



Aplikasi Identifikasi Gaya Bahasa Sarkasme Dalam Lirik Lagu Berbasis Mobile Menggunakan Support Vector Machine Algoritma

Julio Joseph Victor Masengi, Rycko Giovanni Leon Frans, Semmy Wellem Taju*

Fakultas Ilmu Komputer, Sistem Informasi, Universitas Klabat, Manado

Jl. Arnold Mononutu, Airmadidi, Minahasa Utara, Manado, Sulawesi Utara, Indonesia

Email: ¹s11910140@student.unklab.ac.id, ²s21910233@student.unklab.ac.id, ^{3,*}semmy@unklab.ac.id

Email Penulis Korespondensi: semmy@unklab.ac.id

Submitted: 20/10/2024; Accepted: 14/04/2025; Published: 15/04/2025

Abstrak—Di era digital dan maraknya penggunaan media sosial, gaya bahasa sarkasme dalam lirik lagu sering kali menjadi tantangan tersendiri dalam proses interpretasi. Salah satu masalah yaitu sulitnya mendeteksi gaya bahasa sarkasme karena sifatnya yang implisit dan bergantung pada konteks. Metode konvensional sering kali gagal dalam menangkap pola bahasa yang kompleks ini, sehingga berpotensi menimbulkan kesalahpahaman. Penelitian ini bertujuan untuk mengembangkan sebuah aplikasi berbasis mobile yang mampu mengidentifikasi gaya bahasa sarkasme dalam lirik lagu menggunakan algoritma Support Vector Machine (SVM). Aplikasi ini dirancang untuk mendeteksi unsur sarkasme yang terkandung dalam lirik lagu, yang sering kali sulit diidentifikasi secara akurat melalui metode konvensional. Pengembangan aplikasi mencakup beberapa tahapan, mulai dari pengumpulan data, pre-processing data lirik lagu, penerapan metode Term Frequency-Inverse Document Frequency (TF-IDF), hingga feature extraction. Dataset keywords sarkasme yang berisi 600 data dengan elemen sarkasme dan dataset lirik lagu umum yang berisi lirik lagu tanpa elemen sarkasme dikumpulkan dan digunakan untuk pembelajaran model machine learning. Data yang telah diproses kemudian diklasifikasikan menggunakan Support Vector Machine (SVM), yang membagi hasil analisis ke dalam dua kategori utama yaitu sarkasme dan non-sarkasme. Model klasifikasi yang diusulkan menunjukkan performa dengan Accuracy 98,14%, Sensitivity 96,13%, Specificity 100%, dan MCC 0.9645, dimana menandakan kemampuan yang baik dalam membedakan gaya bahasa sarkasme dan non-sarkasme. Penelitian ini bertujuan meningkatkan pemahaman pengguna terhadap lirik lagu, khususnya di media sosial, untuk mengurangi kesalahpahaman terkait sarkasme. Diharapkan penelitian ini dapat berkontribusi pada pengembangan teknologi untuk memahami gaya bahasa sarkasme dalam lirik lagu.

Kata Kunci: Sarkasme; Lirik Lagu; Support Vector Machine; Machine Learning

Abstract—In the digital era, with the widespread use of social media, the use of sarcasm in song lyrics often presents a unique challenge in the interpretation process. One issue is the difficulty in detecting sarcastic language due to its implicit nature and dependence on context. Conventional methods often fail to capture this complex language pattern, which may lead to misunderstandings. This research aims to develop a mobile-based application capable of identifying sarcasm in song lyrics using the Support Vector Machine (SVM) algorithm. The application is designed to detect sarcasm in song lyrics, which is often hard to identify accurately through traditional methods. The development process includes several stages, such as data collection, pre-processing song lyrics data, applying the Term Frequency-Inverse Document Frequency (TF-IDF) method, and feature extraction. A sarcasm keyword dataset containing 600 data points with sarcasm elements and a general song lyrics dataset without sarcasm elements were collected and used for machine learning model training. The processed data is then classified using Support Vector Machine (SVM), which categorizes the analysis results into two main categories: sarcasm and non-sarcasm. The proposed classification model demonstrates performance with an Accuracy of 98.14%, Sensitivity of 96.13%, Specificity of 100%, and MCC of 0.9645, indicating a strong ability to distinguish between sarcastic and non-sarcastic language. This research aims to enhance users' understanding of song lyrics, especially on social media, to reduce misunderstandings related to sarcasm. It is hoped that this research can contribute to the development of technology for understanding sarcastic language in song lyrics.

Keywords: Sarcasm; Song Lyrics; Support Vector Machine; Machine Learning

1. PENDAHULUAN

Bahasa merupakan alat komunikasi yang paling efektif untuk kebutuhan manusia dan sangat berpengaruh pada kehidupan manusia [1]. Bahasa dijadikan sebagai alat untuk berkomunikasi kepada seseorang guna menyampaikan apa dipikirkan dan dirasakan oleh seseorang kepada lawan bicaranya [2]. Bahasa juga disebut sebagai wujud berkomunikasi yang baik, sebab untuk saat ini komunikasi tulis merupakan komunikasi yang mutakhir karena hadirnya media sosial pengguna bisa berbicara dengan banyak orang tanpa wajib bertatap muka secara langsung [3], [4]. Sarkasme adalah jenis gaya bahasa yang mengandung sindiran pedas (atau disebut juga olok-olok) dan bahkan dapat menyakiti hati. Bila dibandingkan dengan gaya bahasa sejenisnya (contohnya yaitu ironi dan sinisme), dapat dikatakan bahwa gaya bahasa sarkasme ini lebih kasar [5]. Dalam era digital yang semakin berkembang seperti saat ini, media sosial sudah menjadi salah satu platform atau aplikasi banyak digunakan oleh masyarakat, karena dengan adanya media sosial, para pengguna bisa berekspresi secara bebas. Video, gambar, lagu, blog, website yang ada di dunia virtual adalah sarana yang digunakan oleh para pengguna media sosial dalam menciptakan sebuah konten [6].

