

Analisis Klasifikasi Ulasan Aplikasi Sirekap 2024 menggunakan Ekstraksi Fitur DistilBert Dan Metode Support Vector Machine

Reno Ridhoi^a Naufal Azmi Verdikha^b Fendy Yulianto^c

^a Universitas Muhammadiyah Kalimantan Timur, Jl. Ir. H. Juanda No.15, Sidodadi, Kec. Samarinda Ulu, Kota Samarinda, 75124, Indonesia

^b Universitas Muhammadiyah Kalimantan Timur, Jl. Ir. H. Juanda No.15, Sidodadi, Kec. Samarinda Ulu, Kota Samarinda, 75124, Indonesia

^c Universitas Muhammadiyah Kalimantan Timur, Jl. Ir. H. Juanda No.15, Sidodadi, Kec. Samarinda Ulu, Kota Samarinda, 75124, Indonesia

INFORMASI ARTIKEL

Sejarah Artikel:

Diterima Redaksi: 02-Februari 2025

Revisi Akhir: 22 Februari 2025

Diterbitkan Online: 15 Maret 2025

KATA KUNCI

DistilBert, Support Vector Machine, SIREKAP, cross-validation

KORESPONDENSI

E-mail: 2111102441165@umkt.ac.id

nav651@umkt.ac.id

fy415@umkt.ac.id

A B S T R A C T

This study aims to automatically classify user reviews of the SIREKAP 2024 application using the DistilBERT feature extraction method and the Support Vector Machine (SVM) algorithm. DistilBERT serves as a lightweight version of BERT, efficiently extracting text representations while preserving contextual meaning, whereas SVM is employed as a supervised learning algorithm to separate reviews into five rating categories available on the Google Play Store. The study utilizes 8,538 user reviews as data, which undergo preprocessing, including text cleaning, tokenization, and feature extraction using DistilBERT embeddings before being fed into the SVM classifier. One of the main challenges identified is class imbalance, where certain rating categories have significantly fewer samples than others, affecting model performance. Through 10-Fold cross-validation, the model achieved an average F1-Score of 36.62%, with the highest performance reaching 37.16%. Further analysis indicates that class imbalance, particularly in minority classes, significantly hinders classification accuracy and leads to biased predictions. The findings conclude that the combination of DistilBERT and SVM produces suboptimal results and requires further optimization, such as applying data balancing techniques or exploring alternative classification models. Therefore, this study provides recommendations to enhance model accuracy and improve the quality of the application based on user reviews..n

1. PENDAHULUAN

Pemilu adalah salah satu pilar penting dalam demokrasi yang menjadi ajang partisipasi masyarakat untuk memilih pemimpin dan wakil rakyat. Setiap lima tahun, Komisi Pemilihan Umum (KPU) bertanggung jawab mengelola pelaksanaan pemilu, termasuk proses rekapitulasi suara yang melibatkan Panitia Pemungutan Suara (KPPS) di berbagai daerah. Dalam proses ini, transparansi, akurasi, dan efisiensi menjadi aspek krusial untuk menjaga kredibilitas hasil pemilu. Seiring perkembangan teknologi, digitalisasi mulai diterapkan dalam pelaksanaan pemilu untuk mengatasi berbagai kendala, seperti rekapitulasi manual yang memakan waktu dan rentan kesalahan. Salah satu inovasi yang dihadirkan adalah aplikasi Sistem Informasi Rekapitulasi (SIREKAP), yang dirancang untuk mendukung pencatatan dan pengunggahan hasil penghitungan suara secara digital langsung dari Tempat Pemungutan Suara (TPS) [1].

SIREKAP bertujuan meningkatkan efisiensi operasional, mengurangi penggunaan kertas, dan meminimalkan potensi sengketa hasil pemilu. Namun, implementasi SIREKAP pada pemilu 2024 menghadapi sejumlah tantangan, seperti kendala login, kesulitan dalam pengunggahan dokumen, serta aksesibilitas aplikasi yang tidak optimal. Berbagai keluhan pengguna aplikasi ini telah tercatat dalam ulasan di *Google Play Store*, yang menunjukkan pentingnya analisis mendalam untuk mengidentifikasi kekurangan aplikasi dan merumuskan solusi perbaikan [2]. Pengumpulan data ulasan dari *platform* digital seperti *Google Play Store* memberikan peluang untuk memahami persepsi dan kepuasan pengguna secara objektif. Namun, analisis manual terhadap ribuan ulasan membutuhkan waktu dan sumber daya yang besar serta rawan bias [3].

