desain antarmuka halaman edit pengaturan. Pada halaman ini user dapat mengedit informasi dari data pribadi dan data kontak darurat.



Gambar 6. Edit Pengaturan

4. Kesimpulan

Berdasarkan dari hasil penelitian yang telah dilakukan, dapat diambil kesimpulan bahwa aplikasi tanda bahaya untuk perlindungan anak adalah sebuah aplikasi yang dirancang dengan berbasis mobile dan menggunakan metode prototype. Tujuan dari dibuatnya aplikasi ini yaitu guna membantu para orang tua dalam melindungi anaknya dan tentu saja mengurangi bertambahnya kasus penculikan anak. Aplikasi ini dibuat dengan sangat sederhana karena mengingat pengguna dari aplikasi ini yaitu anak anak. Hal ini dilakukan untuk menghindari kebingungan pada anak anak dalam menggunakan aplikasi ini. Dengan adanya fitur pengiriman sinyal darurat dan pelacakan lokasi, diharapkan dapat membawa dampak baik dengan berkurangnya kasus penculikan anak dan berkurangnya rasa khawatir orang tua kepada anaknya saat mereke bermain diluar.

Daftar Pustaka

[1] CNN Indonesia, “Kemen PPPA: 28 Anak Jadi Korban Penculikan Sepanjang 2022”, Jan. 05, 2023. https://www.cnnindonesia.com/nasional/20230104152142-12-896122/kemen-pppa-28-anak-jadi-korban-penculikan-sepanjang-2022 (accessed Mei. 23, 2023).
[2] Ayu Isti, “Cara Mencegah Penculikan Anak, Orang Tua Wajib Tahu”, Feb. 01, 2023. https://www.merdeka.com/jateng/cara-mencegah-penculikan-anak-orang-tua-wajib-tahu-kln.html (accessed Mei. 23, 2023)
[6] Terry Cralle, “Missing Children Statistics: How Many Children Are Missing?”, 2023. Available: *https://www.terrycralle.com/missing-children-statistics/.*

JNATIA Volume 1, Nomor 4, Agustus 2023 p-ISSN: 2986-3929
Jurnal Nasional Teknologi Informasi dan Aplikasinya

Implementasi Algoritma A* (Star) dengan Graf untuk Menentukan Rute Terpendek Distributor Kopi

I Putu Andi Wiratama Putra[a1], I Gede Arta Wibawa[a2]

[a]Program Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam,
Universitas Udayana
Jalan Raya Kampus Udayana, Bukit Jimbaran, Kuta Selatan, Badung, Bali Indonesia
[1]andiwp2003@gmail.com
[2]gedearta83@gmail.com

Abstract

In this research, the A Star algorithm is employed to find the most efficient route for goods distribution. Distributors encounter challenges in ensuring timely deliveries due to the presence of multiple destinations spread across different regions. Congestion further adds to the complexity of determining the shortest path. The A Star algorithm utilizes the distance-plus-cost function to prioritize the order of visiting points. This study utilizes primary data, comprising five shop locations in the Tabanan city, as nodes and incorporates the distances between the shops. The implemented program utilizes the A Star algorithm to compute the shortest route and present the path along with its corresponding distance. The objective of this research is to attain the shortest route and calculate the distance covered for the coffee distributor.

Keywords: *A Star, Shortest Route, Graph*

1. Pendahuluan

Dalam kehidupan sehari-hari, ketika berpindah dari satu tempat ke tempat lain, yang menjadi perhatian utama adalah efisiensi waktu dan biaya. Untuk itu, kita memerlukan informasi untuk menentukan rute terpendek antar tempat yang ingin kita jangkau. Algoritma yang digunakan untuk mencari jalur terpendek disebut juga sebagai algoritma jalur terpendek, digunakan untuk menentukan jalur dalam suatu graf [1]. Graf adalah gabungan dari simpul (vertices) dan edge (tepi), dimana setiap edge dihubungkan dengan satu atau dua simpul. Penerapan graf dan keterhubungannya dalam kehidupan nyata sering kita jumpai dalam berbagai bidang seperti komunikasi informasi, jaringan komputer, lalu lintas, jaringan listrik dan saluran air [2]. Distributor adalah badan usaha yang bertanggung jawab untuk mendistribusikan produk dari satu lokasi ke lokasi lain. Peran distributor sangat penting untuk menjamin ketersediaan dan pendistribusian produk yang dibutuhkan masyarakat. Sebagai perantara dalam proses pemasaran, distributor bertanggung jawab atas pengangkutan dan penyimpanan barang dan jasa dari produsen ke konsumen. Dalam peran pemasaran, distributor harus melakukan perjalanan untuk mengirimkan barang ke lokasi yang berbeda. Namun, distributor sering menghadapi tantangan dalam proses pengiriman, seperti waktu tempuh yang lama dan kesulitan dalam mengurutkan kunjungan secara efisien. Faktor tersebut seringkali dikarenakan luasnya wilayah yang harus mereka layani dan seringnya masalah kemacetan saat perjalanan. Algoritma A Star adalah metode yang diperkaya dengan penggunaan heuristik. Algoritma ini menghilangkan langkah-langkah pucat karena langkah-langkah ini tidak mengarah pada solusi yang diinginkan. Algoritma A Star yang digunakan untuk mencari rute terdekat mencakup perhitungan heuristik. Setelah nilai heuristik didapatkan, langkah selanjutnya adalah mencari nilai f(n). Nilai F(n) kemudian dimasukkan ke dalam open list. Kemudian bobot antar node diperhitungkan dan node dengan bobot terendah dipilih sebagai jalur [3]. Algoritme A Star menghitung dan menyimpan semua kemungkinan untuk membandingkan setiap jalur yang dipilih dengan jalur tersimpan lainnya. Hal ini mengarah pada solusi optimal dalam pencarian jarak terpendek. Tujuan utama dari penelitian ini adalah untuk menemukan saluran distribusi barang yang optimal dengan menggunakan algoritma A Star. Penelitian ini berfokus pada pendistribusian barang di jalan khususnya pada beberapa toko yang berada di Kabupaten Tabanan. Pada penelitian ini terdapat 5 toko yang melalui titik saluran distribusi. Pencarian rute terpendek diawali dengan informasi titik awal dan tujuan perjalanan distributor. Lokasi awal dan tujuan yang diberikan digunakan untuk menemukan rute terpendek.

2. Metode Penelitian

2.1. Data Penelitian

Data yang digunakan dalam penelitian ini adalah data primer. Data tersebut mencakup informasi 5 toko yang teridentifikasi di wilayah perkotaan Tabanan. Setiap toko direpresentasikan pada grafik sebagai titik dengan lokasi sebagai simpul dan jarak antara toko sebagai tepi yang menghubungkannya. Jarak antar toko yang dimasukkan ke dalam sistem dan jumlah node bervariasi tergantung tujuan pengiriman yang ditentukan. Evaluasi hasil rute didasarkan pada perhitungan jarak, tanpa memperhitungkan faktor kemacetan atau kondisi geografis rute. Penentuan posisi menggunakan Google Earth sebagai panduan dan menghitung jarak antar titik dengan melihat bujur dan lintang lokasi yang dipilih.

2.2. Graf

Menggunakan grafik untuk memvisualisasikan objek diskrit dan hubungan di antara mereka adalah konsep umum. Dalam representasi visual, setiap objek direpresentasikan sebagai titik pada grafik, sedangkan hubungan antar objek direpresentasikan sebagai garis yang menghubungkan titik-titik tersebut. Grafik dapat digunakan untuk merepresentasikan berbagai informasi, salah satunya adalah peta [5].

2.3. Tahapan Algoritma A Star

Algoritma A Star merupakan metode pencarian jalur terpendek yang menggunakan fungsi heuristik untuk menentukan urutan kunjungan node. Fungsi heuristik ini memberikan estimasi atau skor pada setiap node yang membantu algoritma A Star untuk mencapai solusi yang diinginkan [3]. Menemukan jarak terpendek pada peta, peta harus disajikan sebagai diagram. Setiap simpul grafik mewakili persimpangan atau lokasi khusus di peta, sedangkan tepi grafik mewakili jalur yang dilalui. Tepi grafik memiliki bobot yang mencerminkan panjang jarak antar simpul. Algoritma Star mengevaluasi setiap node dengan menghubungkan g(n), yang merupakan biaya untuk mencapai node tersebut, dan h(n), yang merupakan perkiraan biaya untuk mencapai tujuan node tersebut. Dalam notasi matematika, ini dapat dinyatakan sebagai:

Rumus:

$$f(n) = g(n) + h(n) \quad (1)$$

Keterangan:

a. Nilai **f(n)** adalah hasil penjumlahan **g(n)** dan **h(n)**. Nilai **f(n)** ini merupakan estimasi awal dari jalur terpendek. Sebenarnya, **f(n)** adalah jalur terpendek yang sebenarnya, tetapi tidak sepenuhnya dieksplorasi sampai algoritma A* dipecahkan.
b. Nilai **g(n)** adalah jarak total yang ditempuh dari titik awal ke titik saat ini (titik tujuan). Itu mencerminkan rintangan atau jarak yang ditempuh.
c. Nilai **h(n)** adalah perkiraan jarak dari titik saat ini (yang dikunjungi) ke titik tujuan. Fungsi heuristik digunakan untuk memperkirakan berapa lama waktu yang dibutuhkan untuk mencapai tujuan.

Beberapa istilah dari algoritma A Star [4] dapat dijelaskan sebagai berikut:

a. "Titik awal" mengacu pada posisi awal dalam algoritme.
b. "Node" adalah titik yang mewakili target dan bisa berbentuk persegi, segitiga atau lingkaran.
c. "A" adalah simpul yang digunakan untuk menemukan jalur terpendek.
d. "Open List" adalah sekumpulan node yang dapat diakses dari titik asal atau node yang sedang dievaluasi.
e. "Closed List" adalah kumpulan node yang dipilih sebagai bagian dari jalur terpendek.

f. “Harga” (F) adalah nilai yang dihitung dengan menjumlahkan nilai G yang merupakan penjumlahan dari nilai simpul-simpul jalur terpendek dari titik asal ke simpul A, dan nilai H yaitu estimasi jarak dari node tersebut ke tujuan.

Prinsip dari algoritma ini adalah mencari jalur terpendek dari titik awal (starting point) ke node tujuan, dengan biaya minimum (F).

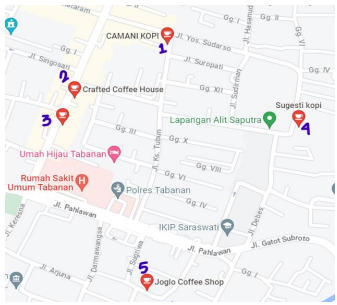
Langkah-langkah algoritma A Star [4] adalah sebagai berikut:

a. Langkah pertama dimulai dengan simpul A sebagai titik awal.
b. Selanjutnya, semua node yang bertetangga dengan A dan tidak memiliki atribut barrier ditambahkan ke open list.
c. Kami kemudian mencari simpul dengan nilai H terkecil di antara simpul dalam daftar terbuka.
d. Setelah itu, A dipindahkan ke node dengan nilai H terkecil. Simpul sebelum A disimpan sebagai induk A dan ditambahkan ke daftar tertutup. Jika ada node lain yang bertetangga dengan A (yang telah dipindahkan) tetapi belum masuk ke dalam open list, node tersebut dimasukkan ke dalam open list.
e. Selanjutnya nilai G yang ada dibandingkan dengan nilai G sebelumnya. Jika nilai G yang baru lebih kecil, A kembali ke posisi semula. Node yang diuji termasuk dalam daftar tertutup. Proses ini diulangi hingga kami menemukan solusi atau tidak ada lagi node dalam daftar terbuka.

3. Hasil dan Diskusi

3.1. Tahapan Penentuan Rute Terpendek

Pada penelitian ini penentuan rute terpendek dilakukan dengan mencari lokasi pada peta menggunakan informasi koordinat geografis (lintang dan bujur). Implementasi algoritma A Star untuk menentukan rute terpendek menuju dealer perusahaan dilakukan dengan menggunakan bahasa pemrograman Python. Tabel 1 di bawah menunjukkan tujuan pengiriman barang-barang yang digunakan dalam penelitian ini.



Gambar 1. Google Map Titik Awal dan Titik Tujuan

Tabel 1. Koordinat Latitude dan Longitude Titik Tujuan

No	Titik Tujuan	Latitude	Longitude
1	Camani Kopi	-8.535638644370819	115.13396379698509
2	Crafted Coffee House	-8.536900448812	115.13244771366291
3	Tan Panama	-8.537571962922883	115.13177912233395
4	Sugesti Kopi	-8.537561631945506	115.13683534425911
5	Joglo Coffee Shop	-8.541133082729948	115.13367888799112

Berdasarkan titik koordinat bujur (Longitude) dan lintang (Latitude) yang didapatkan pada google maps, kemudian akan dilakukan perhitungan jarak antara satu titik ke titik lainnya dengan menggunakan persamaan berikut:

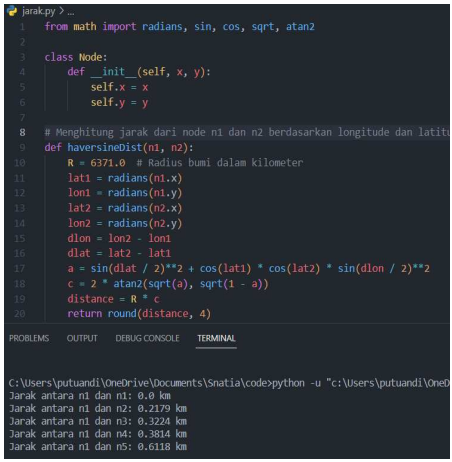
$$d_{ij} = (69 \cdot \sqrt{(lon_i - lon_j)^2 + (lat_i - lat_j)^2}) \cdot 1{,}60934 \quad (2)$$

Dalam konteks ini:

a. "dij" mengacu pada jarak antara dua titik koordinat i dan j, diukur dalam kilometer.
b. "loni" merujuk pada nilai longitude dari titik i.
c. "lonj" merujuk pada nilai longitude dari titik j.
d. "lati" merujuk pada nilai latitude dari titik i.
e. "latj" merujuk pada nilai latitude dari titik j.

Keterangan:

Pada rumus diatas menggunakan rumus Haversine, dimana konstanta 69 yang digunakan bertujuan untuk mengonversi perbedaan dalam derajat lintang (latitude) menjadi jarak dalam mil. Ini didasarkan pada anggapan bahwa pada garis lintang rata-rata di Bumi, jarak sejajar satu derajat lintang sekitar 69 mil (atau 111,045 kilometer). Dengan menggunakan konstanta ini, kita dapat mengaproksimasi jarak dalam mil berdasarkan perbedaan lintang antara dua titik. Setelah kita menghitung jarak dalam mil menggunakan konstanta 69, kita perlu mengonversinya menjadi kilometer. Faktor 1,60934 digunakan untuk mengubah satuan jarak dari mil menjadi kilometer. Faktor ini adalah hasil dari konversi 1 mil menjadi kilometer (1 mil = 1,60934 kilometer). Dengan mengalikan jarak dalam mil dengan faktor ini, kita mendapatkan hasil jarak dalam kilometer.
Melakukan proses jarak masing-masing titik:



Gambar 2. Mencari Jarak Antar Node

Berdasarkan perhitungan jarak antara titik/ tempat yang telah ditentukan sebelumnya. Hasil jarak tersebut akan direpresentasikan dalam satuan km dan dibulatkan 2 angka dibelakang koma sehingga diperoleh hasil sebagai berikut:

Tabel 2. Jarak Masing-Masing Titik Tujuan:

x	1	2	3	4	5
1	0.00	0.22	0.32	0.38	0.61
2	0.22	0.00	0.11	0.49	0.44
3	0.32	0.11	0.00	0.56	0.45
4	0.38	0.49	0.56	0.00	0.53
5	0.61	0.44	0.45	0.53	0.00

Setelah mendapatkan jarak satu titik/ tempat ke titik/ tempat lainnya, maka selanjutkan kita dapt menghitung pencarian rute tercepat/ terbaik dengan mengimplementasikan algoritma A* di dalamnya.

```python
import heapq

class Node:
    def __init__(self, name):
        self.name = name
        self.neighbors = {}
        self.g_score = float('inf')
        self.f_score = float('inf')
        self.came_from = None

    def add_neighbor(self, neighbor, weight):
        self.neighbors[neighbor] = weight

def astar(start, goal):
    open_set = []
    heapq.heappush(open_set, (0, start))
    start.g_score = 0
    start.f_score = heuristic(start, goal)

    while open_set:
        current = heapq.heappop(open_set)[1]

        if current == goal:
            return reconstruct_path(current)

        for neighbor, weight in current.neighbors.items():
            tentative_g_score = current.g_score + weight

            if tentative_g_score < neighbor.g_score:
                neighbor.came_from = current
                neighbor.g_score = tentative_g_score
                neighbor.f_score = tentative_g_score + heuristic(neighbor, goal)
                heapq.heappush(open_set, (neighbor.f_score, neighbor))
```

Gambar 3. Mencari Rute Terpendek Menggunakan Algoritma A*

Output program yang dikembangkan:

```
Masukkan titik awal: Camani Kopi
Masukkan titik akhir: Joglo Coffee Shop

Pengiriman Biji Kopi dari Camani Kopi ke Joglo Coffee Shop

Rute terpendek: Camani Kopi -> Crafted Coffee House -> Tan Panama -> Joglo Coffee Shop
Jarak tempuh: 0.75
```

Gambar 4. Output Yang Dihasilkan

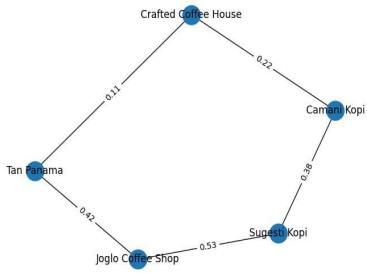
Hasil program yang dijalankan di atas didapatkan berdasarkan proses pencarian rute menggunakan algoritma A* (Star), dimana pada program user dapat mengiputkan titik atau tempat yang akan dijadikan titik awal pengiriman kopi dan titik akhir pengiriman kopi berdasarkan tempat atau titik yang telah ditetapkan sebelumnya pada maps. Proses selanjutkan peneliti akan melakukan inputan agar dapat diimplementasikan oleh graf. Tujuan dari implementasi ini agar memudahkan dalam melihat tampilan arah peta pencarian rute yang lebih jelas dan detail.

Implementasi graf dalam menentukan rute terpendek:

```
import networkx as nx
import matplotlib.pyplot as plt

# Membangun graf
graph = nx.Graph()
graph.add_edge("Camani Kopi", "Crafted Coffee House", weight=0.22)
graph.add_edge("Camani Kopi", "Sugesti Kopi", weight=0.38)
graph.add_edge("Crafted Coffee House", "Tan Panama", weight=0.11)
graph.add_edge("Tan Panama", "Joglo Coffee Shop", weight=0.42)
graph.add_edge("Sugesti Kopi", "Joglo Coffee Shop", weight=0.53)
graph.add_edge("Joglo Coffee Shop", "Sugesti Kopi", weight=0.45)
graph.add_edge("Sugesti Kopi", "Camani Kopi", weight=0.38)
graph.add_edge("Joglo Coffee Shop", "Sugesti Kopi", weight=0.53)
graph.add_edge("Tan Panama", "Crafted Coffee House", weight=0.11)

# Menampilkan graf
pos = nx.spring_layout(graph)
nx.draw(graph, pos, with_labels=True, node_size=500, font_size=12)
edge_labels = nx.get_edge_attributes(graph, 'weight')
nx.draw_networkx_edge_labels(graph, pos, edge_labels=edge_labels)
```

Gambar 5. Pencarian Rute Dengan Tampilan Graf

Output program yang dikembangkan:



Gambar 6. Output Tampilan Graf

Pada pengujian ini, dilakukan variasi pada titik asal dan tujuan untuk mengevaluasi hasil algoritma A Star. Tujuan mengimplementasikan dalam bentuk graf, agar user dapat melihat tampilan dengan lebih jelas dan pasti terkait jarak antar titik/ tempat pendistribusian kopi. Hasil pengujian lengkap dari masing-masing titik atau tempat ke tempat tujuan dilihat pada tabel berikut:

Tabel 4. Hasil Pengujian Sistem

No	Titik Awal	Titik Tujuan	Rute	Jarak Terbaik (km)
1	Camani Kopi	Crafted Coffee House	Camani Kopi → Crafted Coffee House	0.22

No	Titik Awal	Titik Tujuan	Rute	Jarak Terbaik (km)
2	Camani Kopi	Tan Panama	Camani Kopi → Crafted Coffee House → Tan Panama	0.33
3	Camani Kopi	Sugesti Kopi	Camani Kopi → Sugesti Kopi	0.38
4	Camani Kopi	Joglo Coffee Shop	Camani Kopi → Crafted Coffee House → Tan Panama → Joglo Coffee Shop	0.75
5	Crafted Coffee House	Tan Panama	Crafted Coffee House → Tan Panama	0.11
6	Crafted Coffee House	Sugesti Kopi	Crafted Coffee House → Camani Kopi → Sugesti Kopi	0.6
7	Crafted Coffee House	Joglo Coffee Shop	Crafted Coffee House → Tan Panama → Joglo Coffee Shop	0.53
8	Tan Panama	Sugesti Kopi	Tan Panama → Crafted Coffee House → Camani Kopi → Sugesti Kopi	0.71
9	Tan Panama	Joglo Coffee Shop	Tan Panama → Joglo Coffee Shop	0.42
10	Sugesti Kopi	Joglo Coffee Shop	Sugesti Kopi → Joglo Coffee Shop	0.53

Dari hasil pengujian dapat disimpulkan, bahwa program dibuat bisa menerima inputan berupa graph sehingga dapat menghitung rute terpendek, dan menampilkan rute terpendek beserta jaraknya.

4. Kesimpulan

Pada penelitian ini dikembangkan sebuah sistem yang berhasil diuji untuk menentukan rute terpendek menuju distributor kopi dengan menggunakan algoritma A Star. Hasil dari sistem ini menunjukkan bahwa pencarian rute terpendek pada peta menggunakan algoritma A Star dapat diwujudkan dalam bentuk graf. Setiap simpul grafik mewakili lokasi di peta, sedangkan ujungnya mewakili jalur yang menghubungkan lokasi tersebut. Setiap sisi graf memiliki bobot atau jarak yang digunakan untuk menghitung nilai heuristic. Setelah mendapatkan informasi jarak dan nilai heuristic, pencarian rute dilakukan menggunakan algoritma A Star menggunakan pohon pencarian dan antrian prioritas. Pada penelitian ini, dengan menggunakan algoritma A Star, sistem penentuan jalur terpendek dapat mengolah masukan graf, menghitung jalur terpendek, dan menampilkan hasil jalur terpendek beserta jaraknya. Misalnya, jika titik awal adalah Camani Kopi dan titik tujuan adalah Joglo Coffee Shop, maka jarak terpendek adalah 0,75 km pada rute Camani Kopi → Crafted Coffee House → Tan Panama → Joglo Coffee Shop.

Daftar Pustaka

[1] A. T. Putra, M. Rumani, B. Tt, and M. W. Paryasto, “PERBANDINGAN KOMPLEKSITAS ALGORITMA A-STAR, FLOYD-WARSHALL, VITERBI PADA SDN (SOFTWARE

DEFINED NETWORKING)."
[2] R. Wafdan, M. Ihsan, and R. Zuhra, "Connectivity Algorithm in Simple Graphs," Jurnal Natural, vol. 14, no. 1, pp. 30–33, 2014.
[3] I. Bagus, G. Wahyu, and A. Dalem, "PENERAPAN ALGORITMA A* (STAR) MENGGUNAKAN GRAPH UNTUK MENGHITUNG JARAK TERPENDEK," Online, 2018. [Online]. Available: http://jurnal.stiki-indonesia.ac.id/index.php/jurnalresistor
[4] D. Teguh Yuwono and A. Fadlil, "Perbandingan Algoritma Breadth First Search dan Depth First Search Sebagai Focused Crawler," 2016. [Online]. Available: http://ars.ilkom.unsri.ac.id
[5] A. Ananda and N. Siregar, "Fakultas Komputer."
[6] A. Adam Maulana, T. Informatika UDINUS, and F. Ilmu Komputer UDINUS, "Implementasi Algoritma A* Dalam Aplikasi Berbasis Web untuk Menemukan Rute Terpendek sebagai Navigasi Peta Digital Indoor Implementation of A* Algorithm in Web-Based Applications for Finding the Shortest Route as Navigation of Digital Indoor Map," Universitas AMIKOM Yogyakarta, vol. 5, no. 1, 2017.
[7] S. Suhendri, Dede Abdurahman, and Dani Irfan Maulana, "IMPLEMENTASI ALGORITMA A-STAR UNTUK PEMETAAN LOKASI SARANA KESEHATAN KABUPATEN MAJALENGKA BERBASIS GEOGRAPHIC INFORMATION SYSTEM (GIS)," INFOTECH journal, pp. 57–65, Sep. 2021, doi: 10.31949/infotech.v7i2.1512.
[8] S. Purnama, D. Ayu Megawaty, and Y. Fernando, "PENERAPAN ALGORITMA A STAR (A*) UNTUK PENENTUAN JARAK TERDEKAT WISATA KULINER DI KOTA BANDARLAMPUNG," Jurnal TEKNOINFO, vol. 12, no. 1, pp. 28–32, 2018.
[9] A. P. Graf, "Kajian Teori."

Halaman ini sengaja dibiarkan kosong

JNATIA Volume 1, Nomor 4, Agustus 2023 p-ISSN: 2986-3929
Jurnal Nasional Teknologi Informasi dan Aplikasinya

Ekstraksi Fitur pada Nada Gundhul Pacul

Roger Julian Sitorus[1], I Gede Santi Astawa[2]

[a]Program Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam,
Universitas Udayana
Jalan Raya Kampus Udayana, Bukit Jimbaran, Kuta Selatan, Badung, Bali Indonesia
[1]roger.v.sitorus@gmail.com
[2] santi.astawa@unud.ac.id

Abstract

MFCC is an effective method in audio feature extraction, including in song and music analysis. This method involves converting the frequency spectrum of the audio signal into the Mel scale, which is more in line with human auditory perception, and then calculating the cepstral coefficient. The results of the MFCC feature extraction on the song "Gundhul Pacul" show the pattern of the song's spectral and rhythmic characteristics. By using the MFCC representation, it is possible to see changes in energy and frequency patterns in audio signals at various time intervals. The results of the MFCC show that there are 13 resulting cepstral coefficients. However, the number of cepstral coefficients can be adjusted depending on the application and specific needs.

Keywords: *Features Extraction, MFCC, Gundhul Pacul*

1. Pendahuluan

Dalam dunia musik, analisis lagu dapat memberikan wawasan yang berharga tentang karakteristik musik, struktur harmonik, dan perbedaan genre. Salah satu pendekatan yang digunakan untuk menganalisis lagu adalah ekstraksi fitur audio dengan menggunakan Mel-frequency Cepstral Coefficients (MFCCs). Dalam konteks ini, pendahuluan akan fokus pada penggunaan MFCCs untuk ekstraksi fitur pada lagu "Gundhul Pacul". "Gundhul Pacul" adalah salah satu lagu tradisional yang populer di Indonesia. Lagu ini memiliki sejarah dan keunikan tersendiri, baik dalam hal melodi maupun liriknya. Dalam konteks ini, ekstraksi fitur dengan MFCCs dapat membantu dalam mengidentifikasi karakteristik spektral dan ritmis lagu "Gundhul Pacul". MFCC adalah metode ekstraksi fitur audio yang telah terbukti efektif dalam berbagai aplikasi pengenalan suara dan pemrosesan bahasa alami. Metode ini melibatkan konversi spektrum frekuensi sinyal audio ke dalam skala Mel, yang lebih sesuai dengan persepsi pendengaran manusia, dan kemudian menghitung koefisien cepstral. Dengan menggunakan MFCCs, fitur-fitur spektral dan ritmis lagu "Gundhul Pacul" dapat diekstraksi dengan lebih efektif dan menggambarkan ciri khasnya. Penelitian terkait dengan MFCC pada data musik telah dilakukan. Penelitian yang dilakukan adalah. Data yang digunakan adalah file musik instrumental sebanyak 200 dataset. 200 file musik tersebut memiliki 4 kategori mood, yaitu relax, sad, happy, angry. Fitur yang akan diektraksi ada 13 koefisien MFCC. setiap file akan berdurasi 60s dengan panjang frame 40ms dan overlap 40%. kemudian hasil akurasi cross validation knn sebesar 85,5%, precision 87,34% dan recall 85,5% dan nilai k=5 dan metode jarak yang digunakan adalah manhattan disctance [1]. Penelitian lainnya dilakukan oleh. Peneliti melakukan Deteksi Nada Dasar Alat Musik Panting. Metode yang digunakan dalam penelitian ada 3 ,yaitu compressive sensing, Mel-Frequency Cepstral Coefficient dan Support Vector Machine. Data yang digunakan adalah nada dasar do, re, mi, fa, sol, la, si yang didapatkan dari alat musik Panting. Sampel data yang digunakan dalam penelitian ada 140 nada dengan 20 sampel per nada. Durasi audio untuk setiap nada adalah 1 detik. Sampel audio dibagi kedalam 6 rasio, antara lain, 0.625%, 1.25%, 2.5%, 5%, 10%, dan rasio normal. Ada 4 parameter ciri yang digunakan dalam ekstraksi ciri MFCC, yaitu zero crossing rate (ZCR), spectral rolloff, ciri MFCC sebanyak 20, dan mel-spectogram. Hasil akurasi terbaik diperoleh oleh audio terkompresi 2.5% yang memiliki akurasi dan f1-score 98%. Hasil MAE 0.05 dalam waktu klasifikasi 0,062 detik [2]. Dalam penelitian ini, tujuan utama adalah menerapkan ekstraksi fitur dengan MFCCs pada lagu "Gundhul Pacul".

Kami berharap bahwa dengan menggunakan teknik ini, kita dapat menggali lebih dalam tentang pola karakteristik dan kesamaan antara nada-nada dalam lagu ini. Dengan demikian, kita dapat memahami struktur musik "Gundhul Pacul" secara lebih mendalam, serta mengidentifikasi perbedaan atau variasi dalam ekspresi vokal dan instrumen yang digunakan. Melalui penelitian ini, diharapkan kita dapat memberikan sumbangan kepada pemahaman kita tentang lagu "Gundhul Pacul" secara musikal dan memperkaya budaya musik tradisional Indonesia. Dengan menggunakan ekstraksi fitur MFCCs, kita dapat memberikan wawasan baru tentang lagu ini, memungkinkan analisis yang lebih dalam dan pemahaman yang lebih luas tentang karakteristik dan keunikan musik tradisional ini.

2. Metode Penelitian

Dalam melakukan ekstraksi fitur pada musik gundhul pacul diperlukan berbagai tahapan yang harus dilakukan untuk bisa mendapatkan luaran yang memiliki tingkat akurasi yang cukup tinggi. Adapun tahapan-tahapan tersebut, antara lain:

a. Studi Literatur

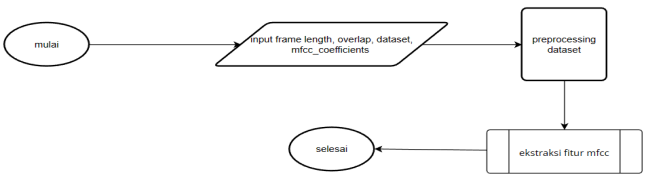
Pada tahap ini dilakukan pencarian referensi terkait penelitian yang dilakukan. Tahapan ini penulis melakukan studi mengenai topik yang akan dibahas. Hal ini ditujukan agar penulis dapat memahami secara mendalam mengenai ekstraksi fitur pada kasus yang akan dilakukan. Dengan hasil studi tersebut, penulis mendapatkan metode ataupun kombinasi metode yang cocok untuk digunakan dalam penelitian. Dalam pembahasan ekstraksi fitur pada nada lagu gundhul pacul diperoleh beberapa referensi terkait. Metode MFCC telah banyak digunakan untuk ekstraksi fitur pada data audio,termasuk data musik.

b. Pengumpulan Data

Pada penelitian kali ini, data yang akan digunakan adalah gabungan data primer dan sekunder. Data tersebut berupa WAV audio lagu gundhul pacul. Data mp3 tersebut diperoleh dari youtube video yang dikonversi menjadi format WAV dan terdapat data yang diperoleh dari memainkan nada lagu dari gundhul pacul.

c. Perancangan Sistem

Sistem yang akan dibangun adalah ekstraksi fitur lagu gundhul pacul. Data yang akan digunakan pada sistem adalah data mentah. Jadi, akan dilakukan pemotongan sinyal audio menjadi beberapa bagian frame. Hasil dari pemotongan tersebut akan dilakukan ekstraksi fitur dengan metode MFCC.



Gambar 1. Flowchart Pembuatan Sistem

d. Ekstraksi Fitur

Ekstraksi fitur suara adalah proses yang bertujuan mendapatkan vektor ciri dari sebuah suara. Pada penelitian ini ekstraksi fitur suara dilakukan menggunakan algoritma MFCC yang didasarkan pada persepsi sistem pendengaran manusia. Secara umum MFCC mempunyai lima

proses yaitu frame blocking, windowing, FFT, mel-frequency wrapping, dan cepstrum. Flowchart ekstraksi fitur menggunakan algoritma MFCC dapat dilihat pada gambar 2 [3].

a. Frame Blocking

Pada tahap pertama metode MFCC adalah melakukukan frame blocking. Frame blocking adalah proses dimana sinyal pada audio akan dipotong menjadi beberapa yang disebut frame. Proses pertama MFCC adalah frame blocking di mana sinyal dipotong menjadi beberapa bagian yang disebut dengan frame. Sinyal yang diterima dipisahkan dalam Frames dan kemudian diambil sampelnya. Setiap frame berisi Sampel 'N' dan berlanjut hingga sampel N-M dan berlanjut hingga sinyal ucapan selesai. Sampel 'N' memiliki nilai 256 dan 'M' memiliki nilai 1000 [4].

b. Windowing

Terakhir, meruncingkan sampel di setiap jendela untuk mengurangi diskontinuitas di tepi jendela sangat bermanfaat. Itu melakukan transformasi berikut pada sampel di jendela sn, n = 1, N [5].

$$S'_n = \left\{0.54 - 0.46\ \cos\left(\frac{2\pi(n-1)}{N-1}\right)\right\} S_n \quad (1)$$

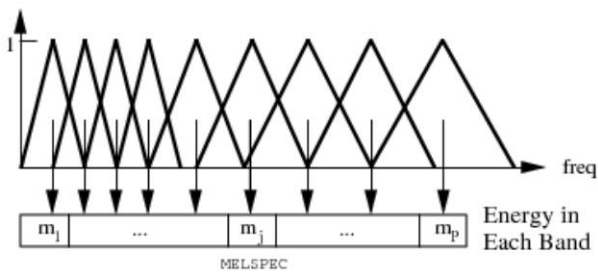
Gambar 2. Rumus

c. Fast Fourier Transform (FFT)

Proses selanjutnya adalah fast fourier transform (FFT) merupakan metode analisis fourier. Analisis fourier adalah metode yang berguna untuk melakukan analisa kepada sinyal yang dimasukkan berupa spectrogram [3].

d. Mel-Frequency Wrapping

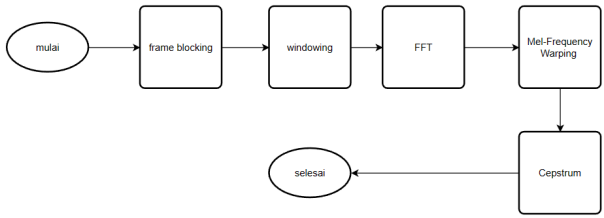
Untuk setiap frame ini, spektrum magnitudo dihitung menggunakan Fast Fourier Transform (FFT) dan diubah menjadi kumpulan output dari bank filter skala mel. Analisis bank filter membuat mendapatkan resolusi frekuensi non-linear yang sesuai jauh lebih mudah. Amplitudo bank filter, di sisi lain, sangat berkorelasi, membuat penggunaan transformasi cepstral hampir wajib dalam skenario ini. Pada skala mel, bank filter berbasis transformasi Fourier sederhana dirancang untuk memberikan resolusi yang setara [5].



Gambar 3. Mel Scale Filter Bank

e. Cepstrum

Proses terakhir adalah cepstrum yaitu sebuah teknik untuk meningkatkan kualitas pengenalan sinyal suara.



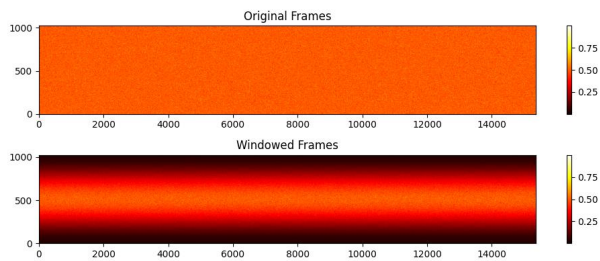
Gambar 4. Flowchart MFCC

3. Hasil dan Pembahasan

a. Frame Blocking

Pada tahap frame blocking, didapatkan hasil jumlah frame pada audio gundhul pacul adalah 15360. Ukuran frame yang digunakan pada pengujian adalah 1024 dan jarak antara frame adalah 512.

b. Windowing



Gambar 5. Hasil Dari Windowing

1. Sumbu X (Time): Sumbu horizontal pada gambar menunjukkan waktu dalam unit yang sesuai (misalnya, detik atau milidetik). Setiap titik pada sumbu ini merepresentasikan waktu relatif di sinyal audio.

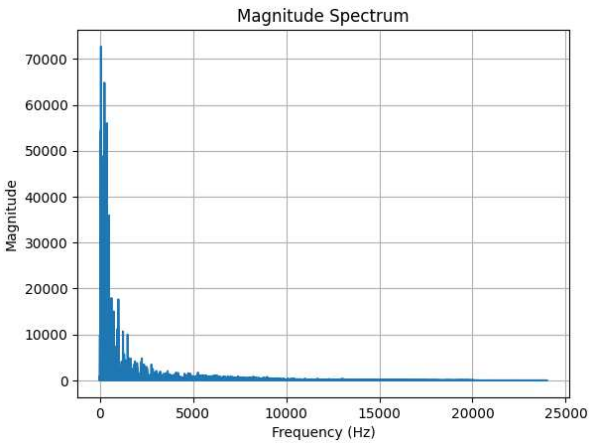
2. Sumbu Y (Amplitudo atau Energi): Sumbu vertikal menunjukkan amplitudo atau energi sinyal audio pada setiap frame waktu. Setiap titik atau garis pada sumbu ini merepresentasikan amplitudo atau energi yang terkait dengan frame-frame tersebut.

3. Frame-frame: Pada gambar, frame-frame sinyal audio ditampilkan sebagai titik-titik atau garis-garis yang berurutan dalam sepanjang sumbu X. Setiap frame mewakili segmen waktu pendek dari sinyal audio yang akan dianalisis secara terpisah.

4. Windowing: Proses windowing telah diterapkan pada setiap frame untuk mengurangi efek sisi atau transisi yang tajam di batas-batas frame. Fungsi window yang umum digunakan, seperti Hamming atau Hann, memberikan bentuk kurva yang halus pada setiap frame.

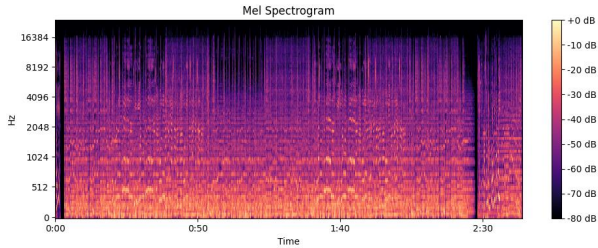
5. Overlap: Biasanya, frame-frame tersebut tumpang tindih satu sama lain untuk mempertahankan kontinuitas informasi waktu di sepanjang sinyal audio. Dengan melakukan tumpang tindih, analisis dapat melibatkan informasi dari beberapa frame sebelumnya dan setelahnya, sehingga memberikan gambaran yang lebih lengkap tentang sinyal audio.

c. Fast Fourier Transform (FFT)



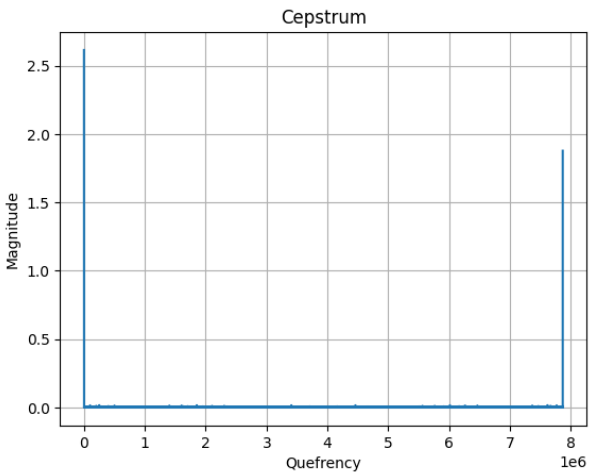
Gambar 6. Fast Fourier Transform (FFT)

d. Mel-Frequency Wrapping



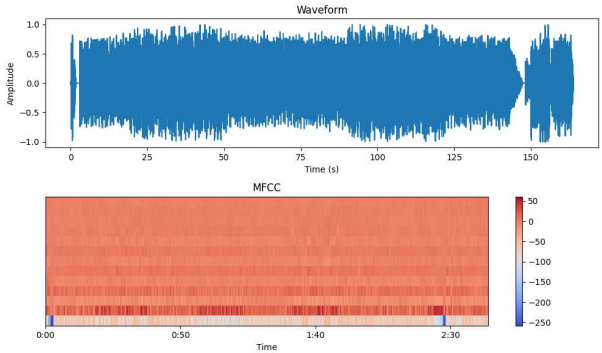
Gambar 7. Mel-Frequency Wrapping

e. Cepstrum



Gambar 8. Cepstrum

f. MFCC



Gambar 9. MFCC

Dengan menggunakan representasi MFCC, gambar ini memberikan informasi tentang fitur akustik dari sinyal audio. Misalnya, perbedaan intensitas warna antara koefisien MFCC pada berbagai waktu dapat menggambarkan perubahan energi atau pola frekuensi dalam sinyal audio. Pada gambar ini, Anda dapat melihat pola spektral yang ditampilkan oleh koefisien MFCC pada setiap frame waktu. Hal ini memberikan informasi tentang distribusi energi frekuensi dalam sinyal audio pada berbagai interval waktu. Juga, perubahan pola atau bentuk dalam representasi MFCC dapat mengindikasikan perubahan dalam karakteristik audio, seperti perubahan nada, timbre, atau ucapan dalam sinyal audio. Dalam gambar MFCC, terlihat bahwa terdapat 13 koefisien cepstral yang dihasilkan. Namun, jumlah koefisien cepstral dapat disesuaikan tergantung pada aplikasi dan kebutuhan spesifik.

6. Kesimpulan

Pada penelitian ini, ekstraksi fitur MFCC akan digunakan untuk mengekstraksi fitur pada lagu daerah gundhul pacul. Hasil dari MFCC menunjukkan bahwa ada 13 koefisien cepstral yang dihasilkan. Dihasilkan juga grafik amplitude dan intensitas warna pada grafik MFCC. Grafik tersebut menggambarkan distribusi energi frekuensi pada sinyal audio.

Daftar Pustaka

[1] F. F. Surenggana, A. Aranta, and F. Bimantoro, "Klasifikasi Mood Musik Menggunakan K-Nearest Neighbor Denganmel Frequency Cepstral Coefficients," Jurnal Teknologi Informasi, Komputer dan Aplikasinya, vol. 4, No. 2, September 2022.
[2] A. S. Sankoh, A. R. Musthafa, M. I. Rosadi, and A. Z. Arifin, "Klasterisasi Jenis Musik Menggunakan Kombinasi Algoritma Neural Network, K-Means dan Particle Swarm Optimization," Jurnal Buana Informatika, Volume 6, Nomor 3, Juli 2015: 183-194.
[3] H. Lana, N. A. Sanjaya, I. D. W. B. Atmaja, A. A. I. N. E. Karyawati, I. M. Widiartha, I. G. A. G. A. Kadyanan, "Aplikasi Identifikasi Nada Darbuka Dengan Onset Detection, MFCC, Dan KNN," Jurnal Elektronik Ilmu Komputer Udayana, vol. 11, No.1, Agustus 2022.
[4] H. M. O. A. Marzuqi, S. M. Hussain, Dr. A. Frank, "Device Activation based on Voice Recognition using Mel Frequency Cepstral Coefficients (MFCC's) Algorithm," International

Research Journal of Engineering and Technology (IRJET), vol. 6, No. 3, Maret 2019.
[5] P. Aurchana, S. Prabavathy, "Musical Instruments Sound Classification using GMM," London Journal of Social Sciences, vol. 1, No. 1, 2021.

JNATIA Volume 1, Nomor 4, Agustus 2023 p-ISSN: 2986-3929
Jurnal Nasional Teknologi Informasi dan Aplikasinya

Perancangan Prototype Aplikasi Deteksi dan Pelacakan Manusia Pada Video

Valentin Gea Affila Pradika[1], I Gusti Agung Gede Arya Kadyanan[2]

[a]Program Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam,
Universitas Udayana
Jalan Raya Kampus Udayana, Bukit Jimbaran, Kuta Selatan, Badung, Bali Indonesia
[1] geaaffila057@student.unud.ac.id
[2] gungde@unud.ac.id (Corresponding Author)

Abstract

The detection and tracking of humans in videos have gained significant attention due to their applications in various fields such as surveillance, activity recognition, and human-computer interaction. This article presents the design and development of a prototype application for human detection and tracking in videos. The use of the prototype methodology allows for early feedback, demonstrates the functionality and features, facilitates effective collaboration, and helps save time and costs in the development process. By following this methodology, the prototype application for human detection and tracking in videos is expected to provide accurate and reliable results, meeting the needs of users in various domains.

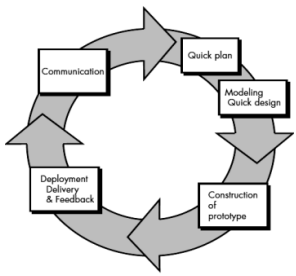
Keywords: *Prototype, Human, Detection, Tracking*

1. Pendahuluan

Dalam beberapa tahun terakhir, teknologi deteksi dan pelacakan manusia telah mengalami kemajuan yang signifikan. Kemampuan untuk secara akurat mendeteksi manusia dalam video dan melacak gerakan mereka telah membuka pintu bagi berbagai aplikasi yang luas. Dari keamanan dan pemantauan hingga realitas virtual dan pengenalan gerakan, penggunaan teknologi ini semakin meluas. Pada artikel ini, penulis akan membahas tentang perancangan prototype sebuah aplikasi deteksi dan pelacakan manusia pada video. Prototype atau prototipe adalah sebuah metode dalam pengembangan produk dengan cara membuat rancangan, sampel, atau model dengan tujuan pengujian konsep atau proses kerja dari produk [1]. Tujuan dari perancangan prototype adalah untuk memberikan sebuah gambaran dari aplikasi yang akan dibuat nantinya. Gambaran yang diberikan berupa desain tampilan antarmuka serta alur kerja aplikasi yang terdiri dari penginputan video, melakukan proses deteksi dan pelacakan manusia pada video inputan, kemudian menampilkan hasil (output) berupa informasi hasil ada atau tidaknya manusia dan video yang telah berisi penanda pada manusia yang terdeteksi. Penelitian mengenai deteksi dan pelacakan manusia telah banyak dilakukan sebelumnya. Penelitian [2] yang berjudul "Monitoring Ruangan Untuk Deteksi Manusia Berbasis CNN Dengan Fitur Push Notification" membuat system yang dapat mendeteksi manusia menggunakan perangkat Rasberry Pi 3 dengan metode pendeteksi berbasis CNN framework YOLO dan akan mengirimkan notifikasi dengan teknolgi Firebase Cloud Messaging jika sistem mendeteksi manusia pada video. Penelitian [3] berjudul "Penerapan Metode Convolutional Neural Network (Cnn) Dan Long Short Term Memory (Lstm) Untuk Pengenalan Aktivitas Manusia Pada Cctv Di Area Tambak Udang" melakukan pendeteksian manusia dari cctv yang ada di area tambak udang, pengenalan aktivitas manusia berupa berjalan dan berlari, serta mengirimkan notifikasi terhadap hasil yang didapatkan. Serta beberapa artikel internasional [4] – [7] yang telah melakukan penelitian serupa, yaitu deteksi dan pelacakan manusia dari data video namun dengan metode yang berbeda-beda. Dengan perancangan prototype aplikasi deteksi dan pelacakan manusia pada video ini, diharapkan dapat memberikan gambaran dengan lebih mudah dan nyata atas aplikasi yang akan dikembangkan nantinya. Umpan balik berupa kritik dan saran dari calon pengguna yang mengetahui maupun mencoba prototype yang dihasilkan nantinya dapat digunakan sebagai bahan pertimbangan atas pengembangan aplikasi yang sesungguhnya.

2. Metode Penelitian

Dalam penelitian ini digunakan metode prototyping teknik pengembangan sistem yang menggunakan prototype untuk menggambarkan sistem sehingga klien atau pemilik sistem mempunyai gambaran jelas pada sistem yang akan dibangun oleh tim pengembang [8]. Metode ini dipilih karena beberapa alasan, yaitu umpan balik dari pengguna atau pemangku kepentingan didapatkan diawal sehingga memungkinkan evaluasi konsep, fitur, dan kinerja aplikasi sebelum melanjutkan ke tahap pengembangan lebih lanjut; dapat mendemonstrasikan fungsi dan fitur dari aplikasi; dapat melakukan pengujian dan perbaikan lebih awal; efektif dalam pengembangan aplikasi dalam sebuah tim; serta dapat menghemat waktu dan biaya. Gambar 1. merupakan gambaran dari cara kerja metode prototype.



Gambar 1. Metode Prototype

2.1 Prototyping

Berikut adalah tahapan dalam metode prototyping [8]:

a. Tahap 1: Analisis Kebutuhan

Pada analisis kebutuhan, dibutuhkan pemahaman mendalam mengenai kebutuhan pengguna dan tujuan aplikasi. Hal ini dapat dilakukan dengan mengadakan diskusi antar pengembang dan pengguna atau pemangku kepentingan.

b. Tahap 2: Desain Cepat

Setelah analisis kebutuhan aplikasi didapatkan, dilakukan pembuatan desain cepat, yaitu desain sederhana yang dapat memberikan gambaran singkat mengenai aplikasi yang akan dikembangkan namun tetap memenuhi seluruh analisis kebutuhan yang diinginkan.

c. Tahap 3: Membangun Prototype

Pembangunan prototype ini berdasarkan pada desain cepat yang telah dibuat dan disepakati sebelumnya. Dapat berupa tampilan statis maupun interaktif yang menggambarkan antarmuka pengguna dan alur aplikasi.

d. Tahap 4: Mengevaluasi Awal

Dari prototype yang telah dibuat, dilakukan evaluasi awal mengenai kinerja dan fungsionalitas dari prototype aplikasi. Biasanya pada tahap ini pengembang melakukan demonstrasi kepada pengguna atau pemangku kepentingan ataupun pengguna diminta

mencoba menjalankan prototype aplikasi secara mandiri. Dari sinilah didapatkan umpan balik berupa komentar, kritik, mupun saran dari pengguna terhadap prototype tersebut yang akan dijadikan sebagai bahan evaluasi.

e. Tahap 5: Memperbaiki Prototype

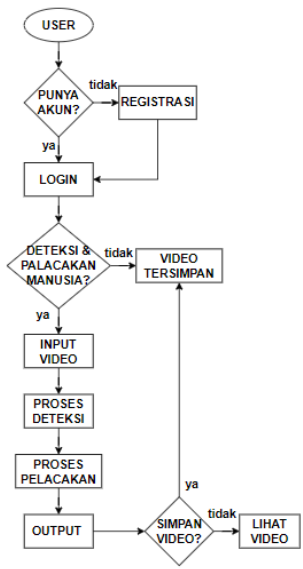
Bahan evaluasi yang didapatkan pada tahap sebelumnya dilakukan sebagai acuan perbaikan prototype. Tahap 4 dan tahap 5 akan terjadi secara berulang berdasarkan umpan balik dari pengguna atau pemangku kepentingan. Perulangan kedua tahap ini akan berhenti ketika pengguna atau pemangku kepentingan telah benar-benar menyetujui dan tidak memberikan revisi terhadap prorotype tersebut.

f. Tahap 6: Implementasi dan Pemeliharaan

Tahap 6 merupakan tahap terakhir dari metode prototyping. Setelah prototype telah disetujui, selanjutnya desain tersebut akan diimplementasikan secara nyata dalam pengembangan aplikasi. Pemeliharaan juga dilakukan agar aplikasi yang telah selesai dikembangkan dapat terus berjalan dengan baik sesuai dengan fungsionalitasnya.

2.1 Alur Kerja Aplikasi

Berikut ini adalah alur kerja dari aplikasi:



Gambar 2. Alur Kerja Aplikasi

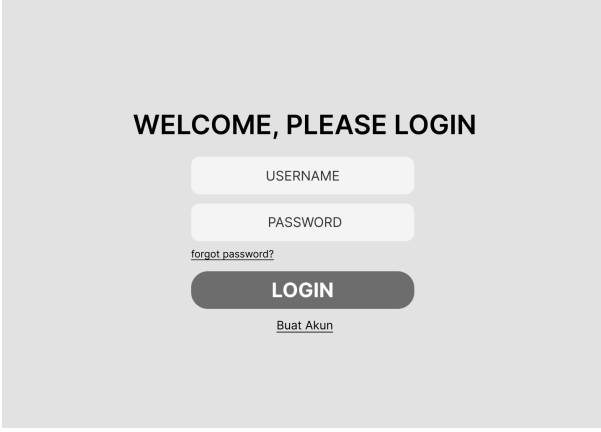
Pengguna harus melakukan login terlebih dahulu sebelum menggunakan aplikasi. Hal ini dilakukan untuk menjaga privasi atau kerahasiaan atas video yang akan digunakan sebagai input nantinya agar hanya diketahui oleh pemilik akun. Selain itu, video yang dihasilkan sebagai output dari aplikasi juga dapat disimpan dalam akun tadi. Jika pengguna belum memiliki akun, maka akan melakukan proses registrasi untuk pembuatan akun. Setelah masuk dalam akun, pengguna dapat melakukan proses deteksi dan pelacakan manusia dengan menginputkan sebuah video maupun dapat melihat video hasil proses deteksi dan pelacakan manusia yang telah disimpan sebelumnya.

3. Hasil dan Pembahasan

Berikut adalah tampilan antarmuka dari prototype aplikasi deteksi dan pelacakan manusia pada video:

3.1 Halaman Login

Pada halaman login, pengguna diminta untuk memasukkan username dan password dari akun yang telah terdaftar. Terdapat pula tombol jika pengguna lupa terhadap password akunnya yang nanti akan diarahkan pada prosedur perubahan password. Jika pengguna belum memiliki akun, maka dapat menekan tombol 'Buat Akun' untuk diarahkan pada halaman register.



Gambar 3. Halaman login

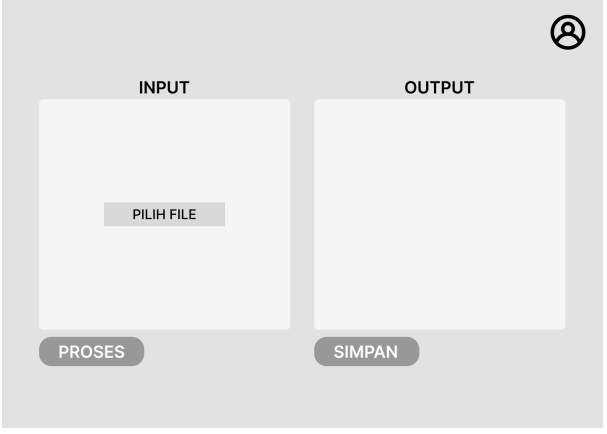
3.2 Halaman Register

Halaman register merupakan halaman pendaftaran akun bagi pengguna baru yang belum memiliki akun. Pengguna akan diminta memasukkan nama, baik itu nama lengkap maupun nama panggilan, tanggal lahir, username yang belum pernah dipakai oleh pengguna lain, password, dan konfirmasi password (pengulangan password). Setelah akun berhasil terdaftar, pengguna akan Kembali diarahkan pada halaman login agar dapat masuk ke dalam aplikasi.

 p-ISSN: 2986-3929

Gambar 4. Halaman Register

3.3 Halaman Utama



Gambar 5. Halaman Utama

Halaman ini berisi proses utama aplikasi, yaitu deteksi dan pelacakan manusia pada video. Pengguna dapat melakukan input dengan memilih file video yang terdapat pada device atau perangkat pengguna. Setelah video berhasil diinput, pengguna dapat menekan tombol 'proses' untuk menjalankan deteksi dan pelacakan manusia. Hasil dari proses akan ditampilkan pada bagian 'output' yang terletak di sebelah kanan. Pengguna dapat hanya memainkan dan/atau menyimpan video yang dihasilkan.

3.4 Halaman Penyimpanan Video

Gambar 6. Halaman Penyimpanan Video

Halaman terakhir merupakan halaman penyimpanan video. Disini terdapat video-video yang disimpan dari hasil proses deteksi dan pelacakan video sebelumnya. Dengan menyimpan video ini, pengguna dapat melihat Kembali hasil pendeteksi dan pelacakan manusia pada video tanpa melakukan proses ulang.

4. Kesimpulan

Perancangan prototype aplikasi deteksi dan pelacakan manusia pada video dalam penelitian ini menggunakan metode prototyping agar kinerja dan fungsionalitas dari aplikasi dapat sesuai dengan keinginan pengguna/pemangku kepentingan dan pengembang. Pembuatan prototype dapat memberikan gambaran dari aplikasi yang akan dikembangkan nantinya. Diharapkan perancangan prototype aplikasi deteksi dan pelacakan manusia pad video ini dapat menjadi salah satu acuan atau referensi dalam mengembangkan aplikasi serupa.