Sosial media merujuk kepada platform-platform digital online yang digunakan oleh individu, organisasi bahkan pelaku bisnis untuk berinteraksi, berbagi informasi, konten, dan pemikiran dengan pengguna lainnya di seluruh dunia [7]. Media sosial biasanya dibangun di atas teknologi web dan mobile. Situs web media sosial seperti Facebook, Twitter, Instagram, LinkedIn, TikTok dan YouTube adalah platform media sosial yang paling banyak



digunakan sekarang ini dan berpotensi memiliki dampak dalam transformasi pemasaran digital [8]. Media sosial memungkinkan pengguna untuk membuat profil, berbagi konten, berkomunikasi dengan orang lain melalui pesan pribadi atau publik, bergabung dengan kelompok atau komunitas, dan berpartisipasi dalam berbagai jenis aktivitas online. Salah satu sarana dari media sosial yang bisa menjadi tempat untuk mengekspresikan diri dan pikiran adalah Musik atau Lagu [9]. Musik atau lagu memiliki peran penting dalam kehidupan manusia karena ini dapat menjadi salah satu sarana yang dapat digunakan oleh setiap individu untuk mengekspresikan diri, contohnya dengan mengungkapkan perasaan, isi batin dan bahkan gagasan pikiran juga dapat digunakan sebagai saran didik anak [10]. Setiap lirik lagu yang kita dengarkan dapat menjadi sarana untuk menyampaikan pesan, mengungkapkan emosi, maupun untuk menghibur pendengar. Namun sekarang ini, banyak pendengar lirik lagu dari musik ataupun media sosial sering mendengar video singkat dengan lagu tanpa mengetahui makna dari lirik lagu tersebut. Tidak semua orang mampu mengenali dan memahami sarkasme, sarkasme adalah gaya bahasa yang sering terlihat seperti lelucon atau sindiran pedas yang dapat menyakiti perasaan orang lain, sehingga diperlukan penyuluhan dan pengarahan bagi seluruh lapisan masyarakat untuk memahami sarkasme, dampaknya, serta penyimpangan yang bisa terjadi jika tidak ditindaklanjuti, terutama dalam konteks lirik lagu yang kerap mengandung gaya bahasa sarkasme secara tidak langsung [11]. Bisa dikatakan bahwa salah satu masalah utama adalah kesulitan dalam mendeteksi gaya bahasa sarkasme yang implisit dan kontekstual, di mana metode konvensional sering gagal menangkap pola bahasa yang kompleks, berisiko menimbulkan kesalahpahaman. Oleh karena itu, pemahaman yang lebih dalam terhadap makna lagu dan konteks penggunaannya sangatlah penting dalam meminimalisir potensi kesalahpahaman atau interpretasi yang salah. Peneliti menganggap bahwa perlu adanya sebuah aplikasi yang dapat membantu pendengar lirik lagu maupun pengguna media sosial dalam mengidentifikasi gaya bahasa sarkasme dalam lirik lagu yang mereka dengarkan melalui media player.

Dengan adanya penelitian ini, para pendengar lagu maupun pengguna media sosial akan dimudahkan dalam memahami gaya bahasa dari sebuah lagu sebelum membagikannya, terutama ketika lagu tersebut mengandung unsur sarkasme. Dalam hal ini, keberadaan dari aplikasi pendeteksi ini memungkinkan pengguna untuk lebih mudah mengidentifikasi dan memahami gaya bahasa sarkasme yang terdapat dalam lirik lagu. Peneliti sangat berharap kiranya sebagai hasil dari penelitian ini, diharapkan akan tercipta kesadaran akan pentingnya komunikasi publik [12] di antara para pengguna media sosial akan pentingnya memahami konteks dan makna di balik lagu yang akan dibagikan pada media sosial, sehingga dapat meminimalisir kemungkinan kesalahpahaman atau interpretasi yang salah. Pada penelitian sebelumnya [13] [14] sudah pernah ada peneliti yang menganalisis dampak penggunaan gaya bahasa sarkasme di beberapa platform media sosial seperti TikTok [3], Twitter [15], Facebook [16], Youtube [17] dan Instagram [4] menggunakan metode machine learning dengan beberapa algoritma yang ada seperti Naïve Bayes, K Nearest Neighbours, Decision Tree, Random Forest dan SVM [13]. Penelitian ini memberikan kontribusi baru di mana akan dirancang sebuah aplikasi untuk mendeteksi gaya bahasa sarkasme dengan data yang digunakan adalah lirik lagu dan sebelumnya belum ada yang menggunakan platform berbasis mobile.

Berdasarkan penelitian yang dilakukan oleh Sarli dkk. (2023), dimana peneliti mencoba mengidentifikasi masalah penggunaan bahasa sarkasme oleh pengguna media sosial TikTok yang dapat menimbulkan kesalahpahaman [3], solusi yang ditawarkan adalah mendeskripsikan bentuk dan makna sarkasme untuk meningkatkan kebijaksanaan berbahasa, hasil akhirnya menemukan berbagai bentuk dan perubahan makna sarkasme, sementara gap yang ditawarkan adalah pentingnya kesadaran akan konteks dan pengaruhnya terhadap komunikasi efektif di platform sosial. Selanjutnya, Akib dan Dahlan [15] dalam penelitiannya mendeskripsikan bentuk, fungsi dan makna sarkasme di Twitter, dengan solusi melalui analisis deskriptif kualitatif terhadap 50 sampel postingan, menghasilkan temuan bahwa bentuk sarkasme yang dominan adalah sarkasme sifat, fungsi penyampaian pendapat dan makna pengerasan, serta gap yang ditawarkan adalah kurangnya pemahaman dan penerapan analisis sarkasme yang mendalam pada platform media sosial. Sama halnya dengan Paramita dkk [16], mereka mencoba mengidentifikasi penggunaan gaya bahasa sarkasme di Facebook yang sering kali mengarah pada penyampaian pendapat, penolakan dan bentuk komunikasi lainnya, dengan solusi berupa pemahaman yang lebih bijak dalam menggunakan bahasa di media sosial, menghasilkan berbagai bentuk sarkasme yang dapat menyebabkan penyimpangan makna seperti pengkasaran, penyempitan, dan perluasan makna, serta menawarkan gap dalam hal peningkatan pemahaman sarkasme di kalangan pengguna media sosial. Terakhir, penelitian yang dilakukan oleh Oktaviani dkk [17] dalam menganalisis penggunaan berbagai bentuk sarkasme di channel YouTube Deddy Corbuzier untuk menyampaikan kritik sosial, menemukan 10 bentuk sarkasme yang meliputi sifat, tindakan, hasil tindakan, himbauan dan sebutan, dengan gap dalam pemahaman penggunaan sarkasme di media sosial.