Pendekatan berbasis *Natural Language Processing* (NLP) menawarkan solusi untuk menganalisis ulasan pengguna secara otomatis dan efisien. Salah satu teknik NLP yang menonjol adalah DistilBERT, model berbasis transformer yang dirancang untuk menghasilkan representasi teks berkualitas tinggi dengan

efisiensi komputasi yang lebih baik dibandingkan model serupa seperti BERT. DistilBERT mampu memberikan hasil yang lebih baik dibandingkan dengan pendekatan berbasis *Bag-of-Words* (BoW), seperti yang ditunjukkan oleh [4] yang menguji performa DistilBERT dan BERT dalam klasifikasi teks. Kombinasi DistilBERT untuk ekstraksi fitur dan algoritma *Support Vector Machine* (SVM). SVM memisahkan data ke dalam dua kelompok, yaitu kelas positif dan negatif. Data tersebut biasanya direpresentasikan dalam dua atau tiga dimensi. Algoritma SVM dirancang untuk membedakan dua kelas dengan mencari *hyperplane* optimal yang memaksimalkan jarak antara kedua kelas tersebut [5] untuk klasifikasi memungkinkan analisis ulasan aplikasi secara otomatis dengan tingkat akurasi yang tinggi [6].

Berdasarkan latar belakang tersebut, penelitian ini merumuskan masalah utama bagaimana keefektifan kombinasi DistilBERT dan SVM dalam menganalisis ulasan aplikasi SIREKAP secara otomatis, serta bagaimana performa model dalam mengatasi ketidakseimbangan data ulasan dengan kategori rating yang bervariasi. Penelitian ini bertujuan untuk menguji efektivitas metode tersebut dalam menganalisis ulasan pengguna, mengevaluasi performa model menggunakan F1-Score sebagai metrik utama, serta memberikan rekomendasi berbasis hasil analisis untuk pengembangan aplikasi SIREKAP.

Penelitian ini dibatasi pada analisis ulasan aplikasi SIREKAP yang dikumpulkan dari *Google Play Store* pada 6 Februari 2024, pukul 22:00 WITA, dengan lima kategori rating sebagai variabel target. Proses pra-pemrosesan data mencakup konversi teks menjadi huruf kecil, penghapusan karakter tidak relevan, pemeriksaan ejaan, dan *stemming* [7]. Penggunaan DistilBERT dilakukan dengan pengaturan *default*, tanpa modifikasi parameter, dan klasifikasi dilakukan menggunakan algoritma SVM [8].

Tujuan dari penelitian ini adalah untuk menguji efektivitas kombinasi *DistilBERT* dan SVM dalam mengklasifikasikan ulasan aplikasi secara otomatis, serta mengevaluasi kualitas model berdasarkan nilai *F1-Score*. Penelitian ini diharapkan dapat memberikan kontribusi dalam pengembangan metode analisis ulasan yang lebih baik, serta memberikan wawasan bagi pengembang untuk meningkatkan kualitas aplikasi SIREKAP berdasarkan respon pengguna.

2. TINJAUAN PUSTAKA

2.1 Natural Language Processing (NLP)

Natural Language Processing (NLP) merupakan cabang dari kecerdasan buatan yang berfokus pada pemrosesan dan pemahaman bahasa manusia menggunakan komputer. Dalam konteks analisis ulasan aplikasi, NLP sangat berguna untuk mengotomatisasi pemahaman teks dalam ulasan pengguna, mengidentifikasi sentimen, serta mengklasifikasikan ulasan ke dalam kategori tertentu [9]. NLP dapat diterapkan dalam berbagai tahap analisis ulasan, mulai dari pembersihan data, ekstraksi fitur, hingga klasifikasi sentimen.

Salah satu teknik yang sering digunakan dalam NLP untuk ekstraksi fitur adalah metode *word embeddings*, yang memungkinkan representasi kata dalam bentuk vektor. Beberapa

metode yang banyak digunakan dalam NLP adalah *Bag-of-Words* (BoW), *Term Frequency-Inverse Document Frequency* (TF-IDF), dan *Word2Vec* [10]. Namun, seiring berkembangnya model berbasis transformer, metode BERT (*Bidirectional Encoder Representations from Transformers*) dan versi yang lebih ringan DistilBERT semakin populer karena kemampuannya dalam memahami konteks semantik dalam teks dengan lebih baik dan efisien secara komputasi [11]. DistilBERT merupakan versi terdistilasi dari BERT yang lebih efisien dalam penggunaan sumber daya komputasi tanpa mengurangi kualitas hasil analisis teks [6].