Daftar Pustaka

[1] R. 'Setiawan, "Apa Itu Prototype? Kenapa Itu Penting?," dicoding, Aug. 11, 2021. https://www.dicoding.com/blog/apa-itu-prototype-kenapa-itu-penting/ (accessed Jun. 11, 2023).

[2] W. Swastika, A. W. Nur, and O. H. Kelana, “Monitoring Ruangan Untuk Deteksi Manusia Berbasis CNN Dengan Fitur Push Notification,” Teknika, vol. 8, no. 2, 2019, doi: 10.34148/teknika. v8i2.166.
[3] M. A. Zulfikar, M. Somantri, and S. Sudjadi, “PENERAPAN METODE CONVOLUTIONAL NEURAL NETWORK (CNN) DAN LONG SHORT-TERM MEMORY (LSTM) UNTUK PENGENALAN AKTIVITAS MANUSIA PADA CCTV DI AREA TAMBAK UDANG,” Transient: Jurnal Ilmiah Teknik Elektro, vol. 10, no. 1, 2021, doi: 10.14710/transient. v10i1.98-105.
[4] X. Zhang, Z. Xu, and H. Liao, “Human motion tracking and 3D motion track detection technology based on visual information features and machine learning,” Neural Comput Appl, vol. 34, no. 15, 2022, doi: 10.1007/s00521-021-06703-2.
[5] M. Kumar, S. Ray, and D. K. Yadav, “Moving human detection and tracking from thermal video through intelligent surveillance system for smart applications,” Multimed Tools Appl, 2022, doi: 10.1007/s11042-022-13515-6.
[6] X. Zhou, J. Yi, G. Xie, Y. Jia, G. Xu, and M. Sun, “Human Detection Algorithm Based on Improved YOLO v4,” Information Technology and Control, vol. 51, no. 3, 2022, doi: 10.5755/j01.itc.51.3.30540.
[7] Y. Bouafia, L. Guezouli, and H. Lakhlef, “Human Detection in Surveillance Videos Based on Fine-Tuned MobileNetV2 for Effective Human Classification,” Iranian Journal of Science and Technology - Transactions of Electrical Engineering, vol. 46, no. 4, 2022, doi: 10.1007/s40998-022-00512-6.
[8] Sutiono, “Metode Prototype: Pengertian, Kekurangan dan Kelebihan,” dosenit. https://dosenit.com/software/metode-prototype#Tahapan_Metode_Prototype (accessed Jun. 11, 2023).

Halaman ini sengaja dibiarkan kosong

JNATIA Volume 1, Nomor 4, Agustus 2023 p-ISSN: 2986-3929
Jurnal Nasional Teknologi Informasi dan Aplikasinya

Klasifikasi Lirik Lagu Bertema Lingkungan dengan Metode Naive Bayes

Putu Ode Irfan Ardika[1], I Gusti Ngurah Anom Cahyadi Putra[2]

[a]Program Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Udayana
Jalan Raya Kampus Udayana, Bukit Jimbaran, Kuta Selatan, Badung, Bali Indonesia
[1]dirfanardika@gmail.com
[2]anom.cp@unud.ac.id

Abstract

Awareness of the importance of protecting the environment is becoming increasingly important in this modern era. Humans as inhabitants of the earth have a responsibility to protect and maintain the natural environment, they live in. Songs can be one of the important roles that can help awaken people to start protecting the environment they live in. This research makes it easy to find songs that have the theme of protecting the environment by classifying song lyrics. This research will create a system that can classify environment-themed song lyrics using the Naive Bayes method with a Multinomial model. The results of the Naïve Bayes test with the Multinomial model get the best results on the composition of the training data and test data of 10.90 which produces a recall score of 38%, precision of 90.4%, F1 score of 53.5%, while for accuracy it gets the best score on the composition of 90:10 with a yield of 75%.

Keywords: *Text Processing, TF-IDF, Naive Bayes, Lyrics*

1. Pendahuluan

Kesadaran akan pentingnya menjaga lingkungan menjadi semakin penting di era modern ini. Manusia sebagai penghuni bumi memiliki tanggung jawab untuk melindungi dan memelihara lingkungan alam yang ditinggali. Lagu dapat menjadi salah satu peran penting yang bisa membantu menyadarkan untuk mulai menjaga lingkungan yang ditinggali. Penelitian ini dapet memudahkan untuk mencari lagu yang memiliki tema menjaga lingkungan dengan melakukan klasifikasi pada lirik lagu. Penelitian mengenai klasifikasi yang sudah dilakukan oleh peneliti sebelumnya dengan seperti Uji Akurasi Klasifikasi Emosi Pada Lirik Lagu Bahasa Indonesia dengan menggunakan metode SVM dengan hasil akurasi 92,13% [1]. Kemudian penelitian mengenai Klasifikasi Cerita Pendek Berbahasa Bali Berdasarkan Umur Pembaca dengan Metode Naive Bayes yang menghasilkan akurasi 72% dengan feature selection [2]. Lalu penelitian mengenai Klasifikasi Emosi Lagu Berdasarkan Lirik pada Teks Berbahasa Indonesia Menggunakan K-Nearest Neighbor dengan Pembobotan WIDF yang menghasilkan akuras 66% [3]. Selanjutnya ada penelitian Klasifikasi Emosi Pada Lirik Lagu Menggunakan Algoritma Support Vector Machine Dan Optimasi Particle Swarm Optimization menghasilkan akurasi paling besar 90% [4]. Terakhir ada penelitian Klasifikasi Emosi Lirik Lagu Menggunakan Improved K-Nearest Neighbor dengan Seleksi Fitur dan BM25 yang mendapatkan hasil rata-rata paling baik saat k = 55 dengan hasil f-measure 0,6693, recall 0,6582 dan precision 0,7427 [5]. Pada penelitian kali ini akan membangun sebuah sistem yang dapat mengklasifikasikan teks lirik lagu dengan metode Naive Bayes. Kategori yang digunakan adalah lingkungan dan non-lingkungan. Penelitian ini digunakan untuk membantu mencarikan lagu yang bertemakan menjaga lingkungan untuk membantu memberikan kesadaran terhadap pentingnya menjaga lingkungan.

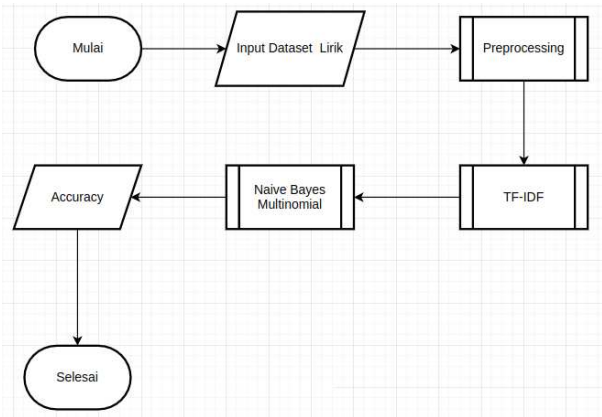
2. Metode Penelitian

2.1. Dataset

Data yang digunakan adalah 58 dokumen lirik lagu berbahasa indonesia. Ada 2 kategori yang digunakan yaitu kategori lingkungan yang dilabel 1 dan kategori non-lingkungan dengan label 0. Data yang digunakan meliputi judul, isi, dan label.

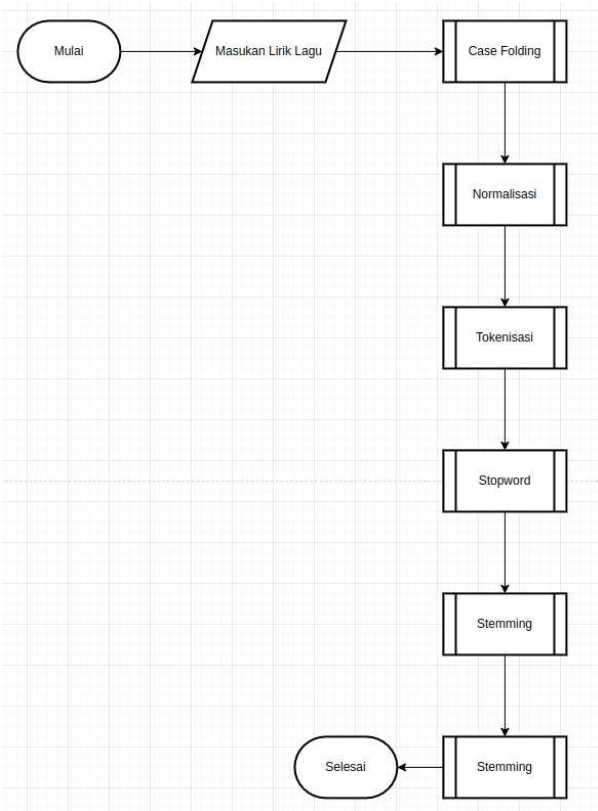
2.2. Alur Sistem

Tahap pertama yang dilakukan adalah pengumpulan data berupa lirik lagu berbahasa indonesia dengan 2 kategori yaitu lingkungan dan non-lingkungan. Setelah terkumpul semua, data kemudian disimpan ke dalam excel. Data kemudian melewati tahap preprocessing kemudian mengalami tahap pembobotan menggunakan TF-IDF. Kemudian hasil dari TF-IDF diproses menggunakan Naive Bayes Multinomial.



Gambar 1. Alur Penelitian

2.3 Preprocessing Data



Gambar 2. Alur Preprocessing

Dalam tahap ini dataset yang berupa lirik lagu diproses terlebih dahulu dimulai dengan tahap case folding untuk merubah menjadi lowercase kemudian menghilangkan kata dan karakter yang tidak diperlukan, seperti tanda baca, emoji, angka. Kemudian ditahap normalisasi menghilangkan dan memperbaiki kata dengan menerjemahkannya ke bahasa indonesia atau memperbaiki kata singkatan. Selanjutnya tahap tokenisasi untuk memisahkan lirik lagu menjadi token atau kata.

Tahap berikutnya stopword untuk menghilangkan kata yang dianggap tidak penting. Terakhir tahap stemming untuk menjadikan kata ke bentuk dasar. Hasil dari preprocessing berupa matriks 2D.

2.4 TF-IDF

TF-IDF adalah salah satu metode pembobotan suatu kata dengan dokumen[1].

Keterangan:

TF-IDF = TF * IDF
TF = Term Frequency
TF = Berapa kali kata muncul dalam sebuah dokumen
IDF = Inverse Document Frequency
IDF = log (total dokumen / jumlah dokumen yang mengandung kata)

2.5 Naive Bayes

Naive Bayes adalah salah satu algoritma machine learning yang digunakan untuk melakukan klasifikasi, salah satunya klasifikasi teks. Model Naive Bayes yang digunakan untuk penelitian ini adalah Multinomial Naive Bayes. Data yang sudah diproses sebelumnya disiapkan untuk diproses menggunakan MNB, dengan X sebagai data hasil TF-IDF dan y sebagai label nya. Kemudian datanya dibagi menjadi data uji dan data latih dengan komposisi 10:90 sampai 90:10. Selanjutnya MNB melakukan prekdiksi dan akan menghasilkan label dari data uji, kemudian label dicocok untuk mendapatkan nilai akurasinya, recall score, prescision dan F1-score.

2.6 Evaluasi

Pada penelitian ini mendapatkan hasil akurasi, recall score, precision, dan F1-score yang dapat dilihat dari tabel dibawah.

Tabel 1. Hasil Pengujian

Komposisi	Akurasi	Recall Score	Precision	F1-score
10:90	68,5%	38%	90,4%	53,5%
20:80	56,9%	13,3%	85,7%	23,1%
30:70	59,7%	18,4%	77,7%	29,7%
40:60	70%	34,4%	83,3%	48,7%
50:50	67,4%	32%	80%	45,7%
60:40	68,1%	30%	85,7%	44,4%
70:30	65,7%	26,6%	80%	40%
80:20	66,6%	30%	75%	42,8%
90:10	75%	33,3%	50%	40%

3. Hasil dan Pembahasan

Pada pengujian metode Naive Bayes menggunakan model Multinominal mendapatkan hasil mendapatkan hasil akurasi yang paling tinggi saat menggunakan data latih dan data uji 90:10

yang mendapatkan hasil 75% sedangkan yang terendah adalah pada komposisi 20:80 yang hanya mendapatkan 56,9% dengan rata-rata hasilnya pengujian akurasinya adalah 66,4%. Kemudian untuk hasil recall score mendapatkan hasil yang tertinggi pada komposisi 10:90 yang mendapatkan hasil 38% sedangkan hasil yang terendah pada komposisi 20:80 yang hanya mendapatkan hasil 13,3% dengan rata-rata hasilnya adalah 28,4%. Selanjutnya untuk hasil precision hasil teringgi didapatkan dengan komposisi 10:90 dengan hasil 90,4% sedangkan hasil terendah pada komposisi 90:10 yang hanya mendapatkan 50% dengan rata-rata hasilnya adalah 78,6%. Terakhir untuk F1-score mendapatkan hasil tertinggi pada komposisi 10:90 dengan 53,5% sedangkan hasil terendah terdapat pada komposisi 20:80 dengan hasil 23,1% dan rata-rata 40,8%.

4. Kesimpulan

Pada penelitian ini dilihat dari hasil pengujian. Hasil dari pengujian Naïve Bayes dengan model Multinomial mendapatkan hasil terbaik pada komposisi data latih dan data uji 10:90 yang menghasilkan recall score 38%, precision 90,4%, F1-score 53,5%, sedangkan untuk akurasi mendapatkan nilai terbaik pada komposisi 90:10 dengan hasil 75%. Sehingga bisa disimpulkan mengurangi data latihnya dapat meningkat hasil recall score, precision, dan F1-score, sedangkan untuk mendapatkan hasil akurasi yang tinggi bisa dengan mengurangi data uji nya.

Daftar Pustaka

[1] Mega Noveanto, Helen Sastypratiwi, dan Hafiz Muhardi, “Uji Akurasi Klasifikasi Emosi Pada Lirik Lagu Bahasa Indonesia” Jurnal Sistem dan Teknologi Informasi, vol. 10, no. 3, Juli, 2022.
[2] Luh Ristiari, AAIN Eka K., I P.G. Hendra S., Agus M., I D.M. Bayu A. D., dan I M. Widiartha “Klasifikasi Cerita Pendek Berbahasa Bali Berdasarkan Umur Pembaca dengan Metode Naive Bayes” Jurnal Elektronik Ilmu Komputer Udayana., vol. 10, no. 4, May 2022.
[3] Diajeng Ninda Armianti, Indriati dan Sigit Adinugroho “Klasifikasi Emosi Lagu Berdasarkan Lirik pada Teks Berbahasa Indonesia Menggunakan K-Nearest Neighbor dengan Pembobotan WIDF” Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer., vol. 3, no. 10, Oktober 2019.
[4] Wahyudi Hermanto, Budhi Irawan dan Casi Setianingsih “Klasifikasi Emosi Pada Lirik Lagu Menggunakan Algoritma Support Vector Machine Dan Optimasi Particle Swarm Optimization” e-Proceeding of Engineering, vol. 8, no. 5, Oktober 2021.
[5] Febrina Sarito Sinaga, Indriati dan Bayu Rahayudi “Klasifikasi Emosi Lirik Lagu Menggunakan Improved K-Nearest Neighbor dengan Seleksi Fitur dan BM25” Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer, vol. 3, no. 6, Juni 2019.

Halaman ini sengaja dibiarkan kosong

JNATIA Volume 1, Nomor 4, Agustus 2023 p-ISSN: 2986-3929
Jurnal Nasional Teknologi Informasi dan Aplikasinya

Klasifikasi Kematangan Buah Apel dengan Ekstraksi Fitur Haralick dan KNN

I Kadek Bagus Deva Diga Dana Putra[a1], I Ketut Gede Suhartana[a2]

[a]Program Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Udayana
Jalan Raya Kampus Udayana, Bukit Jimbaran, Kuta Selatan, Badung, Bali Indonesia
[1]bgsdeva999@gmail.com
[2]ikg.suhartana@unud.ac.id

Abstract

This research aims to classify the ripeness level of apple fruits based on texture features using the Haralick method and color features using histograms. A dataset of 76 apple fruit images was collected. In the preprocessing stage, the apple images were converted to grayscale, followed by the application of a median filter to remove salt and pepper noise, and histogram equalization to enhance image contrast. Texture features were extracted using the Haralick method to obtain contrast, correlation, energy, homogeneity, and entropy features. Color features were extracted using histograms to obtain mean, standard deviation, skewness, and kurtosis. A K-Nearest Neighbor (KNN) model with k = 6 was used for classification. The evaluation results showed an accuracy of 89.47%, precision of 93.75%, recall of 93.75%, and F1-score of 93.75%. This research indicates that texture and color features can effectively classify the ripeness level of apple fruits. Future research can explore more diverse datasets and parameter adjustments to further improve model performance.

Keywords*: apple fruit, ripeness classification, texture features, color features.*

1. Pendahuluan

Tingkat kematangan buah apel merupakan salah satu faktor penting yang mempengaruhi kualitas dan rasa buah apel. Tingkat kematangan buah apel dapat ditentukan berdasarkan warna dan tekstur kulitnya. Buah apel yang masih mentah biasanya memiliki warna hijau atau kuning dan tekstur kulit yang keras. Buah apel yang sudah matang biasanya memiliki warna merah atau oranye dan tekstur kulit yang lembut [1].

Salah satu penelitian yang dilakukan oleh [2], yaitu klasifikasi tingkat kematangan apel manalagi merupakan varietas apel yang populer di Indonesia. Salah satu keunikan apel manalagi adalah warna kulitnya ketika mentah dan matang cukup sulit dibedakan karena perbedaannya tidak terlalu signifikan. Apel yang masih mentah berwarna hijau, sedangkan apel yang sudah matang berwarna hijau kekuningan. Hal ini membuat orang awam cukup sulit untuk membedakannya [2].

Mendeteksi tingkat kematangan buah apel secara manual dapat menyebabkan kesalahan dan ketidakkonsistenan karena bergantung pada persepsi manusia. Oleh karena itu, diperlukan suatu sistem yang dapat mengklasifikasikan tingkat kematangan buah apel secara otomatis dan akurat berdasarkan citra digital buah apel. Sistem ini dapat membantu petani, pedagang, atau konsumen dalam menentukan dan memilih buah apel yang berkualitas dan sehat.

Salah satu cara untuk mengklasifikasikan tingkat kematangan buah apel berdasarkan citra digital adalah dengan menggunakan metode ekstraksi fitur haralick dan metode klasifikasi KNN. Metode ekstraksi fitur haralick adalah metode yang menghitung fitur tekstur dari citra berdasarkan matriks ko-kemunculan abu-abu (GLCM). Fitur tekstur dapat merepresentasikan karakteristik permukaan citra, seperti kehalusan, kekasaran, atau kehomogenan [3]. Metode klasifikasi KNN adalah metode yang mengklasifikasikan citra berdasarkan jarak terdekat antara citra uji dan citra latih. Metode KNN dapat menyesuaikan diri dengan data yang tidak linear dan memiliki kompleksitas

perhitungan yang rendah [4].

Penelitian ini bertujuan untuk mengimplementasikan dan menguji akurasi, presisi, recall, dan F1-score dari metode ekstraksi fitur haralick dan metode klasifikasi KNN untuk mengklasifikasikan tingkat kematangan buah apel berdasarkan citra digital buah apel. Penelitian ini diharapkan dapat memberikan kontribusi bagi pengembangan sistem pengenalan pola citra, khususnya dalam bidang pertanian.

2. Metode Penelitian

2.1. Pengumpulan Data

Data yang digunakan pada penelitian ini adalah data sekunder yang diambil dari situs Kaggle[5]. Data ini berisi 240 data latih dan 60 data uji citra buah-buahan yang terdiri dari buah apel, pisang, dan jeruk. Namun, akan diambil data buah apelnya saja untuk penelitian ini yang terdiri dari 76 data latih dan 19 data uji. Data ini memiliki resolusi 100x100 dan berformat jpg.

2.2. Pra-pemrosesan data

Pra-pemrosesan data dilakukan untuk meningkatkan kualitas citra dan mengurangi noise atau gangguan yang ada pada citra. Tahapan pra-pemrosesan data yang dilakukan adalah sebagai berikut:

a. Konversi citra RGB menjadi citra abu-abu (grayscale) dengan rumus:

$$\mathrm{I} = 0.299\mathrm{R} + 0.587\mathrm{G} + 0.114\mathrm{B} \quad (1)$$

b. Penerapan filter median untuk menghilangkan noise salt and pepper dengan ukuran kernel 3x3 piksel.

c. Penerapan histogram equalization untuk meningkatkan kontras citra dengan menggunakan fungsi bawaan dari OpenCV

2.3. Ekstraksi Fitur

Ekstraksi fitur dilakukan untuk mendapatkan fitur-fitur yang merepresentasikan karakteristik citra, seperti warna, tekstur, bentuk, dll. Fitur-fitur tersebut akan digunakan sebagai input untuk proses klasifikasi. Tahapan ekstraksi fitur yang dilakukan adalah sebagai berikut:

a. Ekstraksi fitur tekstur menggunakan metode Haralick. Metode ini menghitung fitur tekstur berdasarkan matriks ko-kemunculan abu-abu (GLCM) yang mengukur hubungan spasial antara piksel-piksel pada citra. Fitur tekstur yang dihasilkan oleh metode ini adalah:

   1. Contrast: mengukur variasi intensitas antara piksel-piksel yang berdekatan.
   2. Correlation: mengukur ketergantungan linear antara piksel-piksel yang berdekatan.
   3. Energy: mengukur keseragaman distribusi intensitas pada citra.
   4. Homogeneity: mengukur kemiripan antara piksel-piksel yang berdekatan.
   5. Entropy: mengukur ketidakteraturan atau kompleksitas pada citra.

b. Ekstraksi fitur warna menggunakan metode Color Histogram. Metode ini menghitung frekuensi kemunculan setiap nilai intensitas pada citra. Fitur warna yang dihasilkan oleh metode ini adalah:

   1. Mean: rata-rata nilai intensitas pada citra.
   2. Standard deviation: simpangan baku nilai intensitas pada citra.
   3. Skewness: ukuran kemiringan distribusi nilai intensitas pada citra.
   4. Kurtosis: ukuran keruncingan distribusi nilai intensitas pada citra.

2.4. Klasifikasi

Klasifikasi dilakukan untuk membedakan tingkat kematangan buah apel berdasarkan fitur-fitur yang telah diekstraksi. Metode klasifikasi yang digunakan adalah metode KNN. Metode ini mengklasifikasikan sebuah citra uji berdasarkan jarak terdekat dengan k citra latih yang memiliki label kelas yang sama. Jarak antara dua citra dihitung menggunakan rumus jarak Euclidean:

$$d = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2} \quad (2)$$

dimana d adalah jarak, n adalah jumlah fitur, xi adalah nilai fitur ke-i dari citra uji, dan yi adalah nilai fitur ke-i dari citra latih.

2.5. Evaluasi

Evaluasi dilakukan untuk mengukur kinerja dari metode klasifikasi yang digunakan. Metrik evaluasi yang digunakan adalah akurasi, presisi, recall, dan F1-score. Rumus-rumus untuk menghitung metrik evaluasi tersebut adalah sebagai berikut:

a. Akurasi: rasio antara jumlah citra yang diklasifikasikan dengan benar dengan jumlah total citra.

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (3)$$

b. Presisi: rasio antara jumlah citra positif yang diklasifikasikan dengan benar dengan jumlah total citra positif yang diprediksi.

$$\text{Precision} = \frac{TP}{TP+FP} \quad (4)$$

c. Recall: rasio antara jumlah citra positif yang diklasifikasikan dengan benar dengan jumlah total citra positif yang sebenarnya.

$$\text{Recall} = \frac{TP}{TP+FN} \quad (5)$$

d. F1-score: rata-rata harmonik dari presisi dan recall.

$$\text{F1-score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (6)$$

3. Hasil dan Diskusi

3.1. Persiapan Data

Pada dataset penelitian ini, penulis melakukan labeling secara manual, tetapi mengikuti literatur yang ada[2]. Menurut penelitian [2], apel yang tidak matang adalah apel yang masih berwarna kehijauan (tidak merah pekat). Oleh karena itu, penulis melakukan labeling manual dengan melihat warna apel. Warna apel yang kehijauan dilabeli dengan tidak matang (ditulis 0) dan warna apel yang sudah merah dilabeli dengan matang (ditulis 1).

Gambar 1. Contoh Apel Matang Pada Dataset

Gambar 2. Contoh Apel Yang Tidak Matang Pada Dataset

3.2. Prapemrosesan

Sebelum melakukan ekstraksi fitur dengan metode Haralick, diperlukan menghilangkan informasi warna pada dataset gambar terlebih dahulu. Hal tersebut karena metode Haralick adalah metode yang mengoperasikan tekstur pada gambar sehingga informasi warna tidak terlalu diperlukan pada pengolahan.

```
# List untuk menyimpan hasil prapemrosesan dari setiap gambar
preprocessed_images_test = []

# Loop melalui 76 gambar dalam dataset Anda
for i in range(1, 20):
    # Load image
    try:
        image_path = f"C:/Users/hp/OneDrive/Documents/Belajar Pemrograman/SNATIA/dataset/test_zip/test/apple_{i+76}.jpg"
        image = io.imread(image_path)

        # Periksa jika citra memiliki 4 channel, lalu konversi menjadi RGB
        if image.shape[2] == 4:
            image = color.rgba2rgb(image)

        # Konversi citra RGB menjadi citra abu-abu (grayscale)
        gray_image = color.rgb2gray(image)

        # Penerapan filter median
        median_filtered = median(gray_image, selem=np.ones((3, 3)))

        # Konversi gambar menjadi tipe data ubyte (0-255)
        median_filtered = img_as_ubyte(median_filtered)

        # Penerapan histogram equalization menggunakan OpenCV
        equalized_image = cv2.equalizeHist(median_filtered)

        # Tambahkan hasil prapemrosesan ke dalam list
        preprocessed_images_test.append(equalized_image)
    except:
        print('Seth happen')
```

Gambar 3. Kode Untuk Melakukan Penghilangan Informasi Warna Pada Data Test

Gambar 4. Contoh Salah Satu Gambar Yang Sudah Dihilangkan Informasi Warnanya

3.3. Ekstraksi Fitur

a. Ekstraksi Fitur Haralick

Setelah dataset dihilangkan informasi warnanya, sekarang bisa dilakukan pengekstraksian fitur dengan metode Haralick dan akan diekstrak fitur berupa contrast, correlation, energy, homogenety, dan entropy.

```python
# List untuk menyimpan fitur-fitur yang diekstraksi
features_haralick_test = []

# Loop melalui gambar-gambar hasil prapemrosesan
for image in preprocessed_images_test:
    # Hitung matriks GLCM dengan jarak 1 dan sudut 0 derajat
    glcm = greycomatrix(image, distances=[1], angles=[0], levels=256, symmetric=True, normed=True)

    # Ekstraksi fitur contrast
    contrast = greycoprops(glcm, 'contrast')[0, 0]
    features_haralick_test.append(contrast)

    # Ekstraksi fitur correlation
    correlation = greycoprops(glcm, 'correlation')[0, 0]
    features_haralick_test.append(correlation)

    # Ekstraksi fitur energy
    energy = greycoprops(glcm, 'energy')[0, 0]
    features_haralick_test.append(energy)

    # Ekstraksi fitur homogeneity
    homogeneity = greycoprops(glcm, 'homogeneity')[0, 0]
    features_haralick_test.append(homogeneity)

    # Ekstraksi fitur entropy
    entropy = -np.sum(glcm * np.log2(glcm + 1e-10))
    features_haralick_test.append(entropy)
```

Gambar 5. Kode Melakukan Ekstraksi Fitur Haralick Pada Data Test

Digunakan jarak 1 karena mempertimbangkan piksel yang berdekatan secara langsung. Ini mengakibatkan matriks GLCM mencerminkan hubungan spasial antara piksel-piksel yang saling berdekatan. Digunakan sudut 0 derajat karena mempertimbangkan pasangan piksel yang berada dalam satu arah horizontal (piksel di sebelah kanan).

	Contrast	Correlation	Energy	Homogeneity	Entropy
0	413.874867	0.961636	0.021384	0.279138	12.216877
1	20.967311	0.998223	0.135205	0.562925	9.079305
2	56.970746	0.996978	0.462538	0.749005	5.908961
3	336.483045	0.980058	0.499982	0.877345	3.893027
4	28.213824	0.998236	0.340881	0.773430	6.578070
5	222.498216	0.988583	0.490225	0.739776	6.005709

Gambar 6. Hasil Ekstraksi Fitur Haralick Pada 6 Gambar Data Test

b. Ekstraksi Fitur Color Histogram

Pada tahap ini penulis melakukan ekstraksi fitur warna dengan metode Color Histogram. Setelah fitur warna diekstraksi, fitur akan digabungkan antara fitur Haralick dan fitur warna dari Color Histogram yang nantinya akan dilakukan proses pemodelan dengan KNN.

```
# List untuk menyimpan fitur-fitur yang diekstraksi
features_color_test = []

# Loop melalui gambar-gambar dalam dataset Anda
for i in range(1, 20):
    try:
        # Load image
        image_path = f"C:/Users/hp/OneDrive/Documents/Belajar Pemrograman/SNATIA/dataset/test_zip/test/apple_{1+76}.jpg"
        image = cv2.imread(image_path)

        # Ekstraksi histogram warna
        hist = cv2.calcHist([image], [0, 1, 2], None, [8, 8, 8], [0, 256, 0, 256, 0, 256])
        hist = cv2.normalize(hist, hist).flatten()

        # Ekstraksi fitur mean
        mean = np.mean(hist)
        features_color_test.append(mean)

        # Ekstraksi fitur standar deviasi
        std = np.std(hist)
        features_color_test.append(std)

        # Ekstraksi fitur skewness
        skewness = moment(hist, moment=3)
        features_color_test.append(skewness)

        # Ekstraksi fitur kurtosis
        kurtosis = moment(hist, moment=4)
        features_color_test.append(kurtosis)
    except:
        print("ada error di nomor ",i)
```

Gambar 7. Kode Melakukan Ekstraksi Fitur Warna Data Test Dengan Color Histogram

	Mean	Standart	Skewness	Kurtosis
0	0.015879	0.041243	0.000290	0.000069
1	0.005670	0.043829	0.001061	0.000730
2	0.003889	0.044023	0.001568	0.001405
3	0.002635	0.044116	0.001908	0.001892
4	0.004438	0.043971	0.001637	0.001509
5	0.003311	0.044070	0.001831	0.001782

Gambar 8. Hasil Ekstraksi Fitur Warna Dengan Color Histogram Pada 6 Gambar Data Test

```
X_train = pd.concat([df_haralick_train, df_color_train], axis=1)
X_test = pd.concat([df_haralick_test, df_color_test], axis=1)
```

Gambar 9. Kode Untuk Melakukan Penggabungan Fitur Pada Data Test Dan Train

3.4. Pemodelan

Setelah fitur dari setiap data testing dan data training sudah digabungkan. Sekarang data siap untuk dimodelkan. Pemodelan akan dilakukan dengan metode KNN dengan k adalah 6. Pemilihan k bernilai 6 karena itu adalah hasil yang paling optimal pada kasus penelitian ini. Didapatkan hasil akurasi sebesar 89,47%, presisi 93,75%, recall 93,75%, dan F1-score sebesar 93,75%.

```
# Membuat model KNN
knn = KNeighborsClassifier(n_neighbors=6)

# Melatih model KNN
knn.fit(X_train, y_train)

# Memprediksi kelas untuk data uji
y_pred = knn.predict(X_test)

# Menghitung akurasi
accuracy = accuracy_score(y_test, y_pred)

# Menghitung presisi
precision = precision_score(y_test, y_pred)

# Menghitung recall
recall = recall_score(y_test, y_pred)

# Menghitung F1-score
f1 = f1_score(y_test, y_pred)
```

Gambar 10. Kode Melakukan Pemodelan dengan KNN

```
Akurasi: 0.8947368421052632
Presisi: 0.9375
Recall: 0.9375
F1-score: 0.9375
```

Gambar 11. Hasil Akurasi, Presisi, Recall, Dan F1-Score

4. Kesimpulan

Dari penelitian ini, dapat diambil kesimpulan sebagai berikut:

a. Akurasi: Model KNN dengan menggunakan fitur-fitur Haralick dan histogram warna menghasilkan tingkat akurasi sebesar 89,47%. Hal ini menunjukkan bahwa model dapat melakukan klasifikasi dengan akurasi yang cukup baik.
b. Presisi: Presisi sebesar 93,75% menunjukkan bahwa model KNN memiliki kemampuan untuk mengidentifikasi dengan baik buah apel yang matang dari dataset yang digunakan.
c. Recall: Recall sebesar 93,75% menunjukkan bahwa model memiliki kemampuan untuk secara efektif mengenali dan menemukan buah apel yang matang dalam dataset.

d. F1-score: F1-score sebesar 93,75% merupakan ukuran gabungan antara presisi dan recall. Skor ini mengindikasikan bahwa model KNN memiliki keseimbangan yang baik antara kemampuan untuk mengenali dan mengklasifikasikan dengan benar buah apel yang matang.

Dalam penelitian ini, model KNN dengan k = 6 digunakan untuk klasifikasi. Hasil metrik yang diperoleh menunjukkan bahwa model ini berhasil dalam mengklasifikasikan kematangan buah apel berdasarkan fitur tekstur dan fitur warna yang diekstraksi. Namun, perlu diingat bahwa kesimpulan ini berdasarkan pada dataset yang digunakan dalam penelitian ini, yaitu dataset 76 gambar. Untuk menggeneralisasikan kesimpulan ini, penelitian lebih lanjut dengan menggunakan dataset yang lebih besar dan beragam dapat dilakukan. Selain itu, juga perlu mempertimbangkan parameter-parameter lain seperti pemilihan fitur, pemrosesan citra, dan pengaturan parameter pada model KNN untuk memperbaiki dan meningkatkan performa model lebih lanjut

Daftar Pustaka

[1] Hanif, A. I., Razka, H., & Wiratomo, D. S. (2021). Klasifikasi tingkat kematangan buah apel berdasarkan fitur warna menggunakan algoritma K-Nearest Neighbor dan ekstraksi warna HSV. Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer E-ISSN: 2548-964X Vol. 5 No. 6 Mei 2021 Hlm: 2188-2199.
[2] Ciputra, A., De Rosal, I. M., Setiadi, I. M., Rachmawanto, E. H., & Susanto, A. (2018). Klasifikasi tingkat kematangan buah apel manalagi dengan algoritma naive bayes dan ekstraksi fitur citra digital. Jurnal Ilmiah Teknik Elektro Komputer Dan Informatika (JITEKI), 4(1), 1-10.
[3] Wijaya, N., & Ridwan, A. (2019). Klasifikasi jenis buah apel dengan metode K-nearest neighbors dengan ekstraksi fitur HSV dan LBP. Jurnal Sisfokom (Sistem Informasi Dan Komputer), 8(1), 1-10.
[4] Yusuf, T., Prahudaya, A., & Harjoko, A. (2017). Metode klasifikasi mutu jambu biji menggunakan KNN berdasar fitur warna dan tekstur. Jurnal Ilmiah Teknologi Informasi Asia, 11(2), 125-136.
[5] Kinaci, M. B. (2018). Fruit images for object detection. Kaggle.com. Retrieved from https://www.kaggle.com/mbkinaci/fruit-images-for-object-detection

JNATIA Volume 1, Nomor 4, Agustus 2023 p-ISSN: 2986-3929
Jurnal Nasional Teknologi Informasi dan Aplikasinya

Identifikasi Salt and Pepper Noise pada Citra dengan Metode Median Filter

Steven Gerald Parsaoran Berutu[a1], I Komang Ari Mogi[a2]

[a]Program Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Udayana
Jalan Raya Kampus Udayana, Bukit Jimbaran, Kuta Selatan, Badung, Bali Indonesia
[1]stevenberutu31@gmail.com
[2]arimogi@unud.ac.id

Abstract

Noise is a common issue in image processing that can degrade the quality and clarity of images. In this research, we propose a method for identifying the levels of salt and pepper noise in images using the median filter technique. The median filter is a non-linear filtering approach that effectively reduces noise by replacing the pixel values with the median value of neighboring pixels. In our approach, we apply the median filter to the image and observe the changes in the pixel values. By analyzing the differences between the original image and the filtered image, we can determine the extent of salt and pepper noise present in the image. The proposed method offers a reliable and efficient way to quantify the level of salt and pepper noise. Experimental results on various images demonstrate the effectiveness of the proposed method in accurately identifying and quantifying salt and pepper noise levels. The method provides valuable insights into the amount of noise present in images, enabling better understanding and further processing of the image data.

Keywords*: Distortion, Smoothing, Enhancement, Detection, Measurement.*

1. Pendahuluan

Noise dalam citra merupakan permasalahan yang umum dalam pemrosesan citra yang dapat mengurangi kualitas dan kejelasan gambar. Identifikasi tingkat noise menjadi langkah penting dalam pemrosesan citra untuk memahami dan mengatasi masalah ini. Dalam penelitian ini, kami fokus pada pengembangan metode identifikasi tingkat salt and pepper noise menggunakan metode median filter. Metode median filter telah terbukti efektif dalam mengurangi noise dengan menggantikan nilai piksel dengan nilai median dari piksel-piksel tetangga. Kami bertujuan untuk mengembangkan pendekatan sederhana namun efektif untuk mengidentifikasi tingkat salt and pepper noise dalam citra dengan akurasi yang tinggi. Penelitian ini diharapkan dapat memberikan kontribusi penting dalam pemrosesan citra dengan menghadirkan metode yang efektif untuk identifikasi tingkat salt and pepper noise. Dengan menerapkan metode median filter pada citra dan menganalisis perubahan nilai piksel, kami berharap dapat menentukan sejauh mana tingkat salt and pepper noise yang terdapat dalam citra. Identifikasi tingkat noise yang akurat akan memberikan dasar yang kuat untuk langkah-langkah selanjutnya dalam penghilangan noise atau analisis lebih lanjut terkait kualitas citra. Hasil penelitian ini dapat digunakan dalam berbagai aplikasi pemrosesan citra, seperti pengolahan citra medis, analisis satelit, dan pengolahan citra komputer, yang membutuhkan identifikasi tingkat salt and pepper noise untuk meningkatkan kualitas citra secara keseluruhan.

1.1 Citra Digital

Citra digital adalah representasi digital dari gambar yang terdiri dari piksel-piksel dengan nilai intensitas yang menggambarkan tingkat kecerahan atau warna pada setiap piksel. [1] Citra ini dikodekan menggunakan bilangan biner dan dapat diproses serta dianalisis menggunakan berbagai teknik pemrosesan citra. Pemahaman tentang citra digital penting dalam mengembangkan metode dan teknik pemrosesan citra yang efektif. Sebuah citra digital terdiri

dari beberapa elemen, yaitu kecerahan yang mengacu pada tingkat cahaya yang dipancarkan oleh piksel dan dapat diterima oleh sistem penglihatan manusia. Kemudia terdapat kontur yang akan terjadi ketika terdapat perbedaan intensitas cahaya yang signifikan antara piksel-piksel yang berdekatan. Lalu kontras yang merupakan ukuran sebaran antara area terang dan gelap dalam sebuah citra. Dan yang terakhir terdapat warna yang merupakan hasil dari persepsi sistem visual terhadap panjang gelombang cahaya yang dipantulkan oleh objek. [2]

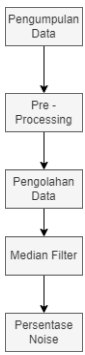
1.2 Noise

Noise adalah gangguan yang terjadi pada data gambar digital akibat penyimpanan dan transmisi, yang dapat merusak kualitas citra. Gangguan ini bisa disebabkan oleh interferensi fisik (seperti debu pada lensa kamera) atau kesalahan dalam proses pengolahan. Konvolusi adalah proses di mana citra dimanipulasi dengan menggunakan masker eksternal atau subwindow untuk menghasilkan citra baru. Di sisi lain, filtering melibatkan modifikasi piksel berdasarkan piksel tetangganya tanpa menggunakan masker eksternal. [3]. Terdapat beberapa jenis noise yang pada umumnya terdapat pada sebuah citra digital. Gaussian noise yang intensitas pikselnya terdistorasi secara acak dengan nilai yang bervariasi. Kemudian terdapat salt and pepper noise yang ditandai dengna munculnya piksel yang memiliki intensitas sangat terang (salt) atau sangat gelap (pepper) secara acak dalam citra. Dan yang terakhir ada speckle noise yang bersifat multiplicative dan dapat menghasilkan titik titik kecil atau butiran diseluruh citra.

1.3 Median Filter

Metode median filter adalah filter non-linear yang dikembangkan oleh Tukey untuk mengurangi noise dan menghaluskan citra. Filter ini bekerja dengan mengurutkan nilai intensitas piksel dalam suatu wilayah tetangga dan menggunakan nilai median sebagai pengganti nilai piksel yang sedang diproses [5]. Proses median filter melibatkan penggeseran sebuah jendela atau wilayah dengan jumlah piksel ganjil pada seluruh daerah citra. Kemudian, nilai intensitas piksel-piksel dalam wilayah tersebut diurutkan secara ascending, dan nilai median (nilai tengah) dihitung. Nilai median tersebut akan menggantikan nilai piksel yang berada di pusat wilayah tersebut. Dengan cara ini, median filter mampu mengurangi noise salt and pepper yang muncul secara acak dalam citra.

2. Metode Penelitian



Gambar 1. Tahapan Penelitian

Berdasarkan gambar 1, tahap-tahap dalam penelitian ini terdiri atas pengumpulan data, pre-processing, pengolahan data, lalu data dimasukkan kedalam metode median filter dan akan menghasilkan persentase noise sebagai output penelitian ini.

a. Pengumpulan Data

Pada penelitian ini, data diambil secara secara sekunder. Dimana, data yang diambil adalah data yang tersedia di open dataset pada kaggle. Data yang digunakan merupakan data citra berwarna maupun grayscale yang berformat .jpg atau .png.

b. Pre-processing Data

Pada tahapan ini, data akan diubah menjadi skala abu abu. Hal tersebut dilakukan agar metode median filter dapat diterapkan dengan baik pada intensitas piksel.

c. Metode Filter Median

Metode median filter dapat digunakan dalam identifikasi dan pengurangan salt and pepper noise dalam citra. Salt and pepper noise adalah jenis noise yang ditandai dengan munculnya piksel-piksel dengan intensitas yang sangat rendah (salt) atau sangat tinggi (pepper) secara acak dalam citra.

Proses identifikasi tingkat salt and pepper noise melibatkan langkah-langkah berikut:

1. Memilih ukuran kernel. Ukuran kernel harus berdimensi ganjil, misalnya 3x3, 5x5, atau 7x7. Agar terdapat piksel tengah yang digunakan sebagai media.
2. Melakukan median filtering. Median filter diterapkan dengan memindai setiap piksel, dan wilayah tetangga yang sesuai dengan kernel akan dipilih.
3. Menghitung nilai median. Nilai median akan dihitung dengan mengurutkan intensitas piksel piksel tetangga dan memilih nilai median sebagai pengganti untuk piksel yang sedang diproses.
4. Mengganti nilai piksel. Nilai piksel yang terganggu oleh salt and pepper noise digantikan oleh nilai median yang telah dihitung.

Setelah melalui proses median filtering, tingkat salt and pepper noise dapat diidentifikasi dengan melihat perbedaan antara citra asli dan citra yang telah difilter. Piksel-piksel yang awalnya terganggu oleh noise akan memiliki intensitas yang lebih seragam dan lebih dekat dengan lingkungan sekitarnya setelah median filtering. Kemudian, tingkat salt and pepper diukur dengan menggunakan metrik seperti persentase piksel yang terkena noise atau menggunakan metide analisis histogram untuk melihat distribusi intensitas piksel.

3. Hasil dan Diskusi

3.1. Persiapan Data

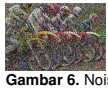
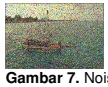
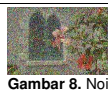
Dataset dari penelitian ini terdiri dari 2 jenis yaitu data latih dan data uji. Untuk data latih akan diambil dari dataset open-source yang tersedia di Kaggle pada link https://www.kaggle.com/datasets/kevinismail/salt-and-pepper-noise-dataset. Data latih yang digunakan merupakan data citra grayscale yang memiliki salt and pepper noise. Data latihnya diambil sebanyak 20 data secara acak dari dataset tersebut. Kemudian untuk data latih nya menggunakan dataset yang bersifat open source juga yang tersedia di Kaggle pada link https://www.kaggle.com/datasets/muhammadmasdar/shiftacademy-flower-classification. Data uji tersebut merupakan data citra bunga RGB yang belum diketahui apakah memiliki salt and pepper noise atau tidak. Data uji akan diambil sebanyak 10 data. Nantinya data uji ini akn diuji terlebih dahulu seberapa sebaran salt and pepper noise yang dimiliki. Kemudian akan dilakukan penambahan salt and pepper noise kedalam data uji tersebut untuk melihat apakah noise yang ditambahkan dapat diidentifikasi oleh metode filter median.

3.2. Implementasi Metode Median Filter

Pada tahapan impelentasi ini akan dilakukan pengujian terhadap 15 data citra yang terdiri dari 7 data citra grayscale yang memiliki salt and pepper noise tetapi belum diketahui intensitas noise nya. Dan 8 data citra RGB yang belum diketahui apakah memiliki salt and pepper noise atau tidak.

a. Data Citra Grayscale

Tabel 1. Intensitas Noise Pada Citra Grayscale

Citra ke-	Citra	Persentase Noise
1.	 **Gambar 2.** Noise 1	33,06 %
2.	 **Gambar 3.** Noise 2	17,59 %
3.	 **Gambar 4.** Noise 3	29,44 %
4.	 **Gambar 5.** Noise 4	23,74 %
5.	 **Gambar 6.** Noise 5	26,12 %
6.	 **Gambar 7.** Noise 6	33,02 %
7.	 **Gambar 8.** Noise 7	33,31 %

b. Data Citra RGB

Tabel 2. Intensitas Noise Pada Citra RGB

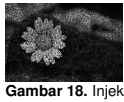
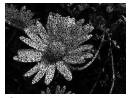
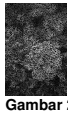
Citra ke-	Citra	Persentase Noise
1.	 **Gambar 9.** RGB 1	6,24 %
2.	 **Gambar 10.** RGB 2	3,78 %
3.	 **Gambar 11.** RGB 3	8,13 %
4.	 **Gambar 12.** RGB 4	10,35 %
5.	 **Gambar 13.** RGB 5	9,79 %
6.	 **Gambar 14.** RGB 6	2,97 %
7.	 **Gambar 15.** RGB 7	17,10 %
8.	 **Gambar 16.** RGB 8	9,16 %

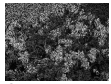
Tabel diatas menunjukkan intensitas salt and pepper noise yang dimiliki oleh citra RGB yang sebelumnya belum diketahui apakah citra tersebut memiliki salt and pepper noise. Selanjutnya akan dilakukan proses injeksi salt and pepper noise. Injeksi ini bertujuan untuk melihat, apakah kadar salt and pepper noise yang terkandung dalam citra RGB akan bertambah setelah dilakukan injeksi tersebut.

c. Data Citra RGB dengan Injeksi Salt and Pepper Noise

Injeksi salt and pepper noise ini dilakukan untuk menambahkan salt and pepper noise tersebut pada citra. Noise yang akan diinjeksi akan dikontrol dengan noise ratio yang menggmabrkan persentase atau proporsi piksel yang akan diubah nilainya menajdi noise dalam citra. Misalnya, jika noise ratio diatur menjadi 0,1 artinya 10% dari piksel pada citra akan diubah nilainya menjadi salt and pepper noise. Pada tahapan kali ini noise ratio yang akan diberikan sejumlah 0,5 untuk setiap citra RGB.

Tabel 3. Intensitas Noise Pda Citra RGB Dengan Injeksi Noise

Citra ke-	Citra	Persentase Noise
1.	 **Gambar 17.** Injeksi 1	23,06 %
2.	 **Gambar 18.** Injeksi 2	12,58 %
3.	 **Gambar 19.** Injeksi 3	16, 14 %
4.	 **Gambar 20.** Injeksi 4	33,37 %
5.	 **Gambar 21.** Injeksi 5	29,73 %
6.	 **Gambar 22.** Injeksi 6	24,59 %

Citra ke-	Citra	Persentase Noise
7.	 **Gambar 23.** Injeksi 7	24,03 %
8.	 **Gambar 24.** Injeksi 8	24,66 %

4. Kesimpulan

Penelitian tentang identifikasi salt and pepper noise pada citra dengan metode median filter ini menunjukkan bahwa metode median filter efektif dalam mengidentifikasi salt and pepper noise pada citra. Dengan menggunakan tabel persentase noise pada citra grayscale yang memiliki salt and pepper noise dan citra RGB yang belum diketahui apakah memngandung salt and pepper noise, metode median filter mampu mengidentifikasi dengan akurasi yang tinggi piksel-piksel yang terkena noise tersebut. Selanjutnya, penelitian dilanjutkan dengan melakukan injeksi salt and pepper noise pada citra RGB dan dilakukan analisis persentase noise. Hasilnya menunjukkan bahwa setelah dilakukan injeksi noise, persentase noise pada citra RGB meningkat, yang mengindikasikan kemampuan metode median filter dalam mengidentifikasi kehadiran salt and pepper noise pada citra RGB. Penelitian ini memberikan kontribusi penting dalam pengembangan metode identifikasi noise pada citra dan memberikan dasar yang kuat untuk langkah-langkah selanjutnya dalam pemrosesan dan analisis citra terkait kualitas citra.

Daftar Pustaka

[1] N.A. Pulung, et al., 2017, “Pengolahan Citra Digital”, Andi, Yogyakarta. Pp. 1.
[2] Simangunsong, P. B. N., 2018, “Peningkatan Kualitas Citra pada Studio Photography Dengan Menggunakan Metode Gaussian Filter”, Jurnal Teknik Informatika Unika St. Thomas (JTIUST), Vol. 3, No. 1.
[3] Simangunsong, P. B. N., 2017, “Reduksi Noise pada Citra Digital Menggunakan Metode Arithmatic Mean Filter”, Jurnal Teknik Informatika Unika St. Thomas (JTIUST), Vol. 2, No. 2, pp. 61.
[4] Wijaya Chandra, M., Tjiharjadi, S., “Mencari Nilai Threshold yang Tepat Untuk Perancangan Pendeteksi Kanker Trofoblas”, Seminar Nasional Aplikasi Teknologi Informasi, pp. C-29, Bandung.
[5] Maulana, I., N. A. Pulung, 2016, “Analisa Perbandingan Adaptif Median Filter dan Median Filter Dalam Reduksi Noise Salt & Pepper”, Cogito Smart Journal, Vol. 2, No.2, pp 160 – 161, Semarang.

Halaman ini sengaja dibiarkan kosong

JNATIA Volume 1, Nomor 4, Agustus 2023 p-ISSN: 2986-3929
Jurnal Nasional Teknologi Informasi dan Aplikasinya

Perbandingan Berbagai Metode Segmentasi dan Mechine Learning pada Makanan Tradisional Sumatera Utara

Anugrah Ignatius Sitinjak[a1], I Made Widiartha[a2]

[a]Program Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Udayana
Jalan Raya Kampus Udayana, Bukit Jimbaran, Kuta Selatan, Badung, Bali Indonesia
[1]ignatiusanugrah03@email.com
[2]madewidiartha@unud.ac.id (Corresponding Author)

Abstract

This study investigates the categorization of traditional North Sumatran dishes using various segmentation methods. The goal is to participate and participate in the preservation of North Sumatran culture. The study covers 34 types of traditional North Sumatran dishes originating from various regions. Food images are processed using segmentation techniques such as Sobel, Prewitt, Robert, Scharr, and Canny filters. The data set is then used in traditional machine learning algorithms, including Random Fortst, Decision Tree, and four SVM algorithms, for classification purposes. Among the algorithms with the highest performance, the Random Forest algorithm with Robert's segmentation method achieves outstanding results on dataset testing, with 85.52% accuracy, 84.63% recall, 83.77% precision, and 82.49% f1 score. The execution time for most of the best performing algorithms is around 1 minute on average. In addition, the Random Forest algorithm with the Canny operator achieves 81.51% accuracy, 84.97% recall, 86.81% precision, and 85.61% f1 score on dataset testing. The Random Forest algorithm with the Sobel operator obtains an accuracy of 78.41%, a recall of 65.28%, a precision of 62.33%, and an f1 score of 63.71%. Among the four SVM algorithms, the Sigmoid SVM with the Scharr operator achieves the highest performance in its category across all classification metrics. The importance of insight into the traditional cuisine of North Sumatra is invaluable. Emphasizing the importance of this research in promoting the preservation and introduction of traditional North Sumatran food.

Keywords: *North Sumatera, food, categorization*

1. Pendahuluan

Sumatera Utara adalah salah satu provinsi di Indonesia memiliki kekayaan budaya yang cukup besar. Salah satu aspek yang memperkaya keunikan provinsi ini adalah ragam makanan tradisionalnya yang lezat dan memikat. Makanan tradisional Sumatera Utara mencerminkan keanekaragaman suku dan etnis yang ada di daerah ini, seperti suku Batak, Melayu, Nias, dan banyak lagi. Dalam era globalisasi dan kemajuan teknologi, promosi budaya dan pariwisata telah menjadi salah satu fokus utama pemerintah daerah untuk meningkatkan kunjungan wisatawan. Dalam hal ini, penggunaan teknologi menjadi sangat relevan dan memiliki potensi besar untuk meningkatkan promosi dan pengenalan makanan tradisional Sumatera Utara. Dalam penelitian ini, kami mengusulkan penggunaan Jaringan Syaraf Konvolusional (Convolutional Neural Network/CNN) dan metode segmentasi untuk memperkenalkan makanan tradisional Sumatera Utara secara efektif kepada khalayak. Jaringan syaraf konvolusional merupakan salah satu metode dalam bidang kecerdasan buatan yang mampu melakukan klasifikasi dan pengenalan pola pada data gambar. Sedangkan metode segmentasi digunakan untuk memisahkan makanan dari latar belakang gambar secara otomatis, memungkinkan gambar makanan tampil dengan lebih jelas dan menarik. Dengan menggunakan teknologi ini, diharapkan promosi budaya dan pariwisata Sumatera Utara dapat meningkat dengan cara yang lebih menarik dan informatif. Dengan melakukan pengenalan makanan tradisional secara

visual, wisatawan dapat memperoleh pemahaman yang lebih baik tentang kekayaan kuliner daerah ini dan terinspirasi untuk menjelajahi lebih lanjut. Untuk mengatasi hal ini, para peneliti telah bekerja untuk mengkategorikan dan mengklasifikasikan makanan ini untuk memberikan pemahaman menyeluruh tentang warisan kuliner Sumatera Utara yang kaya. Melalui penelitian ini, kami berharap bahwa hasil yang diperoleh dapat menjadi sumbangan penting dalam mempromosikan budaya dan pariwisata Sumatera Utara, serta memberikan manfaat bagi pengembangan teknologi pengenalan gambar dan aplikasinya dalam konteks yang lebih luas [1].

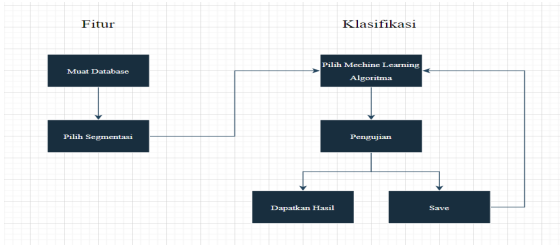
2. Metode Penelitian

2.1. Dataset

Penelitian ini berfokus pada pelatihan algoritma dan evaluasi dataset uji berdasarkan pengujian dari fitur-fitur. Di sini, 34 kelas diperkenalkan: Mie Gomak, Lontong Medan, Soto Medan, Lemang, Bika Ambon, Bihun Bebek Medan, Daun Ubi Tumbuk, Kari Bihun, Mie Balap, Rujak Kolam, Martabak Piring, Gulai Ikan Mas, Laksa Medan, Arsik, Naniura , Pelleng, Lappet, Manuk Napinadar, Babi Panggang Karo, Ayang Pakis, Sayur Gurih Tauco, Putu Bambu, Natinombur, Kolak Durian, Sate Kerang, Sambal Tuktuk, Dali ni Horbo, Kacang Sihobuk, Ssgun, Tanggo-tanggo, Gulai Kuta-Kuta, Manuk Gota, Kidu Kidu, Cimpa, dan Saksang sebanyak 34 label kelas. Kemampuan untuk menggeneralisasi 34 penting untuk algoritma pembelajaran mesin tradisional untuk mengekstraksi karakteristik yang diperlukan.

Gambar 1. Makanan Khas Sumatera Utara

Proses berlanjut ke klasifikasi menggunakan algoritma pembelajaran mesin tradisional. Diagram alir penelitian ini dapat dilihat di bawah ini.