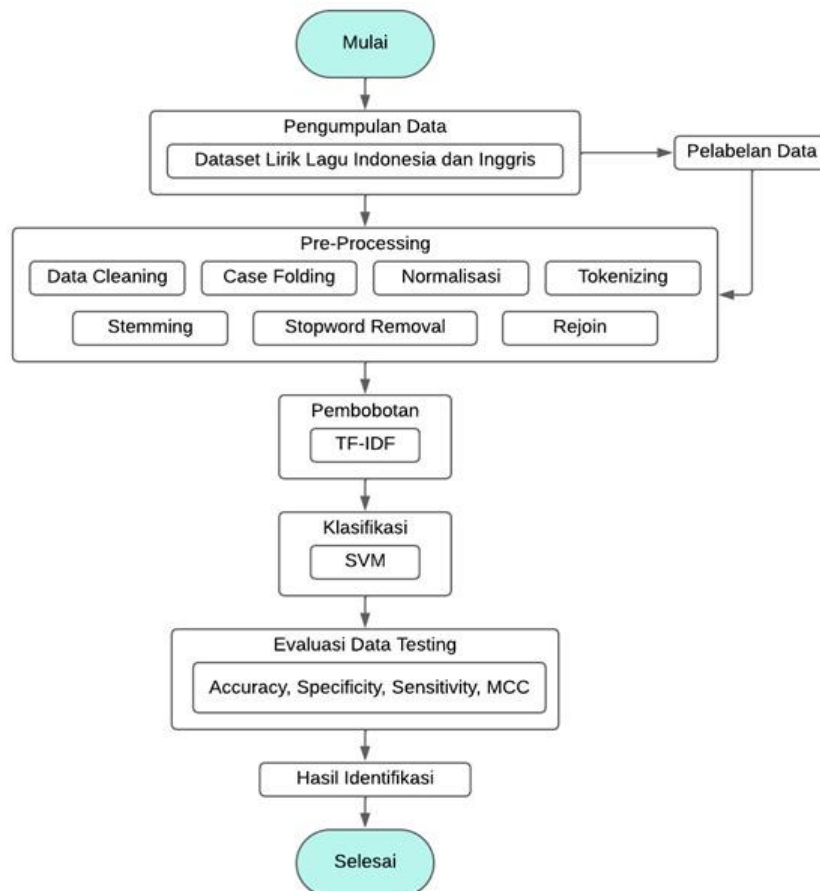
Dalam penelitian ini, dirancang sebuah aplikasi berbasis mobile untuk mengidentifikasi gaya bahasa sarkasme menggunakan algoritma Machine Learning yaitu Support Vector Machine (SVM) [18], dengan pendekatan analisis sentimen di mana terdapat proses untuk preprocessing data dan terdapat dua label kategori data yang digunakan yaitu 1 untuk lirik yang mengandung unsur sarkasme dan 0 untuk lirik yang tidak mengandung unsur sarkasme (non-sarkasme). Data yang digunakan dalam penelitian ini berupa lirik lagu dalam bentuk teks. Beberapa tahapan pemrosesan data dilakukan secara bertahap, mulai dari preprocessing untuk membersihkan dan mempersiapkan data, penggunaan metode TF-IDF [19] untuk menghitung bobot kata yang diinputkan, penerapan lexicon-based feature ini untuk mengekstraksi fitur-fitur penting dari kata-kata dalam data, sampai pada bagian

yang terakhir yaitu algoritma klasifikasi dengan menggunakan SVM algoritma [20]. Setelah keseluruhan proses pengolahan data selesai, hasil klasifikasi (output prediction) akan ditampilkan dalam aplikasi dengan dua kelas, yaitu kelas/label sarkasme dan kelas non-sarkasme.

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Di dalam tahapan penelitian pada Gambar 1, terdapat beberapa tahap yang akan dibuat oleh peneliti. Penelitian ini menarget untuk merancang aplikasi yang mampu mengidentifikasi gaya bahasa sarkasme dalam lirik lagu dengan menggunakan algoritma SVM yang di machine learning khusus untuk proses klasifikasi. Proses diawali dengan pengumpulan dataset lirik lagu berbahasa Indonesia dan Inggris, yang kemudian melalui proses pelabelan untuk memisahkan kategori lirik berdasarkan karakteristik bahasa yaitu kelas/label sarkasme dan kelas non-sarkasme. Tahapan pre-processing diterapkan untuk memastikan data bersih dan siap diolah, mencakup beberapa langkah penting seperti pembersihan data, case folding, normalisasi, tokenisasi, stemming, penghapusan kata-kata yang tidak relevan (stopword removal), dan penggabungan kembali kata-kata (rejoin). Setelah itu, data diberikan bobot menggunakan metode Term Frequency-Inverse Document Frequency (TF-IDF) yang berfungsi untuk menekankan pentingnya kata-kata tertentu dalam lirik lagu. Pada tahap klasifikasi, algoritma SVM digunakan untuk membedakan antara lirik yang mengandung sarkasme dan yang tidak mengandung sarkasme. Evaluasi kinerja model dengan menentukan jumlah positif data dan negatif data yang diuji dalam metode Confusion Matrix [21], metode ini dilakukan melalui pengukuran beberapa metrik seperti accuracy (akurasi), specificity (spesifisitas), sensitivity (sensitivitas) dan Matthew's Correlation Coefficient (MCC) untuk memberikan gambaran menyeluruh tentang efektivitas model machine learning yang dibuat. Bagian yang terakhir pada penelitian ini berupa hasil penelitian. Diharapkan bahwa pendekatan menggunakan machine learning algoritma dengan kombinasi pre-processing yang cermat serta feature extraction yaitu TF-IDF, memungkinkan identifikasi sarkasme dalam lirik lagu secara akurat, sehingga memberikan kontribusi signifikan dalam analisis teks mengandung sarkasme pada lirik lagu.