2.2 DistilBERT

DistilBERT, yang dikembangkan oleh Hugging Face, adalah model bahasa berbasis transformer yang lebih kecil dan lebih cepat dibandingkan BERT, namun tetap mempertahankan 97% dari kemampuan BERT dalam memahami konteks bahasa alami. DistilBERT menggunakan teknik yang disebut *knowledge distillation*, di mana model yang lebih kecil dilatih untuk menyerap pengetahuan dari model yang lebih besar [11]. Beberapa penelitian telah menunjukkan bahwa DistilBERT memiliki performa yang lebih baik dibandingkan model tradisional, seperti BoW atau TF-IDF, dalam tugas klasifikasi teks [4].

Penerapan DistilBERT dalam analisis ulasan dapat membantu mengonversi teks menjadi representasi numerik yang lebih bermakna, yang kemudian dapat digunakan dalam berbagai algoritma klasifikasi. DistilBERT terbukti dapat meningkatkan akurasi analisis teks, termasuk dalam klasifikasi sentimen, yang penting untuk memahami perasaan pengguna terhadap aplikasi tertentu [6].

2.3 Support Vector Machine (SVM)

Support Vector Machine (SVM) adalah salah satu algoritma klasifikasi yang sering digunakan dalam analisis data berbasis teks. SVM bekerja dengan memisahkan data ke dalam dua kelas yang berbeda dengan mencari *hyperplane* optimal yang memaksimalkan jarak antara kedua kelas tersebut [8]. Dalam konteks analisis ulasan, SVM digunakan untuk mengklasifikasikan ulasan pengguna ke dalam kategori rating yang berbeda, seperti positif, negatif, atau netral.

Beberapa penelitian telah menunjukkan bahwa SVM dapat menghasilkan akurasi yang lebih tinggi dalam klasifikasi ulasan dibandingkan algoritma lainnya, seperti Naive Bayes atau K-Nearest Neighbors [12]. Hal ini disebabkan oleh kemampuannya untuk bekerja dengan baik pada data yang memiliki dimensi tinggi, seperti teks yang telah diekstraksi menggunakan model *embedding* seperti DistilBERT.

Dalam penelitian oleh [13], algoritma SVM menunjukkan performa yang unggul dalam analisis sentimen terhadap ulasan pengguna aplikasi *Mobile Legend di Google Play Store*, dengan mencapai akurasi sebesar 0,87. Oleh karena itu, SVM dipilih sebagai metode klasifikasi utama dalam penelitian ini untuk menganalisis ulasan aplikasi SIREKAP.

2.4 Penelitian Terkait

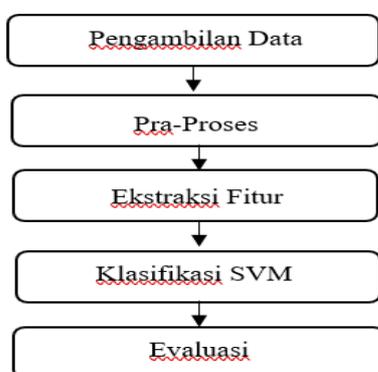
Beberapa penelitian sebelumnya menunjukkan efektivitas kombinasi model berbasis transformer dan algoritma klasik seperti SVM. Misalnya, penelitian oleh [6] menunjukkan bahwa DistilBERT menghasilkan akurasi yang lebih tinggi dibandingkan BERT dalam analisis sentimen pada dataset Twitter berbahasa Indonesia. Dalam penelitian lain, [13] menemukan bahwa SVM mampu menghasilkan akurasi klasifikasi yang lebih tinggi dibandingkan algoritma lain seperti Random Forest dan K-Nearest Neighbors dalam klasifikasi ulasan aplikasi di *Google Play Store*. Penelitian ini mengadopsi pendekatan serupa dengan menggabungkan kemampuan DistilBERT dalam representasi teks dan kekuatan SVM dalam klasifikasi data berdimensi tinggi untuk mengatasi tantangan pada dataset ulasan aplikasi SIREKAP 2024.

2.5 Ketidakseimbangan Data

Ketidakseimbangan data adalah salah satu tantangan utama dalam klasifikasi teks. Dalam dataset dengan distribusi kelas yang tidak merata, model cenderung bias terhadap kelas mayoritas, sehingga performa klasifikasi pada kelas minoritas menjadi rendah. Berbagai teknik telah dikembangkan untuk mengatasi masalah ini, seperti oversampling, undersampling, dan *Synthetic Minority Oversampling Technique* (SMOTE). SMOTE bekerja dengan menghasilkan sampel sintetis pada kelas minoritas, sehingga meningkatkan distribusi data secara keseluruhan [14]. Dalam penelitian ini, ketidakseimbangan data ditemukan pada distribusi rating ulasan aplikasi SIREKAP 2024, di mana rating 1 mendominasi dataset. Solusi seperti SMOTE dapat diterapkan untuk meningkatkan kinerja model, terutama dalam mengklasifikasikan kelas minoritas seperti rating 4 dan 5. Dengan mengatasi ketidakseimbangan data, penelitian ini diharapkan dapat memberikan hasil klasifikasi yang lebih adil dan akurat.