Gambar 2. Flowchart Penelitian

a. Sobel

Operator Sobel, dikenal sebagai operator Sobel-Feldman atau filter Sobel, adalah teknik perbaikan citra yang diciptakan oleh Irwin Sobel dan Gary Feldman. Tujuannya adalah untuk menekankan tepi gambar dengan memperkirakan gradien intensitas gambar. Ini dilakukan dengan menerapkan filter kecil, dapat dipisahkan, dan bernilai bilangan bulat ke gambar dalam arah horizontal (Gx) dan vertikal (Gy).

$$G_x = \begin{bmatrix} +1 & 0 & -1 \\ +2 & 0 & -2 \\ +1 & 0 & -1 \end{bmatrix} * A \qquad G_y = \begin{bmatrix} +1 & +2 & +1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} * A \tag{1}$$

Selain itu, operator Sobel efisien dalam hal sumber daya komputasi, menjadikannya pilihan yang hemat biaya. Jika kita menganggap A sebagai citra, persamaan tersebut mengilustrasikan proses penerapan operasi konvolusi ke setiap piksel dalam citra. Operasi ini melibatkan melakukan perhitungan sepanjang sumbu horizontal dan vertikal, yang diwakili oleh simbol *.

$$G_x = \sqrt{{G_x}^2 + {G_y}^2} \tag{2}$$

Untuk memperoleh perkiraan gradien, persamaan yang disebutkan sebelumnya melibatkan perhitungan akar kuadrat dari jumlah komponen kuadrat dari gradien horizontal dan vertikal. Pemanfaatan operator Sobel telah menguntungkan dalam berbagai aplikasi seperti pengenalan sidik jari, pembuatan peta tepi satelit yang realistis, dan peningkatan visi robotika [4]. Dengan mengintegrasikan filter Sobel ke dalam algoritma deteksi, pemrosesan gambar dapat dipercepat, menghasilkan waktu pengambilan yang lebih cepat. Hal ini dicapai dengan menggunakan Gray Level Co-Occurrence Matrix (GLCCM) di samping filter [5]. Selain itu, filter disarankan untuk analisis citra radar jarak jauh dari berbagai perspektif, sehingga meningkatkan kinerja pendekatan berbasis histogram [5].

b. Prewitt

Metode Prewitt menggunakan operasi konvolusi dengan kernel Prewitt pada citra. Kernel Prewitt terdiri dari dua matriks, yaitu kernel horizontal dan kernel vertikal. Kernel horizontal pada metode Prewitt memiliki elemen matriks yang memberikan bobot yang lebih tinggi pada piksel-piksel di sekitar sumbu horizontal, sedangkan kernel vertikal memberikan bobot yang lebih tinggi pada piksel-piksel di sekitar sumbu vertikal. Sebuah studi baru telah mengusulkan metode yang menggabungkan Prewitt dengan representasi kuantum canggih yang disebut NEQR, yang telah menunjukkan keefektifan yang mengesankan dalam mengekstraksi tepi. Pendekatan ini telah ditemukan menjadi algoritma yang sangat efisien [7]. Ketika diterapkan pada gambar medis yang disimulasikan, khususnya data MRI, Prewitt telah menunjukkan kemampuan pencarian lokal yang kuat dan kecepatan konvergensi yang cepat. Penerapan filter ini telah menghasilkan tingkat akurasi yang luar biasa sebesar 100% saat diuji pada berbagai dataset tumor otak yang dapat diakses secara terbuka seperti BMIBTB, CE-MRI, dan FSB [8].

c. Robert

Pada tahun 1963, Lawrence Roberts mengusulkan operator Robert, sebuah metode untuk memperkirakan gradien citra melalui diferensiasi diskrit. Teknik ini melibatkan evaluasi varians kuadrat antara sumbu horizontal dan vertikal, lalu menjumlahkannya untuk memperkirakan gradien gambar.

$$\begin{bmatrix} +1 & 0 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} 0 & +1 \\ -1 & 0 \end{bmatrix} \tag{3}$$

Pertama, gambar yang diberikan mengalami konvolusi menggunakan filter, yaitu Gx untuk sumbu horizontal dan Gy untuk sumbu vertikal. Setelah itu, gradien dihitung dengan menggunakan persamaan berikut. Nilai gradien yang diperoleh kemudian digunakan sebagai I(x,y) untuk memperkirakan gradien citra. Persamaan yang disarankan Lawrence Robert, ketika digunakan bersamaan dengan Balance Contrast Enhancement Technique (BCET), terbukti efisien dalam meningkatkan karakteristik citra medis [9]. Integrasi operator Robert dengan algoritma aturan fuzzy Mamdani (Tipe-2) untuk deteksi pembuluh darah, bersama dengan Penyetaraan Histogram Adaptif Kontras-Terbatas (CLAHE), membantu dalam partisi gambar retina dengan mengekstraksi informasi gradiennya. Metodologi ini memberikan kontribusi yang signifikan untuk mencapai penentuan ambang batas yang tepat dan pada akhirnya meningkatkan akurasi aplikasi secara keseluruhan [10].

d. Scharr

Scharr adalah metode alternatif dalam pengolahan citra untuk mendeteksi tepi atau perubahan intensitas yang tajam antara piksel-piksel dalam gambar. Metode ini serupa dengan metode Sobel dan Prewitt, tetapi menggunakan kernel Scharr yang memiliki respons lebih sensitif terhadap perubahan intensitas yang halus. Operator filter Scharr digunakan untuk mengidentifikasi dan menyorot tepi atau karakteristik gradien dalam gambar dengan memanfaatkan turunan pertama.

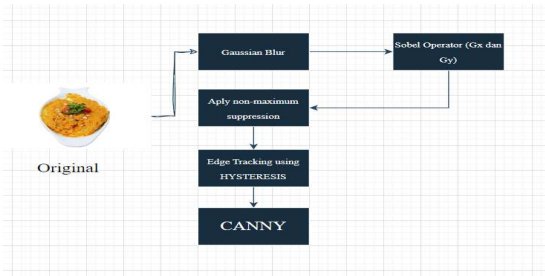
$$\begin{bmatrix} -3 & 0 & 3 \\ -10 & 0 & 10 \\ -3 & 0 & 3 \end{bmatrix} \quad \begin{bmatrix} -3 & -10 & -3 \\ 0 & 0 & 0 \\ 3 & 10 & 3 \end{bmatrix} \tag{4}$$

Mirip dengan teknik Sobel dan Prewitt, algoritma ini memperkirakan gradien gambar dengan menggabungkan gambar dengan filter kecil, dapat dipisahkan, dan berbasis bilangan bulat baik dalam arah horizontal (Gx) maupun vertikal (Gy). Dengan menerapkan filter ini ke nilai intensitas gambar, akan lebih mudah untuk mendeteksi tepi atau fitur gradient. Untuk mendapatkan estimasi gradien tersebut di atas, akar kuadrat dari setiap dimensi dimasukkan ke dalam persamaan sebagai berikut.

$$G = \sqrt{{G_x}^2 + {G_y}^2} \tag{5}$$

e. Canny

Metode Canny adalah salah satu metode yang populer dalam pengolahan citra untuk mendeteksi tepi dengan presisi tinggi. Metode ini dikembangkan oleh John F. Canny pada tahun 1986 dan sering digunakan dalam berbagai aplikasi pengolahan citra dan pengenalan pola. Deteksi tepi melibatkan penggunaan teknik matematis untuk mengidentifikasi titik-titik dalam gambar di mana ada pergeseran intensitas piksel yang signifikan, secara efektif mendeteksi batas objek yang ada dalam gambar. Ini memerlukan pendeteksian perubahan kecerahan dan salah satu pendekatan populer untuk deteksi tepi dikenal sebagai Canny Edge Detection.



Gambar 3. Teknik Canny pada foto yang menampilkan masakan asli Sumatera Utara

Penerapan teknik Gaussian blur melibatkan penggabungan gambar dengan kernel filter Gaussian untuk mengurangi adanya noise. Tingkat kekaburan yang diterapkan pada gambar ditentukan oleh ukuran kernel, yang ditunjukkan dengan dimensi seperti 3x3 atau 7x7. Ukuran kernel yang lebih besar menghasilkan efek buram yang lebih kuat. Prosedur ini membantu mengurangi fluktuasi kecil dalam intensitas piksel yang disebabkan oleh derau, sekaligus mempertahankan tepi penting yang ada dalam gambar.

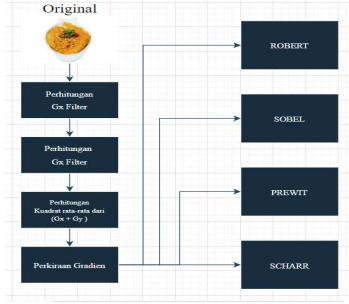
$$G(x,y) = \frac{1}{2\pi\sigma^2} \cdot \exp\left(-\frac{x^2+y^2}{2\sigma^2}\right) \quad (6)$$

Persamaan yang diberikan menunjukkan nilai filter Gaussian pada koordinat (x, y) yang diberikan, di mana σ mewakili standar deviasi distribusi Gaussian.

$$M = \sqrt{G_x^2 + G_y^2} \quad (7)$$

Perhitungan Gradien: Algoritma Canny menggunakan operator Sobel untuk menentukan kekuatan dan orientasi gradien pada setiap piksel. Prosedur ini memerlukan penerapan dua kernel konvolusi: satu untuk mendeteksi gradien horizontal (Gx) dan satu lagi untuk mengidentifikasi gradien vertikal (Gy). Penekanan Non-maksimum: Untuk meningkatkan akurasi tepi yang teridentifikasi, metode ini diterapkan. Teknik ini memerlukan pemeriksaan besaran gradien piksel dan membandingkannya dengan besaran piksel tetangga dalam arah gradien yang sama. Jika piksel memiliki magnitudo tertinggi di antara tetangganya, piksel tersebut dipertahankan. Namun, jika tidak memiliki magnitudo tertinggi, maka akan ditekan dan diberi nilai nol. Ambang Batas Ganda: Teknik ambang ganda mencakup penggunaan dua ambang batas pada besaran gradien. Ambang batas ini terdiri dari ambang tinggi dan ambang rendah. Piksel dengan besaran gradien lebih tinggi dari ambang batas tinggi dikenali sebagai piksel tepi yang kuat, sedangkan piksel dengan besaran lebih rendah dari ambang batas rendah dianggap sebagai piksel non-tepi dan diabaikan. Piksel dengan besaran yang berada di antara dua ambang dikategorikan sebagai piksel tepi lemah.Pelacakan Tepi dengan Histeresis: Tahap selanjutnya melibatkan penautan piksel yang kurang menonjol di sepanjang tepi ke piksel yang lebih menonjol. Ini dimulai dengan piksel tepi yang menonjol dan berlanjut dengan mengikuti jalur piksel tepi yang kurang menonjol yang terhubung di dalam area terdekat. Proses penelusuran ini berlanjut hingga semua piksel tepi yang kurang menonjol yang tersisa terhubung. Akibatnya, kumpulan garis tepi yang tidak terputus tercapai. Algoritma Canny menggabungkan filter median dan filter Gaussian untuk secara efektif menghilangkan noise dari gambar sambil mempertahankan detail penting dan menghadirkan kekaburan halus [15]. Ini banyak digunakan untuk deteksi tepi dan

menawarkan fleksibilitas dalam menyesuaikan kualitas keluaran dengan memanipulasi parameter deteksi tepi [16]. Dalam penelitian medis, seperti analisis tuberkulosis, penggunaan filter Canny untuk mengelompokkan gambar paru-paru telah menunjukkan akurasi sebesar 93,59% dan akurasi sensitivitas sebesar 92,31% [17]. Dalam pengawasan video, mengintegrasikan algoritma Canny dengan transformasi Hough pada platform seperti Hadoop dan Spark memberikan kinerja dan skalabilitas yang sangat baik untuk memproses kumpulan data besar [18]. Integrasi ini memungkinkan deteksi jalur yang efisien, mengaktifkan pelacakan, pemantauan, dan perekaman aktivitas. Metode segmentasi ini sangat penting untuk meningkatkan akurasi algoritma dengan secara akurat mengidentifikasi wilayah yang diminati dalam gambar. Kombinasi deteksi tepi Canny dan transformasi kontur dengan dekomposisi nilai singular digunakan untuk membuat skema watermarking yang kuat dan tidak terdeteksi. Algoritma yang diusulkan telah diuji dan ternyata tahan terhadap serangan umum [19]. Dengan menggabungkan YOLO dan Canny, akurasi 62,60% pada Mean Average Precision (mAP) dicapai dalam mendeteksi kecelakaan di jalan raya dalam kumpulan data tertentu. Model ini memiliki kemampuan untuk menggeneralisasi gambar tersegmentasi dan meningkatkan fitur yang dipelajari [20]. Canny filter digunakan dalam skema komputasi yang menjaga privasi untuk mempertahankan permutasi piksel yang aman selama protokol multiplikasi, memastikan keamanan yang ditingkatkan [21]. Dalam metode segmentasi api, Canny digunakan sebagai alat segmentasi utama untuk mengidentifikasi api pada suatu citra. Model yang diusulkan secara efektif menyoroti tepi dengan menekan kebisingan dan mencapai akurasi 98%, mengungguli operator Sobel [22]. Sebuah sistem parkir cerdas, yang dilatih menggunakan Canny, berhasil mendeteksi kendaraan yang masuk dan keluar dari tempat parkir melalui pemanfaatan servomotors dan sensor [23]. Selain itu, sistem pendeteksi banjir menggabungkan filter Canny dengan metode SCED (Sobel Canny Edge Detection) untuk memantau ketinggian air. Sistem ini menggunakan bola oranye untuk menunjukkan kenaikan permukaan air selama percobaan pendeteksian permukaan air [24].



Gambar 3. Pemanfaatan algoritma Sobel, Prewitt, Robert, dan Scharr pada foto masakan tradisional Sumatera Utara

Menggunakan teknik segmentasi seperti yang ditunjukkan pada gambar 3, algoritma dapat memisahkan dan menekankan wilayah penting, yang mengarah ke peningkatan ketepatan tugas selanjutnya.

2.2. Machine Learning

Untuk menganalisis karakteristik dari gambar yang menggambarkan masakan tradisional Sumatera Utara, digunakan teknik pembelajaran mesin konvensional. Tujuh algoritma yang sudah ada sebelumnya, termasuk Random Forest, Decision Tree, KNN, Linear SVM, Rbf SVM, Polynomial SVM, dan Sigmoid SVM, digunakan untuk mengevaluasi keefektifannya dalam mengklasifikasikan fitur yang relevan dari citra makanan. Algoritma pra-pelatihan ini diperoleh dari API perpustakaan scikit-learn melalui prosedur pengunduhan.

a. Random Forest

Adalah algoritma pembelajaran terawasi yang banyak digunakan dan cocok untuk tugas klasifikasi dan regresi. Ini beroperasi berdasarkan prinsip pembelajaran ansambel, yang melibatkan penggabungan beberapa pengklasifikasi untuk mengatasi masalah yang kompleks dan meningkatkan kinerja algoritma. Random Forest seperti namanya, mempekerjakan banyak pohon keputusan yang dilatih pada himpunan bagian yang berbeda dari kumpulan data yang diberikan, dan kemudian rata – rata prediksi mereka untuk meningkatkan akurasi prediksi secara keseluruhan. Daripada mengandalkan satu pohon keputusan, hutan acak mengumpulkan prediksi dari setiap pohon dan menentukan hasilnya berdasarkan suara mayoritas [26]. Dengan menambah jumlah pohon di hutan, akurasi yang lebih tinggi dapat dicapai sambil mengurangi risiko overfitting [27]. Random Forest sebagai algoritma yang ideal karena mampu menangkap hubungan yang kompleks antara variabel prediktor dan data yang diamati [25]. Ini menguntungkan Random Forest untuk dapat menangani kumpulan data besar dengan jumlah variabel prediktor yang besar. Setiap pohon keputusan dibangun menggunakan himpunan bagian ini, yang bertujuan untuk memaksimalkan perolehan informasi atau meminimalkan ketidakmurnian di setiap node. Selama prediksi, contoh baru dilewatkan melalui setiap pohon keputusan, dan prediksi akhir diperoleh dengan menggabungkan prediksi dari semua pohon melalui pemungutan suara mayoritas (untuk klasifikasi) atau rata-rata (untuk regresi). Secara matematis, prediksi ansambel untuk klasifikasi dapat direpresentasikan sebagai:

$$\hat{y} = \arg\max_{y} \sum_{i=1}^{N} I(y_i = y) \tag{8}$$

di mana y^ adalah kelas prediksi akhir, y mewakili kemungkinan label kelas, N adalah jumlah total pohon keputusan di Hutan Acak, y_i adalah prediksi pohon keputusan ke-i, dan I(y_i=y) adalah fungsi indikator yang mengembalikan 1 jika kondisinya benar dan 0 jika tidak [41]. Hutan acak, seperti namanya, mempekerjakan banyak pohon keputusan yang dilatih pada himpunan bagian yang berbeda dari kumpulan data yang diberikan, dan kemudian rata-rata prediksi mereka untuk meningkatkan akurasi prediksi secara keseluruhan. Daripada mengandalkan satu pohon keputusan, hutan acak mengumpulkan prediksi dari setiap pohon dan menentukan hasilnya berdasarkan suara mayoritas [26]. Dengan menambah jumlah pohon di hutan, akurasi yang lebih tinggi dapat dicapai sambil mengurangi risiko overfitting [27]. Random Forest sebagai algoritma yang ideal karena mampu menangkap hubungan yang kompleks antara variabel prediktor dan data yang diamati [25]. Ini menguntungkan Random Forest untuk dapat menangani kumpulan data besar dengan jumlah variabel prediktor yang besar [26][27].

b. Decision Tree

Algoritma Pohon Keputusan adalah teknik pembelajaran mesin populer yang membangun algoritma seperti pohon untuk membuat keputusan. Ini beroperasi secara top-down, mempartisi dataset secara rekursif berdasarkan fitur yang dipilih untuk membuat struktur pohon. Setiap simpul internal pohon mewakili uji fitur, sedangkan simpul daun sesuai dengan label kelas yang diprediksi atau nilai regresi. Pembangunan pohon keputusan melibatkan pencarian fitur terbaik dan ambang batas untuk pemisahan

pada setiap node internal. Ini dilakukan dengan mengevaluasi kriteria pemisahan yang berbeda, seperti Indeks Gini atau Penguatan Informasi, yang mengukur kualitas pemisahan berdasarkan ketidakmurnian atau pengurangan ketidakpastian yang diberikannya. Indeks Gini adalah kriteria pemisahan umum yang digunakan dalam pohon keputusan. Ini mengukur ketidakmurnian kumpulan data, di mana nilai yang lebih rendah menunjukkan distribusi kelas yang lebih homogen. Untuk masalah klasifikasi biner, Indeks Gini dapat dinyatakan sebagai:

$$Gini(D, c) = 1 - \sum_{i=1}^{c} (p_i)^2 \quad (9)$$

di mana Gini(D) adalah Gini Index dari dataset D, c adalah jumlah kelas, dan p_i mewakili probabilitas sebuah instance milik kelas i dalam dataset D. Indeks Gini dihitung untuk setiap fitur kandidat dan kombinasi ambang batas, dan fitur dan ambang batas yang meminimalkan pengotor dipilih untuk pemisahan [41]. Algoritma pohon keputusan berlanjut secara rekursif, membuat simpul anak untuk setiap kemungkinan hasil pemisahan. Proses ini diulangi sampai kriteria pemberhentian terpenuhi, seperti mencapai kedalaman maksimum, memiliki jumlah minimum contoh per daun, atau mencapai simpul daun murni di mana semua contoh termasuk dalam kelas yang sama.

c. Linear SVM

Algoritma Linear SVM adalah metode pembelajaran mesin populer yang digunakan untuk tugas klasifikasi biner. Ini bertujuan untuk menemukan hyperplane optimal yang memisahkan titik data milik kelas yang berbeda secara linear dipisahkan. Algoritma beroperasi dengan memaksimalkan margin, yaitu jarak antara hyperplane dan titik data terdekat dari masing-masing kelas. dan akurasi validasi yang melebihi 60% disimpan untuk analisis lebih lanjut. Algoritma yang dipilih ini selanjutnya digunakan dalam proses klasifikasi menggunakan algoritma pembelajaran yang diawasi. Diberikan dataset pelatihan dengan vektor fitur x_i dan label kelas yang sesuai y_i (di mana y_i∈\ {-1,1\}), tujuan SVM Linier adalah untuk menemukan vektor bobot optimal w dan suku bias b yang menentukan hyperplane pemisah. Fungsi keputusan dari Linear SVM didefinisikan sebagai:

$$f(\boldsymbol{x}) = \text{sign}(\boldsymbol{w} \cdot \boldsymbol{x} + b) \quad (10)$$

Masalah optimisasi untuk Linear SVM dapat dirumuskan sebagai berikut:

$$[y_i(\boldsymbol{w} \cdot \boldsymbol{x_i} + b) \geq 1, \quad \text{for all } i = 1,2, \dots, N] \quad (11)$$

di mana N adalah jumlah contoh pelatihan. Fungsi tujuan bertujuan untuk meminimalkan norma vektor bobot, mempromosikan margin yang besar antar kelas. Kendala memastikan bahwa titik data diklasifikasikan dengan benar dan berada di atas atau di luar margin. Linear SVM mengatasi masalah overfitting dengan mengidentifikasi hyperplane terbaik untuk secara efektif memisahkan titik data yang termasuk dalam kelas atau kategori yang berbeda [28]. Contoh penting dari kinerjanya ditunjukkan pada kumpulan data UNSW-NB15 baru-baru ini, di mana SVM mencapai tingkat akurasi yang mengesankan sebesar 93,75% [29]. Selain itu, ketika diterapkan pada kumpulan data yang berisi gambar kanker payudara dan paru-paru, algoritma SVM mencapai skor akurasi luar biasa sebesar 99% pada kumpulan data uji [31].

d. Sigmoid SVM

Algoritma Sigmoid SVM adalah varian dari Support Vector Machine yang menggunakan fungsi sigmoid sebagai fungsi keputusan, bukan fungsi linier. Ini biasanya digunakan untuk tugas klasifikasi biner dan dapat menangani data yang tidak dapat dipisahkan secara linier dengan memetakannya ke dalam ruang fitur berdimensi lebih tinggi.

Diberikan dataset pelatihan dengan vektor fitur x_i dan label kelas yang sesuai y_i (di mana y_i∈\{-1,1\}), tujuan dari SVM Sigmoid adalah untuk menemukan vektor bobot optimal w dan istilah bias b yang menentukan batas keputusan. Fungsi keputusan dari SVM Sigmoid didefinisikan sebagai:

$$[f(x) = \text{sign}\left(\sum_{i=1}^{N} \alpha_i y_i \sigma(x_i \cdot x + b_0) + b\right)] \quad (12)$$

di mana N adalah jumlah instance pelatihan, α_i adalah pengali Lagrange yang diperoleh selama proses pengoptimalan, σ (·) adalah fungsi sigmoid, x_i adalah vektor pendukung, dan b_0 dan b adalah suku bias. Masalah pengoptimalan untuk SVM Sigmoid melibatkan meminimalkan fungsi tujuan sehubungan dengan α_i dan b, tunduk pada kendala:

$$[0 \leq \alpha_i \leq C, \quad \sum_{i=1}^{N} \alpha_i y_i = 0] \quad (13)$$

di mana C adalah parameter regularisasi yang mengontrol trade-off antara memaksimalkan margin dan meminimalkan kesalahan pelatihan [41]. Algoritma SVM menunjukkan kinerja yang luar biasa, mencapai skor akurasi yang mengesankan sebesar 99% saat digunakan pada kumpulan gambar kanker payudara dan paru-paru [31]. Dalam bidang penelitian gangguan otak, SVM umumnya digunakan dalam analisis pola multivoxel (MVPA) karena kapasitasnya untuk menangani data pencitraan dimensi tinggi secara efektif, menghindari overfitting. Baru-baru ini, SVM menjadi semakin relevan dalam psikiatri presisi, khususnya dalam memprediksi diagnosis dan prognosis penyakit otak seperti penyakit Alzheimer, skizofrenia, dan depresi [32]. Di bidang ilmu ekonomi, penggunaan SVM untuk peramalan keuangan pada dataset deret waktu telah mendapat perhatian yang signifikan. Telah diamati bahwa algoritma SVM yang diusulkan melampaui algoritma lain dalam memprediksi nilai tukar dalam pasar keuangan, menunjukkan keefektifannya dalam mencapai kinerja peramalan yang unggul, yang sangat penting mengingat krisis ekonomi global baru-baru ini.

e. RBF SMV

Algoritma RBF SVM merupakan varian dari Support Vector Machine yang menggunakan kernel Radial Basis Function (RBF) sebagai dasar klasifikasi. Ini adalah algoritma yang kuat yang mampu menangani data yang dapat dipisahkan secara linear dan non-linear dengan memetakannya ke dalam ruang fitur berdimensi lebih tinggi. Diberikan dataset pelatihan dengan vektor fitur x_i dan label kelas yang sesuai y_i (di mana y_i∈\{-1,1\}), tujuan RBF SVM adalah untuk menemukan hyperplane optimal yang secara maksimal memisahkan titik data milik kelas yang berbeda. Fungsi keputusan dari RBF SVM didefinisikan sebagai:

$$[f(x) = \text{sign}\left(\sum_{i=1}^{N} \alpha_i y_i K(x_i, x) + b\right)] \quad (14)$$

di mana N adalah jumlah instance pelatihan, α_i adalah pengali Lagrange yang diperoleh selama proses optimasi, σ (·) adalah fungsi sigmoid, x_i adalah vektor pendukung, dan b_0 dan b adalah suku bias [41]. Masalah pengoptimalan untuk RBF SVM melibatkan meminimalkan fungsi tujuan sehubungan dengan α_i dan b, tunduk pada kendala:

$$[0 \leq \alpha_i \leq C, \quad \sum_{i=1}^{N} \alpha_i y_i = 0] \quad (15)$$

di mana C adalah parameter regularisasi yang mengontrol trade-off antara memaksimalkan margin dan meminimalkan kesalahan pelatihan. Menerapkan SVM ke set data COVID-19, para peneliti mencapai akurasi sensitivitas rata-rata 95,76% menggunakan gambar sinar-X format jpeg 512x512 piksel [35]. Dengan menggabungkan SVM, menerapkan high pass filter menggunakan Sobel untuk mendapatkan edge, mereka mampu mencapai akurasi 99,76% dan akurasi sensitivitas 100% pada dataset COVID-19 [36][37]. Di bidang pencitraan medis, metode baru yang menggabungkan SVM dengan Depthwise Separable Convolution Neural Network telah menunjukkan kinerja yang sangat baik pada dataset citra Chest X-ray (CXR). Pendekatan ini mencapai akurasi 98,54% untuk klasifikasi biner dan 99,06% untuk klasifikasi multikelas, menawarkan janji dalam mendiagnosis dan mengklasifikasikan kasus COVID-19 secara akurat menggunakan data pencitraan medis

f. Polinomial SVM

Algoritma Polynomial SVM adalah varian dari Support Vector Machine yang menggunakan fungsi kernel polinomial sebagai dasar klasifikasi. Ini biasanya digunakan untuk tugas klasifikasi biner dan dapat menangani data yang tidak dapat dipisahkan secara linier dengan memetakannya ke dalam ruang fitur berdimensi lebih tinggi. Diberikan dataset pelatihan dengan vektor fitur x_i dan label kelas yang sesuai y_i (di mana y_i∈\{-1,1\}), tujuan SVM Polinomial adalah untuk menemukan hyperplane optimal yang memisahkan titik data milik kelas yang berbeda. Fungsi keputusan dari Polynomial SVM didefinisikan sebagai:

$$[f(x) = \text{sign}\left(\sum_{i=1}^{N} \alpha_i y_i (x_i \cdot x + c)^d + b\right)] \tag{16}$$

di mana N adalah jumlah instance pelatihan, α_i adalah pengali Lagrange yang diperoleh selama proses optimasi, y_i adalah label kelas, c adalah suku konstan, d adalah derajat kernel polinomial, x_i adalah vektor pendukung, dan b adalah istilah bias. Fungsi kernel polinomial didefinisikan sebagai:

$$K(x_i, x) = (x_i \cdot x + c)^d \tag{17}$$

di mana c adalah suku konstan dan d adalah derajat kernel polinomial. Kernel polinomial menghitung kesamaan antara dua vektor berdasarkan produk titiknya yang dinaikkan ke derajat yang ditentukan. Masalah pengoptimalan untuk Polinomial SVM melibatkan meminimalkan fungsi tujuan sehubungan dengan α_i dan b, tunduk pada kendala:

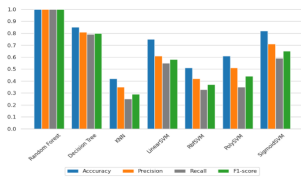
$$[0 \leq \alpha_i \leq C, \quad \sum_{i=1}^{N} \alpha_i y_i = 0] \tag{18}$$

di mana C adalah parameter regularisasi yang mengontrol trade-off antara memaksimalkan margin dan meminimalkan kesalahan pelatihan [41]. Poly SVM digunakan untuk menganalisis kumpulan data tweet COVID-19, menunjukkan tingkat akurasi 80% dan skor f1 81,84% [30]. Studi lain menyelidiki perbedaan antara dataset gambar X-ray normal, pneumonia, dan COVID-19 dengan menggunakan SVM sebagai kerangka dasar, menghasilkan akurasi yang mengesankan sebesar 97,33% di ketiga kelas [38]. Seperti kumpulan data sebelumnya, SVM digabungkan dengan model VGG16 dengan parameter yang disesuaikan untuk mengekstraksi fitur dari citra CT. Penggabungan ini menyebabkan model mencapai waktu pemrosesan cepat 385ms, sambil mempertahankan akurasi tinggi 95,7% pada 208 citra uji menurut kurva AUC [39]. Selain itu, metode yang disebut CSVM, yang menggabungkan SVM dengan Convolutional Neural Network, mengusulkan akurasi rata-rata 94% dan akurasi 96,09% yang mengesankan dalam penentuan sensitivitas untuk gambar CT [40]. Setelah menerapkan algoritma pembelajaran mesin tradisional ini untuk mengklasifikasikan data,

keefektifan algoritma akan dinilai menggunakan berbagai metode evaluasi. Penilaian akan mencakup metrik seperti akurasi, presisi, daya ingat, dan skor F1, yang secara kolektif mengukur seberapa baik performa algoritma. Selain itu, waktu yang dibutuhkan oleh setiap algoritma untuk mengklasifikasikan data akan diukur dan dilaporkan, memberikan informasi berharga tentang efisiensi komputasi algorithm.

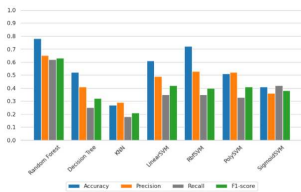
3. Hasil dan Pembahasan

Berdasarkan gambar. 4, terbukti bahwa Random Forest menunjukkan akurasi pelatihan mendekati 1, sedangkan KNN tampil buruk di set ini.



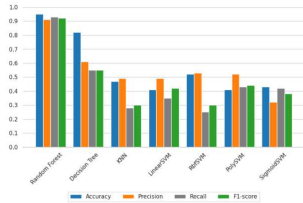
Gambar 4. Pelatihan akurasi dari setiap grafik algoritma pembelajaran mesin tradisional menggunakan Sobel

Berdasarkan gambar. 5, Random Forest melakukan yang terbaik pada pengujian dataset dengan metode segmentasi Sobel. Adapun sisanya rata-rata buruk kecuali KNN yang berkinerja paling buruk pada dataset ini.



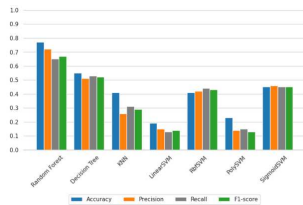
Gambar 5. Menguji akurasi setiap grafik algoritma pembelajaran mesin tradisional menggunakan Sobel

Seperti yang digambarkan pada gambar 6, ini menunjukkan bahwa 4 keluarga SVM tidak berkinerja terlalu baik pada dataset pengujian dengan segmentasi Prewitt Sementara itu, algoritma Random Forest melakukan rata-rata yang layak pada 0,9 atau 90% di seluruh metriknya sementara sebagian besar keluarga SVM tidak berkinerja terlalu baik.



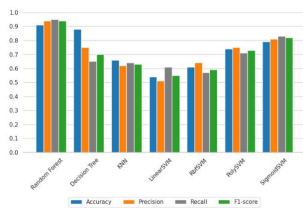
Gambar 6. Pelatihan akurasi dari setiap grafik algoritma pembelajaran mesin tradisional menggunakan Prewitt

Hasil yang ditunjukkan pada gambar. 7 dengan jelas menunjukkan bahwa Random Forest mempertahankan algoritma berkinerja tertinggi teratas dalam dataset ini menggunakan metode segmentasi Prewitt, memberikan kinerja tertinggi yang menguntungkan. Selain itu,perlu dicatat bahwa Pohon Keputusan mengungguli algoritma lain juga.



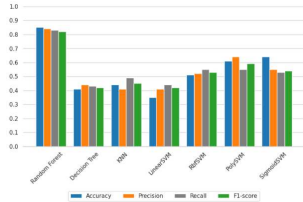
Gambar 7. Menguji akurasi setiap grafik algoritma pembelajaran mesin tradisional menggunakan Prewitt

Berdasarkan gambar. 8, algoritma menunjukkan kinerja yang luar biasa pada dataset pelatihan terutama Hutan Acak yang mencapai rata-rata 0,9 atau 90% pada metriknya. Namun, ketika menggunakan algoritma Linear SVM, tampaknya algoritma tersebut tidak bekerja dengan baik jika dibandingkan dengan algoritma lainnya.



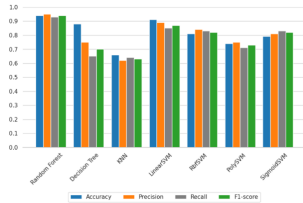
Gambar 8. Pelatihan akurasi setiap grafik algoritma pembelajaran mesin tradisional menggunakan Robert

Berdasarkan gambar. 9, Random Forest mencapai metrik rata-rata tertinggi pada dataset tak terlihat, sehingga dianggap sebagai algoritma terbaik untuk dataset ini menggunakan metode segmentasi Robert.



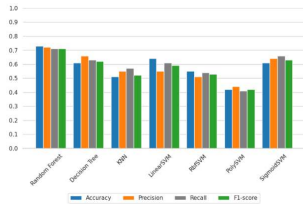
Gambar 9. Menguji akurasi setiap grafik algoritma algoritma mesin tradisional menggunakan Robert

Berdasarkan gambar. 10, di antara algoritma yang dievaluasi, Random Forest mencapai akurasi pelatihan tertinggi saat digabungkan dengan operator Scharr untuk segmentasi gambar. Algoritma seperti Linear dan Rbf SVM juga mencapai hasil yang memuaskan.



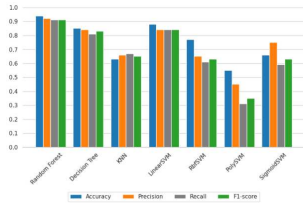
Gambar 10. Pelatihan akurasi dari setiap grafik algoritma pembelajaran mesin tradisional menggunakan Scharr

Berdasarkan gambar. 11, algoritma tidak menghasilkan hasil yang memuaskan pada kumpulan data pengujian, menunjukkan kemampuan terbatas untuk menggeneralisasi secara efektif dalam hal segmentasi Scharr.



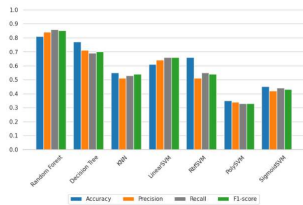
Gambar 11. Menguji akurasi setiap grafik algoritma pembelajaran mesin tradisional menggunakan Scharr

Berdasarkan gambar. 12, di antara algoritma yang diuji, Random Forest mencapai akurasi pelatihan tertinggi saat digabungkan dengan operator Canny, juga, Linear SVM mencapai hasil yang memuaskan selain Random Forest. Namun, beberapa algoritma tidak sepuas yang seharusnya.



Gambar 12. Pelatihan akurasi dari setiap grafik algoritma pembelajaran mesin tradisional menggunakan Canny

Berdasarkan gambar. 13, algoritma terkait akurasi pengujian yang dicapai oleh operator Canny tidak memuaskan karena beberapa menunjukkan hasil rata-rata dibawah 0,7 atau 70% kecuali untuk Hutan Acak yang rata-rata 0,8 atau 80%.



Gambar 13. Menguji akurasi setiap grafik algoritma pembelajaran mesin tradisional menggunakan Canny

Tabel 1: Waktu Eksekusi Random Forest

Algorithms	Time
Random Forest (Sobel)	1min 34s
Random Forest (Prewitt)	1min 29s
Random Forest (Scharr)	1min 27s
Random Forest (Robert)	1min 51s
Random Forest (Canny)	1min 37s
Sigmoid SVM (Scharr)	7min 44s

6 (enam) algoritma terbaik pada dataset pengujian:

a. **Hutan Acak dengan operator Sobel**
   1. Akurasi : 0,7841
   2. Ingat : 0,6528
   3. Presisi : 0,6233
   4. Skor F1 : 0,6371

b. Hutan Acak dengan operator Prewitt

   1. Akurasi : 0,7711
   2. Ingat : 0,7262
   3. Presisi : 0,6578
   4. Skor F1 : 0,6791

c. Hutan Acak dengan operator Robert

   1. Akurasi : 0,8552
   2. Ingat : 0,8463
   3. Presisi : 0,8377
   4. Skor F1 : 0,8249

d. Hutan Acak dengan operator Scharr

   1. Akurasi : 0,7346
   2. Ingat : 0,7241
   3. Presisi : 0,7123
   4. Skor F1 : 0,7149

e. **Hutan Acak dengan operator Canny**
   1. Akurasi : 0,8151
   2. Ingat : 0,8497
   3. Presisi : 0,8681
   4. Skor F1 : 0,8561

f. Sigmoid SVM dengan operator Scharr

   1. Akurasi : 0,6445
   2. Ingat : 0,5578
   3. Presisi : 0,6172
   4. Skor F1 : 0,5991

Berdasarkan hasil yang telah disebutkan sebelumnya, algoritma Random Forest dengan menggunakan metode segmentasi Scharr menunjukkan waktu eksekusi yang jauh lebih cepat dibandingkan dengan algoritma lainnya. Terlepas dari keunggulan waktu ini, algoritma hutan acak mengungguli sebagian besar metode segmentasi, dengan skor rata-rata 0,70 atau 70% di semua metrik. Sebaliknya, algoritma Sigmoid SVM membutuhkan waktu paling lama untuk dieksekusi di antara semua algoritma dan menghasilkan skor yang lebih rendah, tetapi masih dianggap sebagai salah satu dari 6 algoritma dengan kinerja terbaik.

4. Kesimpulan

Menurut temuan yang disajikan, algoritma Random Forest dengan operator Robert menonjol sebagai yang terbaik. Meskipun berjalan sedikit lebih lama dibandingkan dengan algoritma Random Forest lainnya menggunakan metode segmentasi yang berbeda, algoritma Random Forest secara konsisten menunjukkan kinerja yang unggul di semua enam algoritma. Perbedaan waktu eksekusi berkisar 10-15 detik lebih lambat.

Daftar Pustaka

[1] Wibisono, A., Wisesa, H.A., Rahmadhani, Z.P. et al. *Traditional food knowledge of Indonesia: a new high-quality food dataset and automatic recognition system*. J Big Data 7, 69 (2020). https://doi.org/10.1186/s40537-020-00342-5
[2] D. Sarwinda et al., "*Automatic Multi-class Classification of Indonesian Traditional Food using Convolutional Neural Networks*," 2020 3rd International Conference on Computer and Informatics Engineering (IC2IE), Yogyakarta, Indonesia, 2020, pp. 43-47, doi: 10.1109/IC2IE50715.2020.9274636.
[3] Dian Ade Kurnia et al 2021 J. Phys.: Conf. Ser. 1783 012047
[4] Chen, G., Jiang, Z., & Kamruzzaman, M. M. (2020). *Radar remote sensing image retrieval algorithm based on improved Sobel operator*. Journal of Visual Communication and Image Representation, 71, 102720.
[5] Ravivarma, G., Gavaskar, K., Malathi, D., Asha, K. G., Ashok, B., & Aarthi, S. (2021). *Implementation of Sobel operator-based image edge detection on FPGA*. Materials Today: Proceedings, 45, 2401-2407.
[6] Zhou, R. G., Yu, H., Cheng, Y., & Li, F. X. (2019). Quantum image edge extraction based on improved Prewitt operator. Quantum Information Processing, 18, 1-24.
[7] Song, Y., Ma, B., Gao, W., & Fan, S. (2019). *Medical Image Edge Detection Based on Improved Differential Evolution Algorithm and Prewitt Operator*. Acta Microscopica, 28(1).
[8] Liao, B., Chen, X., Yu, Y., & Li, Y. *Few-Shot Brain Tumor MRI Image Classification Using Graph Isomorphic Network and Prewitt Operator. Xiaokun and Yu, Yang and Li, Yong, Few-Shot Brain Tumor MRI Image Classification Using Graph Isomorphic Network and Prewitt Operator*.
[9] Abdel-Gawad, A. H., Said, L. A., & Radwan, A. G. (2020). *Optimized edge detection technique for brain tumor detection in MR images*. IEEE Access, 8, 136243-136259.
[10] Orujov, F., Maskeliūnas, R., Damaševičius, R., & Wei, W. J. A. S. C. (2020). *Fuzzy based image edge detection algorithm for blood vessel detection in retinal images*. Applied Soft Computing, 94, 106452.
[11] G. N. Chaple, R. D. Daruwala and M. S. Gofane, "*Comparisions of Robert, Prewitt, Sobel operator-based edge detection methods for real time uses on FPGA*," 2015 International Conference on Technologies for Sustainable Development (ICTSD), Mumbai, India, 2015, pp. 1-4, doi: 10.1109/ICTSD.2015.7095920.
[12] Z. A. Seghir and F. Hachouf, "*Image quality assessment scheme based on gradient similarity and color distortion*," 2015 12th International Symposium on Programming and Systems (ISPS), Algiers, Algeria, 2015, pp. 1-8, doi: 10.1109/ISPS.2015.7244985.
[13] Wang, Y., Xiao, Z., & Cao, G. (2022). *A convolutional neural network method based on Adam optimizer with power-exponential learning rate for bearing fault diagnosis*. Journal of Vibroengineering, 24(4), 666-678.
[14] Y. A. Sari et al., "*Indonesian Traditional Food Image Identification using Random Forest Classifier based on Color and Texture Features*," 2019 International Conference on Sustainable Information Engineering and Technology (SIET), Lombok, Indonesia, 2019, pp. 206-211, doi: 10.1109/SIET48054.2019.8986058.
[15] Sekehravani, E. A., Babulak, E., & Masoodi, M. (2020). *Implementing canny edge detection algorithm for noisy images*. Bulletin of Electrical Engineering and Informatics, 9(4), 1404-1410.
[16] M. Kalbasi and H. Nikmehr, "*Noise-Robust, Reconfigurable Canny Edge Detection and its Hardware Realization*," in IEEE Access, vol. 8, pp. 39934-39945, 2020, doi: 10.1109/ACCESS.2020.2976860.
[17] Hwa, S. K. T., Bade, A., Hijazi, M. H. A., & Jeffree, M. S. (2020). *Tuberculosis detection using deep learning and contrastenhanced canny edge detected X-Ray images*. IAES International Journal of Artificial Intelligence, 9(4), 713.
[18] Iqbal, B., Iqbal, W., Khan, N., Mahmood, A., & Erradi, A. (2020). *Canny edge detection and Hough transform for high resolution video streams using Hadoop and Spark*. Cluster Computing, 23(1), 397-408.
[19] Gong, L. H., Tian, C., Zou, W. P., & Zhou, N. R. (2021). *Robust and imperceptible watermarking scheme based on Canny edge detection and SVD in the contourlet domain*. Multimedia tools and applications, 80, 439-461.

[20] Chung, Y. L., & Lin, C. K. (2020). *Application of a model that combines the YOLOv3 object detection algorithm and canny edge detection algorithm to detect highway accidents*. Symmetry, 12(11), 1875.
[21] Li, B., He, F., & Zeng, X. (2021). *A novel privacy-preserving outsourcing computation scheme for Canny edge detection*. The Visual Computer, 1-19.
[22] Malbog, M. A. F., Lacatan, L. L., Dellosa, R. M., Austria, Y. D., & Cunanan, C. F. (2020, August). *Edge detection comparison of hybrid feature extraction for combustible fire segmentation: a Canny vs Sobel performance analysis*. In 2020 11th IEEE Control and System Graduate Research Colloquium (ICSGRC) (pp. 318-322). IEEE.
[23] Trivedi, J., Devi, M. S., & Dhara, D. (2020). *Canny edge detection based real-time intelligent parking management system*. Zeszyty Naukowe. Transport/Politechnika Śląska.
[24] Utomo, S. B., Irawan, J. F., & Alinra, R. R. (2021). *Early warning flood detector adopting camera by Sobel Canny edge detection algorithm method*. Indonesian Journal of Electrical Engineering and Computer Science, 22(3), 1796-1802.
[25] Jara, J. D. Z., & Bowen, S. (2022). *Learning Curve Analysis on Adam, Sgd, and Adagrad Optimizers on a Convolutional Neural Network Model for Cancer Cells Recognition*. ADCAIJ: Advances in Distributed Computing and Artificial Intelligence Journal, 11(3), 263-283.
[26] He, S., Wu, J., Wang, D., & He, X. (2022). *Predictive modeling of groundwater nitrate pollution and evaluating its main impact factors using random forest*. Chemosphere, 290, 133388.
[27] Speiser, J. L., Miller, M. E., Tooze, J., & Ip, E. (2019). *A comparison of random forest variable selection methods for classification prediction modeling*. Expert systems with applications, 134, 93-101.
[28] Yeşilkanat, C. M. (2020). *Spatio-temporal estimation of the daily cases of COVID-19 worldwide using random forest machine learning algorithm*. Chaos, Solitons & Fractals, 140, 110210.
[29] Kurani, A., Doshi, P., Vakharia, A., & Shah, M. (2023). *A comprehensive comparative study of artificial neural network (ANN) and support vector machines (SVM) on stock forecasting*. Annals of Data Science, 10(1), 183-208.
[30] Gu, J., & Lu, S. (2021). *An effective intrusion detection approach using SVM with naïve Bayes feature embedding*. Computers & Security, 103, 102158.
[31] Prastyo, P. H., Sumi, A. S., Dian, A. W., & Permanasari, A. E. (2020). *Tweets responding to the Indonesian Government's handling of COVID-19: Sentiment analysis using SVM with normalized poly kernel*. J. Inf. Syst. Eng. Bus. Intell, 6(2), 112.
[32] Reddy, M. R. (2021). *Implementation of SVM machine learning Algorithm to predict lung And Breast Cancer*. Turkish Journal of Computer and Mathematics Education (TURCOMAT), 12(12), 3050-3060.
[33] Pisner, D. A., & Schnyer, D. M. (2020). *Support vector machine*. In Machine learning (pp. 101-121). Academic Press.
[34] Altan, A., & Karasu, S. (2019). *The effect of kernel values in support vector machine to forecasting performance of financial time series*. The Journal of Cognitive Systems, 4(1), 17-21.
[35] Le, D. N., Parvathy, V. S., Gupta, D., Khanna, A., Rodrigues, J. J., & Shankar, K. (2021). *IoT enabled depthwise separable convolution neural network with deep support vector machine for COVID-19 diagnosis and classification*. International journal of machine learning and cybernetics, 1-14.
[36] Mahdy, L. N., Ezzat, K. A., Elmousalami, H. H., Ella, H. A., & Hassanien, A. E. (2020). *Automatic x-ray covid-19 lung image classification system based on multi-level thresholding and support vector machine*. MedRxiv, 2020-03.
[37] Sharifrazi, D., Alizadehsani, R., Roshanzamir, M., Joloudari, J. H., Shoeibi, A., Jafari, M., ... & Acharya, U. R. (2021). *Fusion of convolution neural network, support vector machine and Sobel filter for accurate detection of COVID-19 patients using X-ray images*. Biomedical Signal Processing and Control, 68, 102622.
[38] Le, D. N., Parvathy, V. S., Gupta, D., Khanna, A., Rodrigues, J. J., & Shankar, K. (2021). *IoT enabled depthwise separable convolution neural network with deep support vector*

machine for COVID-19 diagnosis and classification. International journal of machine learning and cybernetics, 1-14.

[39] Novitasari, D. C. R., Hendradi, R., Caraka, R. E., Rachmawati, Y., Fanani, N. Z., Syarifudin, A., ... & Chen, R. C. (2020). *Detection of COVID-19 chest X-ray using support vector machine and convolutional neural network*. Commun. Math. Biol. Neurosci., 2020, Article-ID.

[40] Özkaya, U., Öztürk, Ş., Budak, S., Melgani, F., & Polat, K. (2020). *Classification of COVID-19 in chest CT images using convolutional support vector machines*. arXiv preprint arXiv:2011.05746.

[41] Singh, M., Bansal, S., Ahuja, S., Dubey, R. K., Panigrahi, B. K., & Dey, N. (2021). *Transfer learning–based ensemble support vector machine model for automated COVID-19 detection using lung computerized tomography scan data*. Medical & biological engineering & computing, 59, 825-839.

JNATIA Volume 1, Nomor 4, Agustus 2023 p-ISSN: 2986-3929
Jurnal Nasional Teknologi Informasi dan Aplikasinya

Perbandingan Algoritma Forward Chaining dalam Sistem Pakar Rekomendasi Peminatan Bidang Teknologi

Putu Agus Dharma Kusuma[a1], I Made Widiartha[a2]

[a]Program Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Udayana
Jalan Raya Kampus Udayana, Bukit Jimbaran, Kuta Selatan, Badung, Bali Indonesia
[1]agusdharma48@gmail.com
[2]madewidiartha@unud.ac.id

Abstract

This research aims to compare the forward chaining algorithm with the Backward Chaining, Breadth-First Search (BFS), and Depth-First Search (DFS) algorithms in the context of an expert system for recommending specialization in the field of technology. The primary focus of this study is to analyze the runtime performance of each algorithm and determine the algorithm that provides the fastest runtime. The research methodology involves implementing the four algorithms in an expert system that provides recommendations for technology field specialization based on rules and user responses. The data used in the study consists of specialization rules in the technology field and user responses related to their interests in those fields. The results of the study demonstrate that the forward chaining algorithm outperforms the Backward Chaining, BFS, and DFS algorithms in terms of runtime performance. This indicates that the forward chaining algorithm is more efficient in generating technology field specialization recommendations. Based on the findings of this research, it is recommended to use the forward chaining algorithm in the development of expert systems for technology field specialization. This algorithm can assist users in obtaining recommendations quickly and efficiently, thereby enhancing user experience and the effectiveness of the expert system in providing suitable technology field specializations based on user interests.

Keywords: *Forward Chaining, Expert System, Expertise, Comparative Study, Runtime Performance*

1. Pendahuluan

Sistem pakar adalah sebuah program komputer yang pintar dan mengandalkan pengetahuan serta metode penalaran untuk menyelesaikan masalah yang kompleks dan membutuhkan keahlian khusus dari seorang ahli. Tujuan utama sistem pakar adalah meniru kemampuan pengambilan keputusan seorang ahli dalam suatu bidang tertentu. Dalam melakukan tugasnya, sistem pakar menggunakan pengetahuan khusus yang dimiliki oleh seorang ahli untuk memberikan solusi dan rekomendasi dalam pemecahan masalah [1]. Dalam penerapannya terdapat banyak sekali algoritma dalam pembangunan sistem pakar seperti *forward chaining*. Metode *forward chaining* memiliki banyak kekurangan bahkan bila dibandingkan algoritma serupa seperti backward chaining. Dari segi runtime, metode *forward chaining* memiliki kekurangan yang cukup signifikan yaitu sebesar 72,5% pengguna menunggu lebih lama dibandingkan dengan menggunakan algoritma backward chaining yaitu sebesar 87,5% [2]. Hal tersebut memberikan pengaruh yang cukup signifikan sehingga diperlukannya modifikasi pada algoritma *forward chaining* khususnya dalam runtime dan akurasi. Pada penelitian ini akan berfokus pada algoritma *forward chaining* dari segi implementasi sistem informasi.

2. Metode Penelitian

2.1. Analisis Environment

Pada fase ini akan dilakukan proses analisis environment khususnya dalam melakukan perbandingan algoritma forward chaining dengan algoritma lainnya. Penyamaan environment akan memperjelas perbedaan runtime antar algoritma sehingga perbedaan dapat diamati.

2.2. Analisis Algoritma

Tahapan ini merupakan tahapan untuk melakukan perhitungan algoritma dan melakukan perbandingan dengan algoritma lainnya. Analisis ini akan membandingkan beberapa kali runtime yang dilakukan dan menganalisis berdasarkan hasil perbandingan tersebut.

2.3. Metode Pengembangan SIstem

Pada fase ini akan dilakukan proses pengembangan sistem berdasarkan environment yang sudah ditetapkan. Dalam tahap ini akan dilakukan proses *Software Development Life Cycle* atau SDLC dalam pengembangan sistem.

2.4. Dokumentasi

Pada tahap ini akan dilakukan dokumentasi hasil jadi sistem dengan algoritma yang cukup efisien diterapkan pada sistem tersebut.

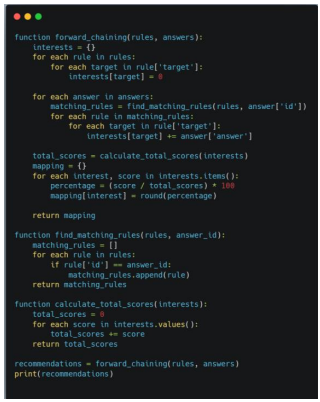
3. Hasil dan Pembahasan

3.1. Analisis Environment

Pada penelitian ini menggunakan beberapa tools dan framework yang akan membantu proses jalannya penelitian. Penelitian ini akan berbasis web menggunakan bahasa utama typescript dan framework next js sebagai sisi frontend. Dan pada sisi algoritma akan menggunakan python khususnya dalam perhitungan runtime algoritma yang telah dibuat. Penggunaan typescript pada proses pengembangan aplikasi ini dikarenakan typescript memiliki kemampuan dalam meminimalisir error dalam proses pengembangan. Hal ini dikarenakan pada bahasa typescript, wajib melakukan definisi tipe variabel sehingga setiap value akan secara tepat mengisi variabel yang ditentukan. Hal ini cukup berbeda dengan bahasa javascript yang secara bebas dapat memasukkan variabel tertentu sehingga akan rentan dalam memicu error pada proses developing bahkan production [3]. Pada proses komunikasinya akan menggunakan sistem rest API dengan menggunakan JSON sebagai perantara komunikasi antara algoritma dengan aplikasi frontendnya. Dari sisi client akan bertugas dalam mengambil input dan hasil inputan tersebut akan di Post ke dalam server untuk dilakukan perhitungan minat yang diinginkan [5]. Pada saat yang bersamaan server akan menghitung runtime algoritma sehingga penulis akan melakukan perbandingan runtime algoritma dengan algoritma lainnya.

3.2. Analisis Algoritma

Algoritma *forward chaining* yang diimplementasikan akan dilakukan proses sedikit modifikasi yaitu pada variabel “target” di rules. Pada variabel tersebut akan menggunakan tipe data list sehingga dalam 1 question kepakaran akan dapat menjurus ke 2 target minat yang berbeda atau lebih.



```
function forward_chaining(rules, answers):
    interests = {}
    for each rule in rules:
        for each target in rule['target']:
            interests[target] = 0

    for each answer in answers:
        matching_rules = find_matching_rules(rules, answer['id'])
        for each rule in matching_rules:
            for each target in rule['target']:
                interests[target] += answer['answer']

    total_scores = calculate_total_scores(interests)
    mapping = {}
    for each interest, score in interests.items():
        percentage = (score / total_scores) * 100
        mapping[interest] = round(percentage)

    return mapping

function find_matching_rules(rules, answer_id):
    matching_rules = []
    for each rule in rules:
        if rule['id'] == answer_id:
            matching_rules.append(rule)
    return matching_rules

function calculate_total_scores(interests):
    total_scores = 0
    for each score in interests.values():
        total_scores += score
    return total_scores

recommendations = forward_chaining(rules, answers)
print(recommendations)
```

Gambar 1. *Forward chaining*

Pseudocode tersebut merupakan hasil modifikasi dari algoritma f*orward chaining* yang akan diimplementasikan, pada algoritma tersebut memiliki input rules dan answers yang merupakan list map/dictionary dari kumpulan aturan yang telah ditetapkan dan kumpulan jawaban dari user.

```
Rule = Dict[str, List[str]]
Answer = Dict[str, int]
InterestMapping = Dict[str, int]


def forward_chaining(rules: List[Rule], answers: List[Answer]) -> InterestMapping:
    interests = {}
    for rule in rules:
        for target in rule['target']:
            interests[target] = 0

    for answer in answers:
        matching_rules = [rule for rule in rules if rule['id'] == answer['id']]
        for rule in matching_rules:
            for target in rule['target']:
                interests[target] += answer['answer']

    total_scores = sum(interests.values())
    mapping = {}
    for interest, score in interests.items():
        percentage = (score / total_scores) * 100
        mapping[interest] = round(percentage)

    return mapping


# Mulai menghitung runtime
start_time = time.time()

recommendations = forward_chaining(rules, answers)
print(recommendations)
```

Gambar 2. *Forward chaining implementation*

Source Code berikut merupakan hasil implementasi algoritma *forward chaining* pada sistem rekomendasi minat bidang informatika. Fungsi utama pada implementasi sistem pakar rekomendasi minat bidang informatika fungsi *forward_chaining()* dalam bahasa pemrograman Python. Fungsi ini menerima dua parameter, yaitu rules dan answers, dan mengembalikan sebuah *InterestMapping* yang berisi hasil rekomendasi peminatan berdasarkan aturan dan jawaban dari pengguna. Pertama, fungsi ini membuat sebuah dictionary interests yang akan menyimpan jumlah nilai awal untuk setiap peminatan. Setiap target peminatan dari setiap aturan diiterasi, dan nilai awalnya diinisialisasi dengan 0. Kemudian, fungsi melakukan iterasi pada setiap jawaban yang diberikan. Mencari aturan-aturan yang cocok dengan id jawaban, dan mengupdate nilai peminatan yang sesuai dalam dictionary interests dengan menambahkan nilai jawaban. Setelah itu, total skor peminatan dihitung dengan menjumlahkan semua nilai dalam interests. Selanjutnya, fungsi membuat dictionary mapping yang akan menyimpan persentase peminatan berdasarkan skor yang dihitung sebelumnya. Setiap peminatan diiterasi, dan persentase peminatan dihitung dengan membagi skor peminatan dengan total skor, kemudian dikalikan dengan 100. Nilai persentase tersebut dibulatkan ke angka terdekat dan disimpan dalam mapping. Akhirnya, fungsi mengembalikan mapping yang berisi rekomendasi peminatan berdasarkan hasil perhitungan. Selain fungsi *forward_chaining()*, dalam source code juga terdapat definisi variabel rules. Variabel ini berisi daftar aturan-aturan peminatan bidang teknologi yang digunakan dalam sistem pakar. Setiap aturan memiliki id, deskripsi, dan target peminatan yang terkait. Dalam penggunaan source code tersebut, perlu diperhatikan bahwa variabel rules dan answers harus didefinisikan dengan nilai yang sesuai sebelum pemanggilan fungsi *forward_chaining()*. Proses perbandingan akan dilakukan dari sisi server menggunakan bahasa Python sehingga akan memperlihatkan efektivitas dari aplikasi yang sedang dibangun. Selain itu rules yang akan digunakan akan sama dengan pembanding lainnya sehingga akan terlihat perbedaan runtime dari masing-masing algoritma. Adapun algoritma pembanding yang akan dipakai yaitu backward chaining, BFS, dan DFS.