Gambar 1. Alur Penelitian Untuk Mengidentifikasi Gaya Bahasa Sarkasme

2.2 Pengumpulan Data

Dalam pembuatan model SVM dari machine learning digunakan dua dataset sebagai data pembelajaran untuk machine learning, yaitu dataset keyword bahasa inggris dan bahasa indonesia yang mengandung gaya bahasa



sarkasme (positif data) dan dataset lirik lagu umum (negatif data). Peneliti memahami bahwa untuk membuat suatu model machine learning, data merupakan hal yang sangat diperlukan sebagai bahan pembelajaran sehingga pada penelitian ini, peneliti mengumpulkan 2 dataset sebagai bahan pembelajaran machine learning, yaitu dataset keywords sarkasme dan dataset lirik lagu umum. (1) Dataset keywords sarkasme adalah dataset yang berisi kata-kata atau keywords yang mengandung gaya bahasa sarkasme. Pengumpulan data dilakukan secara manual sebanyak 600 data, yaitu sumber data online sebagai referensi dalam pembuatan dataset. (2) Dataset lirik lagu umum adalah dataset yang berisi lirik lagu. Dataset ini dikumpulkan secara manual melalui sumber-sumber online yang terpercaya. Kedua dataset ini menjadi bahan pembelajaran bagi SVM algoritma agar dapat memahami konteks gaya bahasa sarkasme dan bukan gaya bahasa sarkasme dalam lirik lagu.

2.3 Pelabelan Data

Dalam penelitian ini, proses pelabelan data dilakukan dengan menggunakan bahasa pemrograman Python. Dataset lirik lagu yang digunakan dalam sistem ini dilabeli secara otomatis ke dalam dua kategori, yaitu 1 untuk lirik yang mengandung unsur sarkasme dan 0 untuk lirik yang tidak mengandung unsur sarkasme (non-sarkasme). Pendekatan pelabelan ini bertujuan untuk memfasilitasi proses klasifikasi pada tahap selanjutnya, di mana algoritma Support Vector Machine (SVM) akan digunakan untuk membedakan kedua kategori tersebut berdasarkan fitur-fitur yang telah diekstraksi dari lirik lagu.

2.4 Preprocessing Data

Pre-processing data merupakan tahapan selanjutnya yang akan dilalui dalam memproses dataset sesudah proses pelabelan. Pada penelitian tentang pengembangan aplikasi identifikasi gaya bahasa sarkasme dalam lirik lagu ini, peneliti akan dilakukan pre-processing data dengan beberapa tahapan yaitu pertama-tama (1) Tahapan Data Cleaning merupakan proses untuk membersihkan data dari angka dan karakter-karakter yang tidak ada hubungannya dengan informasi yang ada pada dokumen contohnya (0-9, !@#%&* _+={ }[];:;''/?<>.). Selanjutnya, (2) Tahapan Case Folding digunakan untuk mengubah seluruh huruf yang ada pada data menjadi huruf kecil. (3) Tahapan Normalisasi adalah proses dimana kata-kata yang tidak baku menjadi baku sesuai dengan atauran penulisan pada Kamus Besar Bahasa Indonesia. Contoh tahapan ini yaitu 'jangan kau bilang' menjadi 'jangan kamu bilang'. (4) Tahapan Tokenisasi adalah tahapan dimana sebuah kalimat akan dipisahkan menjadi kata-kata. Contoh tahapan ini yaitu 'jangan buang sampah sembarangan' menjadi ['jangan', 'buang', 'sampah', 'sembarangan']. (5) Tahapan Stopword Removal adalah metode yang digunakan untuk menghapus kata yang dianggap tidak terlalu penting untuk di jadikan kata kunci dalam melakukan klasifikasi. (6) Tahapan Stemming adalah proses dimana akan dihilangkan kata awalan dan kata imbuhan, dan menjadi kata dasar seperti (be-, di-, ke-, -ku, -mu, atau -nya) kata tersebut akan dihilangkan dan diubah menjadi kata dasarnya. Contoh tahapan ini yaitu 'berikan', 'cintaku' menjadi 'beri', 'cinta'. Terakhir, (7) Tahapan Rejoin adalah tahapan terakhir dalam pre-processing data dimana hasil Stemming yang sebelumnya masih dalam bentuk token akan digabungkan kembali menjadi kalimat utuh. Contoh tahapan ini yaitu 'beri', 'cinta' menjadi 'beri cinta'.

2.5 Metode TF-IDF (Term Frequency-Inverse Document Frequency)

Metode TF-IDF (Term Frequency-Inverse Document Frequency) [19] digunakan dalam penelitian ini untuk memberikan bobot yang lebih tinggi pada kata-kata dalam lirik lagu berdasarkan frekuensi kemunculan dalam dokumen lirik lagu serta memperhitungkan seberapa jarang kata tersebut muncul di dokumen lainnya dalam corpus. Pada penelitian ini, pendekatan TF-IDF diadopsi untuk mengidentifikasi kata-kata yang memiliki relevansi lebih besar dalam mendeteksi unsur sarkasme dalam lirik lagu. Semakin sering kata muncul dalam lirik yang dianalisis dan semakin jarang ditemukan di lirik lain, semakin tinggi bobot yang diberikan oleh metode TF-IDF, yang pada akhirnya berkontribusi dalam meningkatkan efektivitas analisis dan klasifikasi sarkasme pada aplikasi yang akan dibuat nantinya.

2.6 Pemodelan Machine Learning

Algoritma Support Vector Machine (SVM) digunakan sebagai metode utama dalam proses klasifikasi lirik lagu untuk membedakan/mengklasifikasikan antara dua kelas, yaitu sarkasme dan non-sarkasme. Algoritma SVM digunakan untuk mempelajari pola-pola yang terkandung dalam lirik lagu setelah melalui tahap pre-processing, yang melibatkan pembersihan dan ekstraksi fitur dengan metode TF-IDF dengan pendekatan lexicon-based. Melalui proses training, algoritma SVM membangun sebuah model klasifikasi yang mampu mengidentifikasi secara akurat apakah suatu lirik mengandung unsur sarkasme atau tidak mengandung unsur sarkasme. Pemanfaatan algoritma SVM memungkinkan aplikasi berbasis mobile untuk menghasilkan model prediktif yang baik, dengan kemampuan generalisasi yang tinggi dalam mendeteksi sarkasme pada teks input lirik lagu yang akan diinputkan oleh pengguna.