3. METODOLOGI

Metode yang digunakan dalam penelitian ini terdiri dari beberapa tahapan utama, yaitu pengambilan data, pra-pemrosesan, ekstraksi fitur, klasifikasi menggunakan Support Vector Machine (SVM), dan evaluasi. Setiap tahap memiliki peran penting dalam memastikan keakuratan serta efektivitas model yang dikembangkan. Alur penelitian secara keseluruhan dapat dilihat pada Gambar 1.



Gambar 1. Alur Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan metode eksperimen untuk menganalisis ulasan aplikasi

SIREKAP 2024. Penelitian ini terdiri dari beberapa tahapan yang saling berkaitan, dimulai dengan pengumpulan data ulasan pengguna yang diambil dari *platform Google Play Store*. Data yang digunakan dalam penelitian ini mencakup 8.538 ulasan dengan lima kategori rating (1 hingga 5) yang diberikan oleh pengguna aplikasi. Pengambilan data dilakukan dengan menggunakan metode *web scraping*, yang memanfaatkan pustaka Python *google_play_scraper* untuk mengekstrak ulasan, beserta informasi terkait seperti nama pengguna, waktu ulasan, isi komentar, dan jumlah dukungan.

Setelah data terkumpul, tahapan selanjutnya adalah pra-pemrosesan data untuk memastikan kualitas data yang siap dianalisis. Tahapan ini meliputi konversi teks menjadi huruf kecil (*lowercasing*), penghapusan karakter-karakter yang tidak relevan seperti angka dan simbol (*remove characters*), serta koreksi ejaan menggunakan *spell checker* untuk memperbaiki kata-kata yang salah eja, terutama untuk istilah-istilah yang sering digunakan dalam ulasan. Proses berikutnya adalah *stemming* menggunakan pustaka Sastrawi untuk mengubah kata berimbuhan menjadi kata dasar, sehingga teks yang digunakan lebih konsisten dan sesuai dengan kaidah bahasa Indonesia.

Setelah pra-pemrosesan, ekstraksi fitur dilakukan dengan menggunakan model DistilBERT, yang merupakan versi terdistilasi dari BERT. DistilBERT digunakan untuk mengonversi teks ulasan menjadi vektor numerik berkualitas tinggi yang memuat informasi semantik dari setiap teks ulasan. DistilBERT dipilih karena kemampuannya dalam memahami konteks teks dengan akurasi tinggi dan efisiensi komputasi yang lebih baik dibandingkan model-model sebelumnya. Dalam proses ekstraksi, teks ulasan di-tokenisasi dan diproses menjadi representasi vektor dengan dimensi 768, yang kemudian digunakan sebagai input untuk tahap klasifikasi.

Pada tahap klasifikasi, *Support Vector Machine* (SVM) digunakan sebagai algoritma utama. SVM dipilih karena kemampuannya dalam mengklasifikasikan data dengan dimensi tinggi seperti teks dan telah terbukti efektif dalam berbagai penelitian sebelumnya [12]. Model SVM akan dilatih menggunakan data latihan yang terdiri dari fitur-fitur yang diekstraksi oleh DistilBERT, dan hasil klasifikasinya digunakan untuk mengelompokkan ulasan pengguna ke dalam kategori rating yang sesuai.

Untuk memastikan evaluasi model yang akurat, penelitian ini menggunakan metode *K-Fold Cross Validation* dengan 10 lipatan (*folds*). Teknik ini membagi data menjadi 10 bagian yang digunakan bergantian sebagai data latihan dan data uji. Dengan demikian, seluruh data digunakan dalam proses pelatihan dan evaluasi, yang membantu mengurangi bias dan memastikan kestabilan hasil. Metrik evaluasi utama yang digunakan untuk menilai kinerja model adalah F1-Score, yang menggabungkan *precision* dan *recall* untuk memberikan gambaran yang lebih seimbang tentang akurasi model dalam mengklasifikasikan ulasan.

Setelah evaluasi dilakukan, hasil F1-Score dianalisis untuk mengetahui keefektifan kombinasi DistilBERT dan SVM dalam mengklasifikasikan ulasan aplikasi SIREKAP. Selain itu, penelitian ini juga memberikan rekomendasi terkait langkah-

langkah perbaikan yang dapat diambil oleh pengembang aplikasi SIREKAP berdasarkan hasil analisis ulasan pengguna.