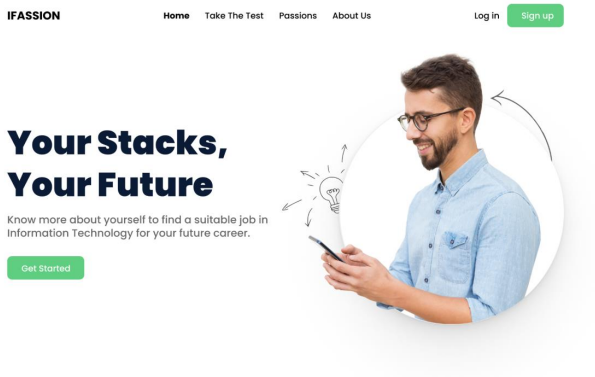
Tabel 1. Karakteristik Basis Data

	Forward Chaining (s)	**Backward Chaining (s)**	**BFS (s)**	**DFS (s)**
Runtime 1	6.675720214843 75e-05	0.00010180473327 636719	0.00010013580322 265625	0.00069403648376 46484
Runtime 2	6.484985351562 5e-05	0.00022697448730 46875	9.89437103271484e-05	0.00072669982910 15625
Runtime 3	6.198883056640 625e-05	9.17911529541015 6e-05	8.89301300048828 1e-05	0.00064277648925 78125
Runtime 4	6.508827209472 656e-05	9.77516174316406 2e-05	9.70363616943359 4e-05	0.00066423416137 69531
Runtime 5	6.675720214843 75e-05	9.70363616943359 4e-05	9.79900360107421 9e-05	0.00066995620727 53906

Melihat tabel tersebut, menunjukkan bahwa algoritma *forward chaining* tersebut memiliki nilai runtime yang secara berturut-turut jauh lebih kecil dibandingkan dengan algoritma lainnya. Algoritma *forward chaining* memiliki waktu runtime dengan skala terlama yaitu sebesar 6.67572021484375e-05 second dan tercepat yaitu 6.198883056640625e-05 second. Kemudian pada posisi kedua memiliki akumulasi kedudukan yang berubah ubah antara backward chaining dengan BFS. Hal ini dikarenakan kedua algoritma tersebut memiliki perbedaan runtime yang cukup dekat. Sedangkan pada algoritma DFS selalu menduduki tempat terakhir dengan waktu runtime terbesar yaitu 0.0007266998291015625 second dan waktu tercepat sebesar 0.0006642341613769531 second

3.3. Metode Pengembangan Sistem

Pengembangan sistem ini akan mengikuti pendekatan Software Development Life Cycle (SDLC) yang terstruktur dan terencana. SDLC adalah serangkaian langkah yang dilakukan oleh tim pengembang perangkat lunak, termasuk programmer dan analis sistem, dalam proses pembuatan sistem informasi atau perangkat lunak. SDLC terdiri dari beberapa tahap utama, yaitu perencanaan, analisis kebutuhan, desain, pengembangan, pengujian, implementasi, evaluasi, dan peluncuran [4]. Tahap perencanaan merupakan langkah awal dalam SDLC, tim pengembang akan mengidentifikasi tujuan dan kebutuhan aplikasi yang akan dikembangkan. Mereka akan menganalisis kebutuhan pengguna, mempertimbangkan aspek teknis dan anggaran yang tersedia, serta merencanakan jadwal dan sumber daya yang dibutuhkan. Tahap berikutnya adalah analisis kebutuhan, di mana tim akan mengumpulkan informasi lebih lanjut tentang kebutuhan pengguna dan membuat spesifikasi yang jelas untuk aplikasi. Setelah itu, tim akan memasuki tahap desain, di mana mereka akan merancang struktur sistem, antarmuka pengguna, dan komponen-komponen lainnya. Dalam pengembangan aplikasi ini, framework frontend Next.js dipilih untuk membangun antarmuka pengguna yang responsif dan interaktif. Dengan bantuan framework CSS seperti Tailwind CSS, tim dapat dengan mudah mengatur tata letak dan gaya desain yang sesuai. Setelah tahap desain, tim akan memulai tahap pengembangan, di mana mereka akan mulai mengkodekan aplikasi berdasarkan desain yang telah ditetapkan. Dalam kasus ini, bahasa pemrograman Python dan framework Flask digunakan sebagai teknologi server-side. Data yang diterima dari klien dalam bentuk JSON akan diproses, dan nilai jawaban pengguna akan disimpan dalam database.



Gambar 3. Sistem Pakar Rekomendasi Peminatan

Setelah tahap pengembangan, tahap pengujian akan dilakukan untuk memastikan bahwa aplikasi berfungsi dengan baik dan sesuai dengan kebutuhan yang telah ditetapkan. Tim pengembang akan melakukan pengujian fungsionalitas, pengujian integrasi, dan pengujian kinerja untuk memastikan bahwa aplikasi memberikan hasil yang akurat dan responsif. Setelah melewati tahap pengujian, aplikasi siap untuk diimplementasikan. Tahap implementasi melibatkan proses peluncuran aplikasi ke lingkungan produksi, memastikan bahwa aplikasi dapat diakses dan digunakan oleh pengguna sesuai dengan rencana yang telah ditentukan. Setelah aplikasi diimplementasikan, tahap evaluasi dilakukan untuk mengevaluasi kinerja aplikasi dan merespon umpan balik dari pengguna. Tim pengembang akan memantau dan menganalisis data

penggunaan aplikasi, mengidentifikasi area perbaikan atau peningkatan yang diperlukan, dan melakukan pemeliharaan rutin untuk memastikan kelancaran aplikasi. Dengan pendekatan pengembangan yang terstruktur menggunakan SDLC dan penggunaan algoritma forward chaining yang efisien, diharapkan aplikasi ini dapat memberikan rekomendasi minat bidang teknologi secara cepat, akurat, dan dapat diandalkan kepada pengguna. Selain itu, dengan menggunakan teknologi seperti Next.js, Tailwind CSS, Flask, dan Python, aplikasi ini dapat memberikan pengalaman pengguna yang baik dan antarmuka yang menarik.

4. Kesimpulan

Berdasarkan pembahasan yang dijabarkan maka dapat disimpulkan bahwa pada segi implementasi sistem informasi algoritma sistem pakar yang memiliki nilai runtime terkecil adalah *forward chaining* dengan skala terlama yaitu sebesar 6.67572021484375e-05 second dan tercepat yaitu 6.198883056640625e-05 second. Sehingga pada pengembangan aplikasi sistem pakar peminatan bidang informatika sangat dianjurkan menggunakan *forward chaining* dari segi kecepatan runtime sistem.

Daftar Pustaka

[1] R. Rosnelly and U. P. Utama, *Sistem Pakar: Konsep dan Teori*. Penerbit Andi.
[2] R. Apriliyani1, F. Tyas Ayuning, and E. Permatasari Kristi, “Perbandingan Metode Forward Chaining dan Backward Chaining pada Sistem Pakar Identifikasi Gaya Belajar,” *Jurnal Informatika dan Teknologi Komputer*, vol. 03, no. 02, pp. 84–92, 2022.
[3] S. bin Uzayr, “TypeScript,” in *TypeScript for Beginners*, Boca Raton: CRC Press, 2022, pp. 1–46. Accessed: Jun. 11, 2023. [Online]. Available: http://dx.doi.org/10.1201/9781003203728-1
[4] Y. Dwanoko Seby, “Implementasi software development life cycle (sdlc) dalam penerapan pembangunan aplikasi perangkat lunak,” *Jurnal Teknologi Informasi*, vol. 7, no. 2, 2016.
[5] F. Doglio, “Developing Your REST API,” in *Pro REST API Development with Node.js*, Berkeley, CA: Apress, 2015, pp. 123–166. Accessed: Jun. 11, 2023. [Online]. Available: http://dx.doi.org/10.1007/978-1-4842-0917-2_7

JNATIA Volume 1, Nomor 4, Agustus 2023 p-ISSN: 2986-3929
Jurnal Nasional Teknologi Informasi dan Aplikasinya

Analisis Celah Keamanan Jaringan WPA dan WPA2 Dengan Menggunakan Metode Penetration Testing

Albert Okario[1], I Putu Gede Hendra Suputra[2]

[a]Program Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam,
Universitas Udayana
Jalan Raya Kampus Udayana, Bukit Jimbaran, Kuta Selatan, Badung, Bali Indonesia
[1]okarioalbert@gmail.com
[2]hendra.suputra@gmail.com

Abstract

With the rapid development of communication and information technology, wireless local area network (WLAN) security has become crucial and a major concern, as data traffic is transmitted without the need for cables. Internet-connected network devices are inherently insecure and can be exploited by crackers or hackers. When data communicates or connects in the data traffic, where data is sent and passes through a series of terminals to reach its destination, an irresponsible user has the opportunity to modify or intercept the data. Therefore, designing a WLAN network connected to the internet must be carefully planned to minimize undesirable incidents. The weakness of the IEEE 802.11 network that uses WEP encryption tends to make the encryption code more easily discoverable by hackers. Based on the aforementioned background, we conducted this research to identify vulnerabilities or security flaws in WPA and WPA2-PSK networks using penetration testing methods.

Keywords: *WPA2-PSK Network Security Analysis, Penetration Testing*

1. Pendahuluan

Dalam era digital yang terus berkembang pesat, keamanan jaringan menjadi isu yang semakin penting. Kehadiran jaringan Wi-Fi telah mempermudah akses internet tanpa perlu menggunakan kabel, namun juga menimbulkan risiko keamanan yang perlu diperhatikan. Protokol keamanan Wi-Fi seperti WPA dan WPA2-PSK (Wi-Fi Protected Access 2 - Pre-Shared Key) dikembangkan untuk melindungi jaringan dari ancaman yang mungkin timbul. Namun, tidak ada sistem keamanan yang sepenuhnya tak terkalahkan. Celah keamanan dapat ada dan sering kali ditemukan oleh peneliti keamanan atau peretas. Oleh karena itu, penting untuk melakukan analisis yang komprehensif terhadap celah keamanan yang ada dalam protokol jaringan Wi-Fi. Dalam penelitian ini, kami akan melakukan analisis celah keamanan pada protokol

WPA dan WPA2-PSK dengan menggunakan metode penetration testing. Penetration testing adalah proses evaluasi keamanan yang bertujuan untuk mengidentifikasi kerentanan dalam sistem dengan mensimulasikan serangan yang dilakukan oleh pihak yang tidak berwenang. Melalui metode penetration testing, kami akan mencoba mengeksplorasi celah keamanan yang mungkin ada dalam protokol WPA dan WPA2-PSK. Tujuan utama dari penelitian ini adalah untuk mengetahui sejauh mana keamanan protokol tersebut dan menyediakan rekomendasi perbaikan untuk mengatasi celah yang ditemukan.

Data yang diperoleh dari penelitian ini akan dianalisis secara menyeluruh untuk mengidentifikasi celah keamanan yang ada dan memberikan rekomendasi yang tepat untuk meningkatkan keamanan jaringan Wi-Fi. Diharapkan bahwa hasil penelitian ini akan memberikan pemahaman yang lebih baik tentang celah keamanan dalam protokol WPA dan WPA2-PSK serta memberikan kontribusi untuk meningkatkan keamanan jaringan Wi-Fi secara umum.

2. Metode Penelitian

2.1 Metode Pengumpulan Data

Metode penelitian adalah kerangka kerja yang digunakan untuk mengarahkan langkah-langkah dalam menyusun hipotesis dan gagasan yang sesuai dengan tujuan penelitian. Metode yang tepat akan mempengaruhi proses penelitian dan hasil yang diperoleh. Pada prosedur penelitian yang dilakukan untuk mendapatkan hasil yang relevan yang sesuai dengan tujuan dari penelitian, terdapat empat langkah utama yang dijalankan:

a. Analisa: Langkah pertama dalam penelitian ini adalah melakukan analisis terhadap rancangan jaringan yang ada pada lokasi penelitian. Analisis ini bertujuan untuk memahami struktur jaringan yang sedang diteliti serta mengidentifikasi potensi celah keamanan yang mungkin ada
b. Perancangan: Langkah selanjutnya adalah merancang spesifikasi kebutuhan perangkat lunak sistem operasi Kali Linux yang akan digunakan dalam metode analisa. Dalam tahap ini, akan dilakukan perencanaan terkait konfigurasi dan persyaratan perangkat lunak yang diperlukan untuk melakukan analisis keamanan.
c. Pengujian: Tahap ini melibatkan pengujian menggunakan metode penetration testing untuk mendapatkan hasil dan menemukan celah keamanan yang ada. Dalam tahap pengujian ini, akan dilakukan serangkaian uji penetrasi untuk menguji efektivitas keamanan jaringan dan mengungkap potensi kerentanan yang mungkin ada.
d. Dokumentasi: Metode terakhir adalah dokumentasi, di mana langkah ini melibatkan studi pustaka, mempelajari jurnal-jurnal yang relevan, serta sumber-sumber lain yang berkaitan dengan topik penelitian. Proses dokumentasi ini penting untuk menggambarkan dan mengkomunikasikan temuan penelitian, langkah-langkah yang dilakukan, analisis data yang telah dilakukan, serta kesimpulan yang diperoleh dari penelitian tersebut.

Dengan melakukan langkah-langkah tersebut, diharapkan dapat menghasilkan data yang relevan dan sesuai dengan tujuan penelitian. Pengumpulan data yang sistematis dan terarah ini akan membantu dalam mencapai hasil penelitian yang sesuai dengan maksud dan tujuan yang telah ditetapkan sebelumnya.

2.2. Metode Penetration Testing

Penetration testing, juga dikenal sebagai pentesting, adalah proses yang disimulasikan untuk menemukan kerentanan, ancaman, dan risiko dalam sistem komputer, jaringan, atau aplikasi perangkat lunak. Dalam keamanan jaringan nirkabel, pentesting sering digunakan untuk menambahkan lapisan keamanan, seperti firewall, pada router. Kerentanan atau vulnerability adalah kelemahan atau celah yang dapat dieksploitasi oleh penyerang untuk mengganggu atau mendapatkan akses ke sistem dan data yang ada di dalamnya. Kerentanan umumnya disebabkan oleh kesalahan desain, konfigurasi, atau perangkat lunak. Tujuan utama dari penetration testing adalah untuk menemukan dan mengidentifikasi potensi kerentanan dan risiko keamanan yang ada dalam sistem. Hal ini memungkinkan pemilik sistem untuk mengambil tindakan yang tepat untuk memperbaiki celah keamanan tersebut sebelum penyerang yang jahat memanfaatkannya. Kerentanan yang sering ditemui meliputi kesalahan konfigurasi, kesalahan perangkat lunak, dan kerentanan lainnya. Dibawah ini merupakan langkah-langkah yang kami lakukan dalam melakukan penetration testing:

a. Perencanaan dan Persiapan

Dalam langkah perencanaan dan persiapan, dilakukan penentuan tujuan dan lingkup pentesting serta memperoleh izin tertulis dari pemilik sistem atau jaringan yang akan diuji, sekaligus mengumpulkan informasi yang diperlukan tentang sistem yang akan diuji, seperti alamat IP, jenis sistem operasi, aplikasi yang digunakan, dan kebijakan keamanan yang ada.

b. Pengumpulan Informasi

Pada langkah pengumpulan informasi, dilakukan pemetaan jaringan untuk mengidentifikasi host yang aktif, menentukan port yang terbuka, serta melakukan pengumpulan informasi lebih lanjut tentang sistem atau jaringan yang akan diuji, termasuk versi perangkat lunak yang digunakan, pengguna yang terdaftar, dan konfigurasi sistem yang relevan.

c. Analisis Kerentanan

Dalam tahap analisis kerentanan, dilakukan analisis kerentanan otomatis dengan menggunakan alat pemindai kerentanan untuk mengidentifikasi kerentanan umum yang terdapat dalam sistem atau jaringan yang diuji, serta dilakukan analisis kerentanan manual yang melibatkan pemeriksaan lebih mendalam secara manual terhadap kode, konfigurasi, dan pengujian yang lebih cermat untuk mencari kerentanan yang mungkin tidak terdeteksi secara otomatis.

d. Eksploitasi dan Mendapatkan Akses

Pada langkah eksploitasi dan mendapatkan akses, dilakukan upaya untuk mengeksploitasi kerentanan yang telah teridentifikasi untuk mendapatkan akses ke sistem atau jaringan yang diuji, dan dalam kasus berhasil mendapatkan akses awal, langkah selanjutnya adalah mencoba mendapatkan akses yang lebih tinggi, seperti akses administrator atau hak istimewa lainnya, dengan tujuan mengevaluasi sejauh mana sistem dapat melindungi data sensitif atau kritis.

e. Pemeliharaan Akses dan Penetrasi Lanjutan

Pada tahap pemeliharaan akses dan penetrasi lanjutan, dilakukan upaya untuk mempertahankan akses yang telah diperoleh agar dapat melakukan evaluasi lebih lanjut terhadap sistem atau jaringan yang diuji, serta dilakukan penetrasi lanjutan untuk mengeksplorasi lebih dalam sistem atau jaringan yang diuji guna mengidentifikasi kerentanan atau celah keamanan yang mungkin terlewatkan sebelumnya.

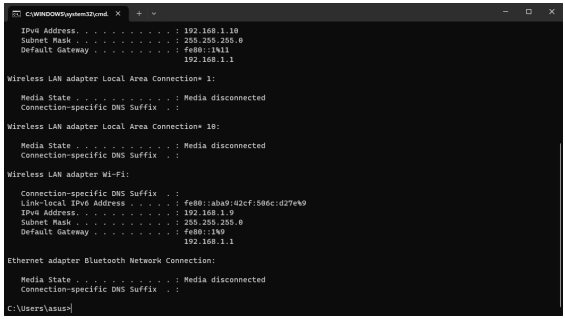
f. Pelaporan dan Rekomendasi

Pada langkah pelaporan dan rekomendasi, dibuat laporan yang berisi temuan secara rinci, termasuk kerentanan yang ditemukan, metode yang digunakan, serta hasil dari pengujian, sekaligus memberikan rekomendasi perbaikan kepada pemilik sistem atau jaringan tentang tindakan yang harus diambil untuk memperbaiki kerentanan dan meningkatkan keamanan.

3. Hasil dan Diskusi

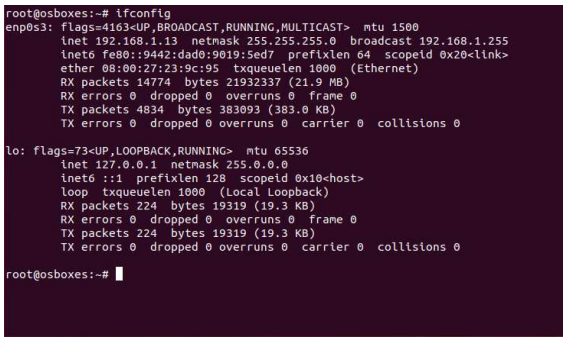
3.1 Tahapan Penelitian

Tahapan awal yang kami lakukan adalah pengumpulan data, yang dilanjutkan dengan penerapan metode penetration testing. Penetration testing dilakukan dengan melakukan pengecekan alamat IP (Internet Protocol) pada setiap perangkat yang terlibat dalam sistem. Selanjutnya, kami melakukan proses scanning dan discovering untuk mengidentifikasi port-port yang terbuka dan layanan-layanan yang berjalan pada port tersebut, menggunakan protokol TCP dan UDP. Dalam proses ini, kami menggunakan alat seperti Nmap untuk mengenali status port, seperti open, open|filtered, closed, closed|filtered, filtered, dan unfiltered. Kami juga melakukan pengecekan IP Windows dan melakukan identifikasi IP router yang terlibat dalam sistem. Cek ip yang digunakan untuk mengatahui ip windows dan juga router:



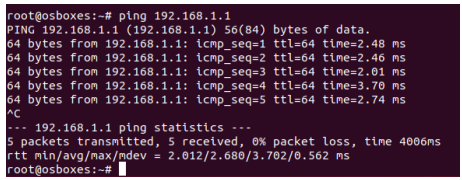
Gambar 1. Ip Windows dan Router

Setelah melakukan pengecekan IP menggunakan perintah ipconfig pada sistem operasi Windows, ditemukan bahwa IP Windows adalah 192.168.1.9 dengan subnet mask 255.255.255.0, dan IP router adalah 192.168.1.1. Selanjutnya, kami akan melakukan pengecekan IP pada sistem operasi Kali Linux:



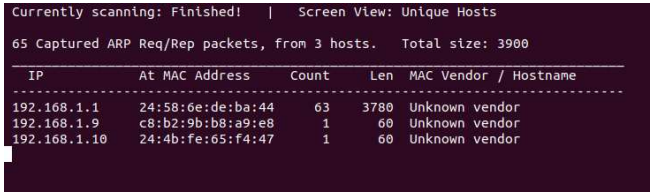
Gambar 2. Ip Linux

Melalui pengecekan tersebut, ditemukan bahwa IP dari Kali Linux adalah 192.168.1.13 dengan netmask 255.255.255.0. Informasi ini mengindikasikan bahwa jaringan yang digunakan adalah jaringan kelas C. Untuk menguji koneksi, kami dapat menggunakan perintah ping sebagai bukti bahwa komputer Kali Linux dapat terhubung ke jaringan dengan mengirimkan paket ke alamat IP yang dituju dan menerima respons balik.

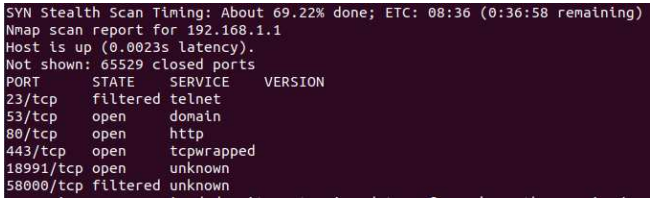


Gambar 3. Ping Gateway/Ip Router

Setelah mengetahui IP dari perangkat-perangkat dan jaringan yang terhubung, langkah selanjutnya adalah melakukan proses scanning dan discovering jaringan. Untuk melakukan scanning jaringan, kami menggunakan perintah "sudo netdiscover -r 192.168.1.0/24" yang akan melakukan scanning pada rentang IP 192.168.1.0 hingga 192.168.1.255 untuk menemukan perangkat-perangkat yang terhubung dalam jaringan tersebut. Selain itu, kami juga menggunakan perintah "nmap -p- -sV -O 192.168.1.1" yang akan melakukan scanning pada IP router 192.168.1.1 dengan melihat semua port yang terbuka (option -p-) dan mengidentifikasi layanan-layanan yang berjalan pada port tersebut (option -sV) serta melakukan pendeteksian sistem operasi (option -O).



Gambar 4. Proses Discovery



Gambar 5. Proses Scanning

Setelah itu gunakan tool meterpreter, Meterpreter adalah sebuah alat yang umumnya digunakan oleh peneliti keamanan dan pengujian penetrasi untuk mendapatkan akses ke sistem yang rentan, ia merupakan sebuah payload yang digunakan dalam kerangka kerja Metasploit. Meterpreter memungkinkan pengguna untuk mengendalikan sistem yang diserang secara jarak jauh dan melakukan berbagai tindakan, termasuk mengambil alih akses administrator. Dalam pengujian yang telah dilakukan, sistem operasi Metasploitable memang diketahui memiliki kerentanan yang dapat dieksploitasi. Metasploitable sebenarnya adalah distribusi sistem operasi

khusus yang dirancang untuk tujuan pengujian keamanan dan rentan terhadap serangan yang diketahui. Dengan menggunakan Metasploit dan payload seperti Meterpreter, pengguna dapat mengidentifikasi, mengeksploitasi, dan menguji kerentanan sistem tersebut. Penting untuk dicatat bahwa penggunaan Meterpreter atau alat serupa untuk tujuan ilegal, seperti mencuri data atau merusak sistem tanpa izin pemiliknya, adalah kegiatan yang melanggar hukum dan tidak etis. Penggunaan alat-alat tersebut harus selalu dilakukan dengan persetujuan dan dalam lingkungan pengujian keamanan yang sah. Dan dari hasil pengujian yang kami lakukan sistem operasi metasploitable dapat dieksploitasi.

4. Kesimpulan

Berdasarkan hasil pengujian terhadap sistem operasi pada jaringan yang menggunakan keamanan WPA dan WPA2, maka dapat disimpulkan seperti berikut ini:

a. Jalur lalu lintas data pada jaringan, jika dilakukannya proses scanning dan discovering selalu ada kemungkinan didapatkannya vulnerability atau suatu celah melalui port yang terbuka, namun proses scanning dan discovering sendiri bisa berjalan cukup lama tergantung dari jenis proteksi jaringan yang digunakan.]
b. port http yang terbuka menjadi salah satu bahaya yang dimana dapat dieksploitasi oleh meterpreter melalui msfconsole pada kali linux
c. orang-orang yang tidak bertanggung jawab sendiri dapat menggunakan tool seperti aircrack-ng dalam meretas password suatu jaringan yang semakin mempermudah dalam melakukan eksploitasi

Daftar Pustaka

[1]. Adiguna, M. A., & Widagdo, B. W “Analisis Keamanan Jaringan Wpa2-Psk Menggunakan Metode Penetration Testing (Studi Kasus : Router Tp-Link Mercusys Mw302r),” Jurnal Sistem Komputer dan Kecerdasan Buatan., vol. 5, No. 2, pp. 1-8, 2022.
[2]. Arif, K, “THESIS WPA2-PSK NETWORK SECURITY ANALYSIS USING THE PENETRATION TESTING METHOD (CASE STUDY: TP-LINK ARCHER A6),” 2021
[3]. Daulay, M. I, “ANALISIS PERBANDINGAN KEAMANAN WEP, WPA, WPA2, PADA ACCESS POINT,” 2019.
[4]. Fauzan, M. F., & dan Irawan, A. S. Y, “Wireless Attack : Menggunakan Tools Aircrack Pada Kali Linux Untuk Melakukan WPA Attack,” Jurnal Lentera., vol. 20, No. 1, pp. 63-74, 2021.
[5]. Haeruddin, & Kurniadi, A, “Analisis Keamanan Jaringan WPA2-PSK Menggunakan Metode Penetration Testing (Studi Kasus : TP-Link Archer A6),” Conference on Management, Business, Innovation, Education and Social Science., vol. 1, no. 1, pp. 508-515, 2021.
[6]. Setyawan, F., & Amnur, H, “Keamanan Jaringan Wireless Dengan Kali Linux,” Jurnal Ilmiah Teknologi Sistem Informasi., vol. 3, no. 1, pp. 16-22, 2022.

JNATIA Volume 1, Nomor 4, Agustus 2023 p-ISSN: 2986-3929
Jurnal Nasional Teknologi Informasi dan Aplikasinya

Optimasi SVM untuk Klasifikasi Warna: Investigasi Terhadap Pengaruh Fungsi Kernel dan Penyetelan Parameter

Pande Gede Dani Wismagatha[a1], I Wayan Santiyasa[a2]

[a]Program Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Udayana
Jalan Raya Kampus Udayana, Bukit Jimbaran, Kuta Selatan, Badung, Bali Indonesia
[1]pandedani@student.unud.ac.id
[2]santiyasa@unud.ac.id

Abstract

Color plays a crucial role in visual applications such as object recognition, image processing, computer vision, and computer graphics. Support Vector Machine (SVM) algorithms have gained attention for color classification due to their ability to handle complex data. SVM, a machine learning algorithm for classification and regression, aims to find optimal decision boundaries. In color classification using SVM, color data is represented by feature vectors, and SVM learns patterns to classify colors accurately. The SVM algorithm demonstrates a high accuracy rate, with an average accuracy of approximately 85% in color detection. This indicates the SVM's ability to effectively separate and classify colors with precision. SVM is proven to be effective in handling non-linear color data by utilizing kernel functions to transform the feature space into higher dimensions, enabling accurate classification of complex color data. The outstanding performance of the SVM algorithm in color detection presents vast potential applications in color recognition, image processing, computer vision, and computer graphics. SVM offers accurate and reliable solutions for object classification based on color characteristics in various contexts.

Keywords: *Color classification, SVM, Image processing, Machine Learning*

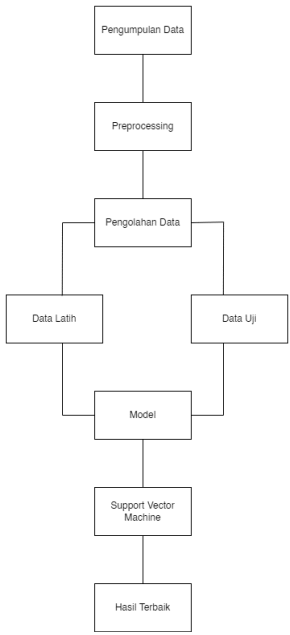
1. Pendahuluan

Warna adalah salah satu aspek penting dalam dunia visual dan memiliki peran yang signifikan dalam berbagai aplikasi, termasuk pengenalan objek, pengolahan citra, visi komputer, dan grafika komputer. Klasifikasi warna adalah suatu proses untuk mengelompokkan objek atau piksel berdasarkan karakteristik warnanya. Penggunaan algoritma Support Vector Machine (SVM) dalam klasifikasi warna menjadi subjek penelitian yang menarik karena kemampuannya dalam membedakan dan mengklasifikasikan data yang kompleks[1]. Support Vector Machine (SVM) merupakan algoritma machine learning yang dapat digunakan untuk klasifikasi ataupun regresi. SVM bertujuan untuk menemukan batas keputusan optimal antara dua atau lebih kelas dengan menggunakan konsep margin maksimal. SVM mampu mengklasifikasikan data dengan baik, terutama dalam kasus-kasus di mana data tidak terpisah secara linear di ruang fitur[2]. Dalam klasifikasi warna menggunakan SVM, data warna diwakili oleh vektor fitur yang terdiri dari komponen warna dalam model warna tertentu. Contohnya, dalam model warna RGB, vektor fitur terdiri dari intensitas merah (R), hijau (G), dan biru (B). SVM kemudian mempelajari pola-pola dalam vektor fitur tersebut untuk memisahkan dan mengklasifikasikan warna dengan akurasi yang tinggi. Metode klasifikasi warna menggunakan SVM memiliki beberapa keuntungan. Pertama, SVM mampu mengatasi data yang tidak linier, dengan menggunakan fungsi kernel untuk mengubah ruang fitur menjadi ruang dimensi yang lebih tinggi. Hal ini memungkinkan SVM untuk mengklasifikasikan data warna yang kompleks dengan presisi yang tinggi. Kedua, SVM mampu menangani data dengan dimensi tinggi, yang sering terjadi dalam representasi warna. SVM dapat memilih fitur-fitur yang relevan dan mengabaikan fitur-fitur yang tidak penting, sehingga meningkatkan efisiensi dan akurasi klasifikasi[3]. Penelitian ini bertujuan untuk mengklasifikasikan warna dengan pend Ekatan SVM Bagdasarian fitur-fitur warna dalam model

warna tertentu, seperti RGB. Metode SVM akan digunakan untuk mempelajari pola-pola dalam vektor fitur warna dan menghasilkan model klasifikasi yang dapat mengklasifikasikan warna dengan akurasi tinggi. Selain itu, penelitian ini juga akan melakukan analisis perbandingan antara beberapa metode terkait dalam klasifikasi warna menggunakan SVM untuk mengetahui kesamaan dan perbedaan dari metode-metode yang digunakan[4]. Diharapkan penelitian ini dapat memberikan kontribusi dalam pengembangan teknik klasifikasi warna dengan menggunakan algoritma SVM. Hasil penelitian ini diharapkan dapat meningkatkan akurasi dan efisiensi dalam pengenalan warna dalam berbagai aplikasi, seperti pengolahan citra, visi komputer, dan grafika komputer.

2. Metode Penelitian

Tahapan-tahapan dalam pelaksanaan penelitian dapat dilihat pada gambar berikut.



Gambar 1. Tahapan Pelaksanaan Penelitian

Berdasarkan gambar 1, tahap-tahap yang dilakukan dalam penelitian ini terdiri atas pengumpulan data, preprocessing, pengolahan data, pembagian data menjadi data latih (train) dan data uji (test), memasukkan data ke model kemudian mencari parameter terbaik untuk menghasilkan hasil yang terbaik.

a. Pengumpulan Data

Data diambil dari dataset pada laman kaggle yang dapat diakses menggunakan link berikut https://www.kaggle.com/datasets/ayanzadeh93/color-classification.
Dataset ini merupakan dataset 9 Jenis warna dengan total gambar berjumlah 107 data latih dan 96 data uji.

b. Pre-processing Data

Pada proses ini data diubah menjadi sebuah matriks dengan representasi angka. Kemudian, data yang telah diolah dapat digunakan untuk melatih dan menguji model klasifikasi.

c. Pengolahan Data

Pengolahan data dilakukan dengan memisahkan dataset yang telah melewati tahap sebelumnya/Pre-processing menjadi 2 (dua) yaitu data uji (train) dan data latih (test) lalu dijalankan pada algoritma Support Vector Machine (SVM). Berikut adalah cara kerja algoritma Support Vector Machine (SVM).

1. Support Vector Machine

Support Vector Machine (SVM) yaitu sistem pembelajaran yang menggunakan fungsi-fungsi linier dalam sebuah fitur yang berdimensi tinggi kemudian dilatih dengan menggunakan algoritma pembelajaran yang didasarkan pada teori optimasi. Akurasi yang dihasilkan oleh model pada algoritma ini sangatlah bergantung dengan penentuan parameter dan fungsi kernel yang digunakan. Algoritma SVM dapat dibedakan menjadi 2 (dua) yaitu SVM linear dan SVM non-linear. SVM linear digunakan untuk mengolah data yang dapat dipisahkan secara linear sedangkan SVM non-linear digunakan untuk data yang tidak bisa dibedakan secara linear sehingga menggunakan kernel untuk memisahkannya.

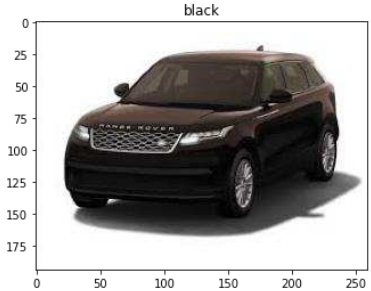
3. Hasil dan Diskusi

3.1. Persiapan Data

Dataset yang digunakan pada penelitian ini diperoleh dari salah satu dataset kaggle yang dibuat oleh Aydin Ayanzadeh dalam bentuk image (gambar). Pada dataset terdapat 107 gambar yang dibagi menjadi 9 klasifikasi gambar berbeda dan 96 tanpa label. Berikut adalah contoh gambar data uji. Peneliti menggunakan 107 gambar yang memilki label untuk memprediksi akurasi dari algoritma SVM.

3.2. Pre-Processing

Dataset yang digunakan dalam penelitian ini masih harus diubah menjadi sebuah nilai numerik agar mesin dapat mengerti akan masukan yang digunakan. Serta dataset harus diperiksa kembali guna mencari kesalahan klasifikasi pada data uji. Berikut ini contoh data yang telah diubah menjadi nilai numerikal yang dapat dimengerti mesin.



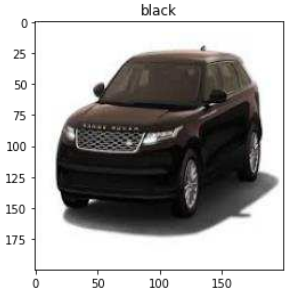
Gambar 2. Data Gambar

```
[[[255 255 255]
  [255 255 255]
  [255 255 255]
  ...
  [255 255 255]
  [255 255 255]
  [255 255 255]]

 [[255 255 255]
  [255 255 255]
  [255 255 255]
  ...
  [255 255 255]
  [255 255 255]
  [255 255 255]]

 [[255 255 255]
  [255 255 255]
  [255 255 255]
  ...
  [255 255 255]
  [255 255 255]
  [255 255 255]]

 ...
...
  ...
  [255 255 255]
  [255 255 255]
  [255 255 255]]]
```

Gambar 3. Data Numerik

Setelah melaksanakan konversi citra menjadi representasi data numerik, diperlukan tahapan penyesuaian ukuran pada seluruh data uji untuk memastikan konsistensi dimensi data. Resize (penyesuaian ukuran) pada gambar dalam machine learning merujuk pada proses mengubah dimensi fisik (ukuran) gambar menjadi ukuran yang ditentukan sebelum digunakan dalam model machine learning. Proses resize ini biasanya dilakukan untuk memastikan bahwa semua gambar memiliki dimensi yang seragam, sehingga mempermudah pemrosesan dan analisis dalam algoritma machine learning. Resize pada gambar melibatkan perubahan ukuran gambar secara proporsional, baik secara peningkatan (upscaling) maupun penurunan (downscaling) ukuran

gambar. Hal ini dapat dilakukan dengan menggunakan metode atau algoritma yang berbeda, seperti metode bilinear, metode nearest neighbor, atau metode bicubic. Tujuan utama dari resize gambar dalam konteks machine learning adalah untuk mencapai konsistensi ukuran gambar, sehingga memungkinkan model machine learning untuk memproses dan mempelajari fitur-fitur yang terdapat pada gambar dengan cara yang seragam. Resize juga dapat membantu mengurangi beban komputasi dan memori yang diperlukan dalam proses pelatihan model, terutama ketika menghadapi dataset gambar dengan variasi ukuran yang berbeda-beda[5]. Berikut adalah contoh gambar setelah proses resize.



Gambar 4. Gambar setelah proses resize.

3.3. Pre-Processing

Pada tahap selanjutnya data uji akan dibagi menjadi 2 tipe data yaitu data latih (train) dan data uji (test). Data yang digunakan sebagai data latih (train) sebanyak 80% dan data uji (test) sebanyak 20%. Namun sebelum itu, data warna dan label dipisah terlebih dahulu ke dalam sebuah variabel agar lebih mudah digunakan kedepannya. Berikut adalah potongan kode untuk memisahkan data warna dan label.

Tabel 1. Kode untuk memisahkan data warna dan label

```
X = []
y = []

for data_warna,label in training_list:
    X.append(data_warna)
    y.append(label)
```

Kemudian, data warna yang sebelumnya berupa matriks dengan dimensi 200x200 (setelah resize) perlu diubah menjadi sebuah matriks 1 dimensi. Berikut adalah potongan kode yang dapat digunakan untuk mengubah dimensi matriks.

Tabel 2. Kode untuk mengubah dimensi matriks menjadi 1 dimensi

```
X = np.array(X).reshape(len(X),-1)
```

Setelah data warna menjadi 1 dimensi, kedua data warna dan label dipisah menjadi 2 (dua) tipe data yang telah disebutkan sebelumnya yaitu 80% dan 20%. Berikut potongan kode untuk membagi dataset.

Tabel 3. Kode untuk memisahkan data menjasi data latih dan data uji

```
X_train, X_test, y_train, y_test = train_test_split(X, y, random_state=42, train_size=0.8,
test_size=0.2)
```

3.4. Proses Klasifikasi

Kemudian proses dilanjutkan ke proses klasifikasi, proses klasifikasi memerlukan beberapa modules antara lain:

Tabel 4. Import modules untuk proses klasifikasi

```
from sklearn.svm import SVC
from tqdm import tqdm
from sklearn.metrics import accuracy_score
from sklearn.model_selection import train_test_split
```

Kemudian, proses dilanjutkan dengan memanggil algoritma pertama yaitu SVC (Support Vector Classifier). Peneliti membuat sebuah fungsi yang akan memanggil SVC dengan kernel, C (Parameter Regularisasi), gamma (koefisien kernel) dan decision function shape (bentuk fungsi keputusan) yang beragam dengan parameter awal kernel = linear, C = 0.1, gamma = scale dan decision function shape = ovo.

Tabel 5. Fungsi SVC (Support Vector Classifier)

```
def svm_train(kernel:str='linear', C:float=1.0, gamma:str='scale', literal:str='ovo'):
    svc = SVC(kernel=kernel, C=C)
    svc.fit(X_train, y_train)
    y2 = svc.predict(X_test)
    return accuracy_score(y_test, y2)
```

Kemudian peneliti membuat sebuah fungsi baru untuk memanggil kernel, C, gamma, dan decision function shape yang berbeda-beda guna membandingkan akurasi yang dihasilkan. Berikut adalah potongan kode yang dapat digunakan.

Tabel 6. Fungsi pemanggilan SVC

```
def get_svm(gamma:str='scale', literal:str='ovo'):
    df = pd.DataFrame(columns=['kernel', 'C', 'accuracy','gamma','decision_function_shape'])
    for x in tqdm([0.1, 1, 10, 100, 1000]):
        for y in ['linear', 'poly', 'rbf', 'sigmoid']:
            acc = svm_train(y, x, gamma, literal)
            df = df.append({'kernel': y, 'C': x, 'accuracy': acc, 'gamma': gamma,
'decision_function_shape':literal}, ignore_index=True)
    return df
```

Peneliti menjalankan fungsi dengan beberapa variasi yang ada. Berikut adalah hasil yang didapatkan.

C	gamma	decision_function_shape	kernel	accuracy
0.1	scale	ovo	linear	0.851852
			poly	0.592593
			rbf	0.333333
			sigmoid	0.185185
1.0	scale	ovo	linear	0.851852
			poly	0.703704
			rbf	0.777778
			sigmoid	0.074074
10.0	scale	ovo	linear	0.851852
			poly	0.703704
			rbf	0.814815
			sigmoid	0.000000
100.0	scale	ovo	linear	0.851852
			poly	0.740741
			rbf	0.814815
			sigmoid	0.037037
1000.0	scale	ovo	linear	0.851852
			poly	0.777778
			rbf	0.814815
			sigmoid	0.148148

C	gamma	decision_function_shape	kernel	accuracy
0.1	auto	ovo	linear	0.851852
			poly	0.592593
			rbf	0.333333
			sigmoid	0.185185
1.0	auto	ovo	linear	0.851852
			poly	0.703704
			rbf	0.777778
			sigmoid	0.074074
10.0	auto	ovo	linear	0.851852
			poly	0.703704
			rbf	0.814815
			sigmoid	0.000000
100.0	auto	ovo	linear	0.851852
			poly	0.740741
			rbf	0.814815
			sigmoid	0.037037
1000.0	auto	ovo	linear	0.851852
			poly	0.777778
			rbf	0.814815
			sigmoid	0.148148

Gambar 5. Perbandingan hasil akurasi SVC dengan mengubah parameter gamma dan decision function shape bernilai ovo (one-vs-one)

C	gamma	decision_function_shape	kernel	accuracy
0.1	scale	ovr	linear	0.851852
			poly	0.592593
			rbf	0.333333
			sigmoid	0.185185
1.0	scale	ovr	linear	0.851852
			poly	0.703704
			rbf	0.777778
			sigmoid	0.074074
10.0	scale	ovr	linear	0.851852
			poly	0.703704
			rbf	0.814815
			sigmoid	0.000000
100.0	scale	ovr	linear	0.851852
			poly	0.740741
			rbf	0.814815
			sigmoid	0.037037
1000.0	scale	ovr	linear	0.851852
			poly	0.777778
			rbf	0.814815
			sigmoid	0.148148

C	gamma	decision_function_shape	kernel	accuracy
0.1	auto	ovr	linear	0.851852
			poly	0.592593
			rbf	0.333333
			sigmoid	0.185185
1.0	auto	ovr	linear	0.851852
			poly	0.703704
			rbf	0.777778
			sigmoid	0.074074
10.0	auto	ovr	linear	0.851852
			poly	0.703704
			rbf	0.814815
			sigmoid	0.000000
100.0	auto	ovr	linear	0.851852
			poly	0.740741
			rbf	0.814815
			sigmoid	0.037037
1000.0	auto	ovr	linear	0.851852
			poly	0.777778
			rbf	0.814815
			sigmoid	0.148148

Gambar 6. Perbandingan hasil akurasi SVC dengan mengubah parameter gamma dan decision function shape bernilai ovr (one-vr-rest)

Dengan menjalankan program tersebut, peneliti dapat mengambil hasil terbaik yang ada pada tiap perubahan parameter. Berikut adalah tabel parameter yang mendapat akurasi paling tinggi.

	kernel	C	accuracy	gamma	decision_function_shape
0	linear	0.1	0.851852	scale	ovo
1	linear	1.0	0.851852	scale	ovo
2	linear	10.0	0.851852	scale	ovo
3	linear	100.0	0.851852	scale	ovo
4	linear	1000.0	0.851852	scale	ovo
5	linear	0.1	0.851852	auto	ovo
6	linear	1.0	0.851852	auto	ovo
7	linear	10.0	0.851852	auto	ovo
8	linear	100.0	0.851852	auto	ovo
9	linear	1000.0	0.851852	auto	ovo
10	linear	0.1	0.851852	scale	ovr
11	linear	1.0	0.851852	scale	ovr
12	linear	10.0	0.851852	scale	ovr
13	linear	100.0	0.851852	scale	ovr
14	linear	1000.0	0.851852	scale	ovr
15	linear	0.1	0.851852	auto	ovr
16	linear	1.0	0.851852	auto	ovr
17	linear	10.0	0.851852	auto	ovr
18	linear	100.0	0.851852	auto	ovr
19	linear	1000.0	0.851852	auto	ovr

Gambar 7. Akurasi terbaik saat menjalankan fungsi SVC

Semua data yang memiliki akurasi terbaik berasal dari kernel yang sama walaupun dengan menggunakan parameter yang berbeda. Maka dari itu, peneliti mencoba mencari laporan klasifikasi yang dijalankan oleh mesin dan mendapatkan hasil berkut.

	precision	recall	f1-score	support
Black	1.00	1.00	1.00	3
Blue	1.00	0.67	0.80	3
Brown	0.62	1.00	0.77	5
Green	1.00	1.00	1.00	2
Violet	1.00	0.67	0.80	3
White	0.80	1.00	0.89	4
orange	1.00	0.50	0.67	2
red	1.00	0.67	0.80	3
yellow	1.00	1.00	1.00	2
accuracy			0.85	27
macro avg	0.94	0.83	0.86	27
weighted avg	0.90	0.85	0.85	27

Gambar 8. Classification Report kernel linear

4. Kesimpulan

Algoritma SVM menunjukkan tingkat akurasi yang tinggi dalam mendeteksi warna. Berdasarkan pengujian yang dilakukan, algoritma SVM mencapai tingkat akurasi rata-rata sekitar 85%. Hal ini menunjukkan kemampuan SVM untuk memisahkan dan mengklasifikasikan warna dengan presisi yang baik. Algoritma deteksi warna berbasis Support Vector Machine (SVM) menunjukkan hasil yang menjanjikan dengan tingkat presisi, recall, dan F1-score yang tinggi. Presisi sebesar 90% mengindikasikan kemampuan algoritma dalam mengklasifikasikan warna dengan akurasi tinggi, dengan minimalisasi hasil positif palsu. Recall sebesar 85% menunjukkan kemampuan algoritma dalam mengidentifikasi dan menangkap instansi positif sebenarnya dari warna. F1-score sebesar 85% juga menunjukkan kinerja yang baik secara keseluruhan dalam deteksi warna. SVM terbukti efektif dalam menangani data warna yang tidak terpisah secara linier. Dengan menggunakan fungsi kernel, SVM dapat mentransformasikan ruang fitur menjadi ruang dimensi yang lebih tinggi, sehingga memungkinkan SVM untuk mengklasifikasikan data warna yang kompleks dengan akurasi yang tinggi. Dengan kinerja yang baik dalam mendeteksi warna, algoritma SVM memiliki potensi aplikasi yang luas dalam bidang pengenalan warna, pengolahan citra, visi komputer, dan grafika komputer. Dalam berbagai konteks, SVM dapat memberikan solusi yang akurat dan handal dalam mengklasifikasikan objek berdasarkan warnany

Daftar Pustaka

[1] M. F. Naufal, "ANALISIS PERBANDINGAN ALGORITMA SVM, KNN, DAN CNN UNTUK KLASIFIKASI CITRA CUACA", doi: 10.25126/jtiik.202184553.
[2] W. Styorini and dan Wahyuni Khabzli, "Jurnal Politeknik Caltex Riau Analisis Perbandingan Machine Learning SVM Dan Adaboost Face Detection Dengan Metode Viola Jones," 2018. [Online]. Available: http://jurnal.pcr.ac.id|
[3] M. Fahmi and I. Suhartana, "Perbandingan Algoritma Decision Tree Dan Support Vector Machine Dalam Prediksi Kualitas Udara," 2022. [Online]. Available: https://data.jakarta.go.id/.
[4] R. L. Thiosdor, K. Gunadi, and L. P. Dewi, "Implementasi Program Presensi Mahasiswa dengan menggunakan Face Recognition."
[5] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks." [Online]. Available: http://code.google.com/p/cuda-convnet/

Halaman ini sengaja dibiarkan kosong

JNATIA Volume 1, Nomor 4, Agustus 2023 p-ISSN: 2986-3929
Jurnal Nasional Teknologi Informasi dan Aplikasinya

Dampak Penggunaan Anotasi Penamaan yang Berbeda Pada Kinerja NER

I Made Widi Arsa Ari Saputra[a1], I Wayan Supriana[a2]

[a]Program Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Udayana
Jalan Raya Kampus Udayana, Bukit Jimbaran, Kuta Selatan, Badung, Bali Indonesia
[1]widiarsa.sama@email.com
[2]wayan.supriana@unud.ac.id.com

Abstract

In developing the NER model, naming annotations are used as an important part of the training process. The impact of using different naming annotations on NER performance has been a concern in the research community. So, the writer wants to once again, test the impact of using different naming annotations using the spaCy library on English documents. Using 2 naming schemes namely BIO and IOBES, using the spaCy model to get 0.96 accuracy for BIO and 0.95 for IOBES.

Keywords: *NER, Person Entity, spaCy, BIO, IOBES, Named Entity Annotation*

1. Pendahuluan

Named Entity Recognition (NER) adalah salah satu tugas penting dalam pemrosesan bahasa alami (Natural Language Processing, NLP) yang bertujuan untuk mengidentifikasi entitas yang signifikan dalam teks, seperti orang, tempat, organisasi, tanggal, dan lain-lain. NER memiliki berbagai aplikasi, termasuk analisis teks, sistem tanya-jawab, dan ekstraksi informasi.

Dalam pengembangan model NER, anotasi penamaan digunakan sebagai bagian penting dalam proses pelatihan. Anotasi penamaan melibatkan memberikan label entitas yang benar kepada kata-kata dalam teks. Namun, masalah muncul ketika penggunaan anotasi penamaan yang berbeda-beda digunakan dalam dataset pelatihan, yang dapat berdampak pada kinerja model NER.

Dampak penggunaan anotasi penamaan yang berbeda pada kinerja NER telah menjadi perhatian dalam komunitas penelitian. Penelitian oleh Chen et al. [1] menunjukkan bahwa inkonsistensi dalam anotasi penamaan dapat menyebabkan penurunan signifikan dalam kinerja model NER. Mereka menemukan bahwa variasi dalam anotasi penamaan dapat menghasilkan ambiguitas dan kesulitan dalam melatih model secara efektif.

Selain itu, penelitian oleh Yang et al. [3] juga mengungkapkan bahwa perbedaan dalam anotasi penamaan dapat mempengaruhi konsistensi dan keandalan model NER. Ketika dataset pelatihan mengandung anotasi penamaan yang tidak konsisten, model cenderung menghasilkan hasil yang tidak dapat diandalkan dan mengalami kesulitan dalam mengenali entitas dengan benar.

Oleh karena itu, penting untuk mempertimbangkan penggunaan anotasi penamaan yang konsisten dalam pengembangan model NER. Dengan memastikan konsistensi dalam anotasi penamaan, dapat meningkatkan kinerja model dan menghasilkan hasil yang lebih akurat. Sehingga penulis ingin sekali lagi, menguji dampak penggunaan anotasi penamaan yang berbeda menggunakan model spaCy pada dokumen berbahasa inggris.

2. Metode Penelitian

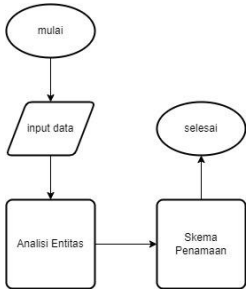
Dalam penelitian ini, akan menggunakan library spaCy untuk melakukan analisis

2.1 Pengumpulan Data

Pada penelitian ini, data yang digunakan adalah data teks yang didapatkan dari website bahasainggris.xyz berupa cerita pendek dari film harry potter menggunakan bahasa inggris. Data ini dibersihkan agar keseluruhan teks hanya mengandung bahasa inggris, sehingga proses identifikasi menggunakan library spaCy dapat lebih efisien.

2.2 Perancangan Sistem

Data input adalah teks mentah yang akan di analisis menggunakan objek nlp dengan model “en_core_web_sm” pada library spaCy. Skema penamaan akan digunakan uji coba terhadap 2 skema yaitu BIO dan IOBES.



Gambar 1. Rancangan Sistem

2.3 Named Entity Recognition

Named Entity Recognition (NER) adalah tugas dalam pemrosesan bahasa alami (Natural Language Processing, NLP) yang bertujuan untuk mengidentifikasi dan mengklasifikasikan entitas yang signifikan dalam teks, seperti orang, tempat, organisasi, tanggal, dan lain-lain. NER sangat penting dalam berbagai aplikasi NLP, termasuk pemahaman teks, analisis sentimen, ekstraksi informasi, dan sistem tanya-jawab[2].

Salah satu pendekatan yang umum digunakan dalam NER adalah dengan menggunakan anotasi penamaan dengan skema BIO (Begin, Inside, Outside) atau IOBES (Inside, Outside, Beginning, End, Single) [4]. Skema ini memungkinkan untuk memberikan label entitas pada setiap kata dalam teks.

Pada skema BIO, setiap kata dalam teks diberikan label yang menunjukkan apakah itu bagian dari entitas atau tidak. Label yang digunakan meliputi:

a. B-ENTITY: Menandakan awal entitas.
b. I-ENTITY: Menandakan kata di dalam entitas.
c. O: Menandakan kata di luar entitas.

Misalnya, dalam kalimat "John Mayer adalah seorang penyanyi yang terkenal", kata "John" diberi label "B-PERSON" (awal entitas orang), kata "Mayer" diberi label "I-PERSON" (di dalam entitas orang), dan kata-kata lain diberi label "O" (di luar entitas).

Sementara itu, skema IOBES memperluas skema BIO dengan menambahkan label untuk menunjukkan akhir entitas ("E") dan entitas tunggal ("S"). Label yang digunakan dalam skema IOBES meliputi:

a. I-ENTITY: Menandakan kata di dalam entitas.
b. O: Menandakan kata di luar entitas.
c. B-ENTITY: Menandakan awal entitas.
d. E-ENTITY: Menandakan akhir entitas.
e. S-ENTITY: Menandakan entitas tunggal.

Penerapan skema IOBES memungkinkan penanganan yang lebih baik untuk entitas yang melibatkan beberapa kata, seperti entitas yang terdiri dari dua kata atau lebih.

2.4 Library SpaCy

SpaCy adalah salah satu library pemrosesan bahasa alami (Natural Language Processing, NLP) yang populer dan kuat yang menyediakan berbagai fitur untuk analisis teks, termasuk Named Entity Recognition (NER). Library ini menawarkan berbagai objek dan fungsi yang dapat digunakan untuk melakukan NER dengan mudah dan efisien.

Objek utama dalam spaCy yang berhubungan dengan NER adalah Doc dan Token. Doc merepresentasikan dokumen teks yang akan diproses, sedangkan Token mewakili setiap kata dalam dokumen tersebut.

Berikut adalah beberapa objek dan fungsi yang relevan untuk melakukan NER menggunakan spaCy:

a. nlp: Objek ini adalah inti dari spaCy dan digunakan untuk memproses teks. Ia akan membuat objek Doc dari teks yang diberikan dan memasukkannya ke dalam aliran pemrosesan spaCy.
b. Doc: Objek ini merepresentasikan dokumen teks yang telah diproses. Ia berisi kumpulan Token yang mewakili setiap kata dalam teks.
c. Token: Objek ini mewakili setiap kata dalam dokumen. Ia menyimpan informasi seperti teks kata, posisi dalam dokumen, informasi morfologis, serta label entitas yang dikenali.
d. Token.ent_type_: Atribut ini menyimpan label entitas yang dikenali untuk setiap Token. Jika sebuah Token bukan bagian dari entitas, nilai atribut ini adalah '' (string kosong).
e. Token.ent_iob_: Atribut ini menyimpan tag IOB (Inside, Outside, Beginning) untuk setiap Token. Nilai atribut ini dapat berupa 'I' (Inside) jika Token berada di dalam entitas, 'O' (Outside) jika Token bukan bagian dari entitas, atau 'B' (Beginning) jika Token adalah awal entitas.