2.7 Evaluasi Data Testing

Pada penelitian ini, data testing akan dievaluasi dengan menggunakan Confusion Matrix [21]. Confusion Matrix digunakan sebagai alat evaluasi performa metode klasifikasi yang diterapkan. Dalam mengevaluasi model pada

aplikasi pendeteksi sarkasme ini, performa diukur menggunakan empat metrik berbeda untuk menilai kemampuan model dalam mengidentifikasi gaya bahasa sarkasme dalam lirik lagu, yaitu Sensitivity (Sens), Specificity (Spec), Accuracy (Acc) dan Matthews Correlation Coefficient (MCC). Metrik-metrik evaluasi ini memberikan gambaran yang umum mengenai efektivitas model dari algoritma SVM, mulai dari kemampuan aplikasi dalam mendeteksi sarkasme dengan benar hingga tingkat keseimbangan klasifikasi secara keseluruhan.

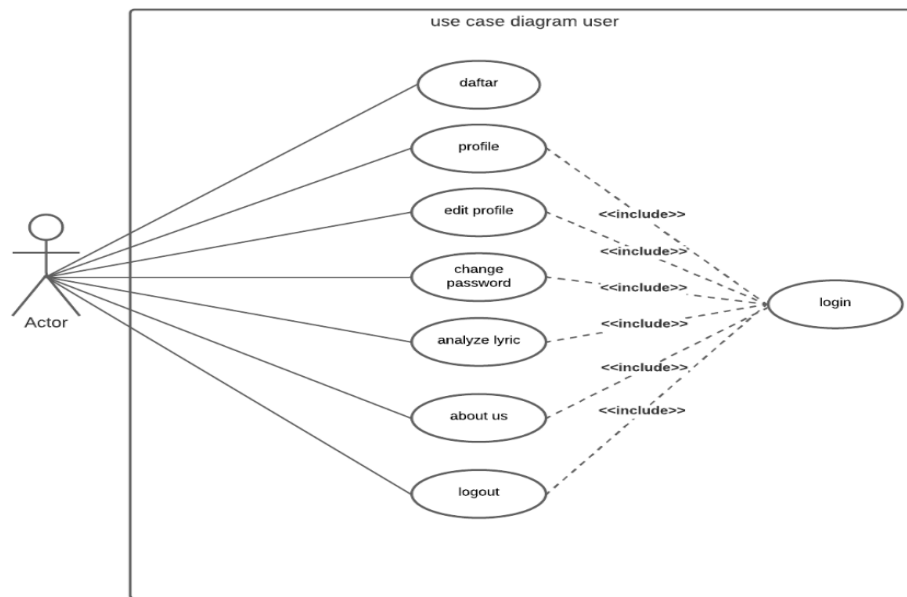
$$\text{Sensitivity} = \left(\frac{TP}{TP+FN} \right) \times 100\% \tag{1}$$

$$\text{Specificity} = \left(\frac{TN}{TN+FP} \right) \times 100\% \tag{2}$$

$$\text{Accuracy} = \left(\frac{TP+TN}{TP+FP+TN+FN} \right) \times 100\% \tag{3}$$

$$\text{MCC} = \left(\frac{(TP \times TN) - (FP \times FN)}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}} \right) \tag{4}$$

Dari empat metrik evaluasi yang digunakan dalam penelitian ini, Sensitivity digunakan untuk mengukur kemampuan model SVM untuk mendeteksi dengan benar data positif yaitu lirik lagu yang mengandung sarkasme. Sedangkan, Specificity (Spec) digunakan untuk mengukur kemampuan model SVM untuk mendeteksi dengan benar data negatif yaitu lirik lagu yang tidak mengandung sarkasme. Selanjutnya, Accuracy (Acc) digunakan untuk mengukur proporsi prediksi yang benar dalam keseluruhan dataset, dimana ini dapat memberikan gambaran umum tentang performa dari model SVM dalam mengidentifikasi lirik lagu secara akurat. Sementara itu, metrik MCC akan memberikan skor antara -1 hingga +1, dengan nilai positif menunjukkan performa dari model SVM yang sangat baik dalam membedakan sarkasme dan non-sarkasme dalam lirik lagu.



Gambar 2. Use Case Diagram dari Aplikasi Pendeteksi Sarkasme

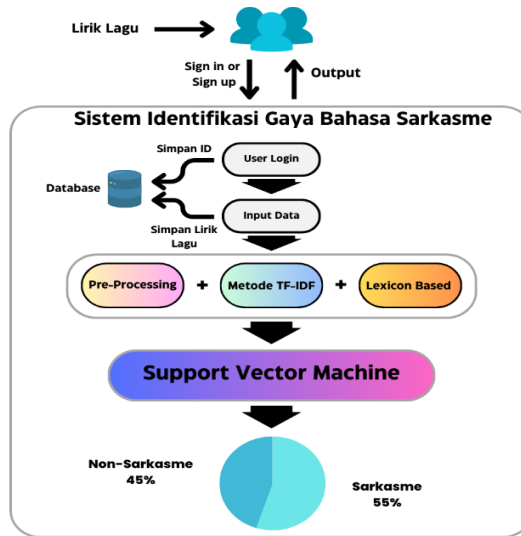
2.8 Use Case Diagram

Pada Gambar 2, user case diagram dari aplikasi pendeteksi sarkasme yang di mana akan menggambarkan setiap kejadian yang dapat dilakukan oleh pengguna aplikasi. User akan masuk pada aplikasi dan harus terlebih dahulu untuk mendaftar dengan cara membuat akun. Pertama user akan membuat akun sebagai user di aplikasi tersebut, setelah berhasil membuat akun user dapat login pada aplikasi, setelah berhasil masuk ke dalam aplikasi user akan langsung masuk pada menu utama aplikasi dan pada menu utama tersebut user dapat menggunakan fitur - fitur di dalam aplikasi tersebut seperti, profile yang dimana pada fitur tersebut berisi data data dari user, lalu ada fitur history yang akan menampilkan hasil analisa yang user gunakan, lalu ada fitur edit profile fitur ini gunanya untuk mengganti data data yang dimiliki oleh user, lalu ada fitur change password fitur ini berfungsi untuk mengganti password, ada fitur Analyze lyric fitur ini adalah fitur utama pada aplikasi ini yakni digunakan untuk menganalisa lirik lagu, lalu ada fitur About us fitur ini berisi informasi mengenai pengembang aplikasi dan ada fitur Logout untuk keluar dari aplikasi.