4. HASIL DAN PEMBAHASAN

Bagian ini menyajikan hasil dari penelitian yang telah dilakukan serta pembahasan untuk menganalisis lebih lanjut temuan yang diperoleh. Hasil penelitian meliputi proses pengambilan data, tahapan pra-pemrosesan, ekstraksi fitur menggunakan model DistilBERT, dan klasifikasi menggunakan metode Support Vector Machine (SVM). Pembahasan difokuskan pada analisis hasil evaluasi model serta kendala yang ditemukan selama penelitian.

4.1. Hasil Pengambilan Data

Data ulasan aplikasi SIREKAP 2024 diperoleh melalui proses *scraping* menggunakan library Python seperti *google_play_scraper*. Sebanyak 8.538 ulasan pengguna berhasil dikumpulkan dari *Google Play Store*. Data ini mencakup kolom seperti *userName*, *score*, *at*, *content*, dan *thumbsUpCount*. Analisis awal menunjukkan distribusi ulasan sebagai berikut:

Tabel 1. Distribusi jumlah ulasan berdasarkan *rating*

Rating	Jumlah Ulasan
1	5.361
2	736
3	478
4	255
5	1.528

Distribusi ini menunjukkan bahwa ulasan dengan *rating* 1 mendominasi dataset, menandakan banyaknya keluhan dari pengguna terhadap aplikasi.

4.2. Pra-Proses Data

Pra-pemrosesan dilakukan untuk memastikan data yang bersih dan konsisten. Hasil dari setiap tahapan pra-pemrosesan adalah sebagai berikut:

1. Penambahan Kolom ID : Memberikan identitas unik pada setiap ulasan untuk mempermudah analisis.
2. *Lowercase* : Mengubah semua huruf menjadi huruf kecil untuk konsistensi data.
3. Penghapusan Karakter Tidak Relevan : Menghapus angka, simbol, dan tanda baca yang tidak relevan.
4. Pemeriksaan Ejaan (*Spell Checker*) : Memperbaiki kesalahan ejaan menggunakan pustaka yang sesuai.
5. *Stemming* : Mengubah kata menjadi bentuk dasar menggunakan library Sastrawi.

Tahapan pra-pemrosesan ini bertujuan untuk memastikan bahwa data yang digunakan dalam analisis lebih bersih, terstruktur, dan konsisten. Setiap tahapan berkontribusi untuk meningkatkan kualitas data, sehingga mampu menghasilkan representasi yang lebih baik pada proses ekstraksi fitur.

ID	Nama	Rating	Waktu	Komentar	Like
0 d0001	Nency Miranda	5	05/02/2024 14:07	Mantap	0.0
1 d0002	ichal sabiel	1	05/02/2024 14:06	Sudah tau pemakai nya masyarakat biasa Malah sp...	0.0
2 d0003	Evandra Aditya	1	05/02/2024 14:06	Apk nya nggk bisa buat log in	0.0
3 d0004	Ganesh insann	1	05/02/2024 14:05	Susah masuk dih	0.0
4 d0005	Yohana Frediana	1	05/02/2024 14:04	Tidak bisa masuk inisialisasi	0.0

Gambar 2. Contoh Sebelum Pra-Pemrosesan Data

```

komentar_stemming
0          mantap
1  sudah tahu maka masyarakat biasa malah spesifi...
2          aplikasi nya tidak bisa buat masuk ini
3          susah masuk dih
4          tidak bisa masuk inisial
...
8353 moga tambah baik mudah dalam masuk dan bagi data
8354 cara masuk nya bagaimana pakai email kok tidak...
8355          tidak tahu
8356 moga tidak seperti sirekap 2020 yang saat guna...
8357 moga dengan ada aplikasi ini mlw akan lebih b...

[8358 rows x 2 columns]
Waktu proses: 41.335997343063354 detik
    
```

Gambar 3. Contoh Hasil Sesudah Pra-Pemrosesan Data

Tahapan ini menghasilkan dataset yang siap digunakan dalam proses analisis lebih lanjut. Dataset yang bersih ini juga membantu dalam mengurangi risiko error saat dilakukan pengolahan data menggunakan algoritma machine learning.

4.3. Ekstraksi Fitur Menggunakan DistilBert

Ekstraksi fitur dilakukan menggunakan model DistilBERT. Setiap teks ulasan diubah menjadi representasi vektor berdimensi 768, yang mencerminkan karakteristik semantik dari teks tersebut. DistilBERT dipilih karena keunggulannya dalam efisiensi komputasi dan kemampuan memahami konteks teks dengan baik. Proses dan hasil dari ekstraksi fitur ini dapat dilihat Dibawah ini

```

[[-0.22397892 -0.06847578 -0.19891159 ... -0.11078235  0.23855396
  0.20429268]
 [-0.2712006  0.15610412 -0.43394405 ... -0.00049034  0.13641356
  0.43174267]
 [-0.2076754  0.06132379 -0.30698162 ... -0.03759241  0.12059601
  0.35723692]
 ...
 [-0.1461092  0.00464123 -0.2265185 ... -0.14038318  0.16260211
  0.30118847]
 [-0.28399152 -0.00187148 -0.14521487 ... -0.00439866  0.16138701
  0.30761564]
 [-0.31894556  0.10667042 -0.3922788 ... -0.07174631  0.03833859
  0.3534185 ]]
    
```