3. Hasil dan Pembahasan

3.1 Source Code untuk BIO

Table 1. Kode Program BIO

```
import spacy
import os
os.system("cls")

def analyze_entities(text):
    nlp = spacy.load("en_core_web_sm")
    doc = nlp(text)

    entities = []
```

```
    current_entity = []
    current_label = None

    for token in doc:
        if token.ent_type_ == "PERSON":
            if current_label == "B-PERSON":
                current_entity.append(token.text)
                current_label = "I-PERSON"
                entities.append((current_entity, current_label))
            else:
                if current_entity:
                    entities.append((current_entity, current_label))
                current_entity = [token.text]
                current_label = "B-PERSON"
                entities.append((current_entity, current_label))
        else:
            if current_entity:
                entities.append((current_entity, current_label))
                current_entity = []
                current_label = None

            if token.text.strip():
                entities.append(([token.text], "O"))

    if current_entity:
        entities.append((current_entity, current_label))

    return entities

text = "Harry Potter is a orphanage young boy who has lost parent since he was very little
baby. Harry was born to a magician father and mother. They are not ordinary human being
they can do a trick and spell. Harry lives together with Dursley's Family; uncaring Aunt
Petunia, loathsome Uncle Vernon, and spoiled cousin Dudley. The Drusley family barely
tolerate Harry, his aunt really don't care about him, his fatty cousin Dudley always bullies him
everytime he stay at home. They were living in suburb of little Whinging, Surrey, London.
Young Harry Potter never feel loved by parents, his private room at home was small cellar
under the stair where we normally put cleaning stuff such us moob, swab, broom etc. He was
Happy even he didn't grow up in very loving family. One day Harry is astonished to receive a
letter addressed to him in the cupboar under the stairs (where he sleeps). Unfortunatelly
before he can open the letter, Uncle Vernon takes it and tear them all up, since that day the
same letter come to Harry, but he never get any chance to open and read it even once. One
day when the Drusley Family go vacation to miserable shack small island on Harry's 11th
birthday a giant called Hagrid arrives to their home and reveals that Harry is a wizard an he
has been accepted at the Hogwarts School of Withcraft and Wizardry"

entities = analyze_entities(text)

for entity, label in entities:
    for i, word in enumerate(entity):
        print(f"{word}\t{label}")


with open("bio.txt", "w", encoding="utf-8") as f:
    for entity, label in entities:
        for i, word in enumerate(entity):
            text = f.write(f"{word}\t{label}\n")
```

3.2 Source Code untuk IOBES

Table 2. Kode Program IOBES

```
import spacy

def analyze_entities(text):
    nlp = spacy.load("en_core_web_sm")
    doc = nlp(text)

    entities = []
    current_entity = []
    current_label = None

    for token in doc:
        if token.ent_type_ == "PERSON":
            if current_entity:
                if len(current_entity) == 1:
                    entities.append((current_entity[0], "S-PERSON"))
                else:
                    entities.append((current_entity[0], "B-PERSON"))
                    for i in range(1, len(current_entity) - 1):
                        entities.append((current_entity[i], "I-PERSON"))
                    entities.append((current_entity[-1], "E-PERSON"))
                current_entity = []
                current_label = None

            current_entity.append(token.text)
            current_label = "B-PERSON"
        else:
            if current_entity:
                if len(current_entity) == 1:
                    entities.append((current_entity[0], "S-PERSON"))
                else:
                    entities.append((current_entity[0], "B-PERSON"))
                    for i in range(1, len(current_entity) - 1):
                        entities.append((current_entity[i], "I-PERSON"))
                    entities.append((current_entity[-1], "E-PERSON"))
                current_entity = []
                current_label = None

            if token.text.strip():
                entities.append((token.text, "O"))

    if current_entity:
        if len(current_entity) == 1:
            entities.append((current_entity[0], "S-PERSON"))
        else:
            entities.append((current_entity[0], "B-PERSON"))
            for i in range(1, len(current_entity) - 1):
                entities.append((current_entity[i], "I-PERSON"))
            entities.append((current_entity[-1], "E-PERSON"))

    return entities
```

```
text = "Harry Potter is a orphanage young boy who has lost parent since he was very little baby. Harry was born to a magician father and mother. They are not ordinary human being they can do a trick and spell. Harry lives together with Dursley's Family; uncaring Aunt Petunia, loathsome Uncle Vernon, and spoiled cousin Dudley. The Drusley family barely tolerate Harry, his aunt really don't care about him, his fatty cousin Dudley always bullies him everytime he stay at home. They were living in suburb of little Whinging, Surrey, London. Young Harry Potter never feel loved by parents, his private room at home was small cellar under the stair where we normally put cleaning stuff such us moob, swab, broom etc. He was Happy even he didn't grow up in very loving family. One day Harry is astonished to receive a letter addressed to him in the cupboar under the stairs (where he sleeps). Unfortunatelly before he can open the letter, Uncle Vernon takes it and tear them all up, since that day the same letter come to Harry, but he never get any chance to open and read it even once. One day when the Drusley Family go vacation to miserable shack small island on Harry's 11th birthday a giant called Hagrid arrives to their home and reveals that Harry is a wizard an he has been accepted at the Hogwarts School of Withcraft and Wizardry"

entities = analyze_entities(text)

for entity, label in entities:
    print(f"{entity}\t{label}")
with open("IOBES.txt", "w", encoding="utf-8") as f:
    for entity, label in entities:
        text = f.write(f"{entity}\t{label}\n")
```

3.3 Analisis Hasil

Hasil program menggunakan library spaCy sebagai berikut. Hasil dari pengenalan entitas disajikan dalam bentuk tabel

Table 3.Tabel Akurasi

Skema Penamaan	Label Benar	Akurasi
BIO	256	0.96
IOBES	253	0.95

Dari data input didapatkan total 266 entitas yang di recognisi. Skema penamaan BIO berhasil merekognisi sejumlah 256 entitas secara tepat. Sedangkan skema penamaan IOBES berhasil merekognisi sejumlah 253 entitas secara tepat.

4. Kesimpulan

Penggunaan skema penamaan yang berbeda berpengaruh kepada akurasi dari Named Entity Recognition pada library spaCy. Skema penamaan BIO mendapat akurasi tertinggi denga, akurasi 0.96. Library spaCy adalah model yang telah terlatih sebelumnya, melakukan re-trained pada model sesuai dengan kondisi teks akan berpotensi menaikkan akurasi dari model.

Daftar Pustaka

[1] S. Haryati, A. Sudarsono, and E. Suryana, "Implementasi Data Mining Untuk [1] Chen, D., Manning, C. D. & Jurafsky, D. (2019). Simple BERT Models for Relation Extraction and Semantic Role Labeling. Proceedings of the 2019 Conference of the Association for Computational Linguistics (ACL), 3543-3551.
[2] Nadeau, D., & Sekine, S. (2007). A survey of named entity recognition and classification. Lingvisticae Investigationes, 30(1), 3-26.
[3] Yang, Z., Yang, D., Dyer, C., He, X., Smola, A. & Hovy, E. (2020). NERDS: Neural Named

Entity Recognition with Dual-Level Selection. Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (ACL), 5104-5113.
[4] Sang, E. F., & Buchholz, S. (2000). Introduction to the CoNLL-2000 Shared Task: Chunking. In Proceedings of the 2nd Workshop on Learning Language in Logic and the 4th Conference on Computational Natural Language Learning (CoNLL), 127-132.
[5] Explosion AI. (2022). spaCy: Industrial-strength Natural Language Processing in Python. [Online]. Available: https://spacy.io/
[6] Haddi, E., Liu, X., & Shi, Y. (2019). Deep Learning for Natural Language Processing. Springer.

Halaman ini sengaja dibiarkan kosong

JNATIA Volume 1, Nomor 4, Agustus 2023 p-ISSN: 2986-3929
Jurnal Nasional Teknologi Informasi dan Aplikasinya

Perancangan Alat Pemberian Pakan Ikan Otomatis Pada Aquarium Berbasis Mikrokontroller AT89S52

I Gusti Bagus Ngurah Agung Brian Wijaya[a1], Ida Ayu Gde Suwiprabayanti Putra [a2]

[a]Program Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Udayana
Jalan Raya Kampus Udayana, Bukit Jimbaran, Kuta Selatan, Badung, Bali Indonesia
[1]Management.brianw@gmail.com
[2]iagsuwiprabayantiputra@unud.ac.id

Abstract

An important factor in keeping fish in an aquarium is the timeliness of feeding fish. Most of those who have a hobby of raising fish are worried about the feeding that must be done every day. Based on this, this final project designed and manufactured an automatic fish feeding device based on the AT89S52 microcontroller. So, a tool was designed that makes it easier to feed the fish automatically according to a predetermined schedule. The supporting components for scheduling fish feed include making a minimum circuit for the AT89S52 system as the brain of this tool which will later be loaded with a program using assembler language, RTC (Real Time Clock) as a timer, DC motor to rotate the valve opener for fish feed.

Keywords: *Feed, Fish, Microcontroller AT89S52*

1. Pendahuluan

Memelihara ikan dalam akuarium memang menjadi salah satu hobi yang populer dan diminati oleh banyak masyarakat. Meskipun menawarkan kemudahan dalam perawatan dan memberi pakan, namun bagi mereka yang memiliki kesibukan padat dalam pekerjaan, tantangan akan muncul saat harus meninggalkan rumah untuk waktu yang cukup lama. Hal ini dapat menyebabkan kesulitan dalam memenuhi kebutuhan ikan, seperti menjaga suhu air yang tepat dan memberikan pakan dengan konsistensi. Untuk mengatasi masalah tersebut, dirancanglah sebuah alat dengan menggunakan teknologi terkini yang dapat diaplikasikan dalam kehidupan sehari-hari. Alat ini dirancang dengan sederhana namun sangat efektif dan tidak memerlukan biaya yang terlalu banyak. Tujuannya adalah memberikan alternatif solusi bagi para penghobi memelihara ikan mas koki agar mereka tidak perlu khawatir ketika harus meninggalkan rumah untuk jangka waktu yang cukup lama. Dengan alat ini, diharapkan masyarakat yang memiliki kesibukan tinggi dapat lebih tenang dan yakin bahwa ikan-ikannya akan tetap terjaga kesehatannya ketika mereka tidak berada di rumah. Alat ini dapat membantu menjaga suhu air dalam akuarium dan memberikan pakan ikan secara otomatis sesuai dengan jadwal yang telah ditentukan sebelumnya. Dengan demikian, kebutuhan ikan akan terpenuhi dengan baik tanpa perlu bergantung pada bantuan orang lain, seperti saudara atau tetangga. Sebagai hasilnya, alat ini dapat menjadi solusi yang efektif dan praktis bagi para penghobi memelihara ikan mas koki yang memiliki mobilitas tinggi. Mereka dapat dengan tenang meninggalkan rumah tanpa khawatir akan kesejahteraan ikan-ikannya, karena alat ini akan secara otomatis merawat ikan dan menjaga kondisi akuarium selama mereka tidak berada di rumah.

2. Metode Penelitian

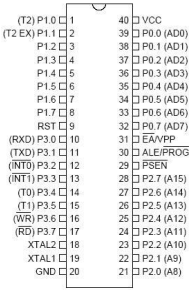
2.1. Ikan Mas (Carassius Auratus)

Ikan mas koki (Carassius auratus) adalah anggota dari keluarga Cyprinidae dan menjadi salah satu spesies ikan hias yang paling umum dipelihara di seluruh dunia. Ikan ini populer di kalangan penggemar ikan hias karena memiliki keindahan dan keunikan bentuk tubuh serta warnanya. Ikan mas koki sering menjadi peserta dalam kompetisi kecantikan ikan hias. Dalam perawatannya,

ikan mas koki dapat hidup cukup lama jika diberikan perawatan yang baik, dengan umur rata-rata antara 10 hingga 20 tahun, yang tergantung pada faktor-faktor seperti perawatan yang diberikan, faktor genetik, dan lingkungan hidupnya. Menurut penelitian oleh ilmuwan Cina bernama Shisan Chen, tercatat paling tidak ada 126 strain baru dari ikan mas koki yang telah tersebar di seluruh dunia (Lingga & Susanto, 1999). Awalnya, ikan mas koki telah mulai diternakkan oleh masyarakat Cina sejak tahun 960-1279, dan kepopulerannya semakin meningkat pada masa pemerintahan Dinasti Ming tahun 1368-1644, karena bentuk tubuhnya yang unik. Ikan mas koki kemudian diekspor dan banyak dijual ke negara-negara lain (Liviawaty & Aprianto, 1990). Dari Cina, perkembangan ikan mas koki kemudian menyebar hingga ke negeri Jepang. Keindahan dan keunikan ikan mas koki telah membuatnya menjadi favorit di dunia hobi ikan hias. Kepopulerannya terus berkembang dan menjadi salah satu ikan hias yang paling dicari oleh penggemar di berbagai belahan dunia.

2.2 Mikrokontroller AT89S52

Mikrokontroler AT89S52 adalah sebuah mikrokomputer dengan arsitektur CMOS 8-bit yang memiliki berbagai keunggulan, seperti daya rendah, performa yang tinggi, dan kapasitas memori flash sebesar 4 kbyte yang dapat diprogram dan dihapus secara cepat dengan menggunakan tegangan rendah. Selain itu, mikrokontroler ini menggunakan teknologi nonvolatile, yang artinya data yang tersimpan di dalamnya tidak akan hilang ketika terjadi pemadaman listrik. Dengan kombinasi CPU 8-bit dan memori flash, AT89S52 memungkinkan untuk dioperasikan secara serpih tunggal (single chip), yang berarti sebuah sistem dapat dibangun hanya dengan menggunakan mikrokontroler ini sebagai otak utama. Selain itu, mikrokontroler ini juga memiliki kemampuan untuk diperluas dengan menambahkan 4 jalur masukan/keluaran 8-bit, sehingga memungkinkan untuk menghubungkan berbagai perangkat eksternal dan memperluas fungsionalitas sistem.



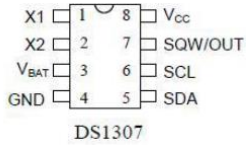
Gambar 2.1. Konfigurasi Pin AT89S52

2.3 Real Time Clock (RTC)

RTC yang dimaksud disini adalah real time clock biasanya berupa IC yang mempunyai clock sumber sendiri dan internal battery untuk menyimpan data waktu dan tanggal. Sehingga jika sistem komputer / mikrokontroler mati waktu dan tanggal didalam memori RTC tetap up to date. Salah satu RTC yang populer dan mudah penggunaannya adalah DS1307, apalagi pada codevision sudah tersedia fungsi-fungsi untuk mengambil data waktu dan tanggal untuk RTC DS1307 ini. Fiturfitur yang dimiliki oleh RTC ini antara lain.

a. RTC menghitung detik, menit, jam, tanggal, bulan serta tahun valid sampai tahun 2100.
b. Ram 56-byte, non-volatile untuk menyimpan data.

c. 2 jalur serial interface (I2C).
d. Output gelombang kotak yang diprogram.
e. Automatic power-fail detection and switch.
f. Konsumsi arus hanya 500nA pada battery internal.
g. Mode dengan oscillator running.
h. Temperature range: -40°C sampai +85°C



Gambar 2.2. Konfigurasi Pin RTC DS1307

2.4 Liquid Crystal Display (LCD)

Liquid Crystal Display (LCD) adalah sebuah modul penampil yang sering digunakan dalam berbagai perangkat elektronik karena kemampuannya menampilkan informasi dengan tampilan yang menarik. Salah satu jenis LCD yang paling banyak digunakan saat ini adalah tipe M1632, karena harganya yang terjangkau dan performa yang memadai. LCD M1632 memiliki ukuran tampilan 2x16, artinya memiliki 2 baris dan masing-masing baris memiliki 16 kolom karakter. Modul ini membutuhkan konsumsi daya yang rendah, membuatnya cocok untuk aplikasi dengan sumber daya terbatas. Selain itu, LCD M1632 juga dilengkapi dengan mikrokontroler yang khusus didesain untuk mengendalikan LCD tersebut. Mikrokontroler ini bertugas untuk mengatur tampilan karakter pada layar LCD, sehingga pengguna dapat dengan mudah menampilkan teks, angka, atau informasi lainnya dengan antarmuka yang lebih sederhana.

Gambar 2.3. Modul LCD 2x16

2.5 Motor DC

Motor DC berperan dalam mengubah energi listrik menjadi energi mekanis, di mana hasil geraknya berbentuk putaran. Prinsip dasar dari motor DC adalah bahwa ketika sebuah kawat yang mengalirkan arus ditempatkan di antara dua kutub magnet (biasanya ditandai sebagai U dan S), maka kawat tersebut akan mengalami gaya yang menyebabkannya bergerak. Ketika arus listrik mengalir melalui kawat, medan magnet akan terbentuk di sekitar kawat tersebut. Apabila arah medan magnet tersebut tegak lurus dengan arah arus listrik, maka akan muncul gaya Lorentz yang mempengaruhi kawat. Sebagai hasilnya, kawat akan mengalami gerakan tertentu sesuai dengan prinsip aksi-reaksi hukum Newton. Dalam motor DC, arus listrik yang mengalir melalui belitan kawat akan diubah secara periodik dengan memutar belitan atau mengubah arah arus dengan bantuan komutator. Proses ini menyebabkan putaran pada motor DC dan menghasilkan energi mekanis yang dapat digunakan dalam berbagai aplikasi seperti pada mesin industri, kendaraan listrik, atau perangkat lainnya.

Gambar 2.4. Motor DC 6000rpm

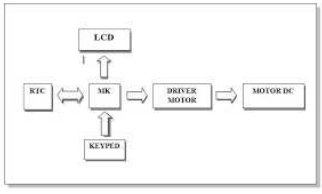
2.6 Keypad

Keypad matriks 4x4 adalah sebuah perangkat input yang terdiri dari 4 baris dan 4 kolom, sehingga menyediakan total 16 tombol yang dapat digunakan. Perangkat ini dapat diaplikasikan dengan mikrokontroler atau mikroprosesor untuk menerima masukan dari pengguna. Keypad matriks ini sering digunakan dalam berbagai aplikasi, seperti pada sistem keamanan, kontrol akses, atau perangkat elektronik lain yang memerlukan interaksi dengan pengguna. Konfigurasi dari keypad matriks 4x4 ini terdiri dari 4 baris input scanning dan 4 kolom output scanning. Ketika sebuah tombol ditekan, salah satu baris akan dihubungkan dengan salah satu kolom, dan hal ini akan menghasilkan kode unik yang mengidentifikasi tombol yang ditekan. Mikrokontroler atau mikroprosesor kemudian dapat membaca kode ini dan mengetahui tombol mana yang telah ditekan oleh pengguna.

Gambar 2.5. Keypad

3. Hasil dan Pembahasan

Dalam pembuatan ataupun perancangan alat diperlukan dahulu bagian-bagian pendukungnya, salah satunya ialah blok diagram sistem. Blok diagram sistem adalah gambaran untuk mempermudah sistem bekerja berserta fungsi dan tugas masing-masing komponen yang digunakan, blok diagram dapat dilihat pada gambar 3.1.



Gambar 3.1. Blok Diagram

Fungsi dari tiap blok diagram antara lain:

a. RTC digunakan untuk memberikan informasi waktu dan tanggal.
b. Mikrokontroler berfungsi memproses input dari data dan membandingkan jam dengan jadwal 4 jika sama akan menggerakan katup makanan.
c. Keypad berfungsi memberikan masukan setting jam dan setting jadwal pakan ikan koki.
d. Driver motor berfungsi sebagai komponen untuk mengendalikan arah putaran dan kecepatan motor dc.
e. Motor dc berfungsi untuk merubah arus menjadi putaran.
f. LCD berfungsi untuk menampilkan jam dan jadwal pakan ikan koki.

3.1 Perancangan Hardware

Sistematika perancangan perangkat keras dilakukan untuk setiap blok garis besar meliputi:

a. perancangan rangkaian minimum mikrokotroler AT89S52.
b. Perancangan rangkaian RTC.
c. Perancangan rangkaian power supplay.
d. Perancangan rangkaian keypad.
e. Perancangan rangkaian LCD.
f. Perancangan rangkaian motor dc.

3.2 Perancangan Software

Dalam perancangan perangkat keras (hardware), diperlukan dukungan dari suatu perangkat lunak (software). Tanpa adanya perangkat lunak yang mendukung, perangkat keras tersebut tidak akan memiliki nilai atau fungsi yang berarti. Dalam konteks perancangan ini, bahasa pemrograman Assembly digunakan sebagai media untuk mengembangkan perangkat lunak. Perangkat lunak ini bertindak sebagai pengatur kerja dari perangkat keras yang telah dibuat. Dengan adanya perangkat lunak ini, perangkat keras dapat berfungsi sesuai dengan tujuan dan kebutuhannya.

3.3 Pengujian

Untuk mengetahui hasil hasil kerja perangkat maka dilakukan pengujian pada tiap-tiap blok rangkaian dan pengujian kerja alat secara keseluruhan.

a. Prosedur Pengujian Keypad

Pada Prosedur ini, langkah pertama yang harus dilakukan adalah. Menghubungkan Keypad dengan rangkaian minimum sistem. Lalu selanjutnya menghidupkan catur daya

b. Prosedur Pengujian LCD

Pada prosedur pengujian LCD, langkah pertama yang dilakukan adalah menghubungkan bagian keypad dengan Rangkaian minimum sistem. Setelah program pengujian LCD diunduh ke mikrokontroller, layar LCD akan menampilkan tampilan sesuai yang diinginkan untuk pengujian LCD tersebut.

c. Prosedur Pengujian Motor DC

Pengujian ini bertujuan untuk mengetahui apakah driver motor DC bekerja secara optimal. Waktu yang disiapkan untuk menjalankan perintah dapat diatur sesuai kebutuhan. Jika waktu sudah habis, maka motor DC akan bergerak untuk menjatuhkan pakan sesuai dengan skenario atau kondisi yang telah ditentukan sebelumnya. Hal ini memungkinkan kita untuk menguji kinerja motor DC dan memastikan bahwa driver motor berfungsi dengan baik serta dapat mengatur gerakan motor sesuai dengan waktu yang telah ditentukan.

3.4 Pengujian Keseluruhan

Pengujian secara keseluruhan sesuai dengan urutan yang telah disiapkan:

a. Menghidupkan Power Supply
b. Input setting waktu dengan jadwal pemberian pakan
c. Melihat hasil pengukuran pada isplay LCD
d. Hasil Pakan yang jatuh ke akuarium

Berdasarkan hasil pengujian alat secara keseluruhan dapat disimpulkan bahwa sistem akan bekerja ketika sudah menginputkan setting jam, setting jadwal, dan diproses, maka secara otomatis alat ini akan mengeluarkan pakan ikannya

4. Kesimpulan

Berdasarkan permasalahan yang telah dibahas dan diselesaikan melalui laporan ini maka terdapat beberapa kesimpulan:

a. Alat aplikasi pakan ikan otomatis ini memberikan kemudahan bagi para pemelihara ikan koki dalam memberi pakan sesuai dengan jadwal yang diinginkan. Dengan adanya alat ini, pemelihara tidak perlu secara manual memberi pakan secara berkala, sehingga dapat menghemat waktu dan tenaga.
b. Rancangan dan pengujian alat ini menunjukkan bahwa alat tersebut dapat bekerja dengan baik. Hal ini terbukti dari hasil pengujian yang telah dilakukan, dimana saat melakukan pemilihan waktu melalui keypad, program bekerja dengan baik dan memberikan perintah kepada rangkaian output/keluaran dengan tepat.
c. Dengan kesimpulan ini, dapat disimpulkan bahwa alat aplikasi pakan ikan otomatis ini berhasil mencapai tujuan yang diinginkan, yaitu membantu memudahkan proses pemberian pakan pada pemeliharaan ikan koki dan berfungsi sesuai dengan harapan. Namun, tetap perlu diperhatikan untuk terus mengembangkan dan memperbaiki alat agar dapat lebih optimal dan efisien dalam penggunaannya.

Daftar Pustaka

[1] Abdurohman, Maman. 2010. Pemograman Bahasa Assembly. ANDI Publisher: Jakarta
[2] Budiharto, Widodo. 2005. Perancangan Sistem dan Aplikasi Mikrokontroller. Penerbit PT. Elex Media Komputindo. Jakarta...
[3] Budiharto, Widodo. 2007. 12 Proyek Mikrokontroller Untuk Pemula. Penerbit PT. Elex Media Komputindo Kelompok Gramedia. Jakarta
[4] Eko Putra, Agfianto. 1999. Belajar Mikrokontroller AT89C51/52/55. Yogyakarta: Penerbit Gava Media
[5] Norman, D. A. 2002, The Design of Everyday Things. The Electronic Journal of Communication. https://doi.org/10.1002/hfm.20127

.

JNATIA Volume 1, Nomor 4, Agustus 2023 p-ISSN: 2986-3929
Jurnal Nasional Teknologi Informasi dan Aplikasinya

Pemodelan Ontologi Semantik: Padmasana Umat Hindu Bali

I Gede Ngurah Arya Wira Putra[a1], Ida Bagus Gede Dwidasmara[a2]

[a]Program Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam,
Universitas Udayana
Jalan Raya Kampus Udayana, Bukit Jimbaran, Kuta Selatan, Badung, Bali Indonesia
[1]gedengurah2112@gmail.com
[2]dwidasmara@unud.ac.id

Abstract

Padmasana and Pelinggih are sacred structures that hold significant importance in the religious and cultural practices of the Balinese Hindu community. Padmasana refers to the elevated throne or shrine dedicated to the highest deity in Hinduism, while Pelinggih represents small sanctuaries or shrines found in various locations. Padmasana, often located in the innermost sanctum of temples or in open spaces, symbolizes the presence of the Supreme God in Hindu belief. It is characterized by an elevated platform adorned with intricate carvings and decorations, usually in the form of a lotus or a throne-like structure. Padmasana serves as a focal point for prayers, offerings, and rituals, where devotees seek spiritual connection and guidance from the divine. Through the ontological modeling of Padmasana and Pelinggih, this research contributes to the preservation and dissemination of the cultural heritage of the Balinese Hindu community. The ontological model serves as a knowledge repository, capturing the intricate details and interrelationships of these sacred structures, fostering a deeper understanding of their spiritual significance. Ultimately, this research promotes cross-cultural understanding and encourages the preservation and appreciation of Balinese Hindu traditions.

Keywords: *Ontology, Padmasana, Umat Hindu Bali, Web Semantik, Methontology*

1. Pendahuluan

Bangunan suci memegang peran sentral dalam kehidupan agama dan budaya umat Hindu Bali. Dalam tradisi Hindu Bali, bangunan suci merupakan tempat penting untuk melakukan pemujaan, ritual keagamaan, dan berinteraksi dengan dewa-dewa yang dihormati. Mereka menjadi simbol kehadiran spiritual dan menjadi pusat kegiatan keagamaan serta kehidupan sosial masyarakat Hindu Bali. Bangunan suci umat Hindu Bali, seperti pura dan padmasana (tempat suci kecil), memiliki karakteristik yang unik dan kompleks. Mereka dirancang dengan menggunakan prinsip-prinsip arsitektur tradisional Bali yang khas, dengan perhatian terhadap detail dan ornamen yang rumit. Bangunan-bangunan ini juga mencerminkan simbolisme dan filosofi yang dalam, mencakup konsep kosmologi dan hierarki spiritual dalam tradisi Hindu. Pemodelan ontologi semantik telah menjadi salah satu pendekatan yang penting dalam pengorganisasian dan representasi pengetahuan di berbagai domain. Ontologi semantik adalah kerangka konseptual yang menggambarkan entitas dan hubungan antara entitas tersebut dengan cara yang lebih terperinci dan bermakna. Dalam konteks ini, pemodelan ontologi semantik bangunan suci umat Hindu Bali menjadi subjek yang menarik untuk diteliti. Pemodelan ontologi semantik bangunan suci umat Hindu Bali bertujuan untuk menyajikan pengetahuan yang sistematis tentang struktur, entitas, hubungan, dan makna yang terkandung dalam bangunan suci tersebut. Dalam penelitian ini, akan digunakan metode pemodelan ontologi semantik yang melibatkan analisis literatur dan konsultasi dengan para ahli budaya Hindu Bali. Tujuannya adalah untuk menghasilkan ontologi semantik yang akurat dan konsisten dengan pemahaman tradisi Hindu Bali. Melalui pemodelan ontologi semantik bangunan suci umat Hindu Bali, diharapkan dapat memperluas pemahaman kita tentang kebudayaan, spiritualitas, dan tradisi Hindu Bali. Ontologi semantik ini dapat menjadi alat penting dalam melestarikan pengetahuan dan keaslian tradisi, serta mendorong penelitian lebih lanjut dalam bidang ini.

2. Metode Penelitian

Disini akan menjelaskan tentang rancangan metodelogi penelitian yang berisikan tahap-tahapan dari penelitian secara umum, berikut beberapa tahapan yang dilakuakan :

2.1. Metode Ontologi

Pada penelitian ini menggunakan metode Methontology. Metode ini adalah suatu pengembangan ontologi dalam basis pengetahuan. Dipilihnya metode pengembangan ontologi yang disebut Methontology ini, karena memiliki kelebihan dalam menjelaskan secara rinci dalam setiap aktivitas yang dilakukan selama membangun pemodelan ontologi ini.

2.2. Data Penelitian

Terdapat dua sumber data yang digunakan dalam penelitian ini, berikut penjelasan:

a. Data Primer

Data primer diperoleh dari studi lapangan yang dilakukan melalui wawancara langsung dengan ahli agama Hindu dan kebudayaan Bali. Dari hasil penelitian mendapatkan arti dan jenis padmasana serta pelinggih-pelinggih yang ada.

b. Data Skunder

Data skunder bersumber dari sumber literatur terdahulu yang berkaitan dengan objek penelitian.

2.3. Teknik Pengumpulan Data

Berasal dari wawancara yang dilakukan dengan ahli dan studi literatur untuk memperoleh jawaban tentang objek penelitian.

2.4. Perancangan Ontologi

Pada tahap perancangan model ontologi dari padmasana dan pelinggih, peneliti menggunakan aplikasi Protégé 4.3.0 untuk mengembangkan ontologi. Protege adalah sebuah aplikasi yang digunakan untuk membangun dan mengelola ontologi. Ontologi adalah representasi formal yang menggambarkan entitas, konsep, hubungan, dan aturan dalam suatu domain pengetahuan tertentu. Protege menyediakan lingkungan pengembangan yang intuitif dan kuat untuk merancang, mengedit, dan mengelola ontologi.

2.5. Evaluasi

Dalam tahap evaluasi, penulis melakukan evaluasi terhadap penelitian. Tahap ini dilakukan dengan cara menguji ontologi yang sudah dibangun dengan menggunakan query SPARQL. Pertanyaan akan diubah menjadi query SPARQL sehingga dapat menampilakan hasil yang ada didalam ontologi yang dibangun.

3. Hasil dan Diskusi

Methontology diterapkan dalam pembangunan ontology terkain domain padmasana, maka proses yang dilakukan seperti proses spesifikasi, akusisi pengetahuan, implementasi, evaluasi.

3.1. Spesifikasi

Spesifikasi dilakuka Langkah identifikasi segala aktivitas terkait degan penyusunan ontologi yang akan digunakan sebagai dokumen identitas terkait dengan domain yang akan di rancang. Berikut adalah spesifikasi dari ontology padmasana:

a. Domain : Padmasana
b. Tanggal : 11 Juni 2023
c. Konseptualisasi oleh : I Gede Ngurah Arya Wira Putra
d. Implementasi oleh : I Gede Ngurah Arya Wira Putra
e. Tujuan : Membangun sebuah ontologi yang digunakan dalam sistem informasi bangunan bali
f. Tingkatan formaliitas : Semi-formal
g. Batasan : Padmasana Utama
h. Sumber Pengetahuan : Internet (babadbali.com)

3.2. Akusisi Pengetahuan

Tahap akuisisi pengetahuan adalah kegiatan independen dalam proses pengembangan ontologi. Dalam tahap ini, teknik-teknik yang penulis gunakan untuk mengakuisisi pengetahuan ontologi padmasana umat Hindu Bali adalah sebagai berikut.

a. Berdiskusi dengan pembimbing terkait untuk membangun draf awal dokumen spesifikasi persyaratan.
b. Analisis teks informal, untuk mempelajari konsep-konsep utama yang diberikan dalam buku dan studi pegangan.
c. Analisis teks formal. Hal yang dilakukan adalah mengidentifikasi struktur yang akan dideteksi (definisi, penegasan, dan lain-lain) dan jenis pengetahuan yang dikontribusikan oleh masing-masing (konsep, atribut, nilai, dan hubungan).

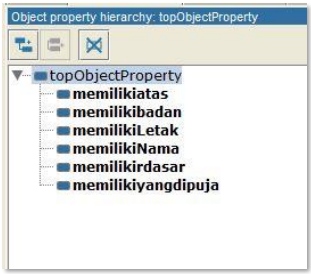
Data yang digunakan untuk membangun model ontologi dalam penelitian ini adalah data padmasana umat Hindu Bali dari beberapa ahli budaya agama dan budaya Bali dan internet.

3.3. Implementasi

Langkah awal yang harus dilakukan yaitu perancangan ontologi. Perancangan ontology dilakukan pembuatan Class dan SubClass dari ontologi. Selanjutnya dilakukan pembuatan Domain, dan Range yang digunakan pada ontologi. Selanjutnya berdasarkan perancangan ontology, maka proses akan dilakukan menggunakan Protégé 4.3.0.

Gambar 1. Struktur Class Ontologi

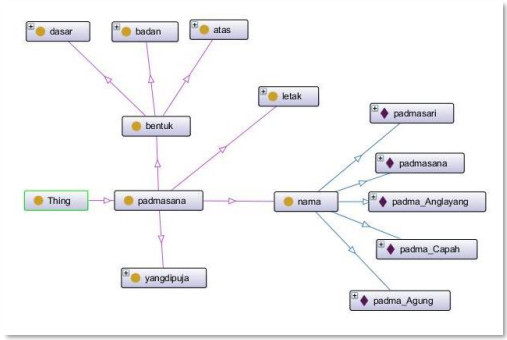
Pada Gambar 1 terdapat 8 class yang ada pada ontologi Padmasana Umat Hindu Bali. Terlihat pada class bentuk dibagi menjadi SubClass atas, badan, dasar. Pada Class bentuk dibagi menjadi tiga subclass yaitu atas, badan, dasar. Ketiga subclass yang dibagun dibawah bentuk dapat dikembangkan dikemudian hari untuk disesuaikan dengan kebutuhan nantinya.



Gambar 2. Object Property pada Ontologi Padmasana Umat Hindu Bali

Langkah berikutnya dilakukan pembuatan Object Property pada ontologi. Pada gambar 2 terdapat 6 Object Property yang terdapat di ontologi Padmasana Umat Hindu Bali. Masing-masing Object Property akan menghubungkan individuals.

Gambar 3. Individual/Instance dari ontologi



Gambar 4. Ontograf dari ontologi Padmasana Umat Hindu Bali

Pada gambar 4 terdapat ontograf yang dimana dapat disusun dari class dan subclass yang telah di buat. Lalu terdapat penjabaran dari Individual yang sudah masuk sebagai type class nama. Ontograf sendiri dapat di mekarkan menjadi lebih besar, dalam hal ini penulis hanya membuatnya sesuai dengan keperluan.

3.4. Evaluasi

Pada proses ini, penulis melakukan evaluasi terhadap ontology yang telah di bangun. Evaluasi dilakukan dengan melakukan query SPARQL pada aplikasi Protégé 4.3.0 dan menjawab pertanyaan. Berikut beberapa pertanyaan yang ditanyakan:

a. Apa saja nama Padmasana yang memuja Sang Hyang Tripurusa?
b. Apa saja nama Padmasana yang memuja Sang Hyang Tripurusa yang memiliki bentuk atas rongga satu?
c. Bentuk atas apa yang dimiliki oleh Padmasana Anglayang?

Berikut adalah jawaban yang ditampilkan dengan Protégé 4.3.0 menggunakan query SPARQL:

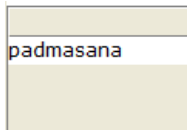
```
SPARQL query:
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX rdd: <http://www.semanticweb.org/aryawira/ontologies/2023/5/snatia#>
SELECT ?nama
                WHERE { ?nama rdd:memilikiyangdipuja rdd:SangHyangTripurusa .}
```

Gambar 5. Query SPARQL pertanyaan a

Gambar 6. Hasil SPARQL pertanyaan a

```
SPARQL query:

PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX rdd: <http://www.semanticweb.org/aryawira/ontologies/2023/5/snatia#>
SELECT ?nama
                WHERE { ?nama rdd:memilikiyangdipuja rdd:SangHyangTripurusa .
?nama rdd:memilikiatas rdd:ronggasatu.}
```

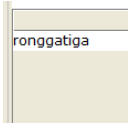
Gambar 7. Query SPARQL pertanyaan b



Gambar 8. Hasil SPARQL pertanyaan b

```
SPARQL query:

PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX rdd: <http://www.semanticweb.org/aryawira/ontologies/2023/5/snatia#>
SELECT?atas
                WHERE { ?atas rdd:memilikiNama rdd:padma_Anglayang .}
```

Gambar 9. Query SPARQL pertanyaan c



Gambar 10. Hasil SPARQL pertanyaan c

4. Kesimpulan

Berdasarkan pembahasan yang telat disampaikan , maka dapat dibuat kesimpulan sebagai berikut. Ontologi Padmasana Umat Hindu Bali yang dibangun dengan aplikasi Protégé 4.3.0 menghasilkan 8 kelas, 6 properti objek, 27 individu. Ontologi Padmasana Umat Hindu Bali ini bertujuan untuk mengumpulkan data dan mendokumentasikan agar mampu menjadi sebuah pengetahuan dan informasi, dan dapat melestarikan budaya bali. Kedepannya penelitian ini akan menggarap kualitas ontologi khususnya dalam unsur-unsur penting menarik yang dikembangkan lagi. Ontologi Padmasana Umat Hindu Bali sebagai dasar pengembangan sistem pengetahuan Bangunan Bali.

Daftar Pustaka

[1] k. putra, "Struktur & Makna Pura Di Bali Berdasarkan Asta Kosala-Kosali," Blogger Bali, [Online]. Available: https://www.komangputra.com/stuktur-makna-pura-di-bali-berdasarkan-asta-kosala-kosali.html/2.

[2] K. P. &Tim, Koleksi Lontar Bali, Bali, Indonesia , 2023.

[3] I. I. N. Sudiarta, "RUMAH TRADISIONAL BALI," Jurusan Teknik Arsitektur Fakultas Teknik Universitas Udayana.

[4] I. B. Idedhyana, media teliti , [Online]. Available: https://media.neliti.com/media/publications/345112-tipologi-bangunan-suci-padma-di-pura-luh-c8dd9c7c.pdf.

Halaman ini sengaja dibiarkan kosong

JNATIA Volume 1, Nomor 4, Agustus 2023 p-ISSN: 2986-3929
Jurnal Nasional Teknologi Informasi dan Aplikasinya

Analisis Data Berbentuk Teks dalam Sistem Diagnosis Penyakit dengan Supervised Learning

I Gusti Ngurah Bagus Ferry Mahayudha[a1], Ida Bagus Gede Dwidasmara[a2]

Program Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam,
Universitas Udayana
Jalan Raya Kampus Udayana, Bukit Jimbaran, Kuta Selatan, Badung, Bali Indonesia
[1]gunggus2367@gmail.com
[2]dwidasmara@unud.ac.id

Abstract

In computer science, text refers to a sequence of characters that can be represented and processed by a computer. It is the basic unit of data for representing human-readable information, such as letters, numbers, symbols, and spaces. In computer programming, text is typically represented as a string of characters. Textual data can be stored in variables, manipulated using various string operations, and displayed to users through input/output operations.Text plays a crucial role in many areas of computer science, including natural language processing, information retrieval, data mining, and text-based communication systems like email, chat applications, and social media. It serves as a fundamental component for storing, analyzing, and processing vast amounts of textual information in various applications.

Keywords: *string of characters, textual data, NLP, information retrieval, data mining*

1. Pendahuluan

Dokter adalah sebutan untuk seorang profesional medis yang memiliki kualifikasi dan pelatihan yang memadai dalam ilmu kedokteran. Mereka memiliki pengetahuan dan keahlian untuk mendiagnosis, merawat, dan mengelola berbagai masalah kesehatan manusia. Tugas utama seorang dokter adalah untuk merawat pasien, mendiagnosis kondisi medis, dan memberikan pengobatan yang sesuai. Dalam penangan pasien yang memerlukan obat atau pengobatan tertentu, dokter akan memberi resep dokter yang dapat digunakan untuk membeli obat di apotek. Sistem pelayanan dokter, atau juga dikenal sebagai sistem perawatan Kesehatan mencakup berbagai komponen, seperti dokter, rumah sakit, klinik, pusat kesehatan, asuransi kesehatan, dan lembaga lain yang terlibat dalam perawatan kesehatan. Sistem pelayanan dokter dapat dilakukan secara otomatis dengan penggunaan machine learning, sehingga dapat membantu dokter dalam memberi pelayanan berupa diagnosis penyakit serta resep dokter kepada pasien selama 24 jam. Rekomendasi obat yang dilakukan dapat melalui proses awal, yaitu mengetahui penyakit yang dialami pasien jika pasien tidak melakukan konsultasi terlebih dahulu ke dokter. Cara agar mengetahui penyakit yang dialami pasien adalah menggunakan machine learning. Penggunaan machine learning dalam mendiagnosis penyakit sudah termasuk ke dalam data mining, karena diagnosis penyakit merupakan proses untuk mencari relasi dan informasi yang terdapat di dalam data.

1.1. Data Mining

Data mining adalah proses pencarian informasi dalam kumpulan data yang memiliki kapasitas yang besar.[6] Tujuan dari data mining adalah untuk mencari pola tersembunyi atau korelasi yang bisa menyediakan pengetahuan dan membantu dalam pengambilan keputusan. Secara garis besar metode dalam data mining dapat dibagi kedalam lima bagian yaitu klasifikasi, prediksi, asosiasi, estimasi dan klastering.[6]

1.2. Natural Language Processing

Natural Language Processing (NLP) adalah cabang dari kecerdasan buatan dan linguistik komputasional yang berfokus pada pemahaman, interpretasi, dan generasi bahasa manusia secara alami. NLP mencakup berbagai teknik dan algoritma yang dirancang untuk memungkinkan komputer untuk berinteraksi dengan, memahami, dan memanipulasi teks atau data bahasa manusia. Tujuan utama NLP adalah untuk memungkinkan komputer untuk memahami dan menganalisis bahasa manusia dalam bentuk teks atau ucapan.

1.3. Supervised Learning

Supervised Learning (Pembelajaran Terawasi) adalah model machine learning yang mempelajari data yang sudah diketahui labelnya (misalnya, klasifikasi gambar dengan label kucing atau anjing). Model ini belajar untuk menghubungkan input dengan output yang diinginkan dan kemudian dapat melakukan prediksi pada data baru.

1.4 Linear SVM

Linear SVM (Support Vector Machine) adalah salah satu algoritma pembelajaran mesin yang digunakan untuk klasifikasi dan pemisahan data. SVM adalah algoritma yang efektif untuk pemisahan dua kelas atau lebih dengan membangun hyperplane (bidang pemisah) dalam ruang fitur. Keuntungan utama dari Linear SVM adalah kemampuannya untuk mengatasi masalah klasifikasi yang kompleks dan penanganan dimensi yang tinggi. Algoritma ini juga efisien dalam penggunaan memori dan memiliki sifat matematis yang kuat.

2. Metode Penelitian

Metode penelitian sangat menentukan hasil penelitian yang akan dilakukan atau dikerjakan, karena terkait cara yang baik dan benar dalam proses pengumpulan data, analisis data dan juga dalam pengambilan keputusan dari hasil penelitian. Adapun metode penelitian yang digunakan adalah sebagai berikut:

2.1. Teknik Pengumpulan Data

Teknik pengumpulan data yang digunakan untuk memperoleh data adalah studi pustaka. Penulis melakukan studi pustaka pada website Kaggle.com untuk mencari data yang akan dipecah menjadi data latih dan data uji yang dapat digunakan untuk membuat data model machine learning. Namun sebelum itu, jumlah data akan dibatasi untuk mengoptimalkan komputasi.

2.2. Model Pengembangan Sistem

Prototyping adalah proses merancang sebuah prototype dimana prototype sendiri adalah sebuah model dari sebuah model produk yang mungkin belum memiliki semua fitur produk sesungguhnya namun sudah memiliki fitur –fitur utama dari produk sesungguhnya dan biasa digunakan untuk keperluan testing/uji coba untuk bahan uji coba sebelum berlanjut ke fase pembuatan produk sesungguhnya. [4]

2.3. Analisis Kebutuhan Sistem

Sistem yang dapat berjalan secara otomatis tentunya memerlukan metode pembelajaran mesin yang mumpuni. Dengan kata lain, sistem yang dibuat berupa model machine learning yang dapat memprediksi secara akurat. Model tersebut akan digunakan untuk memprediksi data uji dari inputan aplikasi. Data user melalui inputan aplikasi akan diproses terlebih dahulu agar mudah diprediksi. Setelah data uji diprediksi, data uji dan hasil prediksi data tersebut akan dimasukkan ke dalam database untuk disimpan. Urutan kebutuhan sistem tersebut meliputi:

a. Aplikasi

Untuk mendapat data dari user, diperlukan adanya aplikasi. Aplikasi yang dibuat berupa web application. Aplikasi ini akan digunakan oleh user untuk mendiagnosis penyakit yang diderita oleh user.

b. Application Programming Interface

API digunakan sebagai alat komunikasi dengan sistem diagnosis. API akan mengirim data dari aplikasi berupa JSON kepada sistem diagnosis yang kemudian akan menghasilkan diagnosis penyakit.

c. Sistem Diagnosis

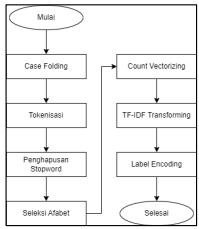
Sistem diagnosis yang dibuat meliputi fungsi pembersihan fitur dan simpanan berupa objek ekstraksi fitur dan model klasifikasi machine learning yang memiliki akurasi dengan skor 75% atau lebih. Dengan data bersih yang sudah melalui pra-pemrosesan dan sistem yang akurat, diagnosis dapat dilakukan dengan baik.

d. Database

Untuk menyimpan data uji beserta hasil prediksi data uji dari API, diperlukan database dengan kapasitas penyimpanan yang cukup. Data dalam database dapat dikirim kembali ke aplikasi sehingga user dapat menikmati sistem diagnosis aplikasi.

2.4. Pra-pemrosesan Data

Setelah data berhasil dikumpul, dilakukan pra-pemrosesan data yang bertujuan untuk mempermudah pembobotan menjadi vektor atau matriks TF-IDF.



Gambar 1. Flowchart Pra-pemrosesan Data

Berbagai tahapan yang dilakukan dalam pra-pemrosesan data adalah:

a. Pembersihan Data

Data yang akan di-preprocessing harus dibersihkan terlebih dahulu agar tidak terjadi kesalahan dalam preprocessing. Alur pembersihan data adalah sebagai berikut:

1. Case-Folding

Case-folding digunakan untuk mengubah huruf kapital menjadi huruf kecil untuk mempermudah tokenisasi.

2. Tokenisasi

Tokenisasi dilakukan untuk mempermudah stemming setiap kata/token dalam suatu kalimat yang sudah menjadi list kata.

3. Penghapusan Stopword

Stopword harus dihilangkan agar meminimalisir komputasi, sehingga pembuatan model klasifikasi dapat dilakukan secara lebih efisien.

4. Seleksi Alfabet

Karena setelah tokenisasi masih ada substring berupa angka, maka dilakukan seleksi alfabet menggunakan metode regular expression sebagai tahap terakhir dalam pembersihan data.

b. Ekstraksi Fitur

Pembobotan data bersih penting dilakukan agar data bersih dapat dikomputasi oleh algoritma machine learning. Alur pembobotan data adalah sebagai berikut:

1. Count Vectorizing

Vektorisasi data teks yang sudah bersih dilakukan dengan metode Count Vectorizer. Vektor yang dihasilkan memiliki besaran panjang yang sama dengan seluruh kata unik dalam data. Selanjutnya, akan dilakukan ekstraksi fitur TF-IDF.

2. TF-IDF Vectorizing

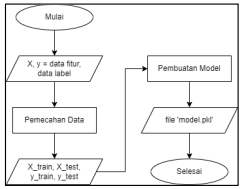
Setelah data mengalami pembobotan dengan metode Count Vectorizer, data akan dibobot dengan metode TF-IDF Vectorizer untuk mendapat ekstraksi fitur berupa vektor TF-IDF yang akan digunakan untuk melatih model.

3. Label Encoding

Label yang masih betipe kategorik harus diubah menjadi numerik dengan metode Label Encoding. Label yang sudah bertipe numerik akan mempermudah komputasi yang dilakukan untuk melatih model.

2.5. Pembuatan Model Klasifikasi

Dalam pembuatan model digunakan modul pickle. Model akan diubah menjadi file dengan ekstensi .pkl. File yang memuat model akan disisipkan ke dalam sistem sehingga proses prediksi atau diagnosis dapat dilakukan.



Gambar 2. Flowchart Pembuatan Model Klasifikasi

Langkah-langkah dalam pembuatan model adalah sebagai berikut:

1. **Pemecahan Data**

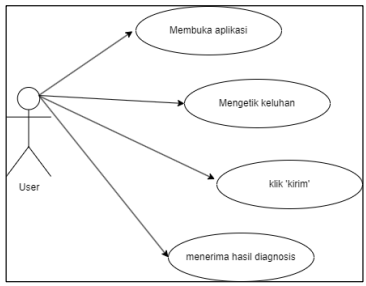
   Dataset dipecah menjadi 2 bagian terlebih dahulu, yaitu data fitur dan data label. Lalu, masing-masing data fitur dan data label akan dipecah menjadi 2, sehingga terdapat 4 bagian dataset, yaitu data fitur latih, data fitur uji, data label latih dan data label uji.

2. **Pembuatan Model**

   Model yang dibuat menggunakan algoritma Linear SVM karena memiliki waktu fitting yang relative cepat untuk di-fitting dengan data dengan banyak fitur serta dengan akurasi yang memuaskan. Fitting model dilakukan menggunakan data fitur latih dan data label latih.

3. Hasil dan Pembahasan

3.1. Use Case Diagram



Gambar 3. Use Case Diagram

Tabel 1. Deskripsi Use Case Diagram

Use Case Diagnosis	
Sasaran	User dapat mendapat diagnosis penyakit
Persyaratan	User memberi keluhan tentang gejala yang dialami
Pasca Kondisi	Website memberi diagnosis penyakit yang benar
Kondisi Akhir yang Gagal	Website memberi diagnosis penyakit yang tidak benar
Aliran Utama/Jalur Dasar	1. User membuka aplikasi 2. User mengetik keluhan pada text input 3. User mengklik tombol 'kirim'

3.2. Hasil Prediksi dan Evaluasi Model

Skor akurasi dari model yang dibuat mencapai angka 77.4 %. Dengan demikian, akurasi dianggap baik dan dapat melakukan diagnosis. Berikut merupakan hasil diagnosis dari model klasifikasi diagnosis penyakit.

Tabel 2. Hasil Prediksi Data Baru

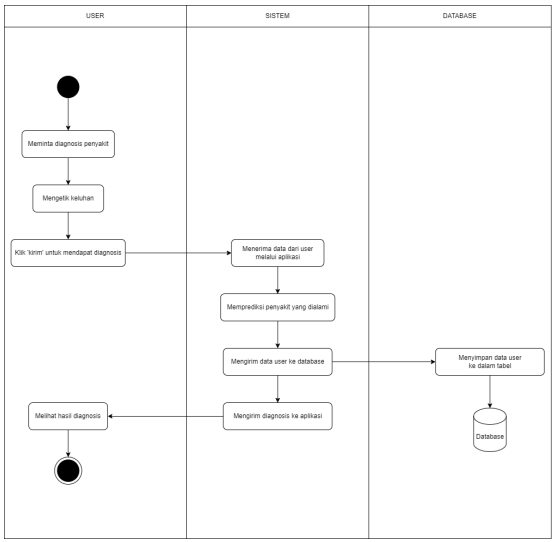
No	Keluhan	Diagnosis
1.	I am so anxious and i often panic	Anxiety
2.	I am so stressed because of the problem which hit me	Anxiety and Stress
3.	dad has been coughing for 4 hours	Cough
4.	my gf can't sleep at night, and she has been sleepless for 2 days	Insomnia
5.	my sister needs to lose some weight	Obesity

Evaluasi model dilakukan dengan metode metric scoring yang meliputi accuracy score, precision score, dan recall score. Berikut merupakan hasil evaluasi model dengan metode metric scoring:

Tabel 3. Hasil Metric Scoring

Accuracy Score	Precision Score	Recall
77.442807	74.504627	71.03638

3.3. Activity Diagram



Gambar 4. Activity Diagram

Dari gambar activity diagram, dapat dijelaskan bagaimana aplikasi bekerja. User yang ingin mendapat diagnosis penyakit dapat menggunakan aplikasi. Dari API aplikasi, sistem dapat mendapat data dari user dan mengolah data sehingga menghasilkan prediksi berupa diagnosis penyakit. Setelah itu, sistem dapat mengirim diagnosis ke aplikasi.

3.4. Rancangan Tabel pada Database

Database yang dibuat hanya memerlukan satu tabel saja untuk menampung data dari user beserta prediksi dari sistem. Data atau atribut yang disimpan di dalam tabel meliputi:

Tabel 4. Deskripsi Atribut Pada Tabel

Nama Atribut	Tipe Data	Status
id_keluhan	Varchar (8)	Unique
keluhan	Text (500)	-
prediksi	Varchar (20)	-

4. Kesimpulan

Berdasarkan penelitian tersebut, model memiliki akurasi prediksi sebesar 77,44%, skor precision sebesar 74,50%, dan skor recall sebesar 71.03%. Dari skor akurasi tersebut, model machine learning dengan algoritma Linear SVM dapat digunakan untuk mendiagnosis penyakit dari pasien. Model tersebut juga dapat diterapkan ke dalam aplikasi berbasis web. Adanya penelitian ini diharapkan dapat membantu pengelola apotek dalam memberikan pelayanan 24 jam tanpa adanya tenaga manusia, sehingga pasien dapat dilayani kapanpun dan dimanapun.

Daftar Pustaka

[1] Muhammadin, A., & Sobari, I. A., “Analisis Sentimen Pada Ulasan Aplikasi Kredivo Dengan Algoritma Svm Dan Nbc”, Reputasi: Jurnal Rekayasa Perangkat Lunak, vol. 2, no. 2, pp. 85-91, 2021.
[2] D. Vonega, A. Fadila, and D. Kurniawan, “Analisis Sentimen Twitter Terhadap Opini Publik Atas Isu Pencalonan Puan Maharani dalam PILPRES 2024”, JAIC, vol. 6, no. 2, pp. 129-135, 2022.
[3] AINIZAR, Muhammad Alif; NISA, Khoirun. “Aplikasi Informasi Pendaftaran Member dan Penjualan Merchandise pada Komunitas Manchester City Supporters Club Indonesia Chapter Purwokerto”. Jurnal Nasional Teknologi Informasi dan Aplikasnya, vol. 1, no. 2, pp. 771-780, 2023.
[4] Fitria Nur Hasanah, M.Pd, Rahmania Sri Untari, M.Pd., Rekayasa Perangkat Lunak, UMSIDA Press, pp. 23, 2020.
[5] Muhammadin, A., & Sobari, I. A., “Analisis Sentimen Pada Ulasan Aplikasi Kredivo Dengan Algoritma Svm Dan Nbc”, Reputasi: Jurnal Rekayasa Perangkat Lunak, vol. 2, no. 2, pp. 85-91, 2021.
[6] Wijaya Kusuma Sandi, Ida Bagus Gede Dwidasmara, “Implementasi Algoritma K-Means Clustering dalam Penentuan Klasifikasi Tingkat Pembangunan Perekonomian di Provinsi Bali”, Jurnal Nasional Teknologi Informasi dan Aplikasnya, vol. 1, no. 2, pp. 761-770, 2023.

Halaman ini sengaja dibiarkan kosong

JNATIA Volume 1, Nomor 4, Agustus 2023 p-ISSN: 2986-3929
Jurnal Nasional Teknologi Informasi dan Aplikasinya

Analisis Sentimen Aplikasi Zenius Menggunakan Metode Logistic Regression

I Made Juniandika[a1], Ida Bagus Made Mahendra[a2]

Program Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam,
Universitas Udayana
Jalan Raya Kampus Udayana, Bukit Jimbaran, Kuta Selatan, Badung, Bali Indonesia
[1]juniandika7072@gemail.com
[2]ibm.mahendra@unud.ac.id

Abstract

E-Learning application is one of the media used to conduct adaptive learning. One of the e-learning applications that is widely used is Zenius. To know the success of an application, sentiment analysis technique is needed. In this research, the data used to perform sentiment analysis is Zenius application user review data taken from Google Play Store with a total of 19892 review data with a score of 1,2,3, and 5. The data is classified using the logistic regression method with two classes of labels, namely positive and negative. The accuracy result obtained using this method is 90% with the classification results of positive labeled data as much as 51.17% and negative labeled data as much as 48.83%.