2.9 Metode Penelitian & Arsitektur Sistem

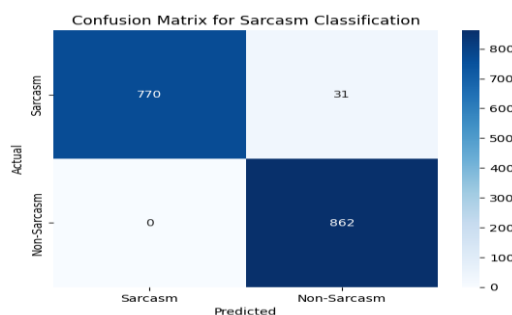
Metodologi penelitian yang kami usulkan dalam pengembangan sistem ini adalah Software Development Life Cycle (SDLC), dimana kerangka kerja ini akan mendeskripsikan proses pengembangan aplikasi identifikasi gaya bahasa sarkasme pada lirik lagu secara terstruktur dan sistematis. Peneliti menggunakan SDLC untuk memastikan

bahwa pengembangan aplikasi pendeteksi gaya bahasa sarkasme mengikuti langkah-langkah yang terdefinisi dengan baik, mulai dari analisis kebutuhan (requirements analysis), desain (design), implementasi (implementation), pengujian (testing) dan pemeliharaan (maintenance). Penerapan SDLC menjadi sangat penting karena memberikan panduan yang jelas dalam proses pengembangan aplikasi pengidentifikasi gaya bahasa sarkasme berbasis mobile. Dengan mengikuti tahapan SDLC, peneliti dapat memastikan bahwa setiap langkah dalam siklus pengembangan diperhatikan secara menyeluruh, sehingga risiko error dalam pengembangan aplikasi akan semakin berkurang. Selain itu, penerapan SDLC juga membantu memastikan bahwa aplikasi yang dihasilkan dapat memenuhi kebutuhan users secara optimal dan memiliki kualitas yang tinggi. Model SDLC yang digunakan dalam penelitian ini adalah model Waterfall, yang memungkinkan proses pengembangan berlangsung secara linier, dimana setiap tahapan dalam pengembangan aplikasi pendeteksi gaya bahasa sarkasme harus diselesaikan sebelum beralih ke tahapan proses selanjutnya.



Gambar 3. Arsitektur Sistem Identifikasi Gaya Bahasa Sarkasme

Seperti yang dilihat pada Gambar 3, kerangka konseptual dari aplikasi identifikasi gaya bahasa sarkasme dalam lirik lagu menggunakan algoritma Support Vector Machine (SVM). Keseluruhan proses dari aplikasi ini akan diawali dengan pengguna yang melakukan sign up atau sign in pada halaman login untuk mengakses aplikasi. Setelah proses autentikasi berhasil, data pengguna akan otomatis disimpan dalam database, bersama dengan lirik lagu yang diinputkan secara manual oleh pengguna. Setiap lirik lagu yang dimasukkan juga disimpan dalam database untuk dianalisis lebih lanjut. Tahapan analisis dimulai dengan proses pre-processing, di mana lirik lagu mengalami sejumlah transformasi untuk membersihkan data, seperti penghapusan tanda baca dan kata-kata yang tidak relevan. Selanjutnya, sistem menggunakan metode TF-IDF (Term Frequency-Inverse Document Frequency) untuk menghitung bobot kata-kata yang ada dalam lirik, guna mengidentifikasi kata-kata yang memiliki signifikansi lebih tinggi dalam konteks sarkasme. Di samping itu, sistem juga menerapkan pendekatan berbasis lexicon, yang menggunakan daftar kata-kata tertentu yang berpotensi mengandung unsur sarkasme. Setelah tahap pengolahan selesai, data yang telah diproses kemudian masuk ke tahap klasifikasi menggunakan algoritma Support Vector Machine. SVM bekerja dengan memisahkan data lirik menjadi dua kelas utama: sarkasme dan non-sarkasme. Hasil akhir dari proses ini ditampilkan dalam bentuk persentase, yang menunjukkan proporsi lirik lagu yang mengandung unsur sarkasme dibandingkan dengan yang tidak mengandung unsur sarkasme. Sebagai contoh, output klasifikasi menunjukkan 55% dari lirik yang dianalisis mengandung sarkasme, sementara 45% sisanya termasuk dalam kategori non-sarkasme.



Gambar 4. Confusion Matrix

3. HASIL DAN PEMBAHASAN

3.1 Hasil Evaluasi Menggunakan Testing Data

Hasil evaluasi data testing menggunakan pendekatan Confusion Matrix. Gambar 4 menunjukkan perbandingan dari nilai prediksi dan dibandingkan dengan actual value dalam testing data. Berdasarkan hasil evaluasi data testing menunjukkan bahwa model Machine Learning menggunakan SVM algoritma yang dibuat memiliki kinerja yang cukup baik dalam mengklasifikasikan sarkasme dengan non-sarkasme. Model klasifikasi sarcasm yang diuji pada dataset testing menunjukkan kinerja yang baik, dengan Accuracy mencapai 98.14%. Ini menandakan bahwa model ini memiliki kemampuan yang baik dalam mengenali dan membedakan antara gaya bahasa yang bersifat sarcasm dan yang tidak. Lebih lanjut, Sensitivity model sebesar 96.13% menegaskan bahwa hampir semua gaya bahasa sarcasm dapat terdeteksi dengan sedikit yang terlewat. Sedangkan, performa sempurna dalam Specificity yaitu 100%, menunjukkan bahwa model tidak pernah keliru dalam mengidentifikasi gaya bahasa non-sarcasm, sehingga tidak ada komentar yang salah diklasifikasikan sebagai sarcasm. Dengan nilai Matthews Correlation Coefficient (MCC) sebesar 0.9645, ini menunjukkan bahwa hasil dari prediksi model SVM sangat baik. Ini juga mencerminkan korelasi yang tinggi antara nilai hasil prediksi dan nilai aktual dengan berkurangnya error prediksi (bias) terhadap salah satu kelas. Hasil evaluasi data testing menggunakan SVM algoritma dapat dilihat pada Gambar 4.

3.2 Implementasi Aplikasi Menggunakan React Native Mobile App

Aplikasi yang dibangun berbasis mobile menggunakan React Native [22] dengan beberapa fitur yang tersedia. Fitur utama aplikasi ini adalah identifikasi lirik lagu yang dimana user dapat mengisi teks lirik lagu kedalam kolom input. Setelah mengisi teks dan menekan tombol “analyze”, maka output akan berupa klasifikasi sarkasme dan non-sarkasme dari teks yang dimasukkan serta menampilkan presentase hasil identifikasi. Tampilan dari aplikasi dapat dilihat pada Gambar 5.