Gambar 4. Nilai Fitur

Berdasarkan hasil ekstraksi fitur yang ditampilkan, setiap baris teks dalam dataset berhasil direpresentasikan ke dalam bentuk vektor berdimensi tinggi. Vektor-vektor ini menunjukkan hasil pemrosesan oleh DistilBERT, di mana setiap nilai dalam vektor mewakili fitur-fitur unik yang menggambarkan karakteristik semantik dari teks tersebut. Sebagai contoh, baris pertama menghasilkan vektor dengan nilai awal seperti [-0.22397892, -0.06847578, -0.19891159], yang menunjukkan bahwa teks telah dikonversi menjadi angka yang dapat dipahami oleh model *machine learning*. Setiap vektor ini memiliki dimensi tetap sebanyak 768, sesuai dengan ukuran *embedding* pada *DistilBERT*.

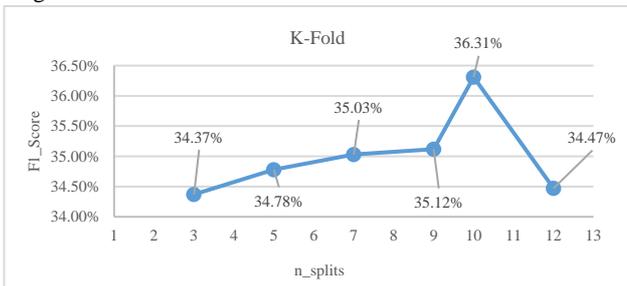
Proses ekstraksi fitur ini mencakup tahapan tokenisasi teks, padding, dan masking sebelum diolah menjadi representasi numerik oleh model DistilBERT. Representasi vektor ini menjadi masukan penting untuk algoritma klasifikasi yang digunakan pada tahap selanjutnya.

4.4. Klasifikasi Menggunakan Support Vector Machine (SVM) dan Evaluasi

Metode klasifikasi dilakukan dengan menggunakan *Support Vector Machine (SVM)*. Proses validasi menggunakan metode *K-Fold Cross Validation* dengan 10 fold untuk memastikan evaluasi

performa model yang konsisten. Dalam setiap iterasi, data dibagi menjadi 10 bagian, di mana 9 bagian digunakan sebagai data latih dan 1 bagian sebagai data uji. Pendekatan ini mengurangi risiko bias dalam evaluasi dan memastikan bahwa setiap data digunakan untuk pelatihan maupun pengujian.

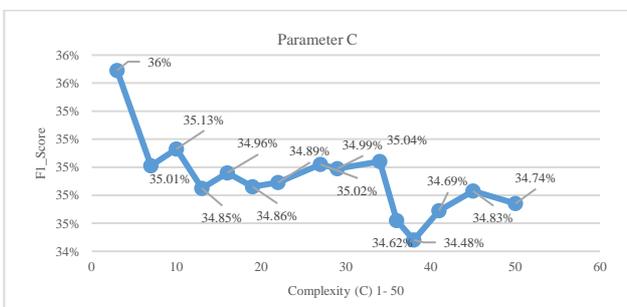
Jumlah *Fold* untuk *Cross Validation* dipilih berdasarkan analisis perbandingan performa model dengan berbagai skenario jumlah *Fold*, yaitu 3, 5, 7, 9, 10, dan 12. Analisis ini dilakukan menggunakan seluruh dataset untuk menentukan konfigurasi jumlah *Fold* yang menghasilkan akurasi terbaik. Berdasarkan hasil analisis, konfigurasi dengan 10 *Fold* menunjukkan performa paling optimal dibandingkan skenario lainnya. Oleh karena itu, 10 *Fold* diputuskan sebagai konfigurasi yang digunakan untuk proses *Cross Validation*. Hasil perbandingan performa model untuk setiap jumlah *Fold* yang diuji divisualisasikan pada diagram di Gambar dibawah ini.



Gambar 5. Nilai hasil perbandingan jumlah Fold

Hasil evaluasi menunjukkan nilai rata-rata F1-Score sebesar 36,19%, dengan performa tertinggi mencapai 37,94% pada salah satu *fold*. Hal ini mengindikasikan bahwa model mampu mengklasifikasikan ulasan dengan cukup baik, meskipun terdapat beberapa tantangan yang perlu diatasi, seperti ketidakseimbangan data antar kelas.