Keywords: *E-Learning, Sentiment Analysis, Zenius, Classifier, Logistic Regression, Accuracy*

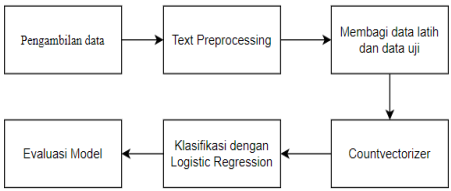
1. Pendahuluan

Kemajuan teknologi informasi memiliki dampak yang sangat besar bagi kehidupan masyarakat, salah satunya yaitu perkembangan pada bidang pendidikan. Perkembangan teknologi informasi menuntut dunia pendidikan untuk dapat mengikuti kemajuan teknologi dengan membuat metode pembelajaran yang adaptif dengan memanfaatkan berbagai media sehingga dapat meningkatkan kualitas hasil belajar. Aplikasi e-learning merupakan salah satu media yang dapat digunakan untuk melakukan proses pembelajaran secara elektronik dengan memanfaatkan internet. Ketersediaan aplikasi e-learning di Indonesia saat ini mengalami perkembangan yang pesat mulai dari yang berbayar hingga tidak berbayar, salah satunya adalah aplikasi Zenius. Zenius merupakan aplikasi yang menyediakan pembelajaran online berbayar yang diciptakan untuk memberikan kemudahan akses belajar bagi pelajar dari kalangan Sekolah Dasar sampai dengan Sekolah Menengah Atas. Zenius menawarkan layanan berupa akses video pembelajaran, forum diskusi, serta ujian online. Saat ini aplikasi Zenius sudah tersedia dalam bentuk web dan mobile sehingga dapat memudahkan siswa untuk mengakses materi pembelajaran di mana saja dan kapan saja. Untuk berbasis mobile, pengguna dapat mengunduh aplikasi secara gratis di google play store. Berdasarkan situs pada google play pada tanggal 23 April 2023, aplikasi Zenius telah diunduh kurang lebih sebanyak 5 juta unduhan, dengan rating 4.6, serta ulasan sebanyak 79.9 ribu ulasan [1]. Saat ini, Zenius bukanlah satu-satunya aplikasi yang menyediakan tawaran pembelajaran online yang tersedia di Indonesia. Terdapat aplikasi lainnya yang menawarkan layanan serupa dengan aplikasi Zenius seperti Ruangguru dan Quipper .[2] Dengan banyaknya aplikasi serupa yang tersedia membuat pengguna untuk lebih selektif dalam menentukan aplikasi yang baik dan cocok untuk digunakan. Ulasan aplikasi pada laman unduhan menjadi salah satu faktor yang memengaruhi pengguna dalam memilih aplikasi karena dijadikan sebagai bahan pertimbangan dalam menilai suatu aplikasi [3]. Ulasan-ulasan tersebut nantinya akan diolah dengan menggunakan teknik analisis sentimen. Teknik analisis sentimen ini akan sangat membantu dalam melakukan pengembangan aplikasi, hal ini dikarenakan analisis sentimen akan mengklasifikasi ulasan-ulasan positif, negatif, dan netral yang nantinya dapat dijadikan sebagai tolak ukur bagi developer dalam meningkatkan performa aplikasi. Pada tahun 2021, hal ini sudah pernah diteliti oleh Nina Ismaya Pangaribuan dkk dengan jurnal yang berjudul "Analisis Sentimen Aplikasi E-Learning Selama Pandemi Covid-19 Degnan Menggunakan Metode Support Vector

Machine dan Convolutional Neural Network" yang bertujuan untuk melakukan evaluasi pada peningkatan hal positif dan memperbaiki hal negatif dari aplikasi e-learning salah satunya yaitu aplikasi Zenius [4]. Penelitian tersebut mengambil ulasan data aplikasi Zenius sebanyak 17.350 [4]. Hasil akurasi yang diperoleh dari aplikasi Zenius dengan menggunakan metode SVM yaitu sebesar 76% sedangkan dengan metode CNN menghasilkan akurasi sebesar 84%. Perbedaan penelitian in dengan penelitian sebelumnya yaitu metode yang digunakan. Pada penelitian ini, penulis melakukan analisis sentimen aplikasi Zenius menggunakan metode logistic regression. Hasil akhir dari penelitian ini yaitu tingkat persentase sentimen positif dan negatif ulasan komentar pengguna yang dihasilkan dengan metode logistic regression

2. Metode Penelitian

Metode penelitian yang digunakan pada penelitian ini menggunakan metode eksperimen dengan urutan langkah-langkah seperti gambar di bawah ini.



Gambar 1. Langkah-langkah Penelitian

2.1 Pengambilan data

Dataset yang digunakan dalam penelitian ini yaitu ulasan komentar pengguna aplikasi Zenius dari tanggal 03 Maret 2019 sampai dengan 26 April 2023 pada google play store yang diambil dengan teknik scrapping. Ulasan yang diambil adalah ulasan berbahasa Indonesia dengan bintang satu, dua, dan tiga yang mempresentasikan ulasan negatif dan bintang lima yang mempresentasikan ulasan positif. Hasil dari data yang diperoleh berjumlah 19892 data dengan label positif sebanyak 16356 data dan label negatif sebanyak 3536 data.

2.2 Preprocessing

Tahap selanjutnya setelah melakukan pengambilan data yaitu melakukan preprocessing data yang diimplementasikan dengan menggunakan bahasa pemrograman python. Adapaun tahapan-tahapan dalam melakukan preprocessing data adalah sebagai berikut:

a. Missing value and duplicated removal

Pada tahapan ini dilakukan pembersihan data berupa penghapusan data duplikat dan menghilangkan missing value setelah proses ini selesai maka diperoleh data bersih sebanyak 16281 data dengan jumlah data berlabel positif sebanyak 12856 data dan berlabel negatif sebanyak 3425.

b. Case folding

Case folding merupakan proses pembersihan data teks dengan mengubah semua bentuk kata menjadi sama seperti mengubah ke dalam lowercase.

c. Punctuations, Special Character, and Emoticon removal

Penghilangan punctuation pada teks yang meliputi tanda baca serta simbol-simbol yang tidak penting seperti!"#$%&'() *+, -. /:;<=>?@[\]^_`{|}~, penghapusan special karakter seperti angka-angka yang terdapat dalam teks, serta penghapusan emoji.

d. Stopwords removal

Stopwords removal merupakan proses pembersihan kata-kata yang dianggap tidak relevan dengan dokumen sehingga dapat meningkatkan efektivitas pemrosesan data. Dalam tahapan ini, kata-kata yang tidak penting serta tidak memberikan kontribusi signifikan terhadap makna dokumen akan dihilangkan

e. Slang words

Slang words adalah tahapan untuk mengubah kata-kata yang mengalami kesalahan ejaan dalam penulisan ke dalam bentuk yang baku.

f. Stemming

Stemming adalah proses untuk mengubah kata ke dalam bentuk dasarnya sehingga dapat mengurangi variasi kata dalam dokumen. Tahapan ini akan menormalisasikan teks dengan menghilangkan imbuhan berupa awalan, akhiran dan sisipan dalam teks.

g. Oversampling

Oversampling adalah sebuah teknik untuk mengatasi ketidakseimbangan kelas label dalam suatu dataset dengan cara menambahkan sampel kepada kelas label yang jumlahnya sedikit. Pada penelitian ini, jumlah data kelas negatif hanya sebanyak 3425 data sehingga diperlukan teknik oversampling untuk kelas berlabel negatif sehingga jumlah data antara label positif dan negatif menjadi seimbang.

2.3 Pembagian data

Pada tahapan ini, dilakukan pembagian data menjadi data latih dan data uji dengan perbandingan 80 :20. Data latih nantinya digunakan untuk melatih model machine learning sedangkan data uji untuk menguji kinerja dari model. Pembagian proporsi yang tepat pada data latih dan data uji akan membantu dalam menciptakan model yang akurat dalam memprediksi kelas label pada data baru.

2.4 Countvectorizer

Countvectorizer adalah sebuah teknik untuk mengubah teks menjadi representasi numerik dengan menghitung frekuensi kemunculan kata dalam setiap dokumen tanpa memperhatikan urutan kata. Pada tahapan ini, setiap huruf dalam kata akan sudah mengalami proses tokenisasi.

2.5 Klasifikasi dengan Logistic Regression

Pada tahapan ini, peneliti menggunakan metode logistic regression untuk melakukan analisis sentimen yang diimplementasikan dengan bahasa pemrograman python. Tujuan digunakan metode ini yaitu untuk mendapatkan model sederhana dan terbaik untuk menjelaskan hubungan antara output dari variabel respon (Y) dengan variabel prediktornya (X) [5].

Rumus Logistic Regression:

$$\sum_{i=1}^{m} \frac{y\_pred\,(i) r== y\_test\,(i)}{M}, \text{dimana m = jumlah dataset} \quad (1)$$

2.6 Evaluasi

Pada tahapan ini, peneliti menggunakan confusion matrix untuk mengetahui performa klasifikasi model.

3. Hasil dan Diskusi

Proses pengambilan data dalam penelitian ini dilakukan dengan teknik scrapping dengan bahasa pemrograman python menggunakan library google-play-scraper dari ulasan aplikasi Zenius pada google play store.

	reviewId	userName	userImage	content	score	thumbsUpCount	reviewCreatedVersion	at	replyContent	repliedAt
0	591c5fb1-059e-45b6-8c06-8687f6b6c8cf	Juwita Anggi Pratiwi	https://play-lh.googleusercontent.com/a-/ACB-R...	masih tahap awal belajar lewat zenius dan udh ...	5	6	2.8.9	2023-04-07 10:31:28	None	NaT
1	7a86c37b-53a0-4a75-b80a-d803577c3f2b	Oktanito Yeremia	https://play-lh.googleusercontent.com/a-/ACB-R...	kenapa size data appnya besar sekali? bukannya...	3	1	2.8.9	2023-04-16 20:08:02	Halo, terimakasih atas kritik dan saran dari k...	2023-04-17 12:15:20
2	6442ab15-e7d3-4a7a-b760-b51e5289c033	Dany AkmalIunTN	https://play-lh.googleusercontent.com/a-/ACB-R...	Aplikasinya banyak bug jadi ga mood belajarnya...	1	60	2.8.7	2023-02-23 12:04:00	Hai, mohon maaf atas kendala yang kamu temui y...	2023-02-26 19:01:24

Gambar 3. Hasil Scrapping Data

Setelah proses scrapping data selesai, maka kita seleksi fitur score dengan mengambul score bernilai 1,2,3 yang nantinya mewakili ulasan negatif dan 5 mewakili ulasan positif lalu disimpan dalam file bernama “Zenius.csv”.

content	score	label
masih tahap awal belajar lewat zenius dan udh ...	5	0
Aplikasinya banyak bug jadi ga mood belajarnya...	1	1
Aplikasi nya bagus, video nya mudah di mengert...	5	0
Dari sekian banyaknya apk belajar,nih aplikasi...	1	1
zenius masih banyak bug, kadang kadang keluar ...	1	1
...	...	...
😍😍😍😍😍😍😍😍😍😍😍😍😍	5	0
💜💜💜💜💜💜💜💜💜💜💜💜💜💜💜💜💜 ...	5	0
Gud Gud Gud	5	0

Gambar 4. Seleksi Fitur Score

3.2 Text Preprocessing

Sebelum data digunakan dalam model, maka data harus mengalami preprocessing terlebih dahulu. Pada tahapan ini dilakukan case folding ke dalam lowercase

content	label	content_lower
masih tahap awal belajar lewat zenius dan udh ...	0.0	masih tahap awal belajar lewat zenius dan udh ...
kenapa size data appnya besar sekali? bukannya...	1.0	kenapa size data appnya besar sekali? bukannya...
Aplikasinya banyak bug jadi ga mood belajarnya...	1.0	aplikasinya banyak bug jadi ga mood belajarnya...
Saat pernah mendapatkan soal terkait tabel dat...	1.0	saat pernah mendapatkan soal terkait tabel dat...
Aplikasi nya bagus, video nya mudah di mengert...	0.0	aplikasi nya bagus, video nya mudah di mengert...

Gambar 5. Hasil Convert Lowercase

Selanjutnya melakukan penghapusan punctuation dengan memanfaatkan library string, special character, serta penghapusan emoticon.

	content	label	content_clean
0	masih tahap awal belajar lewat zenius dan udh ...	0.0	masih tahap awal belajar lewat zenius dan udh ...
1	kenapa size data appnya besar sekali bukannya ...	1.0	kenapa size data appnya besar sekali bukannya ...
2	aplikasinya banyak bug jadi ga mood belajarnya...	1.0	aplikasinya banyak bug jadi ga mood belajarnya...
3	saat pernah mendapatkan soal terkait tabel dat...	1.0	saat pernah mendapatkan soal terkait tabel dat...
4	aplikasi nya bagus video nya mudah di mengerti...	0.0	aplikasi nya bagus video nya mudah di mengerti...

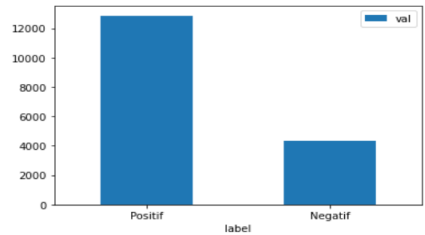
Gambar 6. Hasil Penghapusan Punctuation, Special Character, dan Emoticon

Tahap selanjutnya yaitu melakukan penghapusan stopwords untuk mengurangi kata-kata yang tidak relevan terhadap dokumen, contoh kata yang termasuk ke dalam stopwords bahas indonesia adalah “di”, “ke”, “dari”, “dimana”, “kemana”, dan lain sebagainya. Setelah melakukan penghapusan stopwords,tahap berikutnya yaitu melakukan slang word untuk mengubah kesalahan ejaan dalam penulisan kalimat seperti “7an : tujuan”, “abis : habis”, “aj : saja”, dan lain sebagainya. Dan yang terakhir adalah melakukan stemming untuk mengubah bentuk kata ke dalam bentuk dasarnya seperti “mengajar: ajar”, “memberikan: beri”, dan lain sebagainya.

	content	label
0	tahap ajar zenius udh langgan premium member n...	0.0
1	size data appnya platform bas streaming data v...	1.0
2	aplikasi bug ga mood ajar video kadang henti a...	1.0
3	kait tabel data foto tabel lengkap ngilangin p...	1.0
4	aplikasi nya bagus video nya mudah erti bahas ...	0.0

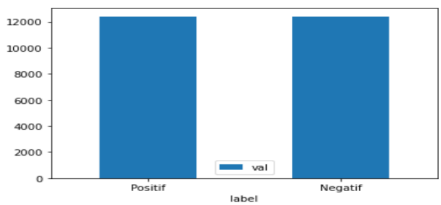
Gambar 7. Hasil Preprocessing Akhir

Setelah itu dilanjutkan dengan tahapan sampling data, karena jumlah data berlabel positif tidak seimbang dengan jumlah data yang berlabel negatif.



Gambar 8. Jumlah Data Tiap Label

Dari gambar di atas terlihat bahwa data pada kelas label tidak seimbang. Data dengan label positif berjumlah 12856 data sedangkan untuk yang berlabel hanya 5324 data sehingga perlu dilakukan oversampling pada data yang berlabel negatif sehingga jumlahnya menjadi sama dengan data berlabel positif



Gambar 9. Jumlah Data Setelah Sampling Data

3.3 Klasifikasi dengan Logistic Regression

Sebelum memasuki pemodelan dengan logistic regression, data harus dibagi menjadi data latih dan data uji. Perbandingan data uji dan data latih yang digunakan adalah 80:20. Tujuan dari data latih yaitu untuk melatih model sedangkan data uji nantinya digunakan sebagai data yang akan diprediksi. Setelah melakukan pembagian data, maka tahapan selanjutnya yaitu melakukan countvectorizer. yang nantinya di dalam tahapan ini setiap kata dalam kalimat akan mengalami proses tokenisasi kemudian akan dirubah ke numerik dalam bentuk vektor. Setelah mengalami proses countvectorizer maka masuk ke tahapan pemodelan dengan menggunakan metode logistic regression. Input yang dibutuhkan adalah hasil feature extraction X yang kemudian menghasilkan output prediksi y dengan nilai antar 0 atau 1. Jika nilai y prediksi lebih dari sama dengan 0.5 maka masuk kategori label positif (1) dan sebaliknya.

3.4 Evaluasi

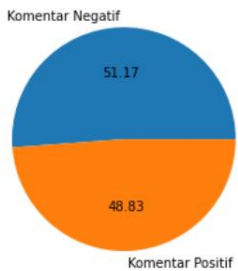
Tahapan terakhir dari pemodelan logistic regression adalah evaluasi. Evaluasi logistic regression dalam penelitian ini yaitu menggunakan confusion matrix

```
[[2186  248]
 [ 238 2292]]
              precision    recall  f1-score   support

           0       0.90      0.90      0.90      2434
           1       0.90      0.91      0.90      2530

    accuracy                           0.90      4964
   macro avg       0.90      0.90      0.90      4964
weighted avg       0.90      0.90      0.90      4964
```

Gambar 10. Hasil Confusion Matrix

Dari gambar diatas dapat dilihat bahwa nilai precision, recall, dan f-1 score untuk label kelas 1 dan label kelas 0 berada di atas 90%, yang menandakan bahwa model logistic regression ini baik dalam melakukan prediksi kelas label. Hasil akurasi yang diperoleh yaitu 90%.

3.5 Hasil



Gambar 11. Hasil Prediksi Data Tes

Berdasarkan gambar, model menghasilkan prediksi label pada data tes dengan jumlah label negatif sebesar 51.17% dan label positif sebesar 48.83%

4. Kesimpulan

Berdasarkan hasil pembahasan di atas, maka disimpulkan bahwa metode logistic regression menghasilkan prediksi yang baik dalam melakukan proses klasifikasi pada ulasan aplikasi Zenius pada google play store. Untuk melakukan pengujian, data hasil scrapping dilakukan preprocessing terlebih dahulu untuk menghasilkan data yang bersih kemudian data hasil preprocessing dibagi menjadi data latih dan data testing dengan perbandingan 80:20. Tahapan selanjutnya mengubah data tersebut ke dalam bentuk vector supaya bisa dilatih pada model logistic regression. Hasil dari pelatihan model pada data uji menunjukkan bahwa pada kelas 0, nilai precision, recall, dan f-1 score yang dihasilkan sebesar 90%, sedangkan untuk kelas 1 nilai precision dan f-1 score yang dihasilkan sebesar 90% dan recall sebesar 91%. Hasil akurasi dari

model logistic regression yaitu sebesar 90% yang menandakan bahwa model baik digunakan untuk melakukan kalsifikasi pada teks. Saran untuk penelitian selanjutnya yaitu diharapkan menggunakan jumlah dataset yang lebih banyak dalam pelatihan model serta mencoba metode baru untuk melakukan klasifikasi data teks seperti penggunaan metode deep learning.

Daftar Pustaka

[1] "Zenius - #GantiCaraBelajar - Aplikasi di Google Play." https://play.google.com/store/apps/details?id=net.zenius.mobile&hl=id&gl=US (accessed May 07, 2023).

[2] M. R. Firdaus, F. M. Rizki, F. M. Gaus, and I. K. Susanto, "Analisis Sentimen Dan Topic Modelling Dalam Aplikasi Ruangguru," J-SAKTI (Jurnal Sains Komput. dan Inform., vol. 4, no. 1, p. 66, 2020, doi: 10.30645/j-sakti.v4i1.188.

[3] E. Fitri, Y. Yuliani, S. Rosyida, and W. Gata, "Analisis Sentimen Terhadap Aplikasi Ruangguru Menggunakan Algoritma Naive Bayes, Random Forest Dan Support Vector Machine," J. Transform., vol. 18, no. 1, pp. 71–80, Jul. 2020, Accessed: Apr. 25, 2023. [Online]. Available: https://journals.usm.ac.id/index.php/transformatika/article/view/2317.

[4] A. S. Simbolon, N. I. Pangaribuan, and N. M. Aruan, "Analisis Sentimen Aplikasi E-Learning Selama Pandemi Covid-19 Dengan Menggunakan Metode Support Vector Machine Dan Convolutional Neural Network," Seminastika, vol. 3, no. 1, pp. 16–25, 2021, doi: 10.47002/seminastika.v3i1.236.

[5] K. Kelvin, J. Banjarnahor, E. I. -, and M. NK Nababan, "Analisis perbandingan sentimen Corona Virus Disease-2019 (Covid19) pada Twitter Menggunakan Metode Logistic Regression Dan Support Vector Machine (SVM)," J. Sist. Inf. dan Ilmu Komput. Prima(JUSIKOM PRIMA), vol. 5, no. 2, pp. 47–52, 2022, doi: 10.34012/jurnalsisteminformasidanilmukomputer.v5i2.2365.

JNATIA Volume 1, Nomor 4, Agustus 2023 p-ISSN: 2986-3929
Jurnal Nasional Teknologi Informasi dan Aplikasinya

Analisis Penggunaan Metode MFCC dalam Mendeteksi Emosi pada Musik Indonesia

I Komang Sutrisna Eka Wijaya[a1], Luh Arida Ayu Rahning Putri[a2]

Program Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam,
Universitas Udayana
Jalan Raya Kampus Udayana, Bukit Jimbaran, Kuta Selatan, Badung, Bali Indonesia
[1]ekawijaya032@student.unud.ac.id
rahningputri[2]@unud.ac.id

Abstract

This research aims to develop a method for detecting emotions in Indonesian music using the MFCC method. The MFCC method is used to identify emotions in music by measuring acoustic features of music, such as tempo, pitch, and intensity. The study uses a dataset of 40 Indonesian music samples from various regions, which are analyzed to detect emotions. The confusion matrix is used to calculate the precision of emotion prediction. The results show that the MFCC method is effective in detecting emotions in Indonesian music. The research also highlights the importance of using a representative dataset to improve the accuracy of emotion identification in music. This study provides insights into the challenges and opportunities of using the MFCC method for emotion detection in Indonesian music.

Keywords: *MFCC, emotion detection, Indonesian music, dataset, confusion matrix.*

1. Pendahuluan

Musik merupakan salah satu bentuk seni yang memiliki pengaruh besar terhadap emosi manusia. Musik dapat membangkitkan perasaan bahagia, sedih, marah, dan lain sebagainya. Oleh karena itu, banyak penelitian yang dilakukan untuk mengidentifikasi emosi dalam musik. Salah satu metode yang digunakan adalah Mel Frequency Cepstral Coefficients (MFCC). Metode ini telah banyak digunakan dalam pengenalan suara dan pengenalan wajah, namun belum banyak digunakan dalam mendeteksi emosi dalam musik Indonesia. Terdapat beberapa tantangan dalam menggunakan metode MFCC dalam mendeteksi emosi dalam musik Indonesia. Pertama, musik Indonesia memiliki ciri khas yang berbeda dengan musik dari negara lain. Hal ini memerlukan penyesuaian dalam penggunaan metode MFCC. Kedua, emosi dalam musik Indonesia dapat bervariasi tergantung pada genre musik dan lirik lagu. Oleh karena itu, diperlukan dataset yang representatif untuk dapat mengidentifikasi emosi dengan akurasi yang tinggi. Beberapa penelitian sebelumnya telah dilakukan dalam mendeteksi emosi dalam musik menggunakan metode MFCC. Salah satu penelitian yang relevan adalah panellation Music Emotion Recognition: A State-of-the-Art Review yang dibuat oleh Yang, Y., & Chen, K [1]. Penelitian ini membahas tentang pengenalan emosi dalam musik dan memberikan tinjauan terhadap teknologi pengenalan emosi dalam musik yang ada pada saat itu. Penelitian ini juga membahas tentang tantangan dan tren pengembangan teknologi pengenalan emosi dalam musik. Penelitian ini memiliki relevansi yang tinggi dalam industri musik Indonesia. Dengan menggunakan metode MFCC, dapat dikembangkan aplikasi yang dapat mengidentifikasi emosi dalam musik Indonesia. Aplikasi ini dapat membantu para musisi dan produser musik dalam memilih lagu yang tepat untuk menciptakan suasana yang diinginkan dalam sebuah karya musik. Selain itu, aplikasi ini juga dapat digunakan dalam industri film dan iklan untuk menciptakan suasana yang tepat dalam sebuah adegan atau iklan.

2. Metode Penelitian

Tahapan yang akan dilakukan oleh peneliti adalah Studi Literatur, Pengumpulan Data, Ekstraksi Fitur, Analisis Emosi dalam Musik dan Evaluasi Tingkat Akurasi

2.1 Studi Literatur

Setelah masalah diidentifikasi, langkah selanjutnya adalah melakukan studi literatur untuk memperoleh informasi tentang penggunaan metode MFCC dalam mendeteksi emosi dalam musik. Studi literatur dilakukan dengan mencari referensi dari jurnal ilmiah, buku, dan sumber-sumber terpercaya lainnya.

a. Emosi dalam musik

Emosi dalam musik adalah perasaan atau pengalaman subjektif yang muncul ketika seseorang mendengarkan musik. Emosi dapat dipengaruhi oleh berbagai faktor seperti melodi, harmoni, ritme, dan lirik. Beberapa penelitian menunjukkan bahwa musik dapat memengaruhi suasana hati dan emosi seseorang [2]. Oleh karena itu, deteksi emosi dalam musik menjadi penting dalam berbagai aplikasi seperti pengembangan musik terapi, pengembangan perangkat lunak pemutar musik yang dapat menyesuaikan diri dengan emosi pengguna, dan pengembangan sistem pengenalan emosi dalam musik. Musik memiliki kekuatan untuk membangkitkan emosi pada pendengarnya. Emosi yang dihasilkan dari musik dapat beragam, tergantung dari jenis musik yang didengarkan dan pengalaman pendengar. Menurut Juslin dan Västfjäll (2008) [6], terdapat enam jenis emosi yang dapat dihasilkan dari musik, dan ini juga lah yang menjadi label dari emosi yang di gunakan pada penelitian ini, yaitu:

1. Kegembiraan (happiness)
2. Kesedihan (sadness)
3. Ketakutan (fear)
4. Kemarahan (anger)
5. Ketenangan (peacefulness)
6. Keindahan (beauty)

Setiap jenis emosi memiliki karakteristik yang berbeda-beda. Kegembiraan biasanya dihasilkan dari musik yang memiliki tempo cepat dan irama yang riang. Sedangkan kesedihan dihasilkan dari musik yang memiliki tempo lambat dan melodi yang sedih. Ketakutan dihasilkan dari musik yang memiliki nada-nada minor dan tempo yang lambat. Kemarahan dihasilkan dari musik yang memiliki tempo cepat dan irama yang keras. Ketenangan dihasilkan dari musik yang memiliki tempo lambat dan irama yang tenang. Sedangkan keindahan dihasilkan dari musik yang memiliki harmoni yang indah dan melodi yang menyentuh. Berdasarkan penelitian yang dilakukan oleh Juslin dan Västfjäll (2008) [6], jenis emosi yang dihasilkan dari musik dapat mempengaruhi suasana hati dan perilaku seseorang. Oleh karena itu, pemilihan jenis musik yang tepat dapat membantu seseorang untuk mengatasi stres, meningkatkan konsentrasi, dan memperbaiki suasana hati.

b. Metode MFCC

Mel-frequency cepstral coefficients (MFCC) adalah metode yang digunakan untuk menganalisis sinyal suara. Metode ini mengubah sinyal suara menjadi serangkaian koefisien cepstral yang merepresentasikan karakteristik sinyal suara. Metode MFCC telah digunakan dalam berbagai aplikasi seperti pengenalan suara, pengenalan pembicara, dan pengenalan musik. Beberapa penelitian telah menggunakan metode MFCC untuk mendeteksi emosi dalam musik. Misalnya, penelitian oleh Yang [1] menggunakan metode MFCC untuk mengidentifikasi emosi dalam musik Tiongkok. Penelitian tersebut menunjukkan bahwa metode MFCC dapat digunakan untuk mendeteksi emosi dalam musik dengan akurasi yang tinggi. Penelitian lain oleh Li et al. (2016) [3] menggunakan metode MFCC untuk mendeteksi emosi dalam musik barat. Penelitian tersebut menunjukkan bahwa metode MFCC dapat digunakan untuk mendeteksi emosi dalam musik barat dengan akurasi yang tinggi.

c. Metode SVM

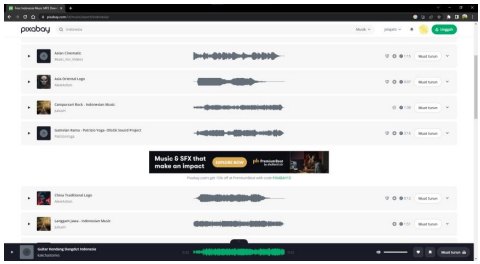
Metode SVM (Support Vector Machine) adalah algoritma pembelajaran mesin yang digunakan untuk klasifikasi dan regresi. SVM bekerja dengan cara memetakan data pelatihan ke dalam ruang fitur yang lebih tinggi menggunakan fungsi kernel, sehingga data yang semula tidak dapat dipisahkan secara linear dapat dipisahkan dengan jelas di ruang fitur yang lebih tinggi. SVM kemudian mencari hyperplane (bidang pemisah) yang optimal untuk memisahkan data ke dalam kelas-kelas yang berbeda. SVM memiliki keunggulan dalam menangani data yang kompleks dan memiliki performa yang baik dalam klasifikasi data dengan dimensi yang tinggi.

d. Python

Python adalah bahasa pemrograman tingkat tinggi yang mudah dipelajari dan dipahami. Python memiliki sintaks yang sederhana dan mudah dibaca, sehingga sangat direkomendasikan bagi pemula yang ingin belajar pemrograman. Python juga memiliki banyak keunggulan, terutama dalam bidang Data Science, seperti kemampuan untuk melakukan analisis data dari database Big Data dan menyediakan dukungan struktur kode yang mempermudah pengembangan software. Python juga digunakan di berbagai bidang pengembangan seperti produk, software, dan aplikasi, termasuk penerapannya pada website dan internet seperti HTML, email processing, dan FTP. Python menjadi salah satu bahasa pemrograman yang paling populer dan diminati oleh praktisi data karena kemudahan penggunaannya pada proses pengolahan data.

2.2 Pengumpulan Data

Pengumpulan data dilakukan dengan cara mengumpulkan data musik yang akan digunakan dalam penelitian ini. Data musik yang digunakan harus memenuhi kriteria tertentu, seperti memiliki variasi emosi yang cukup, memiliki durasi yang cukup, dan memiliki kualitas suara yang baik. Saya mengumpulkan data untuk penelitian ini dengan menggunakan dataset musik Indonesia dari Pixabay. Pixabay merupakan sumber data yang saya pilih karena menyediakan akses publik ke berbagai konten visual, termasuk gambar dan video. Dalam konteks penelitian ini, saya fokus pada penggunaan dataset musik Indonesia yang tersedia di Pixabay. Atas izin dan kebijakan penggunaan yang ditetapkan oleh Pixabay, saya memilih dataset musik Indonesia yang relevan dengan tujuan penelitian saya. Dataset ini mencakup berbagai genre musik Indonesia yang mencerminkan keragaman emosi yang ingin saya analisis. Dalam proses pengumpulan data, saya mencari kata kunci yang sesuai dengan musik Indonesia pada platform Pixabay dan memilih file-file musik yang relevan dengan fokus penelitian saya. Setelah itu, saya mengunduh dataset tersebut dan memprosesnya sesuai dengan langkah-langkah pra-pemrosesan yang telah disebutkan sebelumnya.



Gambar 1. Data yang ada di pixabay

Pemilihan dataset musik Indonesia dari Pixabay sebagai sumber data dalam penelitian ini memungkinkan saya untuk mengakses kumpulan musik yang bervariasi dan mewakili berbagai emosi yang terkait dengan musik Indonesia. Dalam pengumpulan data, saya memastikan kepatuhan terhadap persyaratan penggunaan dan hak cipta yang ditetapkan oleh Pixabay.

2.3 Ekstraksi Fitur

Setelah data musik terkumpul, langkah selanjutnya adalah melakukan ekstraksi fitur menggunakan metode MFCC. Metode ini digunakan untuk mengubah sinyal musik menjadi beberapa parameter yang dapat digunakan untuk mendeteksi emosi dalam musik. Berikut adalah tahapan ekstraksi fitur yang akan dilakukan dalam penelitian ini:

a. Preprocessing Data Musik

Sebelum melakukan ekstraksi fitur, data musik harus di preprocessing terlebih dahulu. Preprocessing ini dilakukan untuk memastikan bahwa data musik siap untuk diekstraksi fiturnya menggunakan metode MFCC. Preprocessing dapat meliputi normalisasi amplitudo, filtering, dan lain-lain.

b. Frame Blocking

Setelah data musik di preprocessing, langkah selanjutnya adalah melakukan frame blocking. Frame blocking dilakukan untuk membagi sinyal musik menjadi beberapa frame yang memiliki durasi tertentu. Durasi frame yang digunakan dapat disesuaikan dengan kebutuhan penelitian.

c. Windowing

Setelah frame blocking, langkah selanjutnya adalah melakukan windowing. Windowing dilakukan untuk memperhalus peralihan antara frame yang berbeda. Pada tahap ini, digunakan windowing dengan jenis Hamming window.

d. Fast Fourier Transform (FFT)

Setelah windowing, langkah selanjutnya adalah melakukan Fast Fourier Transform (FFT). FFT dilakukan untuk mengubah sinyal musik dari domain waktu menjadi domain frekuensi.

e. Mel Frequency Cepstral Coefficients (MFCC)

Setelah FFT, langkah selanjutnya adalah melakukan ekstraksi fitur menggunakan metode MFCC. Metode ini digunakan untuk mengubah sinyal musik menjadi beberapa parameter yang dapat digunakan untuk mendeteksi emosi dalam musik. Tahapan ekstraksi fitur menggunakan metode MFCC meliputi:

1. Mel Filterbank: Mel Filterbank digunakan untuk mengubah domain frekuensi menjadi domain mel.
2. Logarithmic Transformation: Logarithmic Transformation digunakan untuk mengubah domain mel menjadi domain logaritmik.
3. Discrete Cosine Transform (DCT): DCT digunakan untuk mengubah domain logaritmik menjadi domain cepstral.
4. Normalisasi Fitur. Setelah fitur musik diekstraksi, langkah selanjutnya adalah melakukan normalisasi fitur. Normalisasi fitur dilakukan untuk memastikan bahwa setiap fitur memiliki rentang nilai yang sama.

Dengan melakukan tahapan ekstraksi fitur yang sistematis dan terstruktur, diharapkan fitur musik yang diekstraksi menggunakan metode MFCC dapat memperoleh informasi yang cukup

untuk mendeteksi emosi dalam musik. Selain itu, dengan melakukan normalisasi fitur, diharapkan setiap fitur memiliki rentang nilai yang sama dan dapat diinterpretasikan dengan baik.

2.4 Analisis Emosi dalam Musik

Setelah fitur musik diekstraksi, langkah selanjutnya adalah melakukan analisis emosi dalam musik. Analisis ini dilakukan dengan menggunakan metode klasifikasi yakni Support Vector Machine. Analisis emosi dalam musik Indonesia menggunakan metode MFCC dan SVM adalah pendekatan yang efektif dalam mengidentifikasi dan memahami emosi yang terkandung dalam musik. Dalam tahap ekstraksi fitur, koefisien MFCC diekstraksi dari data musik, yang merepresentasikan karakteristik audio yang berhubungan dengan emosi. Selanjutnya, SVM digunakan sebagai algoritme klasifikasi untuk memisahkan dan mengklasifikasikan data berdasarkan fitur-fitur MFCC yang diekstraksi. SVM membantu dalam mengenali emosi seperti kegembiraan, kesedihan, ketakutan, kemarahan, ketenangan, dan keindahan yang terkandung dalam musik Indonesia. Dengan mengintegrasikan MFCC dan SVM, analisis emosi dalam musik memberikan pemahaman yang lebih dalam tentang pengaruh musik terhadap emosi manusia dan dapat mendukung pengambilan keputusan dalam industri musik. Penggunaan metode MFCC dan SVM dalam analisis emosi musik Indonesia memberikan manfaat penting bagi industri musik dan seni. Dengan mengidentifikasi emosi yang terkandung dalam musik, musisi dan produser musik dapat menciptakan pengalaman musik yang lebih kuat dan autentik. Selain itu, analisis emosi dalam musik juga dapat digunakan dalam industri film dan iklan untuk menciptakan suasana yang sesuai dengan pesan yang ingin disampaikan. Dalam beberapa kata, metode MFCC dan SVM memberikan pendekatan yang kuat dalam menganalisis emosi dalam musik Indonesia, membantu meningkatkan pengalaman pendengar, dan mendukung pengambilan keputusan dalam industri kreatif.

2.5 Evaluasi Tingkat Akurasi

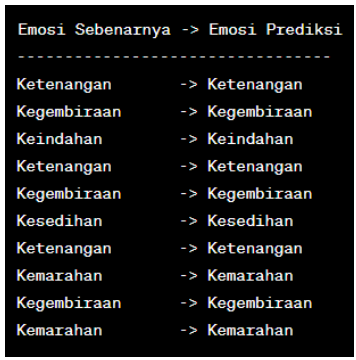
Setelah analisis dilakukan, langkah selanjutnya adalah mengevaluasi tingkat akurasi dari metode yang digunakan. Evaluasi ini dilakukan dengan membandingkan hasil klasifikasi dengan label emosi yang sebenarnya pada data uji. Confusion matrix adalah sebuah metode yang digunakan untuk mengevaluasi kinerja model klasifikasi atau prediksi dengan memperlihatkan jumlah prediksi yang benar dan salah dalam bentuk matriks. Matriks ini memberikan gambaran tentang sejauh mana model dapat mengklasifikasikan data dengan benar berdasarkan emosi yang diuji. Dalam penelitian ini, confusion matrix digunakan untuk mengukur kualitas prediksi emosi dalam musik Indonesia. Confusion matrix membantu dalam mengevaluasi akurasi, presisi, recall, dan F1-score dari model yang digunakan. Dengan melihat matriks ini, peneliti dapat mengidentifikasi seberapa baik model dapat memprediksi emosi yang benar dan di mana model mungkin mengalami kesalahan dalam mengklasifikasikan emosi. Confusion matrix membantu peneliti dalam memahami performa model dan membantu mengambil keputusan yang tepat berdasarkan evaluasi kualitas prediksi. Dengan memperhatikan matriks ini, peneliti dapat mengidentifikasi kelas emosi mana yang lebih sulit diprediksi atau memerlukan peningkatan dalam model. Oleh karena itu, penggunaan confusion matrix dalam penelitian ini dapat memberikan pemahaman yang lebih komprehensif tentang kinerja model dalam mendeteksi emosi dalam musik Indonesia.

3. Hasil dan Pembahasan

Hasil dari penelitian ini adalah pengembangan metode untuk mendeteksi emosi dalam musik dengan menggunakan metode MFCC. Metode ini digunakan untuk mengidentifikasi emosi dalam musik dengan mengukur fitur akustik dari musik, seperti tempo, nada, dan intensitas. Metode MFCC digunakan karena metode ini telah terbukti efektif dalam pengenalan suara dan pengenalan ucapan. Dalam Program, metode yang dihasilkan sudah diterapkan sebagai berikut:

a. Import library yang diperlukan: Program dimulai dengan mengimpor buku arsip seperti numpy, librosa, dan sklearn.

b. Fungsi Ekstraksi Fitur MFCC: Fungsi extract_mfcc_features mengekstraksi fitur MFCC dari file audio. Fungsi ini menggunakan library librosa untuk memuat file audio dan mengekstraksi 13 koefisien MFCC.
c. Data dan label emosi: Fitur MFCC dan label emosi yang terkait disimpan dalam variabel data dan label. Fitur MFCC diekstraksi dari setiap file audio yang mewakili emosi yang berbeda menggunakan fungsi extract_mfcc_features, yang kemudian disimpan dalam variabel data dan label emosi.
d. Peneliti memiliki 40 data musik Indonesia dari berbagai daerah yang dianalisis untuk mendeteksi emosi. Kita dapat menggunakan confusion matrix untuk menghitung presisi dari prediksi emosi.
e. Pembagian data: Fungsi train_test_split dari library sklearn digunakan untuk membagi data fitur dan label menjadi 80% data latih dan 20% data uji.
f. Normalisasi fitur: Untuk memastikan bahwa data berada pada skala yang sama, fitur data latih dan uji dinormalisasi menggunakan StandardScaler dari sklearn.
g. Pelatihan model SVM: Model Support Vector Machine, atau SVM, dilatih dengan menggunakan data latih yang telah dinormalisasi.
h. Pengujian model: Untuk memastikan akurasi model, model yang telah dilatih diuji pada data uji.
i. Hasil: Pernyataan print digunakan untuk mencetak akurasi model pada layar.



Gambar 2. Emosi Sebenarnya dan Emosi Prediksi

Berikut adalah confusion matrix yang dihasilkan:

Tabel 1. Confusion Matrix untuk mengevaluasi kinerja model klasifikasi

		Emosi Prediksi				
		Ketenangan	Kegembiraan	Keindahan	Kesedihan	Kemarahan
Emosi Sebenarnya	Ketenangan	10	0	0	0	0
	Kegembiraan	0	10	0	0	0
	Keindahan	0	0	8	2	0
	Kesedihan	0	0	1	1	8
	Kemarahan	0	0	0	0	0

Dalam konteks Confusion Matrix yang telah diberikan sebelumnya, untuk menghitung presisi, kita dapat menggunakan rumus:

Presisi = TP / (TP + FP)

Dimana:

TP (True Positive) adalah jumlah data yang diprediksi dengan benar sebagai positif.
FP (False Positive) adalah jumlah data yang diprediksi salah sebagai positif.
Berikut adalah contoh perhitungan presisi menggunakan Confusion Matrix sebelumnya:
Presisi Kegembiraan = 10 / (10 + 0) = 1.00
Presisi Kesedihan = 10 / (10 + 1) = 0.91
Presisi Ketakutan = 8 / (8 + 1) = 0.89
Presisi Kemarahan = 8 / (8 + 2) = 0.80
Dalam kasus ini, presisi untuk setiap emosi yang berhasil dihitung adalah sebagai berikut:
Presisi Kegembiraan: 1.00 atau 100%
Presisi Kesedihan: 0.91 atau 91%
Presisi Ketakutan: 0.89 atau 89%
Presisi Kemarahan: 0.80 atau 80%

Presisi menggambarkan sejauh mana model berhasil memprediksi dengan benar untuk setiap kelas emosi secara individu. Semakin tinggi nilai presisi, semakin baik model dalam memprediksi kelas emosi tersebut dengan benar. Dalam contoh ini, semua emosi memiliki presisi yang sempurna (1.0), yang menunjukkan bahwa model berhasil mengklasifikasikan dengan benar semua prediksi untuk setiap emosi. Berdasarkan hasil pengukuran akurasi dan presisi menggunakan confusion matrix pada output yang diberikan sebelumnya, dapat ditarik beberapa kesimpulan:

a. Akurasi Keseluruhan:

Akurasi keseluruhan dapat dihitung dengan menjumlahkan semua prediksi yang benar (diagonal) pada Confusion Matrix, kemudian membaginya dengan jumlah total data yang dievaluasi.
Jumlah prediksi yang benar = TP (True Positive) dari semua emosi = 10 + 10 + 8 + 8 = 36
Jumlah total data yang dievaluasi = Jumlah semua entri pada Confusion Matrix = 40
Akurasi keseluruhan = (Jumlah prediksi yang benar / Jumlah total data yang dievaluasi) * 100%
= (36 / 40) * 100%
= 90%
Dengan demikian, akurasi keseluruhan dari model ini adalah 90%.

b. Presisi Keseluruhan:

Presisi Keseluruhan:
Presisi keseluruhan dapat dihitung dengan menjumlahkan semua TP (True Positive) dari setiap emosi, kemudian membaginya dengan jumlah semua TP dan FP (False Positive) dari setiap emosi.
Jumlah TP dari semua emosi = 10 + 10 + 8 + 8 = 36
Jumlah FP dari semua emosi = 0 + 1 + 1 + 2 = 4
Presisi keseluruhan = (Jumlah TP / (Jumlah TP + Jumlah FP)) * 100%
= (36 / (36 + 4)) * 100%
= (36 / 40) * 100%
= 90%

Dengan demikian, presisi keseluruhan dari model ini adalah 90%. Model klasifikasi yang dihasilkan memiliki akurasi keseluruhan sebesar 90%, yang berarti model tersebut mampu memprediksi dengan benar 90% dari total data yang dievaluasi. Selain itu, presisi keseluruhan

juga sebesar 90%, yang menunjukkan bahwa model ini memiliki kemampuan yang baik dalam memprediksi emosi dengan benar secara keseluruhan. Hal ini menunjukkan bahwa model memiliki performa yang baik dalam mengklasifikasikan emosi pada data yang diberikan.

4. Kesimpulan

Penelitian ini menunjukkan bahwa metode MFCC efektif dalam mendeteksi emosi dalam musik Indonesia. Dalam penelitian ini, peneliti menggunakan dataset musik Indonesia yang representatif dan menghitung presisi prediksi emosi menggunakan confusion matrix. Hasil penelitian menunjukkan bahwa metode MFCC dapat mengidentifikasi emosi dalam musik dengan akurasi yang baik. Penelitian ini juga menyoroti pentingnya penggunaan dataset yang representatif dalam meningkatkan akurasi identifikasi emosi dalam musik. Dengan demikian, penelitian ini memberikan wawasan tentang tantangan dan peluang penggunaan metode MFCC untuk mendeteksi emosi dalam musik Indonesia.

Daftar Pustaka

[1] Yang, Y. dan Chen, K., "Music emotion recognition: A state-of-the-art review," IEEE Transactions on Affective Computing, vol. 10, no. 3, pp. 374-393, Jul. 2018. Tersedia: https://ieeexplore.ieee.org/document/8274792 .

[2] "Mengapa musik memiliki pengaruh dahsyat pada emosi?," BBC News Indonesia, 7 Oktober 2015. [Online]. Available: https://www.bbc.com/indonesia/vert_fut/2015/10/151002_vert_fut_musik_emosi. [Diakses: 13 Juni 2023]

[3] Y. Li, Y. Yang, dan K. Chen, "Music emotion recognition: A survey of recent advances," IEEE Transactions on Human-Machine Systems, vol. 49, no. 4, pp. 377-393, 2019. [Online]. Available: https://ieeexplore.ieee.org/document/8648764

[4] A. S. Hidayat and M. Handayani, "Implementasi Metode Mel-Frequency Cepstral Coefficients (MFCC) pada Pengenalan Suara dengan Backpropagation Neural Network," Jurnal Matematika, Statistika dan Aplikasinya, vol. 18, no. 1, pp. 1-10, 2019. Tersedia: https://ejournal.undip.ac.id/index.php/jmasif/article/download/34875/18385

[5] R. Lestari, "Perancangan Sistem Informasi Musik Tradisional Jawa Tengah," Repertoar: Jurnal Ilmiah Pendidikan Musik, vol. 1, no. 2, pp. 71-78, 2018. [Online]. Tersedia: https://journal.unesa.ac.id/index.php/Repertoar/article/view/18765

[6] P. N. Juslin dan D. Västfjäll, "Emotional responses to music: The need to consider underlying mechanisms," Behavioral and Brain Sciences, vol. 31, no. 5, pp. 559-575, 2008.

JNATIA Volume 1, Nomor 4, Agustus 2023 p-ISSN: 2986-3929
Jurnal Nasional Teknologi Informasi dan Aplikasinya

Perancangan Ontologi Semantik: Representasi Digital Kain Songket Bali

I Made Suma Gunawan[a1], Dr. Dra. Luh Gede Astuti[a2]

Program Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam,
Universitas Udayana
Jalan Raya Kampus Udayana, Bukit Jimbaran, Kuta Selatan, Badung, Bali Indonesia
[1]sumagunawan04@gmail.com
[2]lg.astuti@unud.ac.id

Abstract

Advances in information technology are currently driving an increase in the number of internet users. In the business world, technology plays an important role and can be considered as a basic necessity for today's entrepreneurs. Likewise, in the textile industry, especially Balinese songket craftsmen or entrepreneurs, it is very important to utilize digital platforms to assess the results of Balinese songket fabrics and manage various aspects properly such as quality, color and pattern, innovation, and price. Thus, this knowledge can be shared with the public, passed down from generation to generation, and promoted to both local and foreign tourists, thereby increasing industry recognition and contributing to the country's foreign exchange earnings. To achieve this, the use of ontologies as an information representation technique is highly favored. Ontologies can greatly facilitate the development of semantic applications, especially when dealing with the semantic web. The methodology used for the development of the ontology model in this study is a process that allows the reuse of existing ontologies in subsequent system development. The development of the ontology model in this study uses a software application called Protégé, which has shown promising results during the ontology evaluation stage in providing satisfactory answers to user questions.

Keywords: *Ontologi, Kain Songket Bali, SPARQL, Web Semantik, Methontology*

1. Pendahuluan

Bali merupakan salah satu pulau yang memiliki berbagai warisan budaya yang diwariskan oleh leluhur. Warisan budaya yang ada di Bali sungguh sangat banyak dan beragam [7], dan salah satunya adalah kain songket bali. Namun, sejak akhir tahun 2019 terdapat wabah virus Corona yang datang membawa dampak buruk yang besar bagi seluruh sektor khususnya industri pariwisata di Bali. Hal tersebut menyebabkan turunnya jumlah wisatawan lokal maupun mancanegara yang liburan di Bali dan tentu sangat berdampak pada perekonomian para pengusaha atau pengrajin souvenir khususnya kain songket di Bali. Untuk menjaga kelestarian warisan budaya dan memulihkan sektor pariwisata di Bali, diperlukan suatu solusi yang dapat mengubahnya menjadi bentuk digital yang jelas dan nyata. Penggunaan ontologi juga membuka peluang untuk mengembangkan sistem manajemen pengetahuan dan mengubah paradigma dari orientasi dokumen menjadi pengetahuan yang terhubung, dapat digabungkan, serta dapat dimanfaatkan kembali dengan lebih fleksibel dan dinamis [5]. Ontologi membuka peluang untuk mengembangkan sistem manajemen pengetahuan dan memungkinkan peralihan dari pendekatan berbasis dokumen menjadi pengetahuan yang saling terkait, dapat digabungkan, dan dapat dimanfaatkan kembali dengan lebih fleksibel dan dinamis [5]. Selain itu, ontologi juga memiliki hubungan dengan web semantik. Web semantik merupakan teknologi web yang membantu komputer memahami makna kata atau kalimat yang diberikan oleh pengguna. Dengan menggunakan web semantik, komputer dapat lebih mudah memproses informasi dan memahami informasi yang diinginkan oleh pengguna [3-4]. Melalui penelitian ini, penulis bertujuan untuk memperoleh informasi terkait dengan kain Songket Bali guna melestarikan warisan budaya kita serta memberikan pemahaman kepada generasi muda tentang budaya ini. Selain itu, penelitian ini juga bertujuan untuck mendukung pemulihan sektor pariwisata di Bali. Metode yang digunakan

untuk membangun model ontologi adalah Methontology. Methontology merupakan salah satu pendekatan yang dapat digunakan dalam mengembangkan model ontologi. Keunggulan dari metode ini adalah kemampuannya dalam menggambarkan setiap aktivitas secara detail. Selain itu, metode Methontology juga memungkinkan penggunaan kembali ontologi yang telah dibangun untuk pengembangan sistem yang lebih lanjut [4]. Dalam usulan penelitian ini, tujuannya adalah untuk membangun model ontologi yang merepresentasikan pengetahuan tentang kain songket Bali. Dan diharapkan bahwa melalui penelitian ini, dapat dibangun model ontologi yang berkualitas tinggi dan mampu menggambarkan pengetahuan dengan baik.

1.1 Kain Songket Bali

Kain songket Bali adalah jenis kain tradisional khas yang terbuat dari sutra murni dengan benang emas. Awalnya, kain ini diperuntukkan khusus bagi penduduk Bali yang berasal dari keturunan Brahmana, kasta tertinggi dalam agama Hindu. Kain songket memiliki pola hiasan yang bervariasi dan sangat terkait dengan ritual pemujaan. Penenun songket di lingkungan puri membuat kain ini sesuai dengan fungsi dan tujuan yang diinginkan. Proses pembuatan songket berbeda tergantung pada penggunaannya. Saat ini, kain songket Bali dapat digunakan oleh siapa saja, baik untuk acara resmi maupun keagamaan. Kain ini sering dipakai dalam berbagai upacara penting dalam kehidupan masyarakat Bali, seperti upacara potong gigi, perkawinan, hari raya, kremasi, upacara keagamaan, dan acara adat [8].

1.2 Ontologi

Ontologi merupakan bidang pengetahuan yang berisi informasi tentang keberadaan hal-hal yang nyata maupun tidak nyata. Menurut para ahli, ontologi adalah ilmu yang mempelajari bagaimana menyampaikan informasi dengan cara yang terhubung satu sama lain. Selain itu, ontologi juga dapat diartikan sebagai gambaran konsep yang ada dalam suatu domain dan dapat terhubung antara domain yang berbeda. Salah satu makna lain dari ontologi adalah sebagai representasi perkembangan teknologi dalam sebuah kosakata yang dapat ditemukan melalui resource description framework yang memungkinkan adanya keterhubungan antara informasi. Menggunakan ontologi memiliki beberapa keuntungan, seperti kemampuannya untuk menjelaskan secara eksplisit domain pengetahuan. Ontologi juga menyediakan struktur hierarki konsep yang membantu dalam menjelaskan domain dan hubungan antara konsep-konsep tersebut. Selain itu, ontologi memungkinkan berbagi pemahaman tentang informasi terstruktur dan memungkinkan penggunaan kembali domain pengetahuan [4-6].

1.3 Web Semantik

Semantik web adalah hasil dari penggabungan informasi menggunakan metode dan kemampuan mesin untuk membaca informasi yang luas. Semantik web merupakan konsep yang terkait dengan World Wide Web dan mengacu pada sebagian web yang memiliki makna atau semantik. Dalam konteks tersebut, web semantik tidak bertujuan untuk menggantikan web yang ada saat ini, melainkan untuk meningkatkan kualitas informasi yang disajikan. Tujuannya adalah untuk memperkaya definisi informasi sehingga komputer dapat memahaminya, sehingga komputer dan manusia dapat berkolaborasi. Teknologi web semantik terdiri dari berbagai standar dan teknologi yang memungkinkan dokumen web dapat dibagikan dan digunakan kembali di berbagai aplikasi. Hal ini memungkinkan mesin untuk memproses data yang dipublikasikan dengan makna yang jelas [3].

1.4 SPARQL

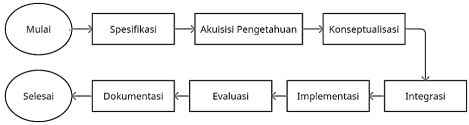
SPARQL adalah sebuah bahasa query yang direkomendasikan oleh W3C untuk mengambil informasi dari graph RDF. Selain itu, SPARQL juga berfungsi sebagai standar protokol untuk mengakses sumber daya di Semantic Web. Dengan menggunakan SPARQL, sebuah web dapat mengambil nilai dari data yang terstruktur maupun semi-terstruktur. Hasil dari query SPARQL dapat dikembalikan dalam berbagai format, seperti XML, JSON, RDF, dan HTML. SPARQL didasarkan pada standar RDF Data Access Working Group (DAWG) dan menyediakan beberapa klausa yang digunakan dalam query SPARQL beserta fungsinya masing-masing [3-6].

1.5 Protégé

Protégé adalah sebuah platform open-source gratis yang menawarkan serangkaian alat kepada komunitas pengguna yang sedang berkembang untuk membangun model domain dan aplikasi berbasis pengetahuan menggunakan ontologi. Protégé menyediakan kemudahan dalam pengembangan prototipe dan digunakan untuk membuat struktur ontologi. Selain itu, Protégé dapat menyimpan ontologi dalam berbagai format, seperti database relasional, UML, XML, dan RDF [6].

2. Metode Penelitian

Metode yang digunakan pada penelitian ini adalah Methontology. Methontology merupakan suatu metodologi pengembangan model ontologi dalam tingkat pengetahuan. Metode ini telah dikembangkan oleh para peneliti dari Universidad Politécnica de Madrid. Methontology memiliki kemampuan dalam melakukan siklus hidup ontologi, yang didasarkan pada pengembangan prototipe yang memungkinkan penambahan, perubahan, dan penghapusan terminologi pada setiap versi ontologi yang baru. Metode pengembangan ontologi yang disebut Methontology dipilih karena memiliki keunggulan dalam memberikan deskripsi yang sangat rinci tentang setiap aktivitas yang dilakukan selama pembangunan ontologi [4]. Terdapat tahapan-tahapan yang harus dilalui dalam membangun sebuah model ontologi menggunakan metode Methontology, sebagai berikut [12]:



Gambar 1. Alur Metode Penelitian

2.1 Spesifikasi

Pada tahap spesifikasi ini memiliki tujuan untuk menghasilkan dokumen spesifikasi ontologi yang dapat berbentuk formal, semi-formal, atau informal, ditulis dalam bahasa alami (natural language). Metode ini menggunakan satu set representasi yang berada di tengah-tengah (menengah) atau menggunakan pertanyaan kompetensi.

2.2 Akuisisi Pengetahuan

Tahap selanjutnya adalah tahap yang berdiri sendiri dalam pembangunan ontologi. Sebagian besar tahap akuisisi telah diselesaikan secara bersamaan dengan tahap spesifikasi, dan secara bertahap mengurangi peranannya seiring dengan berjalannya proses pengembangan ontologi.

2.3 Konseptualisasi

Pada tahap konseptualisasi ini, akan dibentuk sebuah model konseptual yang mewakili pengetahuan domain dan menggambarkan masalah serta solusinya menggunakan istilah-istilah yang telah diidentifikasi pada tahap spesifikasi. Tugas yang harus dilakukan adalah membangun Glossary of Terms (GT) yang lengkap, mencakup konsep-konsep, contoh-contoh, kata kerja, dan properti yang terkait. GT akan mencari dan mengumpulkan semua pengetahuan domain yang potensial dan memberikan arti dari masing-masingnya.

2.4 Integrasi

Tahap selanjutnya adalah tahap integrasi, tahap Integrasi ini melibatkan penggunaan definisi ontologi yang telah ada dan menggabungkannya ke dalam ontologi lain, sehingga pembangunan ontologi dapat dilakukan tanpa perlu memulai dari awal.

2.5 Implementasi

Tahap selanjutnya adalah tahap implementasi yaitu langkah penerapan ontologi yang telah dirancang, mulai dari tahap spesifikasi hingga integrasi. Pada fase ini, dilakukan pengubahan dan implementasi ontologi yang telah dirancang menggunakan perangkat lunak Protégé.