Gambar 5. Tampilan Aplikasi untuk Splash Screen, Login Screen, SignUp Screen

Gambar 5(1) merupakan user interface splash screen dari aplikasi sarcasm identifier. Tampilan ini akan muncul ketika user akan membuka aplikasi. Gambar 5(2) merupakan tampilan user interface dari login aplikasi sarcasm identifier. Pada tampilan ini user akan memasukan email dan password yang sudah terdaftar untuk masuk pada tampilan menu utama aplikasi. Gambar 5(3) merupakan tampilan user interface dari create account aplikasi sarcasm identifier. Pada tampilan ini user akan diminta untuk mengisi data diri untuk proses membuat akun, data yang harus dimasukan oleh user yaitu email, nama lengkap, password, re password dan phone number. Gambar 5(4) adalah contoh gambar user interface dari user telah berhasil login ke dalam aplikasi, dimana user akan mendapatkan notifikasi seperti gambar diatas.



Gambar 6. Tampilan Aplikasi untuk Menu Utama, Input Lirik Lagu, Profile Screen dan Prediksi



Gambar 6(1) merupakan tampilan user interface dari menu utama setelah user berhasil login. Pada tampilan menu utama terdapat fitur-fitur seperti nama pengguna, email, profile, history, about us, get started untuk memulai analisa lirik lagu, dan tombol keluar. Gambar 6(2) diatas merupakan tampilan user interface untuk menganalisa lirik lagu. Pada tampilan tersebut user diminta untuk memasukkan judul lagu dan lirik lagu yang ingin dianalisa. Batas karakter yang ditentukan untuk dianalisa sebanyak 150 karakter. Gambar 6(3) merupakan tampilan user interface fitur edit profile user. Pada fitur ini user dapat mengganti data yang sudah di daftarkan sebelumnya, data data yang dapat di edit oleh users adalah nama, email, dan phone number. Pada aplikasi ini juga terdapat tampilan yang menunjukkan fitur about us. Fitur ini akan menunjukkan mengenai informasi yang terdapat pada aplikasi tersebut. Gambar 6(4) menunjukkan tampilan ketika user menekan tombol “ANALYZE”, maka akan muncul alert hasil identifikasi sarkasme dalam teks lirik lagu yang diinput pada input field lyrics.

Tabel 1. Hasil Pengujian Fitur Aplikasi Identifikasi Sarkasme

No.	Fitur	Output	Hasil
1	Sign Up	Akun baru dibuat	Berhasil
2	Sign in	Masuk pada menu utama aplikasi	Berhasil
3	Input Lirik	Memasukkan lirik lagu	Berhasil
4	Analyze Lyric	Menganalisa lirik lagu	Berhasil
5	Profile	Melihat informasi akun	Berhasil
6	Change Password	Mengganti kata sandi akun	Berhasil
7	Edit Profile	Mengganti informasi akun	Berhasil
8	Password	Mengganti kata sandi akun	Berhasil
9	About-us	Melihat informasi aplikasi	Berhasil
10	Keluar	Keluar dari aplikasi	Berhasil

3.3 Hasil Pengujian Fitur Aplikasi

Aplikasi yang dibangun memiliki beberapa fitur untuk user experience selama menggunakan aplikasi pendeteksi gaya bahasa sarkasme. Fitur tersebut dapat dilihat dari Tabel 1, di mana tabel tersebut menampilkan hasil application feature testing (pengujian fitur aplikasi) yang dirancang untuk mengidentifikasi gaya bahasa sarkasme dalam lirik lagu menggunakan algoritma Support Vector Machine (SVM). Pengujian dilakukan pada sepuluh fitur utama aplikasi, termasuk fitur pendaftaran akun (Sign Up), masuk ke aplikasi (Sign In), penginputan lirik, analisis lirik dan pengelolaan profil pengguna. Setiap fitur menghasilkan keluaran yang sesuai dengan harapan, seperti pembuatan akun baru, memasukkan lirik lagu, dan menganalisis lirik lagu. Hasil pengujian dari aplikasi ini menunjukkan bahwa seluruh fitur yang diuji, termasuk perubahan kata sandi, pengeditan profil dan informasi aplikasi, berfungsi dengan baik, ditunjukkan dengan status "Berhasil" pada setiap fitur. Dari hasil pengujian ini disimpulkan bahwa aplikasi telah memenuhi tujuan fungsional dalam mendukung identifikasi sarkasme pada lirik lagu.

4. KESIMPULAN

Dalam melakukan pengembangan aplikasi pengidentifikasi gaya bahasa sarkasme dalam lirik lagu ini, peneliti menggunakan Support Vector Machine (SVM) Algoritma. Dari hasil yang sudah didapatkan maka diambil beberapa kesimpulan. Pertama, aplikasi pengidentifikasi gaya bahasa sarkasme menggunakan SVM Algoritma sudah dapat digunakan untuk menganalisis lirik lagu yang mengandung sarkasme. Pengembangan aplikasi pendeteksi sarkasme ini juga dapat menjadi langkah positif untuk membantu pengguna atau dalam hal ini yaitu pendengar lagu dalam mengenali gaya bahasa sarkasme yang ada di dalam lagu tersebut. Penelitian ini menunjukkan bahwa algoritma SVM efektif dalam mengidentifikasi sarkasme dengan performa yang baik, yaitu Accuracy 98,14%, Sensitivity 96,13%, Specificity 100%, dan MCC 0.9645. Hal ini juga dapat membuka peluang untuk mengaplikasikan SVM dalam berbagai konteks linguistik lainnya. Meskipun demikian, pengembangan aplikasi pendeteksi gaya bahasa sarkasme dalam lirik lagu belum sepenuhnya sempurna. Oleh karena itu, peneliti memberikan beberapa saran untuk pengembangan sistem aplikasi ini ke depannya, yaitu memperluas cakupan penelitian aplikasi dengan menambah bahasa asing lainnya agar aplikasi ini bisa digunakan oleh lebih banyak orang lagi, menambahkan fitur Song to Text, sebagai pembanding bisa menggunakan algoritma machine learning lainnya dalam mengidentifikasi gaya bahasa sarkasme, mengintegrasikan analisis sentimen untuk mengidentifikasi nuansa emosional dalam lirik yang bisa membantu dalam pemahaman sarkasme, serta disarankan juga untuk dapat bekerjasama dengan ahli linguistik dan musikologi untuk meningkatkan pemahaman terhadap berbagai bentuk sarkasme dan gaya bahasa lainnya dalam lirik lagu.