F1-Score pada setiap Fold diperoleh melalui pelatihan model SVM dengan LinearSVC untuk mengoptimalkan parameter, terutama C, yang mengatur regularisasi, berdasarkan studi sebelumnya (Pravina, Cholissodin, & Adikara, 2019) dengan nilai awal C = 10 dan max_iter = 50. Awalnya, model diuji tanpa class_weight='balanced', menghasilkan akurasi 31,72%, lalu meningkat menjadi 35,13% setelah fitur ini ditambahkan untuk menangani ketidakseimbangan data. Selanjutnya, dilakukan penyesuaian parameter guna menemukan kombinasi terbaik yang menghasilkan akurasi tertinggi, sebagaimana ditampilkan pada Gambar 6.

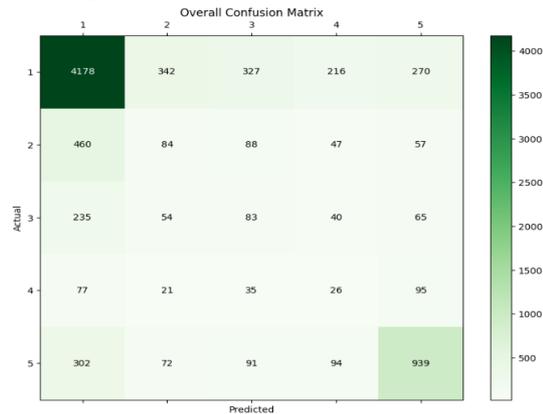


Gambar 6. Perbandingan Parameter C

Gambar diatas menunjukkan bahwa F1-Score tertinggi (36%) dicapai saat C = 3, lalu performa model fluktuatif dengan penurunan pada C = 7 hingga 13, sedikit meningkat di C = 22 hingga 29, dan kembali turun mendekati C = 50, menunjukkan pengaruh signifikan parameter C terhadap model. Selanjutnya,

dilakukan evaluasi terhadap jumlah iterasi maksimum max_iter, yang awalnya 50 berdasarkan penelitian [15]. Namun, untuk memperoleh hasil yang lebih optimal, pengujian dilakukan dengan berbagai nilai iterasi dalam rentang 0 hingga 1000. Akurasi tertinggi 36,62% dicapai saat max_iter = 28, setelah pengujian berbagai kombinasi C dan max_iter, menunjukkan optimasi terbaik untuk performa model.

Evaluasi performa model klasifikasi juga dianalisis melalui *Confusion Matrix* secara keseluruhan, yang memperlihatkan perbandingan antara label sebenarnya dengan label hasil prediksi untuk seluruh dataset. *Confusion Matrix* ini memberikan pandangan yang lebih rinci terkait kemampuan model dalam mengklasifikasikan data secara tepat serta jenis kesalahan prediksi yang terjadi. Gambaran lengkap dari *Confusion Matrix* dapat dilihat pada Gambar 7.



Gambar 7. Overall Confusion Matrix

Berikut hasil revisinya dengan angka yang telah diperbarui: *Confusion Matrix* menggambarkan hasil klasifikasi model secara keseluruhan. Nilai tertinggi terlihat pada kolom dan baris label 1, yaitu sebanyak 4178 data, menunjukkan bahwa data dengan label aktual 1 sebagian besar berhasil diklasifikasikan dengan benar. Sebaliknya, nilai terendah ditemukan pada kolom label 4 dan baris label 4, dengan hanya 26 data yang diklasifikasikan ke label tersebut. Warna dalam matriks menunjukkan intensitas jumlah data, di mana warna lebih gelap menandakan jumlah data yang lebih tinggi.

Secara lebih rinci, jumlah data yang diklasifikasikan dengan benar untuk masing-masing label adalah sebagai berikut: label 1 sebanyak 4178 data, label 2 sebanyak 84 data, label 3 sebanyak 83 data, label 4 sebanyak 26 data, dan label 5 sebanyak 939 data. Berdasarkan hasil ini, evaluasi model menunjukkan Macro F1-Score sebesar 36,62%, Macro Precision 36,41%, dan Macro Recall 36,63%, mencerminkan tingkat akurasi model dalam keseluruhan proses klasifikasi.

Kinerja model menunjukkan bahwa kombinasi DistilBERT dan SVM belum dikategorikan efektif dalam menganalisis data teks, sehingga masih harus mencari peluang lain untuk meningkatkan akurasi, terutama untuk kelas dengan jumlah data yang lebih sedikit.