2.6 Evaluasi

Pada tahap ini, dilakukan penilaian teknis terhadap ontologi, lingkungan perangkat lunak, dan dokumentasi yang berkaitan dengan kerangka referensi pada setiap tahap dan antara tahap siklus kehidupan mereka. Evaluasi terdiri dari dua proses, yaitu verifikasi dan validasi. Verifikasi merujuk pada proses teknis untuk memastikan kebenaran ontologi, lingkungan perangkat lunak, dan dokumentasi yang terkait dengan kerangka referensi pada setiap tahap dan antara tahap siklus kehidupan mereka. Validasi memastikan bahwa ontologi, lingkungan perangkat lunak, dan dokumentasi sesuai dengan sistem yang ingin mereka wakili.

2.7 Dokumentasi

Ini adalah tahap terakhir yaitu dilakukan proses dokumentasi baik dalam kode ontologi, teks bahasa alami yang dilampirkan pada definisi formal, maupun makalah yang diterbitkan dalam proses konferensi dan jurnal yang mengatur pertanyaan-pertanyaan penting dari ontologi yang sudah dibangun.

3. Hasil dan Diskusi

Pada penelitian ini dibangun sebuah ontologi yang berdomain Kain Songket Bali. Berikut merupakan hasil yang diperoleh dari setiap tahapan metode penelitian yang telah dilakukan.

3.1. Spesifikasi

Tujuan dari fase spesifikasi ini adalah untuk menghasilkan dokumen spesifikasi ontologi yang dapat berbentuk formal, semi-formal, atau informal, ditulis dalam bahasa alami (natural language). Metode ini menggunakan satu set representasi yang berada di tengah-tengah (menengah) atau menggunakan pertanyaan kompetensi.

a. Domain : Kain Songket Bali
b. Tanggal : 19 Mei 2023
c. Dikonsep-oleh : I Made Suma Gunawan
d. Dilaksanakan oleh : I Made Suma Gunawan
e. Tujuan : Membangun Model Ontologi untuk memudahkan klasifikasi Kain Songket Bali
f. Tingkat Formalitas : Formal
g. Ruang Lingkup : Kain Songket Bali
h. Sumber Pengetahuan : Internet, jurnal, dan wawancara.

3.2. Akuisisi Pengetahuan

Dalam proses pengembangan ontologi ini, sebagian besar akuisisi pengetahuan dilakukan pada tahap pemrosesan dengan persyaratan spesifikasi saat proses pengembangan ontologi. Pada tahap akuisisi pengetahuan ontologi pariwisata menggunakan teknik sebagai berikut.

a. Melakukan wawancara dengan para ahli atau pengrajin kain songket bali untuk mendapatkan informasi dan dapat merancang ontologi.
b. Melakukan identifikasi pengetahuan dan struktur yang digunakan melalui studi literatur.

Dalam penelitian ini, penulis menggunakan data kain songket bali untuk membangun model ontologi dari berbagai pihak yang memiliki informasi mengenai kain songket bali.

3.3. Konseptualisasi

Tahap konseptualisasi bertujuan untuk mengatur domain pengetahuan menjadi bentuk konseptual serta menjaga dan mengelola pengetahuan yang diperoleh dalam proses akuisisi pengetahuan. Setelah model konseptual dibangun, metodologi akan berubah untuk mengubah model konseptual tersebut menjadi model formal yang akan diimplementasikan dalam Bahasa ontologi. Ontologi ini dibangun untuk domain Kain Songket Bali dan akan disusun dalam bentuk class dan sub-class yang terlihat pada gambar di bawah ini.

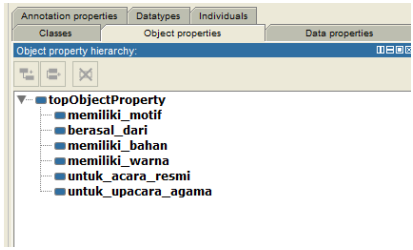
Gambar 2. Kelas Ontologi Kain Songket Bali

3.4. Integrasi

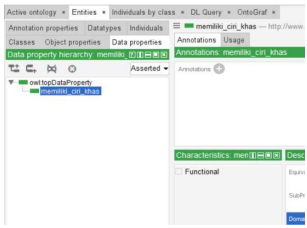
Tahap integrasi ini adalah pertimbangan dalam penggunaan ontologi yang telah dirancang sebelumnya adalah untuk melakukan integrasi agar sesuai dengan domain Kain Songket Bali. Dengan memilih ontologi yang sesuai dengan yang telah dirancang, kita dapat memperoleh hasil yang diharapkan.

3.5. Implementasi

Dalam pelaksanaan implementasi model ontologi, peneliti menggunakan aplikasi Protégé 5.5.0 untuk mengembangkan ontologi. Protégé adalah perangkat lunak yang dikembangkan oleh Stanford Center for Biomedical Informatics Research di Stanford University School of Medicine. Protégé digunakan sebagai alat bantu untuk mengembangkan ontologi berdasarkan sistem pengetahuan dasar. Setiap bagian ontologi didefinisikan sesuai dengan hasil dari setiap tahap tugas dalam metode Methontology. Rancangan konseptual yang telah dilakukan kemudian diformalkan menggunakan aplikasi Protégé 5.5.0 Ontografi, dan dari situ dapat dihasilkan model ontologi yang dibangun dalam laporan ini. Dapat dilihat pada gambar 3 adalah implementasi object properties yang berguna untuk menghubungkan individu satu dengan yang lainnya.

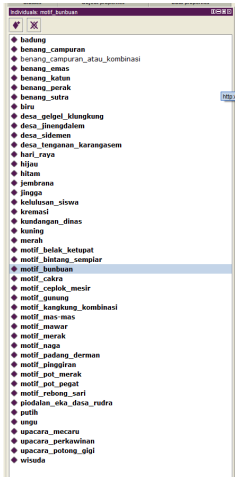


Gambar 3. Object Properties



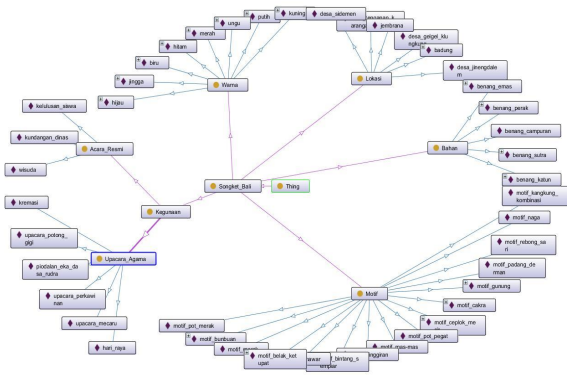
Gambar 4. Data Property

Di sini, hanya ada satu data properti Kain Songket Bali yang terkait dengan individu. Data properti ini memberikan nilai yang lebih spesifik dan menggunakan tipe data string.



Gambar 5. Individual

Terdapat 49 individu yang ditampilkan pada ontologi Kain_Songket_Bali. individu yang diperluas dalam kelas disebut dengan instance.

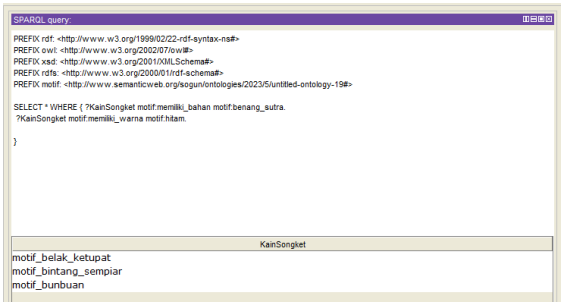


Gambar 6. Ontograph dari Ontologi Kain Songket Bali

Pada gambar 6 ini adalah contoh hubungan semantik yang menggambarkan masing-masing class, object property, dan individual yang dibangun pada ontologi kain songket bali.

3.6. Evaluasi

Pada tahap evaluasi ini, penulis melakukan evaluasi terhadap ontologi yang telah dibuat. Evaluasi dilakukan dengan melakukan pengujian menggunakan query SPARQL yang tersedia dalam aplikasi Protégé 5.5.0. Pertanyaan yang ingin diajukan dapat diubah menjadi query SPARQL, sehingga dapat menampilkan hasil yang ada dalam ontologi yang telah dibuat.



Gambar 7. Hasil SPARQL Query

3.7. Dokumentasi

Hasil dokumentasi dari penelitian pengembangan ontologi semantik Kain Songket Bali berupa tulisan yang tertuang dalam laporan ini.

4. Kesimpulan

Berdasarkan hasil dan pembahasan i, pembangunan ontologi terkait dengan kain songket bali telah selesai. Pembuatan ontologi ini dilakukan menggunakan aplikasi Protégé 5.5.0 dengan metode Methontology. Hasilnya mencakup 9 class, 6 Object Properties, 1 Data Properties, dan 49 individual atau contoh pada setiap class. Dalam tahap evaluasi Metode Methontology dapat digunakan untuk mengembangkan struktur ontologi yang berkualitas tinggi. Ontologi kain songket bali ini dapat digunakan sebagai dasar dalam pengembangan sistem manajemen pengetahuan terkait dengan kain songket bali.

Daftar Pustaka

[1] L. C. Dewi, M. D. Angendari, N. K. Widiartini, “TENUN SONGKET NEGARA (SONGKET TANPA SAMBUNGAN) DARI KELOMPOK TENUN PUTRI MAS DI KECAMATAN JEMBRANA,” Jurnal Bosaparis: Pendidikan Kesejahteraan Keluarga, vol.l 12, no. 1, 2021.
[2] I. M. P. Adiputra, K. N. H. Wardana, “PEMBERDAYAAN PENGRAJIN SONGKET BALI UTARA,” SULUH: Jurnal Abdimas, vol. 2, no. 1, 2020.
[3] M. A. Al’Izza, A. Jazuli, M. Nurkamid, “Implementasi Teknologi Semantik Web untukPencarian Koleksi Perpustakaan Universitas Muria Kudus,” Jurnal Dialektika Informatika (Detika), Vol. 2, No. 2, 2022.
[4] K. D. P. Novianti, R. A. N. Diaz, “SISTEM PENCARIAN PROGRAM STUDI PADAPERGURUAN TINGGI DI BALI BERBASIS SEMANTIK,” Jurnal Sains dan Teknologi, Vol. 6, No. 1, 2017.

[5] Himawan, T. W. Harjanti, R. Supriati, and H. Setiyani, "Evolusi Penggunaan Teknologi Web 3.0: Semantic Web," J. Inf. Syst. Hosp. Technol., vol. 2, no. 02, 2020.
[6] F. Azzahra, C. I. Ratnasari, "Implementasi Ontologi untuk Klasifikasi atau Pencarian: Kajian Literatur, "
[7] Kementrian Agama Republik Indonesia. (2022). Menyemai Kerukunan dan Menjaga Keajegan Budaya Bali. Diakses pada 10 juni 2023 dari https://kemenag.go.id/moderasi-beragama/menyemai-kerukunan-dan-menjaga-keajegan-budaya-bali-z565eq.
[8] 1001Indonesia. (2017). Songket Bali, Kekayaan Kerajinan Tradisional Pulau Dewata. Diakses pada 10 juni 2023 dari https://1001indonesia.net/songket-bali-kekayaan-kerajinan-tradisional-pulau dewata/#:~:text=Saat%20ini%2C%20songket%20Bali%20dapat,keagamaan%20serta%20dalam%20acara%20adat.
[9] bisnisbali.com. (2017). Berkelas, Kain Songket Bali untuk Acara Formal. Diakses pada 10 juni 2023 dari http://bisnisbali.com/berkelas-kain-songket-bali-untuk-acara-formal/.
[10] ksmtour.com. Kain Songket Sidemen Oleh-oleh Mewah Khas Bali. Diakses pada 10 juni 2023 dari https://ksmtour.com/pusat-oleh-oleh/oleh-oleh-khas-bali/kain-songket-sidemen-oleh-oleh-mewah-khas-bali.html.
[11] L. J. E. Dewi, N. K. Kertiasih, I. K. Purnawaman. "Pembuatan Katalog Produk Kerajinan Tenun Songket Desa Jinengdalem Buleleng," Seminar Nasional Vokasi dan Teknologi (SEMNASVOKTEK). 2017.
[12] P. R. Ganeswara, C. R. A. Pramartha, "Ontology-Based Approach for Klungkung Royal Family," Jurnal Elektronik Ilmu Komputer Udayana, Vol. 8, No. 4, 2020.

Halaman ini sengaja dibiarkan kosong

JNATIA Volume 1, Nomor 4, Agustus 2023 p-ISSN: 2986-3929
Jurnal Nasional Teknologi Informasi dan Aplikasinya

Implementasi Random Forest dengan LASSO Dalam Klasifikasi Penyakit yang Ditularkan Melalui Nyamuk

Kadek Dwitya Adhi Pradyto[a1], Made Agung Raharja[a2]

Program Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam,
Universitas Udayana
Jalan Raya Kampus Udayana, Bukit Jimbaran, Kuta Selatan, Badung, Bali Indonesia
[1]adhipradyto32@gmail.com
[2]made.agung@unud.ac.id

Abstract

Several diseases that can attack human health can be transmitted through disease vectors. One of the insects belonging to the disease vector is the mosquito. Diseases that can attack humans due to transmission through mosquitoes include malaria, dengue fever, chikungunya, yellow fever, rift valley fever, and many more. With so many types of diseases that are transmitted by mosquitoes and the symptoms that look quite similar, a classification process is carried out to distinguish the types of diseases. In this study, the classification was carried out using the Random Forest algorithm with the LASSO algorithm for feature selection. It was found that the average accuracy values of the Random Forest before and after carrying out feature selection using LASSO were 88% and 76%, respectively. From the values obtained, it can be concluded that the Random Forest has better performance without feature selection using the LASSO method.

Keywords: *Classification, Random Forest, LASSO, Mosquito-Borne Diseases*

1. Pendahuluan

Beberapa penyakit yang bisa menyerang kesehatan manusia dapat ditularkan melalui vektor penyakit. Vektor penyakit merupakan jenis serangga yang bisa menularkan penyakit. Terjadi pertumbuhan sumber penyakit seperti virus di dalam tubuh vektor dengan jumlah yang cukup untuk menimbulkan penyakit baru [1]. Salah satu serangga yang tergolong ke dalam vektor penyakit adalah nyamuk. Nyamuk betina menghisap darah manusia untuk dijadikan sumber energi bagi perkembangan telurnya. Melalui proses menghisap darah inilah nyamuk dapat menularkan penyakit yang mereka bawa. Penyakit-penyakit yang dapat menyerang manusia akibat penularan melalui nyamuk antara lain yaitu malaria, demam berdarah *dengue* (DBD), *chikungunya*, demam kuning, demam *rift valley*, dan masih banyak lagi.

Dengan banyaknya jenis penyakit yang ditularkan melalui nyamuk serta gejala yang terlihat cukup mirip yaitu adanya demam, maka untuk membedakan jenis-jenis penyakit tersebut dilakukan proses klasifikasi. Klasifikasi merupakan proses pengelompokkan sekumpulan kelas berdasarkan data-data yang ada [2]. Klasifikasi dapat digunakan untuk mengelompokkan dua jenis kelas (*binary class*) ataupun lebih dari dua jenis kelas (*multi-class*). Salah satu algoritma yang bisa digunakan baik untuk klasifikasi *binary class* ataupun *multi-class* yaitu *Random Forest*. Algoritma *Random Forest* merupakan pengembangan dari *Classification and Regression Tree* (CART) dengan menerapkan *random feature selection* dan *bagging* (*bootstrap-aggregating*). Model klasifikasi dari algoritma *Random Forest* merupakan sekumpulan *Decision Tree*. Kelemahan dari penggunaan *Decision Tree* dalam proses klasifikasi yaitu terbentuknya model yang *overfitting* karena terpengaruh *noise* dari data latih. *Overfitting* adalah suatu kondisi dimana model yang dilatih lebih condong dalam memprediksi data latih daripada data uji. Dalam mengatasi *overfitting*, diterapkan algoritma *Random Forest* yang terdiri dari sekumpulan *Decision Tree* yang dilatih dengan keadaan yang berbeda di tiap pohonnya dan masih bisa mendapatkan akurasi yang maksimum [3].

Terdapat banyak cara yang bisa dilakukan dalam mengoptimalkan proses klasifikasi salah satunya yaitu dengan menyeleksi fitur yang tidak ada hubungannya dengan kelas prediksi. Proses seleksi fitur akan menghapus variabel yang tidak bisa diprediksi dan berlebihan. Proses klasifikasi dengan fitur yang optimal akan menghasilkan model klasifikasi yang bekerja lebih cepat serta dapat mengurangi *overfitting* [4].

Pada penelitian ini, dilakukan proses klasifikasi penyakit yang ditularkan melalui nyamuk menggunakan algoritma *Random Forest*. Adapun optimasi yang dilakukan yaitu penyeleksian fitur dengan menggunakan LASSO. Penelitian ini juga akan menguji tingkat akurasi *Random Forest* sebelum dan sesudah seleksi fitur menggunakan LASSO.

2. Metode Penelitian

2.1. Teknik Pengumpulan Data

Data yang akan digunakan dalam penelitian ini merupakan data sekunder yang didapat dari *website* Kaggle (https://www.kaggle.com/datasets/richardbernat/vector-borne-disease-prediction). Data pada Kaggle masih berupa data penyakit yang ditularkan melalui vektor penyakit. Maka dari itu, data akan dibatasi sampai baris ke-184 yang merupakan bagian data penyakit yang ditularkan melalui nyamuk.

	sudden_fever	headache	mouth_bleed	nose_bleed	muscle_pain	...	ulcers	toenail_loss	speech_problem	bullseye_rash	prognosis
0	0	1	1	1	1	...	0	0	0	0	Chikungunya
1	1	1	1	1	1	...	0	0	0	0	Chikungunya
2	0	1	0	1	0	...	0	0	0	0	Chikungunya
3	0	0	0	0	0	...	0	0	0	0	Chikungunya
4	1	0	0	0	1	...	0	0	0	0	Chikungunya
...	...	...	...	...	...	...	...	...	...	...	...
179	1	0	0	1	1	...	0	0	0	0	West Nile fever
180	0	1	0	0	0	...	0	0	0	0	West Nile fever
181	0	0	1	1	1	...	0	0	0	0	West Nile fever
182	1	0	0	1	0	...	0	0	0	0	West Nile fever
183	0	0	1	1	1	...	0	0	0	0	West Nile fever

184 rows × 65 columns

Gambar 1. Data Penyakit Yang Ditularkan Oleh Nyamuk

Pada data tersebut, terdapat 64 jenis fitur untuk mendeteksi penyakit yang ditularkan melalui nyamuk. Fitur-fitur tersebut terdiri dari "*sudden_fever*" sampai "*bullseye_rash*". Terdapat 8 kelas yang akan diklasifikasikan. Kelas-kelas tersebut antara lain "*Chikungunya*", "*Dengue*", "*Rift Valley fever*", "*Yellow Fever*", "*Zika*", "*Malaria*", "*Japanese encephalitis*", dan "*West Nile fever*" Total keseluruhan data yang digunakan sebanyak 184 data dengan pembagian data latih dan data uji adalah 70:30.

2.2. Seleksi Fitur Menggunakan LASSO

LASSO adalah singkatan dari *Least Absolute Shrinkage and Selection Operator* dan merupakan metode yang cukup bagus dalam melakukan seleksi fitur. Metode ini bekerja dengan proses penyusutan nilai koefisien dari fitur yang sedang diuji menjadi nol. Setelah itu, metode ini memilih variabel-variabel yang tidak bernilai nol untuk menjadi fitur yang digunakan untuk proses klasifikasi [4].

2.3. Random Forest

Random Forest merupakan salah satu algoritma *Machine Learning* yang bekerja dengan mengombinasikan sejumlah algoritma *Decision Tree* dalam pengambilan keputusannya [5].

Metode ini digunakan untuk membangun *Decision Tree* yang terdiri dari *root node*, *internal node*, dan *leaf node* dengan mengambil fitur secara acak (*random*). *Root node* merupakan bagian yang terletak paling atas, atau biasa disebut sebagai akar dari pohon. *Internal node* adalah bagian percabangan, dimana bagian ini memiliki keluaran minimal dua dan hanya ada satu masukkan. Sedangkan *leaf node* merupakan bagian terakhir yang hanya memiliki satu masukkan dan tidak memiliki keluaran [6].

Proses klasifikasi menggunakan algoritma *Random Forest* akan dilakukan sebelum dan sesudah melakukan seleksi fitur dengan LASSO. Pencarian model terbaik dari *Random Forest* akan dilakukan dengan menggunakan metode *Random Search*. Metode ini digunakan karena lebih unggul dibandingkan dengan *Grid Search* dan *Bayesian Search* dalam mencari *parameter* dari *Random Forest* [5]. Adapun *parameter* yang dicari menggunakan *Random Search* yaitu "*n_estimators*", "*criterion*", "*max_depth*", "*min_samples_split*", dan "*max_features*".

2.4. Evaluasi

Evaluasi dilakukan dengan menghitung nilai akurasi dan presisi dari model. Pengujian akurasi dilakukan untuk mengetahui seberapa akurat model dalam melakukan prediksi. Sedangkan pengujian presisi dilakukan untuk mengetahui seberapa benar model melakukan prediksi di tiap kelasnya. Nilai akurasi dan presisi dapat ditentukan dengan rumus berikut.

$$Akurasi = \frac{Prediksi\ benar}{\text{Total prediksi}} \quad (1)$$

$$Presisi = \frac{True\ positive}{True\ positive + False\ positive} \quad (2)$$

3. Hasil dan Pembahasan.

3.1. Seleksi Fitur

Proses seleksi fitur, nilai *alpha* yang digunakan pada LASSO adalah 0,1. Adapun nilai koefisien setiap fitur yang didapat setelah melakukan seleksi fitur dengan metode LASSO dapat dilihat pada Gambar 2 di bawah.

```
array([0.14982622, 0.        , 0.        , 0.34365757, 0.26094839,
       0.        , 0.04326063, 0.        , 0.        , 0.        ,
       0.        , 0.06598752, 0.        , 0.        , 0.        ,
       0.        , 0.03846259, 0.        , 0.        , 0.        ,
       0.16111974, 0.0686008 , 0.33595216, 0.05387347, 0.        ,
       0.046126  , 0.3407998 , 0.23243593, 0.        , 0.        ,
       0.15462936, 0.        , 0.20797139, 0.99854777, 0.85869685,
       0.32251722, 0.14780409, 0.21390347, 0.4914114 , 0.        ,
       0.13825577, 0.62696595, 0.29986494, 0.        , 0.        ,
       0.09864327, 0.35127592, 0.        , 0.        , 0.        ,
       0.        , 0.05747193, 0.        , 0.        , 0.58092856,
       0.        , 0.        , 0.        , 0.        , 0.        ,
       0.        , 0.        , 0.        , 0.        ])
```

Gambar 2. Nilai Koefisien Setiap Fitur

Fitur-fitur yang akan digunakan dalam proses klasifikasi memiliki nilai koefisien yang lebih dari nol (koefisien > 0). Berdasarkan nilai koefisien yang didapat di atas, daftar fitur-fitur yang terpilih dan tidak terpilih dapat dilihat pada Gambar 3 dan Gambar 4.

```
array(['sudden_fever', 'nose_bleed', 'muscle_pain', 'vomiting', 'ascites',
       'myalgia', 'stomach_pain', 'orbital_pain', 'neck_pain', 'weakness',
       'weight_loss', 'gum_bleed', 'jaundice', 'inflammation',
       'loss_of_appetite', 'urination_loss', 'slow_heart_rate',
       'abdominal_pain', 'light_sensitivity', 'yellow_skin',
       'yellow_eyes', 'microcephaly', 'rigor', 'bitter_tongue',
       'cocacola_urine', 'hypoglycemia', 'confusion', 'lymph_swells'],
      dtype='<U21')
```

Gambar 3. Hasil Seleksi Fitur Yang Terpilih

```
array(['headache', 'mouth_bleed', 'joint_pain', 'rash', 'diarrhea',
       'hypotension', 'pleural_effusion', 'gastro_bleeding', 'swelling',
       'nausea', 'chills', 'digestion_trouble', 'fatigue', 'skin_lesions',
       'back_pain', 'coma', 'diziness', 'red_eyes', 'facial_distortion',
       'convulsion', 'anemia', 'prostraction', 'hyperpyrexia',
       'stiff_neck', 'irritability', 'tremor', 'paralysis',
       'breathing_restriction', 'toe_inflammation', 'finger_inflammation',
       'lips_irritation', 'itchiness', 'ulcers', 'toenail_loss',
       'speech_problem', 'bullseye_rash'], dtype='<U21')
```

Gambar 4. Hasil Seleksi Fitur Yang Tidak Terpilih

3.2. Hasil Klasifikasi

Pada penelitian ini, proses klasifikasi dilakukan sebanyak 2 kali yaitu sebelum melakukan seleksi fitur dan sesudah melakukan seleksi fitur. Sebelum melakukan klasifikasi dengan *Random Forest*, dicari model terbaik dengan metode *Random Search*. Adapun nilai-nilai dari *hyperparameter* yang akan dicari bisa dilihat pada Tabel 1.

Tabel 1. Nilai *Hyperparameter*

Nama *Parameter*	Nilai
n_estimators	[50, 100, 150, 200]
criterion	[‘gini’, ‘entropy’]
max_depth	[None, 5, 10]
min_samples_split	[2, 4, 6, 8, 10]
max_features	['sqrt', 'log2']

Pada saat sebelum melakukan seleksi fitur didapatkan model dengan nilai parameter optimal seperti pada Tabel 2.

Tabel 2. Nilai *Parameter* Optimal Sebelum Melakukan Seleksi Fitur

Nama *Parameter*	Nilai
n_estimators	200
criterion	‘entropy’
max_depth	10
min_samples_split	4
max_features	'log2'

Berdasarkan *parameter* tersebut, dilakukan proses klasifikasi dan didapatkan rata-rata akurasi dan presisi setiap kelas seperti pada Tabel 3.

Tabel 3. Hasil Akurasi Dan Presisi Sebelum Melakukan Seleksi Fitur

Kelas	Rata-Rata Presisi (%)	Rata-Rata Akurasi (%)
Chikungunya	89	88
Dengue	97	
Rift Valley fever	79	
Yellow Fever	100	
Zika	81	
Malaria	83	
Japanese encephalitis	84	
West Nile fever	85	

Setelah melewati tahap seleksi fitur, pencarian model *Random Forest* yang optimal dilakukan sekali lagi menggunakan metode *Random Search*. Hasil pencarian tersebut dapat dilihat pada Tabel 4.

Tabel 4. Nilai *Parameter* Optimal Setelah Melakukan Seleksi Fitur

Nama *Parameter*	Nilai
n_estimators	150
criterion	'gini'
max_depth	5
min_samples_split	8
max_features	'log2'

Setelah pencarian *parameter* optimal dilakukan, tahap selanjutnya adalah menguji *Random Forest* berdasarkan *parameter* di atas dan fitur yang sudah diseleksi. Adapun hasil dari pengujian tersebut dapat dilihat pada Tabel 5.

Tabel 5. Hasil Akurasi Dan Presisi Setelah Melakukan Seleksi Fitur

Kelas	Rata-Rata Presisi (%)	Rata-Rata Akurasi (%)
Chikungunya	62	76
Dengue	100	
Rift Valley fever	64	
Yellow Fever	71	
Zika	68	
Malaria	81	
Japanese encephalitis	73	
West Nile fever	84	

4. Kesimpulan

Berdasarkan hasil evaluasi yang sudah dilakukan, didapatkan bahwa nilai rata-rata akurasi dari *Random Forest* sebelum dan sesudah melakukan seleksi fitur menggunakan LASSO secara berturut-turut yaitu 88% dan 76%. Dari nilai yang sudah didapat, dapat disimpulkan bahwa *Random Forest* memiliki performa yang lebih baik tanpa adanya seleksi fitur dengan metode LASSO. Hal tersebut dapat terjadi karena beberapa faktor seperti nilai *alpha* dari LASSO yang kurang optimal dalam menentukan fitur-fitur terbaik, adanya fitur yang sebenarnya berpengaruh besar terhadap akurasi dan presisi namun tidak terpilih saat penyeleksian fitur, dan faktor lainnya yang belum bisa penulis dapatkan.

Daftar Pustaka

[1] A. Dinata, *Bersahabat dengan Nyamuk: Jurus Jitu Atasi Penyakit Bersumber Nyamuk*. Arda Publishing House, 2018.
[2] I. Fadilla, P. P. Adikara, dan R. S. Perdana, "Klasifikasi Penyakit Chronic Kidney Disease (CKD) Dengan Menggunakan Metode Extreme Learning Machine (ELM)," *J. Pengemb. Teknol. Inf. dan Ilmu Komput. e-ISSN*, vol. 2548, hlm. 964X, 2018.
[3] R. Wasono, "Perbandingan Metode Random Forest dan naive bayes untuk Klasifikasi Debitur Berdasarkan Kualitas Kredit," 2022.
[4] M. Parzinger, L. Hanfstaengl, F. Sigg, U. Spindler, U. Wellisch, dan M. Wirnsberger, "Comparison of different training data sets from simulation and experimental measurement with artificial users for occupancy detection — Using machine learning methods Random Forest and LASSO," *Build Environ*, vol. 223, hlm. 109313, 2022, doi: https://doi.org/10.1016/j.buildenv.2022.109313.
[5] U. Sunarya dan T. Haryanti, "Perbandingan Kinerja Algoritma Optimasi pada Metode Random Forest untuk Deteksi Kegagalan Jantung," *Jurnal Rekayasa Elektrika*, vol. 18, no. 4, 2022.
[6] V. W. Siburian dan I. E. Mulyana, "Prediksi Harga Ponsel Menggunakan Metode Random Forest," dalam *Annual Research Seminar (ARS)*, 2019, hlm. 144–147.

JNATIA Volume 1, Nomor 4, Agustus 2023 p-ISSN: 2986-3929
Jurnal Nasional Teknologi Informasi dan Aplikasinya

Klasifikasi Emosi Lirik Lagu dengan Long Short-Term Memory dan Word2Vec

I Putu Diska Fortunawan[1], Ngurah Agus Sanjaya ER[2]

Program Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam,
Universitas Udayana
Jalan Raya Kampus Udayana, Bukit Jimbaran, Kuta Selatan, Badung, Bali Indonesia
[1]diskafortunawan@gmail.com
[2]agus_sanjaya@unud.ac.id

Abstract

This research focuses on the classification of emotions in song lyrics using LSTM (Long Short-Term Memory) and Word2Vec embedding. Emotion classification in lyrics plays a crucial role in music recommendation systems, sentiment analysis, and understanding the affective aspects of music. The study explores the effectiveness of LSTM, a type of recurrent neural network (RNN), in capturing the sequential dependencies and patterns in lyrics, combined with Word2Vec embedding to represent the semantic meaning of words.The dataset consists of a collection of song lyrics labeled with 2 emotions. The lyrics are preprocessed and converted into word vectors using the Word2Vec model. The LSTM model is then trained on the preprocessed lyrics data, aiming to predict the corresponding emotion category for a given set of lyrics. Experimental results demonstrate that the proposed approach achieves a maximum accuracy of 72.8% in classifying emotions in song lyrics. The LSTM model leverages the sequential information in the lyrics to capture the emotional context effectively. The Word2Vec embedding enhances the representation of words, allowing the model to understand the semantic relationships between words and better discriminate between different emotional categories.

Keywords: *Text Processing, Classification, LSTM, Word2Vec*

1. Pendahuluan

Kebanyakan orang mendengarkan lagu tidak hanya berdasarkan genre favoritnya saja, namun juga berdasarkan mood atau emosi yang sedang dirasakan. Lagu sering diklasifikasi berdasarkan genrenya, namun belum banyak klasifikasi lagu yang dikategorikan berdasarkan emosi dari lagu menggunakan lirik dari lagu tersebut. Lirik lagu mencerminkan pengalaman dan perasaan manusia yang mendalam, sehingga menganalisis emosi dalam lirik lagu dapat memberikan wawasan yang berharga tentang ekspresi dan persepsi emosi dalam konteks musik. Dalam penelitian ini, penelitian dipusatkan pada penggunaan metode Deep Learning, khususnya Long Short-Term Memory (LSTM), untuk melakukan klasifikasi emosi dalam lirik lagu. LSTM adalah jenis arsitektur jaringan saraf rekuren yang terkenal karena kemampuannya dalam mempelajari dan mengingat informasi sepanjang urutan data [1]. Penelitian ini juga menggabungkan LSTM dengan metode embedding Word2Vec untuk menghasilkan representasi vektor kata yang kaya dalam lirik lagu. Pada penelitian ini, kami ingin mengetahui performa dari klasifikasi emosi lirik lagu dengan menggunakan LSTM dan Word2vec. Dengan penelitian ini diharapkan dapat menjadi landasan dan referensi bagi penelitian-penelitian selanjutnya dalam ruang lingkup yang sama

2. Metode Penelitian

2.1 Pengumpulan Data

Data yang akan digunakan adalah data lirik lagu saja tanpa menggunakan audio dari lagu, data diambil dari Genius, sedangkan untuk kumpulan lagu yang akan diambil liriknya ditentukan melalui playlist Spotify “senang&Bahagia” yang berisikan lagu lagu yang ceria dan playlist “nangis

dah" yang berisikan kumpulan lagu-lagu sedih yang merupakan kebalikan dari emosi Bahagia. Data yang digunakan hanya memiliki label "Bahagia" dan "sedih" dengan nilai 1 pada label jika merupakan lagu Bahagia dan nilai 0 jika merupakan lagu sedih.

2.2 Preprocessing

Setelah data didapatkan dan dikumpulkan, perlu dilakukan preprocessing sebelum data siap untuk dijadikan sebagai data training. Proses ini akan menghilangkan noise atau data yang sekiranya tidak penting dan tidak diperlukan. Tahapan preprocessing yang dilakukan adalah:

a. Case Folding

Pada tahapan ini, jika ditemukan character yang merupakan uppercase maka character ini akan diubah menjadi lowercase

b. Punctuation Removal

Proses ini menghilangkan/menghapus tanda baca yang ada pada lirik lagu seperti koma(,), titik(.) dan tanda baca lain yang tidak diperlukan dalam proses.

c. Normalization

Tahapan ini akan mengubah kata-kata yang merupakan bahasa gaul ataupun bahasa yang tidak baku ke dalam bentuk bakunya.

d. Stemming

Menghilangkan imbuhan pada kata, imbuhan ini termasuk awalan, sisipan, maupun akhiran. Sehingga kata menjadi kata dasar.

e. Stopword Removal

Menghilangkan kata-kata yang sering muncul yang bersifat umum atau tidak memiliki makna terhadap kalimat.

f. Tokenization

Kalimat akan dipecah menjadi bagian yang lebih kecil yaitu kata.

2.3 Word2Vec Embedding

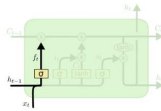
Untuk mengerti lebih dalam mengenai konteks yang ada dalam lirik lagu, kita memerlukan word embedding khususnya Word2Vec. Word2Vec dapat memecahkan masalah mewakili hubungan kata kontekstual dalam ruang fitur yang dapat dihitung [2]. Metode Word2Vec berfokus pada ide bahwa kata-kata yang sering muncul bersama memiliki kesamaan makna atau konteks. Word2Vec memanfaatkan data teks yang besar untuk melatih model yang dapat menghasilkan representasi vektor kata-kata tersebut. Ada 2 argumen yang akan digunakan yaitu max_input_length dan embed dimension.

2.4 Long Short-Term Memory (LSTM)

LSTM merupakan salah satu contoh model RNN (Recurrent Neural Network). LSTM dapat bekerja lebih baik dibandingkan RNN karena LSTM dapat mengatasi masalah vanishing gradient yang ditemukan pada RNN [3]. Masalah ini diatasi dengan menggunakan blok memory-cell yang terdiri dari input gate, forget gate dan output gate untuk mengganti lapisan RNN [4]. LSTM memiliki memiliki koneksi berulang atau struktur yang seperti rantai Fungsi-fungsi tiap gate pada LSTM adalah:

a. Forget Gate

Gerbang ini bertugas untuk melupakan beberapa informasi yang tidak relevan dan sudah tidak diperlukan oleh sebuah sistem. Alhasil, LSTM dapat menyajikan kumpulan informasi yang lengkap, tetapi tetap aktual sesuai dengan kebutuhan.

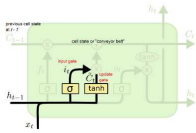


$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] \; + \; b_f\right)$$

Gambar 1. Forget gate [5]

Untuk menentukan apakah informasi dari state sebelumnya disimpan atau tidak, ditentukan dengan fungsi sigmoid, fungsi ini akan menghasilkan f_t antara 0 dan 1. Jika dihasilkan 0 maka semua informasi akan dilupakan, sebaliknya jika dihasilkan 1 maka tidak ada informasi yang dilupakan.

b. Input Gate

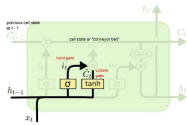


$$i_t = \sigma\left(W_i \cdot [h_{t-1}, x_t] \; + \; b_i\right)$$
$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] \; + \; b_C)$$

Gambar 2. Input Gate [5]

Tugas input gate adalah untuk menambahkan informasi yang sebelumnya telah diseleksi terlebih dahulu melalui gerbang forget gate. Gerbang ini tidak dimiliki oleh RNN yang hanya memungkinkan satu input data untuk satu output data. Dalam input gate kemudian dikenal istilah input modulation gate. Sesuai namanya, input modulation gate berfungsi untuk memodulasi informasi yang ada, sehingga dapat mengurangi kecepatan konvergensi dari data zero-mean. Pada gate ini juga diaplikasi fungsi sigmoid yang juga akan memiliki nilai antara 0 dan 1. Sekarang informasi baru yang perlu diteruskan ke state cell adalah fungsi dari state tersembunyi pada timestamp t-1 sebelumnya dan masukan x pada timestamp t. Fungsi aktivasi di sini adalah tanh. Karena fungsi tanh, nilai informasi baru akan berada di antara -1 dan 1. Jika nilai Nt negatif, informasi dikurangi dari keadaan sel, dan jika nilainya positif, informasi ditambahkan ke state cell pada timestamp saat ini.

c. Output Gate



$$i_t = \sigma\left(W_i \cdot [h_{t-1}, x_t] \; + \; b_i\right)$$
$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] \; + \; b_C)$$

Gambar 3. Output Gate [5]

Output gate menjadi gerbang terakhir untuk menghasilkan informasi data yang komplet dan aktual. Gerbang ini bisa menjadi yang terakhir atas sebuah informasi atau hanya menjadi bagian dari tahap pertama saja, sebelum akhirnya informasi akan diproses lewat input gate di sel berikutnya. Perhitungan pada output gate juga sangat mirip dengan gate-gate sebelumnya, nilainya juga akan berada diantara 0 dan 1. Untuk menghitung status hidden state saat ini, digunakan Ot dan tanh dari cell state yang diperbarui.

3. Hasil dan Pembahasan

3.1. Data Understanding

Data yang digunakan diperoleh dari Genius, data yang digunakan hanyalah data teks, yaitu lirik lagu itu sendiri tanpa menggunakan data audio dari lagu. Dataset yang digunakan berjumlah 90 lagu dengan label 1 yang berarti lagu bahagia, sedangkan label 0 yang merupakan lagu sedih, Setiap label terdiri dari masing masing 45 lagu.Data yang dikumpulkan dapat dilihat pada Gambar 4.

```
Pengambilan Data
                                                lyrics
0    Sandarkan lelahmu dan ceritakan Tentang apapu...
1    Menyesal, tak kusampaikan Cinta monyetku ke K...
2    Kapan terakhir kali kamu dapat tertidur tenan...
3   Salahkah bila ku mencinta Salahkah bila semua ...
4    Kuingin cinta hadir untuk selamanya Bukan han...
..                                                 ...
85  Biar aku sentuhmu Berikan ku rasa itu Pelukmu ...
86  Hujan tak juga reda Ku harus menyaksikan cinta...
87  Belum sempat ku membagi Kebahagia-anku Belum s...
88   Pedihnya tanya yang tak terjawab Mampu menjat...
89   Jika wangimu saja bisa Memindahkan duniaku Ma...
```

Gambar 4. Hasil Pengambilan Data

3.2. Data Preprocessing and Preparation

Data yang telah didapatkan tadi kemudian dikenakan preprocessing sebelum diproses lebih lanjut rincian tahapannya dapat dilihat pada Tabel 1.

Tabel 1. Preprocessing data

Tahapan	Hasil
Data Awal	Menari bersamaku Temani hingga akhir waktu Jalani bersamaku Kau pasti 'kan bahagia Kutemani dirimu Di setiap perjalanan cinta
Case Folding	menari bersamaku temani hingga akhir waktu jalani bersamaku kau pasti 'kan bahagia kutemani dirimu di setiap perjalanan cinta

Tahapan	Hasil
Normalization	menari bersamaku temani hingga akhir waktu jalani bersamaku kau pasti akan bahagia kutemani dirimu di setiap perjalanan cinta
Punctuation Removal	menari bersamaku temani hingga akhir waktu jalani bersamaku kau pasti akan bahagia kutemani dirimu di setiap perjalanan cinta
Stemming	menari bersama teman hingga akhir waktu jalani bersama kau pasti akan bahagia teman kamu setiap jalan cinta
Stopword Removal	menari bersama teman hingga akhir jalani bersama bahagia teman kamu setiap jalan cinta
Tokenization	menari, bersama, teman, hingga, akhir, jalani, bahagia, kamu, setiap, cinta

Setelah dilakukan tokenization ditemukan sebanyak 1686 kamus data atau jumlah kata unik dari lirik lagu yang telah dikumpulkan. Setelah itu, data akan dibagi menjadi data latih dan data uji, data uji diambil dari 20% dataset yang telah didapatkan

3.3. Model LSTM

Pemodelan diawali dengan melakukan penambahan lapisan embedding untuk mengubah kata ke dalam bentuk vector dengan max_input length = jumlah unique word, serta embed dim = 100. Digunakan pula 2 layer LSTM dengan unit masing masing sebanyak 64 unit dan dropout rate 0.1 Detail dari model yang digunakan adalah sebagai berikut:

Tabel 2. Model LSTM

Layer	Jumlah	Addition
Embedding		Max_input_length = 1686, embed dim = 100
LSTM	64 Unit	Dropout Rate = 0.1
LSTM	64 Unit	Dropout Rate = 0.1
Dense	1 Unit	Activation = sigmoid

3.4. Evaluasi Model

Training dilakukan sebanyak 3 kali dengan epoch dan jumlah data yang berbeda-beda, seperti pada tabel berikut:

Tabel 3. Hasil Training

Keterangan	Hasil
Epoch 50 dataset 60	52.6%
Epoch 100 dataset 90	68.3%
Epoch 50 dataset 60	58.8%
Epoch 100 Dataset 90	72.8 %

4. Kesimpulan

Dari Penelitian yang telah dilakukan, didapatkan hasil akurasi tertinggi yaitu 72.8 %, nilai tersebut akan berubah sesuai dengan perubahan hyperparameter. Akurasi dapat ditingkatkan dengan menambah dataset dan cleani

Daftar Pustaka

[1] Wang, J.H., Liu, T.W., Luo, X. and Wang, L., 2018, October. An LSTM approach to short text sentiment classification with word embeddings. In Proceedings of the 30th conference on computational linguistics and speech processing (ROCLING 2018) (pp. 214-223).
[2] Sari, W.K., Rini, D.P., Malik, R.F. and Azhar, I.S.B., 2020. Multilabel Text Classification in News Articles Using Long-Term Memory with Word2Vec. Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi), 4(2), pp.276-285.
[3] Nugroho, K.S., Akbar, I. and Suksmawati, A.N., 2023. Deteksi Depresi dan Kecemasan Pengguna Twitter Menggunakan Bidirectional LSTM. arXiv preprint arXiv:2301.04521.
[4] Fajri, F.N. and Syaiful, S., 2022. Klasifikasi Nama Paket Pengadaan Menggunakan Long Short-Term Memory (LSTM) Pada Data Pengadaan. Building of Informatics, Technology and Science (BITS), 4(3), pp.1625-1633.
[5] Eddine, B.C. (2021) LSTM Model for the Prediction of PM2.5 Concentration in city of Algiers.

JNATIA Volume 1, Nomor 4, Agustus 2023 p-ISSN: 2986-3929
Jurnal Nasional Teknologi Informasi dan Aplikasinya

Deteksi Penyakit Diabetes Menggunakan Gaussian Naive Bayes, Regresi Logistik, dan Random Forest

Kenny Belle Lesmana[a1], I Ketut Gede Suhartana[a2],

Program Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam,
Universitas Udayana
Jalan Raya Kampus Udayana, Bukit Jimbaran, Kuta Selatan, Badung, Bali Indonesia
[1]kennybelle015@unud.ac.id
[2]ikg.suhartana@unud.ac.id

Abstract

Diabetes is a very common health problem in the world. The number of people with diabetes is increasing year to year. Therefore, it is necessary to realize the symptoms of diabetes as early as possible. Diabetes is a chronic disease characterized by high sugar levels in the blood. In this study, a system was made about a diabetes detection system based on numerical data using three methods. That three methods are Gaussian Naive Bayes method, Logistic Regression, and Random Forest by taking a dataset in the form of numerical data. The accuracy value on the data tested in this study using Gaussian Naive Bayes, Logistic Regression, Random Forest is 0.74; 0;78; 078.

Keywords: *Gaussian Naive Bayes, Regresi Logistik, Random Forest*

1. Pendahuluan

Diabetes merupakan salah satu penyakit kronis yang ditandai dengan kadar glukosa (gula) yang tinggi dalam darah. Gejala umum pada diabetes yaitu mulai dari haus yang berlebihan, kelelahan yang berlebihan, sampai penurunan berat badan tanpa alasan yang jelas. Maka dari itu, sangat penting untuk mengecek dan mengelola diabetes dengan baik karena akan memiliki dampak yang sangat serius jika tidak segera diobati. Sistem pendeteksi penyakit diabetes diperlukan untuk penanganan yang lebih cepat. Pada proses pembuatannya dapat menggunakan convolutional neural network (CNN), Decision Tree, pengolahan citra (image processing). Namun, pada penelitian ini, akan dibuat sistem pendeteksi penyakit diabetes pada manusia berdasarkan data numerik menggunakan tiga metode, yaitu metode Gaussian Naive Bayes, Regresi Logistik, dan Random Forest dengan mengambil dataset berupa data numerik.

2. Metode Penelitian

Proses pada penelitian ini dibagi menjadi beberapa bagian yaitu studi pengumpulan data, metode yang digunakan, dan pembuatan sistem.

2.1 Pengumpulan Data

Pada penelitian ini, dataset berupa data sekunder yang diambil dari website kaggle https://www.kaggle.com/datasets/akshaydattatraykhare/diabetes-dataset. Data tersebut merupakan data numerik yang akan dimasukan sebagai input dari sistem yang akan dibuat pada penelitian ini.

2.2 Metode

Metode yang digunakan pada pembuatan sistem ini yaitu Gaussian Naive Bayes, Regresi Logistik, dan Random Forest.

a. Gaussian Naive Bayes

Naive Bayes merupakan salah satu metode pengklasifikasian data menggunakan perhitungan probabilitas yang akan terjadi di masa depan berdasarkan peristiwa yang pernah terjadi sebelumnya. Klasifikasi bayes merupakan salah satu metode klasifikasi yang memiliki akurasi paling tinggi diantara metode klasifikasi lainnya [1]. Pada penelitian ini, data yang dimasukkan berupa data numerik dan menghasilkan data yang numerik. Maka dari itu dapat memakai fungsi Probability Density Function (PDF) [2]. PDF merupakan sebuah konsep yang digunakan teori peluang dalam menggambarkan seberapa besar probabilitas terdistribusi dalam rentang nilai-nilai tertentu. Berikut adalah rumus dari PDF dalam persamaan 1 dan persamaan 2 merupakan rumus dasar standar deviasi atau yang disebut sebagai formula Gaussian Naive Bayes Classifier [3].

$$P(X_1 = x_1 \mid Y = y_j) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(x_j - \underline{x})^2}{2\sigma^2}} \quad (1)$$

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n} (x_1 - \underline{x})^2}{n-1}} \quad (2)$$

Keterangan:

P : Probabilitas
X_i : Atribut
x_i : nilai atribut
Y : Kelas yang berhubungan
y_j : sub kelas yang berhubungan
$\underline{x}$: rata-rata
σ : standar deviasi
n : banyaknya data

b. Regresi Logistik

Regresi Logistik merupakan salah satu metode analisis statistika dengan tujuan mengetahui hubungan antara variabel terkait yang memiliki dua kategori atau lebih dengan satu atau lebih peubah bebas berskala kategori atau kontinu [4]. Regresi Logistik biasanya digunakan pada analisis prediktif dan dalam pemodelan data yang mana variabel target hanya memiliki dua nilai yaitu 0 dan 1 (ya atau tidak). Maka dari itu persamaan rumus mengikuti distribusi Bernoulli yaitu sebagai berikut [5].

$$f(y_i) = \pi_i^{y_i} (1 - \pi_I)^{1-y_i} \quad (3)$$

dimana: π_i = peluang kejadian ke-i
y_i= peubah acak ke-i yang terdiri dari 0 dan 1

Bentuk model regresi logistik dengan satu variabel prediktor adalah [6]:

$$\pi(x) = \frac{\exp(\beta_0 + \beta_1 x)}{1 + \exp(\beta_0 + \beta_1 x)} \quad (4)$$

Untuk mempermudah menaksir parameter regresi, maka $\pi(x)$ pada persamaan diatas ditransformasikan sehingga menghasilkan bentuk logit regresi logistik, sebagai berikut:

$$g(x) = \ln\left[\frac{\pi(x)}{1-\pi(x)}\right] = \beta_0 + \beta_1 x \quad (5)$$

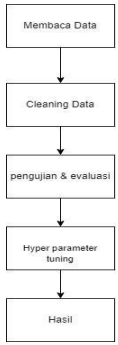
c. Random Forest

Random Forest merupakan salah satu pengembangan dari Decision Tree yang menggunakan beberapa Decision Tree, yang mana setiap decision tree telah dilakukan

pelatihan dari sampel individu dan setiap atribut dipecah pada pohon yang dipilih antara atribut subset yang bersifat acak [6]. Random Forest merupakan salah satu metode yang memiliki tingkat akurasi cukup tinggi dan dapat menangani nilai yang hilang atau fitur yang tidak berkaitan. Metode ini juga dapat bekerja baik pada data yang besar dengan parameter yang kompleks.

2.3 Pembuatan Sistem

Pada pembuatan sistem ini, dilakukan melalui beberapa tahap yaitu membaca dataset, cleaning Dataset, pengujian dan evaluasi, Hyperparameter Tuning, dan hasil. Alur pembuatan sistem dapat dilihat pada gambar 1.



Gambar 1. Alur Pembuatan Sistem

a. Membaca Data

Tahap ini merupakan tahapan import dataset pada sistem, kemudian sistem akan membaca informasi yang ada pada dataset tersebut.

b. Cleaning Data

Pada tahapan ini, dilakukan proses pembersihan sekaligus mempersiapkan dataset sebelum dilakukan analisis ataupun pemodelan. Tujuan dari tahap ini yaitu untuk menghilangkan data yang terduplikat, memperbaiki data yang hilang, tidak lengkap, dan juga tidak relevan [7].

c. Pengujian dan Evaluasi Dataset

Tahapan ini dilakukan proses pengukuran kinerja model terhadap data yang sudah dicari. Tujuan pada tahapan ini yaitu untuk mencari nilai akurasi pada sistem yang dibuat.

d. Hyperparameter Tuning

Hyperparameter Tuning merupakan tahapan dimana nilai-nilai hiperparameter dioptimalisasi untuk mendapatkan kinerja model yang baik. Hiperparameter ditentukan sebelum training pada model [8].

e. Hasil

Pada tahap ini merupakan output dari data yang telah diproses dari sistem.

3. Hasil dan Pembahasan

Pada penelitian sistem pendeteksi penyakit diabetes pada manusia ini dibuat menggunakan bahasa python. Pada penelitian ini mengambil data berupa data numerik. Pengkodean sistem diambil dari

3.1 Hasil menggunakan Metode Gaussian Naive Bayes

Nilai precision, recall, dan f-1 score menggunakan metode ini adalah 0.83; 0.79; 0.81 pada angka 0 (tidak diabetes) dan 0.58; 0.65; 0.61 pada angka 1 (diabetes). Hasil precision, recall, dan f-1 score dari penggunaan Metode Gaussian Naive Bayes menunjukan data diuji negatif terkena diabetes.

	precision	recall	f1-score	support
0	0.83	0.79	0.81	150
1	0.58	0.65	0.61	68
accuracy			0.74	218
macro avg	0.70	0.72	0.71	218
weighted avg	0.75	0.74	0.75	218

Gambar 2. Hasil uji menggunakan metode Naive Bayes

3.2 Hasil menggunakan Metode Regresi Logistik

Nilai precision, recall, dan f-1 score menggunakan metode ini adalah 0.81; 0.88; 0.85 pada angka 0 (tidak diabetes) dan 0.68; 0.56; 0.61 pada angka 1 (diabetes). Hasil precision, recall, dan f-1 score dari penggunaan Metode Gaussian Naive Bayes menunjukan data diuji negatif terkena diabetes.

	precision	recall	f1-score	support
0	0.81	0.88	0.85	150
1	0.68	0.56	0.61	68
accuracy			0.78	218
macro avg	0.75	0.72	0.73	218
weighted avg	0.77	0.78	0.77	218

Gambar 2. Hasil uji menggunakan metode Regresi Logistik

3.3 Hasil menggunakan Metode Random Forest

Nilai precision, recall, dan f-1 score menggunakan metode ini adalah 0.82; 0.87; 0.84 pada angka 0 (tidak diabetes) dan 0.66; 0.57; 0.61 pada angka 1 (diabetes). Hasil precision, recall, dan f-1 score dari penggunaan Metode Gaussian Naive Bayes menunjukan data diuji negatif terkena diabetes.

	precision	recall	f1-score	support
0	0.82	0.87	0.84	150
1	0.66	0.57	0.61	68
accuracy			0.78	218
macro avg	0.74	0.72	0.73	218
weighted avg	0.77	0.78	0.77	218

Gambar 3. Hasil uji menggunakan metode Random Forest

4. Kesimpulan

Pada penelitian ini dapat disimpulkan bahwa tingkat akurasi pada metode Gaussian Naive Bayes, Regresi Logistik, Random Forest pada data yang diuji yaitu sebesar 0.74; 0;78 ; 078. Berdasarkan nilai akurasi pada data yang diuji metode Gaussian Naive Bayes memiliki akurasi paling rendah, sementara Regresi Logistik dan Random Forest memiliki nilai akurasi yang sama.

Daftar Pustaka

[1] C. J. Hinde, R. Stone and X. Daniela, "Naive Bayes vs. Decision Trees vs. Neural Networks in the Classification of Training Web Pages," international Journal of Computer Science, vol. 1, no. 4, September 2009.

[2] D. I. Saputra and D. L. Hakim, "ImplementasiAlgoritma Gaussian Naive Bayes Classifier Untuk Prediksi Potensi Tsunami Berbasis Mikrokontroler," Journal of Electrical Engineering and Information Technology, vol. 2, no. 20, December 2022.

[3] J. P. Sianipar, R. E. S. and C. S. , "Waves with Multi-Sensor System Based on Web Application Using Naive Bayes Algorithm," e-Proceeding of Engineering, vol. 9, no. 5, pp. 6183-6188, 2021.

[4] H. D. and S. L. , "Applied Logistic Regression," vol. 2, 2000.

[5] A. A, "Categorical Data Analysis John Wiley and Sons," 1990.

[6] R. Supriyadi, W. Gata, N. Maulidah and A. fauzi, "Penerapan Algoritma Random Forest Untuk Menentukan Kualitas Anggur Merah," Jurnal Ilmiah Ekonomi dan Bisnis, vol. 13, no. 2, pp. 67-75, December 2020.

[7] L. and J. , "Data Cleaning in Python : the Ultimate Guide," 4 Febuary 2020.

[8] S. Paul, "Hyperparameter Optimization in Machine Learning Models," August 2018.