REFERENCES

- [1] O. Mailani, I. Nuraeni, S. A. Syakila, and J. Lazuardi, “Bahasa sebagai alat komunikasi dalam kehidupan manusia,” *Kampret Journal*, vol. 1, no. 2, pp. 1–10, 2022.



- [2] L. Suryaningsih, “Analisis penggunaan gaya bahasa sarkasme pada lirik lagu Mbojo,” *Ainara Journal (Jurnal Penelitian Dan PKM Bidang Ilmu Pendidikan)*, vol. 2, no. 3, pp. 274–280, 2021.
- [3] S. Sarli, N. NURHADI, and E. S. SARI, “Analisis Penggunaan Gaya Bahasa Sarkasme Netizen di Media Sosial Tiktok,” *KNOWLEDGE: Jurnal Inovasi Hasil Penelitian dan Pengembangan*, vol. 3, no. 1, pp. 84–92, 2023.
- [4] U. Ulfatun, “Analisis Penggunaan Gaya Bahasa Sarkasme Netizen di Media Sosial Instagram,” *Jurnal Onoma: Pendidikan, Bahasa, Dan Sastra*, vol. 7, no. 2, pp. 411–423, 2021.
- [5] Sarli, S., NURHADI, N. and SARI, E.S., “Analisis Penggunaan Gaya Bahasa Sarkasme Netizen di Media Sosial Tiktok,” *KNOWLEDGE: Jurnal Inovasi Hasil Penelitian dan Pengembangan*, vol. 3, no. 1, pp.84-92, 2023.
- [6] D. Yadewani, M. L. Arief, and W. I. Mursalini, “Pengaruh pemanfaatan platform sosial media pada era digital terhadap prestasi mahasiswa,” *Jurnal Apresiasi Ekonomi*, vol. 8, no. 3, pp. 521–527, 2020.
- [7] M. A. N. Zen and A. S. Sitanggang, “Analisis Dampak Sosial Media Dalam Pengembangan Sistem Informasi,” *Cerdika: Jurnal Ilmiah Indonesia*, vol. 3, no. 7, pp. 671–682, 2023.
- [8] U. Ulfia, R. Rahmi, and S. Yana, “Pengaruh Media Sosial Dalam Transformasi Pemasaran Digital,” *JUPEIS: Jurnal Pendidikan dan Ilmu Sosial*, vol. 3, no. 3, pp. 11–17, 2024.
- [9] S. I. A. Cahya and G. G. Sukendro, “Musik sebagai media komunikasi ekspresi cinta (analisis semiotika lirik lagu ‘rumah ke rumah’ karya Hindia),” *Koneksi*, vol. 6, no. 2, pp. 246–254, 2022.
- [10] J. Alimuddin, “Lagu anak sebagai salah satu sarana mendidik anak,” *Jurnal Ilmiah Pendidikan Dasar*, vol. 2, no. 2, pp. 108–116, 2015.
- [11] P. M. Dewi, C. G. Syavica, D. T. Wahyuni, M. Rafi Hanif, and E. Mauzah, “Analisis Penggunaan Gaya Bahasa Sarkasme dan Satire dalam Ragam Aplikasi Media Sosial,” *Jurnal Majemuk*, vol. 1, no. 4, pp. 463–471, 2022.
- [12] N. Riana et al., *Komunikasi Publik: Panduan Praktis untuk Sukses dalam Berkomunikasi*. PT. Sonpedia Publishing Indonesia, 2024.
- [13] M. V. Rao and C. Sindhu, “Detection of sarcasm on amazon product reviews using machine learning algorithms under sentiment analysis,” presented at the 2021 sixth international conference on wireless communications, signal processing and networking (WiSPNET), IEEE, 2021, pp. 196–199.
- [14] M. Noveanto, H. Sastypratiwi, and H. Muhandi, “Uji akurasi klasifikasi emosi pada lirik lagu bahasa indonesia,” *JUSTIN (Jurnal Sistem dan Teknologi Informasi)*, vol. 10, no. 3, pp. 311–318, 2022.
- [15] R. A. Salam, T. Akib, and M. Dahlan, “Analisis Wacana Kritis terhadap Sarkasme dalam Twitter Sejak Bulan September–November 2023,” *Jurnal Onoma: Pendidikan, Bahasa, dan Sastra*, vol. 10, no. 2, pp. 1855–1863, 2024.
- [16] D. Paramita, M. R. Aldiano, K. I. S. Siregar, H. Sazali, and M. Andinata, “Analisis Penggunaan Gaya Bahasa Sarkasme Netizen di Media Sosial Facebook,” *Jurnal Pendidikan Tambusai*, vol. 6, no. 2, pp. 14438–14445, 2022.
- [17] D. S. Oktaviani, H. Suaedi, and M. Afrizal, “Penggunaan Bentuk-Bentuk Gaya Bahasa Sarkasme di Channel Youtube Deddy Corbuzier,” *Jurnal Pendidikan Tambusai*, vol. 8, no. 1, pp. 9357–9366, 2024.
- [18] Irmanda, H. N., & Astriratma, R., “Klasifikasi Jenis Pantun Dengan Metode Support Vector Machines (SVM),” *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 4, no. 5, pp.915-922, 2020.
- [19] O. I. Gifari, M. Adha, I. R. Hendrawan, and F. F. S. Durrand, “Analisis Sentimen Review Film Menggunakan TF-IDF dan Support Vector Machine,” *Journal of Information Technology*, vol. 2, no. 1, pp. 36–40, 2022.
- [20] M. Iqbal, M. Afdal, and R. Novita, “Implementasi Algoritma Support Vector Machine Untuk Analisa Sentimen Data Ulasan Aplikasi Pinjaman Online di Google Play Store: Implementation of Support Vector Machine Algorithm for Sentiment Analysis of Online Loan Application Review Data on Google Play Store,” *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 4, no. 4, pp. 1244–1252, 2024.
- [21] F. Amin and M. Mahmoud, “Confusion matrix in binary classification problems: A step-by-step tutorial,” *Journal of Engineering Research*, vol. 6, no. 5, pp. 0–0, 2022.
- [22] W. Wu, “React Native vs Flutter, Cross-platforms mobile application frameworks,” 2018.