Halaman ini sengaja dibiarkan kosong

JNATIA Volume 1, Nomor 4, Agustus 2023 p-ISSN: 2986-3929
Jurnal Nasional Teknologi Informasi dan Aplikasinya

Klasifikasi Penyakit HepatitisC Menggunakan Algoritma Support Vector Machine

Ni Kadek Trisnawati[a1], I Gede Santi Astawa[a2]

Program Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam,
Universitas Udayana
Jalan Raya Kampus Udayana, Bukit Jimbaran, Kuta Selatan, Badung, Bali Indonesia
[1]kadektrisna646@gmail.com
[2]santiastawa@gmail.com

Abstract

Hepatitis is an inflammation of the liver which is often caused by a virus. Hepatitis that lasts less than 6 months is called "acute hepatitis", hepatitis that lasts more than 6 months is called "chronic hepatitis" Hepatitis consists of various types starting from hepatitis A, B, C, D and E. This study will discuss the classification Hepatitis C is classified as blood based (blood donor), suspected blood donor, fibrosis (moderate to severe hepatitis) and cirrhosis (chronic hepatitis). Classification uses the Support Vector Machines (SVM) algorithm with Confusional Matrix testing. The dataset used was obtained from the kaggle.com site with a total of 590 data consisting of 14 features or attributes. This study produces an accuracy of 93%

Keywords: *Hepatitis, Support Vector Machines (SVM)*

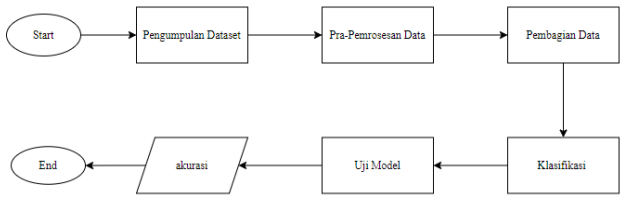
1. Pendahuluan

Penyakit Hepatitis adalah penyakit yang disebabkan oleh beberapa jenis virus yangmenyerang dan menyebabkan peradangan serta merusak sel-sel organ hatimanusia. Namun, hepatitis juga dapat disebabkan oleh faktor lain seperti autoimun, alkohol, dan obat-obatan (Martinez, 1996). Virus hepatitis C adalah virus yang ditularkan melalui darah dan sebagian besar infeksi terjadi melalui paparan darah dari praktik injeksi yang tidak aman, perawatan kesehatan yang tidak aman, transfusi darah yang tidak diskrining, penggunaan narkoba suntikan, dan praktik seksual yang menyebabkan paparan darah [1]. Tingkatan dari hepatitis C ini seperti fibrosis yaitu penyakit hepatitis dengan level sedang berat dan sirosis yaitu penyakit hepatitis dengan level kronis. Dalam masalah ini akan dilakukan klasifikasi jenis penyakit hepatitis C menggunakan algoritma Support Vector Machine (SVM). Algoritma SVM merupakan algoritma yang sudah terbukti efektif dalam klasifikasi datayang kompleks. Dalam kasus klasifikasi sudah terdapat banyak algoritma yang digunakan seperti, Naïve Bayer, Random Forest, K-Nearest Neighbour (KNN) dan Support Vector Machines (SVM). KNN menjadi algoritma yang paling sering digunakan oleh para peneliti dalam bidang klasifikasi. Namun, dalam beberapa penelitian menunjukkan tingkat akurasi SVM lebih tinggi dibandingkan dengan algoritma KNN. Karamoy, Gandis Helen (2021) melakukan tingkat kerusakan hati yang diakibatkan oleh virus hepatitis C. Pada penelitian tersebut menggunakan beberapa metode seperti Support Vector Machine (SVM), Random Forest (RF), Naive Bayes, Logistic Regression dan kNN. Dengan menggunakan dataset dari the University of California Irvine Machine Learning Repository (UCI-MLR), data di dibuat oleh penelitian memaparkan akurasi terbaik dari algoritma Support Vector Machines (SVM) sebesar 0,883 dan akurasi terendah dengan menggunakan algoritma decision tree. Indri Monika Parapat, Muhammad Tanzil Furqon , Sutrisno (2018) melakukan penelitian terhadap penyimpangan tumbuh kembang anak. Pada penelitiannya menggunakan metode SVM dengan 90 data yang dibagi menjadi 3 kelas Kelas penelitian ini mewakilkan 3 jenis penyimpangan tumbuh kembang anak yaitu Down Syndrome, Autisme, dan attention deficit hyperactivity disorder (ADHD). Hasil akhir dari penilitan ini menghasilkan rata-rata akurasi tertinggi sebesar 63,11% λ = 10, C = 1, itermax = 200 dan juga menggunakan kernel polynomial. Penelitian ini dilakukan untuk mengklasifikasi penyakit hepatitis C apakah tergolong ke dalam diagnosis blood donor (memiliki riwayat pen donor darah), suspect blood donor (diduga memiliki potensi menjadi pendonor

darah), hepatitis, fibrosis atau sirosis. Dataset yang digunakan berasal dari dataset kaggle "Hepatitis C virus – Blood Based Detection" dengan 14 atribut di dalamnya.

2. Metode Penelitian

Dalam penelitian ini dilakukan dengan menggunakan algoritma SVM, adapun proses pada sistem ini berjalan seperi berikut



Gambar 1. Flowchart Klasifikasi penyakit hepatitis

2.1 Pengumpulan Data

Penelitian ini menggunakan dataset yang didapatkan dari situs kaggle yaitu "Hepatitis C virus – Blood bases Detection" dan disimpan ke dalam google drive. Dataset ini terdiri dari 590 data csv dengan 14 atribut diantaranya yaitu ID, Category(diagnosis), Age (umur), Sex (jenis kelamin), ALB (serum albumin), ALP (fosfatase alkali), ALT (alanine aminotransferase), AST (aspartate aminotransferase), BIL (bilirubin), CHE (kolinesterase), CHOL (kolesterol), CREA kreatini), GGT (gamma-glutamyl transferase), PROT (protein).

2.2 Pra-Pemrosesan Data

Pada tahapan ini dilakukan seleksi fitur, dimana fitur yang tidak digunakan akan dihilangkan atau dihapus dari dataset. Adapun fitur atau atribut yang digunakan dalam penelitian ini yaitu usia (Age), tingkat albumin (ALB), tingkat fosfatase alkali (ALP), tingkat aspartate aminotransferase (AST), dan tingkat bilirubin (BIL). Setelah atribut atau fitur ditentukan selanjutnya melakukan pemisahan data menjadi data trening dan data testing. Pemisahan ini menggunakan 'train_test_split()' dari scikit-learn yaitu data trening (x_train, y_train) dan data testing (x_test, y_test).

2.3 Pembagian Data

Table 1. Pembagian Data

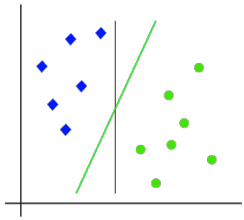
Jumlah Data	Persentase
462	80%
128	20%

2.4 Klasifikasi

a. Support Vector Machines (SVM)

Support Vector Machine atau SVM adalah algoritme pembelajaran mesin yang diawasi yang dapat digunakan untuk klasifikasi dan regresi. Support Vector Machines (SVM)

merupakan algoritma membentuk ruang hitosis berupa fungsi-fungsi linier di sebuah ruang fitur berdimensi tinggi [2]. SVM dibagi menjadi dua jenis yaitu SVM linear dan SVM non-linear, pada penelitian ini menggunakan algoritma SVM linear. SVM linier digunakan pada data yang memiliki kemampuan untuk dipisahkan secara linear. Artinya, jika sebuah dataset dapat dibagi menjadi dua kelas menggunakan sebuah garis lurus tunggal, maka data tersebut dianggap dapat dipisahkan secara linier, dan pengklasifikasi yang digunakan disebut sebagai pengklasifikasi SVM linear. Hyperplane dengan margin maksimum dan margin untuk SVM yang dilatih dengan sampel dari dua kelas. Sampel pada margin disebut vektor pendukung.



Gambar 2. Margin SVM Linear

Adapun rumus Support Vector Machines Linear yang digunakan seperti berikut :

$y(x) = w^T * x + b$

Di mana:

a. y(x) adalah prediksi kelas untuk sampel x.
b. w adalah vektor bobot yang ditemukan selama proses pelatihan SVM linier.
c. x adalah vektor fitur dari sampel yang ingin diprediksi.
d. b adalah bias (intercept) yang ditentukan selama proses pelatihan SVM linier.

b. Uji Model

Pengujian dilakukan dengan menggunankan confusionan matrixs dimana confusion matrix sendiri merupakan alat penting untuk mengevaluasi sejauh mana model klasifikasi mampu melakukan prediksi yang benar. Dengan menganalisis kombinasi dari keempat istilah tersebut, kita dapat memperoleh wawasan tentang seberapa baik model dalam mengklasifikasikan data ke dalam kelas yang benar.

3. Hasil dan Pembahasan

3.1 Confusion Matrix

Pada penelitian ini digunakan untuk memberikan indikasi sejauh mana model klasifikasi SVM dapat memprediksi dengan benar setiap kelas atau kelas yang ada. Dengan menganalisis matriks ini, kita dapat mendeteksi kesalahan klasifikasi yang dibuat oleh model dan mengukur keefektifan dan akurasi model dalam membuat prediksi. Berikut merupakan confusion matrix yang dihasilkan dalam penelitian ini

```
Confusion Matrix:
[[132   0   0   0   0]
 [  1   0   0   0   1]
 [  1   0   1   0   0]
 [  2   0   2   0   0]
 [  2   0   1   0   5]]
```

Gambar 3. Tampilan Matriks Confusion

Penjelasan setiap elemen:

a. Pada baris pertama menunjukkan terdapat 132 data yang diklasifikasikan dengan benar kedalam kelas 0 (Blood based) tanpa ada kesalahan klasifikasi
b. Baris kedua menunjukkan kelas sebenarnya berlabel 1. Pada matriks diatas terdapat 1 sampel diklasifikasikan dengan benar sebagai kelas 1, sementara 1 sampel salah diklasifikasikan sebagai kelas 4 dan tidak ada sampel yang salah diklasifikasikan sebagai kelas 1.
c. Baris ketiga menunjukkan kelas sebenarnya berlabel 2. Pada matriks diatas, ada 1 sampel yang diklasifikasikan dengan benar sebagai kelas 2 tanpa salah klasifikasi.
d. Baris keempat menunjukkan kelas sebenarnya yang diwakili oleh label 3. Pada matriks diatas, ada 2 sampel yang benar diklasifikasikan sebagai kelas 3, sedangkan ada 2 sampel yang salah diklasifikasikan sebagai kelas 4 dan tidak ada sampel yang salah diklasifikasikan sebagai kelas 3.
e. Baris kelima menunjukkan kelas sebenarnya yang diwakili oleh label 4. Pada matriks diatas, ada 5 sampel yang benar diklasifikasikan sebagai kelas 4, sedangkan ada 2 sampel yang salah diklasifikasikan sebagai kelas 3 dan 1 sampel yang salah diklasifikasikan sebagai kelas 4

3.2 Pengujian Model SVM

```
[0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 2 4 0 0 0 0 2 0 2 2 2 0
 0 0 0 4 4 2 0]
```

Gambar 3. Hasil klasifikasi Model SVM

Penjelasan:

a. Sempel 0: Menunjukkan bahwa sampel tersebut diklasifikasikan ke dalam kelas 0.
b. Sempel 2: Menunjukkan bahwa sampel tersebut diklasifikasikan ke dalam kelas 2.
c. Sempel 4: Menunjukkan bahwa sampel tersebut diklasifikasikan ke dalam kelas 4.

Dapat disimpulkan dari hasil diatas mengklasifikasikan penderita sebanyak sempel 0 untuk penderita dengan blood based (akibat donor darah), sebanyak sempel 2 untuk penderita hepatitis dan sebanyak sempel 4 untuk penderita penyakit hepatitis kronis (sirosis).

```
Akurasi: 0.9324324324324325
```

Gambar 4. Memperlihatkan Hasil Akurasi

4. Kesimpulan

Kesimpulan yang dapat dibuat berdasarkan penelitian yang telah dilakukan adalah: Algoritma Support Vector Machines (SVM) dapat diterapkan pada klasifikasi penyakit hepatitis. Cara dalam

menerapkan algoritma ini yaitu dengan perhitungan SVM linear menggunakan atrbut dari dataset. Pengunjian dengan menggunakan confusional matrix menghasilkan akurasi sebesar 93%.

Daftar Pustaka

[1] World Health Organization. 24 June 2022. Hepatitis C. Diakses pada 09 Juni 2023, dari https://www.who.int/news-room/fact-sheets/detail/hepatitis-c
[2] Zuama,RA. 2021. Pembelajaran Mesin untuk diagnosis tingkat kerusakan hati akibat hepatitis C. Sentra Penelitian Engineering dan Edukasi. Volume 13 No 3.
[3] Hearst, M. A., Dumais, S. T., Osuna, E., Platt, J., & Scholkopf, B. (1998). Support vector machines. IEEE Intelligent Systems and their applications, 13(4), 18-28.
[4] Parapat,IM dkk. (2018). Penerapan Metode Support Vector Machine (SVM) Pada Klasifikasi Penyimpangan Tumbuh Kembang Anak. Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer. Vol. 2, No.10
[5] Nurlaily dkk. (2022). Classification of Hepatitis Patients Using Logistic Regression and Support Vector Machines Methods. Jurnal Pendidikan Matematika. Volume 5, Number 2

Halaman ini sengaja dibiarkan kosong

JNATIA Volume 1, Nomor 4, Agustus 2023 p-ISSN: 2986-3929
Jurnal Nasional Teknologi Informasi dan Aplikasinya

Perancangan dan Pembangunan Aplikasi E-Commerce Berbasis Mobile TechTrove

Saifulloh Rahman[a1], I Komang Ari Mogi[a2],

Program Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam,
Universitas Udayana
Jalan Raya Kampus Udayana, Bukit Jimbaran, Kuta Selatan, Badung, Bali Indonesia
[1]saifulloh.rahman028@student.unud.ac.id
[2]arimogi@unud.ac.id

Abstract

This paper presents the design and development of mobile-based e-commerce applications specifically designed for electronic devices. With the growing popularity of smartphones and growing demand for online shopping, the need for a user-friendly and efficient mobile e-commerce application has become imperative. The proposed application aims to provide a seamless shopping experience for users who wish to purchase electronic devices through their mobile devices. The design process covers multiple aspects, including user interface design, navigation flows, and integration of important ecommerce features such as product listings, filtering, and secure payment methods. The development phase involves implementing the design using the appropriate programming languages and frameworks to ensure compatibility across various mobile platforms. The app's success will be evaluated based on user feedback, performance metrics, and its ability to improve the overall shopping experience for electronics enthusiasts. By combining mobile technology and e-commerce, this research contributes to the advancement of the digital shopping landscape and offers a convenient platform for users to easily browse and buy electronic devices.

Keywords: *E-Commerce, Marketpalce, mobile application, Electronic Device, designing.*

1. Pendahuluan

Dalam era digital yang terus berkembang, perdagangan elektronik atau e-commerce telah menjadi salah satu sektor yang paling pesat pertumbuhannya. Penggunaan perangkat seluler, seperti smartphone dan tablet, semakin meluas di seluruh dunia, sehingga mengubah cara orang berbelanja. Perangkat seluler memberikan kenyamanan dan aksesibilitas yang tak tertandingi dalam melakukan transaksi online kapan saja dan di mana saja. aplikasi e-commerce berbasis mobile telah menjadi alat yang sangat penting untuk membantu pengguna dalam menjelajahi dan melakukan pembelian secara online. Dalam artikel ini, saya memfokuskan pada perancangan dan pembangunan aplikasi e-commerce berbasis mobile yang khusus ditujukan untuk alat elektronik. seperti smartphone, laptop, kamera, dan perangkat audio lainnya, menjadi salah satu kategori produk yang paling diminati dalam perdagangan elektronik. Kebutuhan akan aksesibilitas, informasi yang lengkap, dan pengalaman berbelanja yang lancar semakin tinggi di kalangan konsumen alat elektronik. Oleh karena itu, perancangan dan pengembangan aplikasi e-commerce yang dioptimalkan khusus untuk perangkat seluler dapat menjadi solusi yang efektif untuk memenuhi kebutuhan tersebut. penelitian ini akan membahas langkah-langkah yang diperlukan dalam perancangan dan pembangunan aplikasi e-commerce berbasis mobile untuk alat elektronik. menggabungkan prinsip-prinsip desain antarmuka pengguna yang intuitif dengan fitur-fitur penting dalam e-commerce, seperti daftar produk, filter, dan metode pembayaran yang aman. Selain itu, saya akan menjelaskan tentang pengembangan aplikasi menggunakan bahasa pemrograman dan kerangka kerja yang sesuai untuk memastikan kompatibilitasnya dengan berbagai platform seluler. saya berharap dapat memberikan kontribusi untuk kemajuan dunia belanja digital dan memberikan pengalaman berbelanja yang nyaman bagi pengguna alat elektronik. Dengan mengintegrasikan teknologi seluler dan e-commerce, aplikasi yang dihasilkan

akan memberikan platform yang efisien dan efektif bagi para pengguna untuk menjelajahi dan membeli alat elektronik dengan mudah melalui perangkat seluler mereka.

1.1. E-Commerce

E-Commerce secara umum dapat diartikan sebagai transaksi jual beli secara elektronik melalui media internet. Selain itu, E-commerce juga dapat diartikan sebagai suatu proses berbisnis dengan memakai teknologi elektronik yang menghubungkan antara perusahaan, konsumen dan masyarakat dalam bentuk transaksi elektronik dan pertukaran atau penjualan barang, servis, dan informasi secara elektronik.

1.2. Marketplace

Marketplace adalah platform yang menjadi perantara antara penjual dan pembeli di internet. Jadi, website marketplace bertindak sebagai pihak ketiga dalam transaksi online dengan menyediakan tempat berjualan dan fasilitas pembayaran.

1.3. IOS

IOS sebelumnya iPhone OS merupakan sistem operasi seluler yang dibuat dan dikembangkan oleh Apple inc khusus untuk perangkat kerasnya. Ini adalah sistem operasi yang saat ini memberdayakan banyak perangkat seluler perusahaan, termasuk iPhone, dan iPod Touch; itu juga mendukung iPad sebelum pengenalan iPadOS pada 2019. Ini adalah sistem operasi seluler terpopuler kedua di dunia setelah Android.

2. Metode Penelitian

Pada penelitian ini digunakan beberapa metode untuk merancang aplikasi, yaitu:

2.1. Usability testing.

usability testing untuk mengumpulkan data dan mendapatkan wawasan yang mendalam tentang pengalaman pengguna dalam menggunakan aplikasi e-commerce berbasis mobile.

2.2. User Centered Design

User Centered Design (UCD) merupakan metodologi yang digunakan oleh developer dan desainer dalam memastikan mereka membangun produk yang memenuhi kebutuhan pengguna. Produk yang dikembangkan dengan pendekatan UCD dioptimalkan untuk end user serta ditekankan pada bagaimana kebutuhan atau keinginan end user terhadap penggunaan produk.

2.3. System Usability Scale

System Usability Scale (SUS) merupakan metode dalam pengujian suatu aplikasi menggunakan sepuluh skala yang memberikan pandangan pengguna secara global dari sisi kebergunaannya. Pada metode ini, pengujian usability menitikberatkan pada sudut pandang pengguna akhirsehingga hasil evaluasi bisa sesuai dengan keadaan nyata. Kelebihan yang menjadi alas an penggunaan metode ini adalah responden mampu memahami pertanyaan dengan lebih mudah, tidak membutuhkan responden dalam jumlah besar namun memiliki akurasi yang tinggi, dan memperoleh informasi mengenai kebergunaan aplikasi [4]. Berikut adalah langkah-langkah dalam metode SUS:

a. Penyusunan Kuesioner: Pertama, penyusu*n kuesioner mengembangkan kuesioner SUS* yang terdiri dari 10 pernyataan yang berkaitan dengan kegunaan sistem. Pernyataan-pernyataan ini biasanya menggunakan skala Likert dengan pilihan jawaban dari "Sangat Setuju" hingga "Sangat Tidak Setuju".

Tabel 1. Pertanyaan System Usability Scale

No	Pertanyaan
1	Saya berpikir akan menggunakan aplikasi TechTrove lagi
2	Saya merasa aplikasi TechTrove ini rumit untuk digunakan
3	Saya merasa aplikasi TechTrove ini mudah digunakan
4	Saya membutuhkan bantuan dari orang lain atau teknisi dalam menggunakan aplikasi TechTrove ini
5	Saya merasa fitur tampilan aplikasi TechTrove ini berjalan dengan semestinya
6	Saya merasa ada banyak hal yang tidak konsisten (tidak serasi pada aplikasi TechTrove ini)
7	Saya merasa orang lain akan memahami cara menggunakan aplikasi TechTrove ini dengan cepat
8	Saya merasa aplikasi TechTrove ini membingungkan
9	Saya merasa tidak ada hambatan dalam menggunakan aplikasi TechTrove ini
10	Saya perlu membiasakan diri terlebih dahulu sebelum menggunakan aplikasi TechTrove ini

b. Pengujian Pengguna: Kuesioner SUS kemudian diberikan kepada sekelompok pengguna yang telah berinteraksi dengan sistem atau produk yang dievaluasi. Pengguna diminta untuk memberikan tanggapan mereka berdasarkan pengalaman penggunaan mereka.

c. Perhitungan Skor SUS: Skor SUS dihitung dengan menjumlahkan dan mengubah skor jawaban pengguna pada pernyataan-pernyataan dalam kuesioner. Skor SUS dapat berkisar antara 0 hingga 100.

$$\bar{x} = \frac{\sum x}{n} \quad (1)$$

Keterangan:

$\bar{x}$: Rata-rata skor
$\sum x$: Jumlah Skor SUS
n : Jumlah Responden [5].

d. Interpretasi Skor: Skor SUS dapat diinterpretasikan untuk mengukur kegunaan sistem atau produk. Skor rata-rata dianggap sebagai indikator umum kegunaan, dengan skor di atas 68 dianggap sebagai kegunaan yang baik.

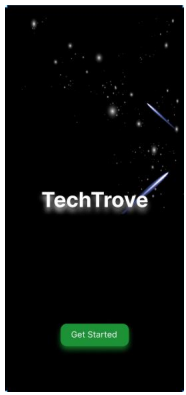
3. Hasil Diskusi

Bagian ini memuat hasil dan pembahasan penelitian dan dapat disajikan dalam bentuk deskripsi, grafik atau gambar.

3.1. Implementasi

a. Tampilan Screan Awal

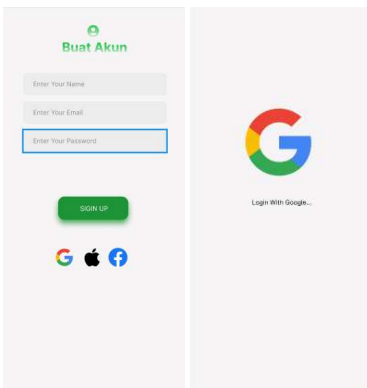
Interface yang menampilkan halaman splash screen yang muncul ketika user baru membuka aplikasi. Selanjutnya user akan diarahkan pada tampilan halaman boarding screen. Pada interface boarding screen terdapat tombol selanjutnya yang akan membawa user ke halaman registrasi. Lebih jelasnya dapat dilihat pada gambar berikut.



Gambar 1. Screen Awal

b. Tampilan Register dan contoh login.

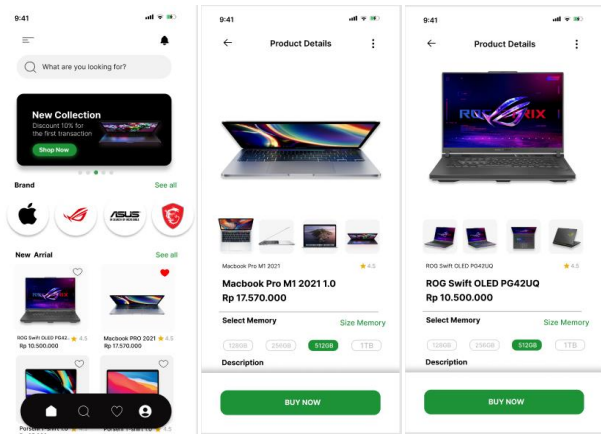
Halaman Login & Register adalah tempat di mana pengguna dapat masuk ke akun mereka yang sudah terdaftar atau membuat akun baru. Pengguna akan diminta untuk memasukkan informasi login seperti email dan kata sandi, atau melakukan proses registrasi dengan mengisi formulir dengan informasi pribadi mereka.



Gambar 2. Register dan Login.

c. Tampilan Home Page

Halaman Beranda adalah tampilan utama aplikasi di mana pengguna dapat melihat konten-konten yang relevan dan terkait dengan barang yang ingin dicari.



Gambar 3. Home Page

4. Hasil Evaluasi

Berikut merupakan table hasil evaluasi dari kuisioner yang telah dibagikan kepada 20 responden

5. Kesimpulan

Perdagangan elektronik atau e-commerce telah mengalami pertumbuhan yang pesat dalam era digital yang terus berkembang. Penggunaan perangkat seluler, seperti smartphone dan tablet, telah mengubah cara orang berbelanja dengan memberikan kenyamanan dan aksesibilitas tak tertandingi dalam melakukan transaksi online. Dalam konteks ini, aplikasi e-commerce berbasis mobile menjadi alat yang sangat penting untuk membantu pengguna menjelajahi dan melakukan pembelian secara online. Dalam penelitian ini, fokus diberikan pada perancangan dan pembangunan aplikasi e-commerce berbasis mobile yang ditujukan khusus untuk alat elektronik. Alat elektronik, seperti smartphone, laptop, kamera, dan perangkat audio lainnya, menjadi salah satu kategori produk yang paling diminati dalam perdagangan elektronik. Kebutuhan akan aksesibilitas, informasi yang lengkap, dan pengalaman berbelanja yang lancar semakin tinggi di kalangan konsumen alat elektronik. Oleh karena itu, perancangan dan pengembangan aplikasi e-commerce yang dioptimalkan khusus untuk perangkat seluler menjadi solusi yang efektif untuk memenuhi kebutuhan tersebut. Metode penelitian yang digunakan dalam penelitian ini meliputi usability testing untuk mendapatkan wawasan mendalam tentang pengalaman pengguna dalam menggunakan aplikasi e-commerce berbasis mobile. Selain itu, pendekatan User Centered Design (UCD) digunakan untuk memastikan aplikasi yang dibangun memenuhi kebutuhanpengguna dengan fokus pada pengguna akhir. Metode System Usability Scale (SUS) juga digunakan untuk menguji kebergunaan aplikasi secara global dari perspektif pengguna. Dengan menggabungkan teknologi seluler dan e-commerce, aplikasi yang dihasilkan dari

penelitian ini memberikan platform yang efisien dan efektif bagi pengguna untuk menjelajahi dan membeli alat elektronik dengan mudah melalui perangkat seluler mereka. Dengan kontribusi ini, diharapkan dapat memberikan kemajuan dalam dunia belanja digital dan memberikan pengalaman berbelanja yang nyaman bagi pengguna alat elektronik

Daftar Pustaka

[1] Clover, Juli (13 Desember 2022). "Apple Releases iOS 16.2 and iPadOS 16.2 With Freeform, Apple Music Sing, Advanced Data Protection and More". MacRumors (dalam bahasa Inggris). Diakses tanggal 09 Juni 2023.

[2] Monalia Mariana, (2012). Apa aitu E-Commerce? Diakses pada 09 Juni 2023, https://www.unpas.ac.id/apa-itu-e-commerce/

[3] Ilham Mubarok ,(2022). Apa Itu Marketplace? Pahami Bedanya dengan Toko Online! Diakses pada 09 Juni 2023, https://www.niagahoster.co.id/blog/marketplace-adalah/#Apa_Itu_Marketplace

[4] M. Prabowo dan A. Suprapto, “Usability Testing pada Sistem Informasi Akademik IAIN Salatiga Mengunakan Metode System Usability Scale,” 2021.

JNATIA Volume 1, Nomor 4, Agustus 2023 p-ISSN: 2986-3929
Jurnal Nasional Teknologi Informasi dan Aplikasinya

Perancangan Aplikasi Sistem Keamanan Checkout Online Shop

Viencent[1], Cokorda Pramartha[2]

[1,2]Program Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Udayana
Jl. Raya Kampus Unud, Jimbaran, Bali, Indonesia
[1]viencentviencent@gmail.com
[2]cokorda@unud.ac.id

Abstract

Online shopping has become a preferred choice for many people to meet their consumer needs. However, security issues in online transactions remain a major concern, especially when unauthorized checkouts occur without the account owner's knowledge. To address this problem, this research aims to design an online shop checkout security application. The application ensures that each checkout is performed by the account owner and not by unauthorized individuals, such as children without their parent's knowledge. Advanced technology approaches and appropriate security methods will be employed in the design of this application, including the use of two-factor authentication and verification through unique security codes. It is expected that this application will provide additional protection to online shop users in maintaining the security and integrity of their transactions.

Keywords: *Online Shop, Checkout Security, 2FA, Unauthorized Transactions*

1. Pendahuluan

Perkembangan zaman yang disertai dengan teknologi yang semakin canggih, membuat segala informasi dapat dengan mudah didapatkan. Perubahan tersebut perlahan-lahan membuat gaya hidup manusia berubah dari melakukan segala sesuatu dengan cara yang rumit dan tradisional menjadi sesuatu yang sederhana dan modern [1]. Contohnya ketika ingin berbelanja, sebelum adanya online shop kita perlu datang secara langsung ke toko yang menjual barang yang kita inginkan. Setelah adanya online shop kita tidak perlu datang secara langsung ke toko tersebut, melainkan hanya memilih barang yang kita inginkan, lalu melakukan pembayaran, dan terakhir menunggu barang yang kita beli di antar ke rumah kita, hal tersebut dapat kita lakukan dari mana saja dan kapan pun. Dalam era digital saat ini, belanja online telah menjadi fenomena yang semakin populer di kalangan masyarakat. Kepraktisan dan kemudahan dalam melakukan transaksi secara online menjadikan online shop sebagai salah satu pilihan utama dalam memenuhi kebutuhan konsumen. Melalui online shop, pengguna dapat dengan melakukan pembelian barang dan melakukan proses checkout untuk menyelesaikan transaksi. Namun, perkembangan teknologi dan kemudahan akses juga membawa implikasi baru dalam hal keamanan transaksi online. Dengan beberapa aturan pada sistem online shop seperti sistem distribusi produk, sistem pembayaran dan sistem informasi, maka hal keamanan adalah faktor penting agar semua sistem in dapati berfungsi seperti yang diharapkan [2]. Salah satu masalah yang sering muncul adalah adanya risiko penyalahgunaan atau transaksi yang tidak sah, terutama ketika terjadi proses checkout tanpa sepengetahuan pemilik akun. Dalam konteks ini, masalah keamanan yang timbul adalah ketika seorang anak melakukan checkout pada aplikasi online shop dengan sistem pembayaran COD (Cash on Delivery) tanpa sepengetahuan orang tua atau pemilik akun. Hal ini tentu dapat menimbulkan kerugian finansial dan ketidaknyamanan bagi pemilik akun yang tidak mengetahui transaksi tersebut. Oleh karena itu, dalam rangka mengatasi masalah ini diperlukan perancangan sebuah aplikasi sistem keamanan checkout online shop yang dapat memberikan perlindungan terhadap penyalahgunaan akun dan transaksi yang tidak sah. Aplikasi ini akan memastikan bahwa setiap proses checkout yang dilakukan

benar-benar dilakukan oleh pemilik akun, dan bukan oleh pihak yang tidak berwenang seperti anak-anak tanpa sepengetahuan orang tua. Pada aplikasi yang diusulkan, pengguna akan diminta untuk melakukan konfirmasi tambahan saat akan melakukan proses checkout. Aplikasi ini akan memastikan bahwa pengguna yang melakukan checkout adalah pemilik akun yang sah, dengan meminta verifikasi tambahan seperti penggunaan kode keamanan unik atau verifikasi melalui otentikasi ganda. Dengan demikian, diharapkan dapat mencegah akses yang tidak sah dan melindungi pemilik akun dari transaksi yang tidak diinginkan. Pada penelitian ini akan dijelaskan secara detail perancangan aplikasi sistem keamanan checkout online shop, termasuk fitur-fitur keamanan yang diimplementasikan dan juga dilakukan evaluasi terhadap efektivitasnya dalam meningkatkan keamanan sistem checkout. Diharapkan hasil dari penelitian ini dapat memberikan kontribusi dalam meningkatkan keamanan sistem checkout online shop dan memberikan perlindugan yang lebih baik bagi pengguna dalam melaksanakan transaksi secara online.

2. Metode Penelitian

Berdasarkan permasalahan yang telah diidentifikasi, maka langkah-langkah metode yang digunakan untuk menyelesaikan masalah tersebut adalah melakukan analisis pengumpulan kebutuhan.

2.1. Kebutuhan Fungsional

Kebutuhan fungsional adalah langkah analisis fitur-fitur apa saja yang akan diimplementasikan pada aplikasi sistem keamanan checkout online shop dan juga memberikan gambaran umum tentang apa saja yang dapat dilakukan pengguna dengan aplikasi [3]. Berikut adalah fitur-fitur yang akan tersedia di aplikasi sistem keamanan checkout online shop:

a. Aplikasi dapat menampilkan halaman persetujuan untuk menggunakan sidik jari yang telah tersimpan di smartphone pengguna.
b. Aplikasi dapat menampilkan halaman untuk meminta fitur keamanan kunci pola atau Personal Identification Number (PIN) dari pengguna.
c. Aplikasi dapat menampilkan halaman utama/dashboard aplikasi.
d. Aplikasi dapat menampilkan menu ubah sandi dan pengguna dapat mengganti kunci pola atau PIN yang telah dibuat saat pertama kali membuka aplikasi.
e. Aplikasi dapat menampilkan menu keamanan dan pengguna dapat mengaktifkan atau menonaktifkan fitur keamanan sidik jari (Fingerprint Identification) dan pengenalan muka (Face Identification).

2.2. Kebutuhan Non-Fungsional

Kebutuhan non-fungsional adalah kebutuhan pendukung kelayakan sistem atau aplikasi yang akan dibangun [3]. Spesifikasi kebutuhan non-fungsional mencakup 2 komponen yaitu perangkat keras (hardware) dan perangkat lunak (software).

a. Kebutuhan Perangkat Keras

Hardware adalah peralatan apa pun yang membantu pengoperasian program perangkat lunak dengan benar. Berikut adalah spesifikasi perangkat keras yang digunakan:

- RAM : 4 GB
- Memori Internal : 64 GB
- Versi Android OS : 10 (Android Q)

b. Kebutuhan Perangkat Lunak

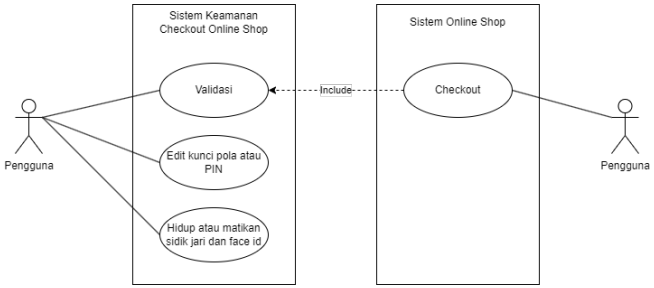
Software adalah perangkat lunak atau bahasa pemrograman yang digunakan untuk membangun aplikasi ini, yang meliputi:

- Sistem operasi Windows 10 Home Single Language 64-bit.
- Android Studio IDE (Integrated Development Environment), Android Software Development Kit (SDK) dan Java Runtime Environment (JRE), dan Android Emulator.

3. Hasil dan Pembahasan

3.1. Use Case Diagram

Diagram use case menggambarkan interaksi yang terjadi antara aktor dan sistem [4]. Pada aplikasi sistem keamanan checkout online shop ini, hanya terdapat satu aktor yaitu pengguna. Pengguna harus melakukan validasi keamanan, pengguna dapat mengubah fitur keamanan kunci pola atau PIN dan menghidupkan atau mematikan fitur keamanan sidik jari (Fingerprint ID) dan pengenalan muka (Face ID).



Gambar 1. Diagram Use Case

3.2. Activity Diagram

Activity Diagram bertujuan untuk mengetahui dengan jelas bagaimana cara sistem tersebut bekerja dan masalah yang dihadapi sistem untuk dapat digunakan sebagai dasar perancangan analisa sistem yang sedang berjalan yang dilakukan secara berurutan atas dasar peristiwa yang ada [4].

Gambar 2. Activity Diagram

4. Kesimpulan

Perancangan aplikasi sistem keamanan checkout online shop merupakan aplikasi yang dapat memberikan perlindungan terhadap penyalahgunaan akun dan transaksi tidak sah. Dengan adanya aplikasi ini dapat dipastikan bahwa setiap proses checkout yang dilakukan benar-benar dilakukan oleh pemilik akun dengan meminta verifikasi keamanan yang sudah ditetapkan pengguna, sehingga diharapkan dapat mencegah kerugian yang dapat diakibatkan oleh penyalahgunaan akun tersebut.

Daftar Pustaka

[1] T. Listiani and A. Wulandari, “Pengaruh Keamanan Bertransaksi, Kemudahan Transaksi dan Citra Merek Terhadap Minat Beli Konsumen pada E-Commerce Tokopedia,” J. Ekonomi Manajemen Akuntansi Kewirausahaan, pp. 50–58, 2022.
[2] Hairullah, C. R. A. Pramartha, and I. A. G. S. Putra, “Aplikasi Keamanan E-Commerce Berbasis Web Menggunakan Metode Algoritma Blowfish,” J. Nasional Teknologi Informasi dan Aplikasinya, vol. 1, no. 1, pp. 79-88, 2022.
[3] D. R. Nalle, C. R. A. Pramartha, and I. G. N. A. C. Putra, “Pengembangan Prototype Aplikasi E-Commerce Berbasis Android Sebagai Media Penjualan Tenun Rote Ndao,” Jurnal Pengabdian Informatika, vol. 1, no. 2, pp. 483-490, 2023
[4] D. N. Huda, A. Saputra, and Yulinda, “Perancangan Aplikasi IT Help Desk Menggunakan Platform Node.js Pada Mittasys,” Jurnal Bangkit Indonesia, 9(1), pp. 137-143, 2020.

JNATIA Volume 1, Nomor 4, Agustus 2023 p-ISSN: 2986-3929
Jurnal Nasional Teknologi Informasi dan Aplikasinya

Penerapan Metode Kompresi Wavelet dalam Pengolahan Data Gambar untuk Mengurangi Ukuran File

Ni Putu Suci Paramita[a1], I Gede Arta Wibawa[a2]

[a]Program Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Udayana
Jalan Raya Kampus Udayana, Bukit Jimbaran, Kuta Selatan, Badung, Bali Indonesia
Mitasuci686@gmail.com
gede.arta@unud.ac.id

Abstract

Currently more and more computer users the increase in the number of computer users has led to an increase digital data user. One of the most used digital data today is digital images. The smallest element of a digital image is called pixels. The higher the number of pixels, the higher the digital resolution picture. The higher the pixel count, the larger the digital image file size. Resulting in fast full data storage capacity.Image data compression is an important process in data processing and storage, especially with the increasing use of images in various applications and platforms. This study aims to apply the Wevalet compression method in processing image data to reduce file size. The Wevalet compression method combines the wavelet transform with an adaptive and efficient compression breaking procedure, to form significant compression without significant sacrifice of image quality.

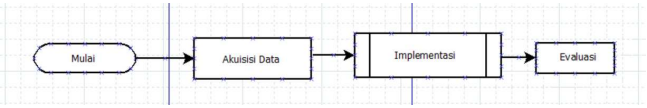
Keyword: *images, compression, Wevalet*

1. Pendahuluan

Penggunaan gambar dalam berbagai aplikasi dan platform telah meningkat secara signifikan dalam beberapa tahun terakhir. Dari media sosial hingga situs net, gambar menjadi bagian fundamental dari pengalaman pengguna. Namun, dengan pertumbuhan ini juga datang tantangan dalam hal penyimpanan dan pengiriman records gambar yang memakan ruang dan bandwidth yang besar. Oleh karena itu, kompresi information gambar menjadi sangat penting untuk mengatasi masalah tersebut. Metode kompresi Wevalet adalah salah satu metode yang mengurangi ukuran file grafik tanpa mengorbankan kualitas secara signifikan. Metode ini menggabungkan konsep transformasi wavelet dengan algoritma kompresi yang adaptif dan efisien, memungkinkan penggunaan sumber daya penyimpanan dan bandwidth jaringan yang lebih efisien. Transformasi ini membagi gambar menjadi langkah-langkah skala dan frekuensi yang berbeda, memungkinkan pola dan struktur yang signifikan untuk diidentifikasi pada gambar. Dengan menggunakan algoritme kompresi adaptif Wevalet, redundansi data gambar dapat dihilangkan dengan mempertahankan informasi yang diperlukan untuk rekonstruksi gambar berkualitas tinggi.

2. Metode Penilitian

2.1. Desain Penelitian



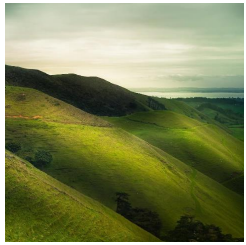
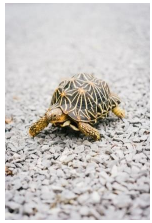
Gambar 1. Desain Penelitian

2.2. Akusisi Data

Pada tahap ini, dilakukan pengumpulan data dengan tehnik scrapping secara manual pada website Kaggle. Penelitian hanya menggunakan data latih dan data uji saja dikarenakan data prediksi tidak diberi label. Dataset ini memiliki 6 label, yaitu flower, mountain, meadow,lake,turtle,apple.Contoh data pada setiap label ditunjukan pada Table 1.

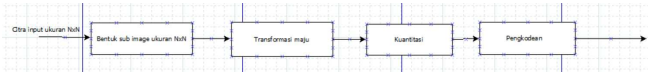
Tabel 1. Contoh Data pada Label Dataset

No	Data	Label
1		Flower
2		Mountain

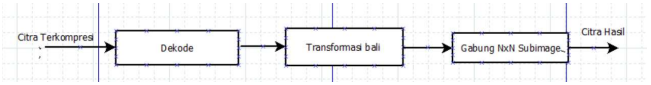
No	Data	Label
3		Meadow
4		Lake
5		Turtle
6		Apple

2.3. Implementasi

Pada Proses kompresi melakukan reduksi ukuran suatu facts untuk menghasilkan representase citra digital yang padat atau mampat namun tetap dapat mewakili kuantitas informasi yang terkandung pada statistics tersebut. Pada citra gambar kompresi mengarah pada minimisasi jumlah bit price untuk representase digital. Pada beberapa literature, istilah kompresi sering disebut juga supply coding, records compression, bandwith compression, dan signal compression.



Gambar 2. Proses Kompresi



Gambar 3. Proses Dekompresi

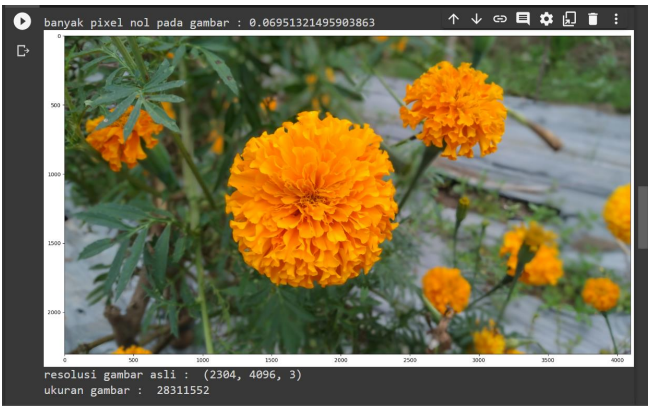
2.4. Evaluasi

Pada tahap evaluasi dilakukan untuk menganalisis performa dari model yang sudah selesai dibuat. Evaluasi pada penelitian ini dilakukan dengan memasukkan data uji ke dalam model lalu akan dihitung akurasi kompresinya.

3. Hasil Pembahasan

Pada proses kompresi menggunakan metode wavelet didapatkan beberapa hasil yang dijabarkan sebagai berikut.

3.1. Hasil



Gambar 4. Data Image Sebelum di Kompresi

Gambar 5. Data Image Setelah di Kompresi

4. Kesimpulan

Berdasarkan hasil kompresi citra gambar dengan menggunakan metode wevelat, didapatkan hasil kompresi yang digunakan pada metode wevelat adalah di ukuran asli data gambar sebesar 28311552 setelah melakukan proses kompresi data gambar menjadi 9437184 dan perbandingan nya sebesar 18874368. Dari metode Wevelat berhasil mengurangi ukuran data gambar dengan efisiensi yang signifikan. Ukuran data gambar berhasil dikurangi lebih dari dua kali lipat, sehingga dapat menghemat ruang penyimpanan. Hal ini menunjukkan bahwa metode Wevelat efektif dalam melakukan kompresi citra gambar.

Daftar Pustaka

[1] D. Putra, “Pengolahan Citra Digital”, Yogyakarta: Andi,2010.
[2] S. Madenda, “Pengolahan Citra & Video Digital”, Jakarta: Erlangga, 2015.
[3] R. Jasmi, M. Perumal and D. M. Rajasekaran, "Comparison of Image Compressio Techniques Usig Huffman Coding, DWT, and Fractal Algorithm," in IEEE Conference Publications, Coimbatore, 2015.
[4] Hermawati, “Pengolahan CitraDigital”. Yogyakarta: Penerbit ANDI,2013
[5] Erwin Fajar Hia. (2006). Kompresi Citra Berbasis Wavelet Menggunakan EZW Dan Trees (SPHIT), Bandung: Telkom University

Halaman ini sengaja dibiarkan kosong

JNATIA Volume 1, Nomor 4, Agustus 2023 p-ISSN: 2986-3929
Jurnal Nasional Teknologi Informasi dan Aplikasinya

Analisis Vulnerability Sistem Informasi di Universitas Udayana Menggunakan Tool Acunetix Web Vulnerability Scanner

I Putu Adi Yuda[1], I Gusti Ngurah Anom Cahyadi Putra[2]

Program Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam,
Universitas Udayana
Jalan Raya Kampus UNUD, Bukit Jimbaran, Kuta Selatan, Badung, Bali, Indonesia
adiyuda418@gmail.com[1]
anom.cp@unud.ac.id[2]

Abstract

Now, information technology is developing rapidly. Information systems provide convenience, especially in the field of education such as information systems at Udayana University. In addition, in universities, information systems are also used in elementary schools and high schools. Information systems must have good security. The requirements that must be met are privacy, integrity, authentication, and availability. However, no information system is completely secure. Like the information system at Udayana University has a vulnerability where it is still possible to be attacked by unauthorized parties. System vulnerability has three levels, namely low level, medium level, and high level. This can disrupt a system, and information can be changed by unauthorized parties. Therefore, it is necessary to scan the Udayana University information system to find out system vulnerabilities. The author uses the Acunetix Web Vulnerability Scanner tool. The results of the scanning can be used as a reference for evaluating system vulnerabilities so that further system security can be improved better than before.

Keywords: *information system, vulnerability, system security*

1. Pendahuluan

Sistem informasi memiliki peran yang penting untuk organisasi. Seperti halnya sistem informasi Universitas Udayana yang dapat memberi kemudahan untuk dapat menjalankan kegiatan dibidang pendidikan. Selain itu, dengan adanya sistem informasi dapat meningkatkan daya kompetitif suatu universitas di era perkembangan teknologi seperti sekarang ini. Di bidang pendidikan, selain universitas bahkan menggunakan juga sebagai sarana peningkatan pelayanan pendidikan seperti sekolah dasar sampai sekolah menengah atas.

Universitas Udayana sebagai perguruan tinggi yang menerapkan penggunaan sistem informasi yang digunakan untuk pendaftaran mahasiswa baru, registrasi KRS, melihat KHS, perpustakaan, dan lain sebagainya. Namun, suatu sistem informasi yang ada pastilah belum bisa sempurna. Masih terdapat suatu kerentanan (*vulnerability*) terhadap bermacam serangan seperti *DDoS*, *SQL Injection*, *ClickJacking*, *Cross Site Scripting (XSS)*, dan lain sebagainya [1]. Hal tersebut dapat menyebabkan ancaman untuk keamanan server yang ada. Kurangnya pemahaman terhadap evaluasi kerentanan suatu sistem informasi membuat ancaman dapat dengan mudah memasuki sistem. Oleh karena itu, perlu dilakukan pengujian penetrasi (*Penetration Testing*) untuk mengetahui apa saja kerentanan yang ada dalam sistem. Untuk melakukan pengujian menggunakan *Penetration Testing Tool* seperti *Acunetix Web Vulnerability Scanner [2]*. Pada *tool* tersebut dapat dideteksi apakah kerentanan tersebut termasuk *level low, medium,* atau *high*. Tujuan dari pengujian penetrasi adalah untuk menemukan kerentanan sistem sehingga dapat dilakukan evaluasi.

2. Metode Penelitian

Metode penelitian ini dilakukan secara sistematik sebagai acuan untuk melakukan penelitian agar memperoleh hasil yang dapat menjadi suatu solusi untuk menyelesaikan permasalah yang akan diteliti.

2.1. Kerangka Kerja Penelitian

Kerangka kerja penelitian adalah langkah-langkah yang digunakan dalam penelitian sampai mendapatkan suatu kesimpulan. Berikut ini merupakan alur dari penelitian yang dilakukan:

Gambar 1. Kerangka Kerja Penelitian

Berdasarkan kerangka kerja penelitian pada gambar 1 dapat dideskripsikan langkah-langkah penelitian sebagai berikut:

- Mengumpulkan Data: Pengumpulan data merupakan suatu hal yang penting sebagai penunjang penelitian. Penunjang tersebut dapat berupa jurnal, buku, media internet, maupun dokumen-dokumen.
- Membuat Skenario Pengujian: Skenario pengujian perlu dibuat agar dapat melakukan tindakan yang tepat dalam melakukan pengujian.
- Melakukan Pengujian: Pengujian dilakukan untuk mengetahui kelemahan dalam suatu sistem sebelum dapat melakukan evaluasi.
- Evaluasi: Evaluasi dilakukan setelah menemukan hasil yang berupa suatu data kerentanan apa saja yang terdapat pada sistem sehingga dapat ditarik suatu kesimpulan.

s

2.2. Pengertian Keamanan Jaringan

Keamanan jaringan merupakan tindakan yang dilakukan sebagai kontrol terhadap jaringan sehingga hanya dapat digunakan oleh orang yang memiliki hak untuk mengaksesnya. Suatu sistem dikatakan aman jika memiliki empat aspek, yaitu kerahasiaan (*privacy*), integritas(*integrity*), *authentication*, dan *availability* [3].

a. *Privacy*
Privacy merupakan suatu aspek dimana suatu sistem informasi hanya dapat dilihat atau diakses oleh pihak yang memiliki hak akses tanpa diketahui pihak siapapun yang tidak memiliki hak akses.

b. *Integrity*
Integritas berkaitan dengan keutuhan suatu informasi. Suatu informasi tersebut jangan sampai diubah oleh pihak yang tidak memiliki hak akses kecuali mendapatkan izin dari pihak yang berwenang.

c. *Authentication*
Authentication adalah mensyaratkan ketika pengiriman suatu informasi bisa untuk diidentifikasi dengan baik dan benar serta memberikan jaminan bahwa identitas yang diperoleh adalah benar atau tidak palsu.

d. *Availability*
Availability berkaitan dengan ketersediaan informasi. Sesuatu yang dibutuhkan oleh pengguna layanan teknologi informasi haruslah dapat dipenuhi sehingga tingkat ketersediaan informasi dapat.

Selain itu, Menurut John D. Howard dalam bukunya "*An Analysis of security incidents on the internet*" menyatakan bahwa: "Keamanan komputer adalah tindakan pencegahan dari serangan pengguna komputer atau pengakses jaringan yang tidak bertanggung jawab." [4].

2.3. Pengujian Penetrasi (*Penetration Testing*)

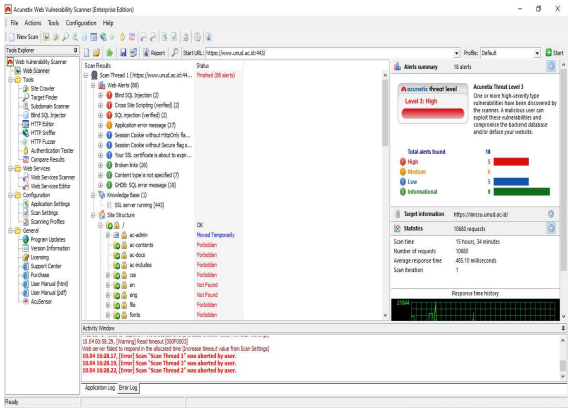
Menurut Georgia Weidman, pengujian penetrasi, atau *pentesting* melibatkan simulasi serangan nyata untuk menilai risiko yang terkait dengan potensi pelanggaran keamanan [5]. Pada *pentest* (sebagai lawan untuk penilaian kerentanan), penguji tidak hanya menemukan kerentanan yang dapat digunakan oleh penyerang tetapi juga mengeksploitasi kerentanan, jika memungkinkan, untuk menilai apa yang mungkin diperoleh penyerang setelah sukses eksploitasi. Tujuan melakukan pengujian penetrasi adalah untuk menemukan kerentanan yang memungkinkan dapat diserang oleh pihak yang tidak memiliki hak akses terhadap suatu sistem. Dengan memperoleh hasil dari pengujian penetrasi sistem dapat dilakukan evaluasi sehingga dapat digunakan untuk meningkatkan keamanan sistem [6].

3. Hasil dan Pembahasan

Pada bagian ini berisikan tentang hasil dari proses *scanning website* Universitas Udayana untuk menguji kerentanan yang ada dan pembahasan mengenai kerentanan yang terdeteksi oleh *tool Acunetix Web Vulnerability Scanner*.

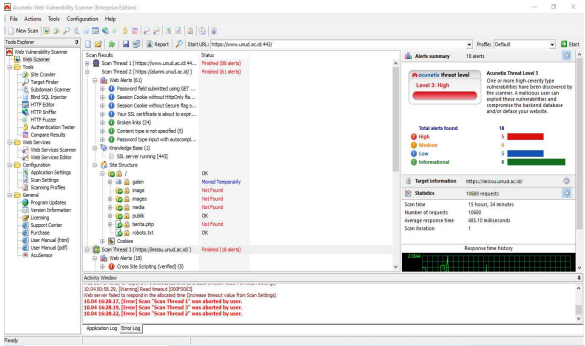
3.1. Vulnerability Scanning

Bagian ini memuat pembahasan dari data hasil penelitian dari kerentanan sistem informasi di Universitas Udayana.Pengujian yang dilakukan penulis adalah *vulnerability scanning*. Untuk melakukan *vulnerability scanning* menggunakan *tool acunetix web vulnerability scanner* untuk mengetahui celah yang terdapat pada sistem informasi di Universitas Udayana.



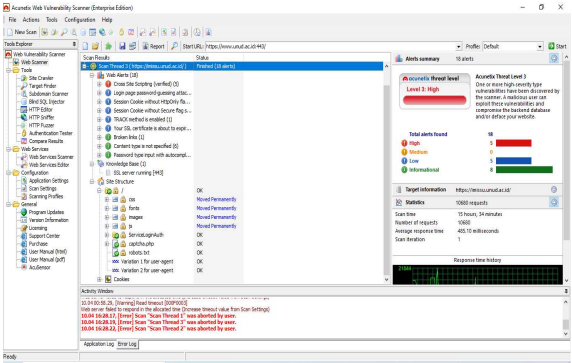
Gambar 2. Hasil *Scanning* dengan *Tool Acunetix Web Vulnerability Scanner* pada https://unud.ac.id

Pada gambar 2 merupakan hasil scanning yang dilakukan pada *https://unud.ac.id* yang terdapat suatu kerentanan *level high* yaitu *SQL Injection* dan *Cross Site Scripting (XSS)*. Selain itu, terdapat kerentanan *level medium* seperti *Application error message* dan kerentanan *level low* yaitu *session cookie without httpOnly flag set* dan *session cookie without secure flag set*. Selanjutnya hasil *scanning* gambar 3 akan ditampilkan pada gambar di bawah ini.



Gambar 3. Hasil *Scanning* dengan *Tool Acunetix Web Vulnerability Scanner* pada https://alumni.unud.ac.id

Pada gambar 3 hasil scanning pada *https://alumni.unud.ac.id* yang terdapat kerentanan *level low* yaitu *password field submitted,SSL certificate is about to expired, session cookie without httpOnly flag set* dan *session cookie without secure flag set*. Selanjutnya hasil *scanning* gambar 4 akan ditampilkan dibawah ini.



Gambar 4. Hasil *Scanning* dengan *Tool Acunetix Web Vulnerability Scanner* pada https://imissu.unud.ac.id

Pada gambar 4 hasil *scanning* menunjukan terdapat kerentanan *level high* yaitu *cross site scripting (XSS)*. Selain itu terdapat kerentanan *level low* yaitu *login page password- quession attack*, *SSL certificate is about to expired*, *session cookie without httpOnly flag set*, *session cookie without secure flag set*, dan *TRACK method is enable*. Berdasarkan hasil scanning yang telah dilakukan dapat diklasifikasikan tingkatan kerentanan yang terdapat pada website Universitas Udayana pada tabel di bawah.

Tabel 1. Klasifikasi Tingkat Kerentanan yang Ditemukan

Tingkat Kerentanan	Keterangan
Level High	*SQL Injection* *Cross Site Scripting (XSS)*
Level Medium	*Application error message*
Level Low	*login page password- quession attack* *SSL certificate is about to expired* *Session cookie without httpOnly flag set* *Session cookie without secure flag set* *TRACK method is enabled*

4. Kesimpulan

Berdasarkan hasil penelitian yang dilakukan dapat disimpulkan bahwa suatu informasi tidak ada yang sempurna. Seorang *web programmer* maupun *server administrator* hanya bisa melakukan pengurangan resiko-resiko kerentanan sistem semaksimal mungkin. Hasil pengujian menunjukan

terdapat kerentanan sistem dengan tiga tingkatan yaitu *level high*, *level medium*, dan *level low*. Pengujian penetrasi dapat dilakukan kembali setelah melakukan evaluasi terhadap kerentanan sistem dengan tujuan untuk meningkatkan keamanan sistem informasi Universitas Udayana.

Daftar Pustaka

[1] Sahren, Dalimuthe, R. A., & Amin, M. (2019). Penetration Testing Untuk Deteksi Vulnerability Sistem Informasi Kampus, 994–1001.
[2] Al Fajar, F. (2020). Analisis Keamanan aplikasi web Prodi Teknik Informatika Uika Menggunakan ACUNETIX web vulnerability. INOVA-TIF, 3(2), 110. https://doi.org/10.32832/inova-tif.v3i2.4127
[3] Retna Mulya, B. W., & Tarigan, A. (2018). Pemeringkatan Risiko Keamanan Sistem Jaringan Komputer Politeknik Kota Malang Menggunakan CVSS Dan Fmea. ILKOM Jurnal Ilmiah, 10(2), 190–200. https://doi.org/10.33096/ilkom.v10i2.311.190-200
[4] Sirait, F., & Putra, S. (2018). Implementasi Metode Vulnerability Dan Hardening Pada Sistem Keamanan Jaringan, 9.
[5] Weidman, G. (2014). Penetration testing: A hands-on introduction to hacking. No Starch Press.
[6] Kamilah, I., Ritzkal, & Hendrawan, A. H. (2019). Analisis Keamanan Vulnerability Pada Server Absensi Kehadiran Laboratorium di Program Studi Teknik Informatika.