4.5. Analisis Dan Pembahasan

Penelitian ini menganalisis klasifikasi 8.538 ulasan pengguna aplikasi SIREKAP 2024 menggunakan metode ekstraksi fitur DistilBERT dan algoritma SVM. Data dikumpulkan melalui web scraping dari Google Play Store, kemudian diseleksi untuk menghapus duplikat dan ulasan tidak relevan. Selanjutnya,

dilakukan pra-pemrosesan seperti lowercasing, penghapusan karakter tidak penting, spell checker, dan stemming agar data lebih bersih dan siap untuk ekstraksi fitur.

Pada tahap ekstraksi fitur, model DistilBERT digunakan untuk mengubah teks ulasan menjadi vektor berdimensi tetap sebanyak 768 dimensi melalui proses tokenisasi, padding, dan masking. Vektor ini kemudian digunakan sebagai input untuk klasifikasi dengan algoritma Support Vector Machine (SVM), yang dipilih karena kemampuannya menangani data berdimensi tinggi dan menghasilkan hyperplane optimal. Validasi model dilakukan menggunakan teknik K-Fold Cross Validation dengan berbagai nilai Fold, di mana Fold ke-10 memberikan rata-rata F1-Score tertinggi sebesar 36,31%, sehingga dipilih sebagai konfigurasi optimal.

Pada tahap optimasi, penelitian ini mengacu pada studi sebelumnya dengan menetapkan parameter C awal sebesar 10 dan iterasi maksimum (max_iter) sebanyak 50. Pada konfigurasi awal, model menghasilkan F1-Score sebesar 31,72% tanpa metode class_weight='balanced', yang kemudian meningkat menjadi 35,13% setelah diterapkan metode tersebut. Selanjutnya, dilakukan pencarian nilai C terbaik dalam rentang 0 hingga 50, dan nilai C = 3 ditemukan sebagai parameter optimal dengan F1-Score sebesar 36%. Secara keseluruhan, penelitian ini mencapai F1-Score akhir sebesar 36,62%, namun hasil ini masih terpengaruh oleh ketidakseimbangan data antar kelas yang menjadi tantangan utama dalam klasifikasi.

5. KESIMPULAN DAN SARAN

Dari hasil penelitian yang dilakukan, diperoleh bahwa hasil akhir dari rangkaian klasifikasi ulasan aplikasi SIREKAP menggunakan metode ekstraksi fitur DistilBERT dan algoritma klasifikasi Support Vector Machine (SVM) adalah sebesar 36,62% berdasarkan nilai F1-Score. Hasil yang belum dapat dikategorikan baik ini menunjukkan bahwa kombinasi DistilBERT dengan SVM tidak cukup efektif dalam menganalisis data dengan distribusi tidak seimbang, khususnya pada kelas minoritas seperti ulasan aplikasi SIREKAP. Hal ini menegaskan bahwa ketidakseimbangan data ulasan menjadi kendala utama yang memengaruhi performa model, sehingga diperlukan pengembangan lebih lanjut baik dari sisi metode ekstraksi fitur maupun strategi klasifikasi untuk meningkatkan kemampuan analisis pada data serupa di masa mendatang.

Penelitian selanjutnya disarankan untuk menggunakan teknik penyeimbangan data seperti SMOTE (*Synthetic Minority Oversampling Technique*) untuk mengatasi ketidakseimbangan kelas. Selain itu, eksperimen lebih lanjut dengan tuning hyperparameter pada DistilBERT dan SVM dapat meningkatkan akurasi model. Mengumpulkan lebih banyak data dari berbagai sumber dapat membantu model untuk belajar lebih baik dan menghasilkan hasil yang lebih akurat. Penelitian masa depan juga dapat mencoba algoritma lain seperti Random Forest, Gradient Boosting, atau deep learning untuk dibandingkan dengan SVM. Saran ini diharapkan dapat menjadi panduan untuk penelitian lanjutan yang bertujuan meningkatkan akurasi klasifikasi ulasan dan memberikan kontribusi nyata dalam pengembangan aplikasi berbasis teknologi NLP. Saran ini diharapkan dapat menjadi

panduan untuk penelitian lanjutan yang bertujuan meningkatkan akurasi klasifikasi ulasan dan memberikan kontribusi nyata dalam pengembangan aplikasi berbasis teknologi NLP.

DAFTAR PUSTAKA

- [1] M. F. Y. Herjanto and C. Carudin, "Analisis Sentimen Ulasan Pengguna Aplikasi Sirekap Pada Play Store Menggunakan Algoritma Random Forest Classifier," *J. Inform. dan Tek. Elektro Terap.*, vol. 12, no. 2, pp. 1204–1210, 2024, doi: 10.23960/jitet.v12i2.4192.
- [2] K. D. Pratama, D. W. Brata, and W. Purnomo, "Analisis Sentimen Ulasan Pengguna Aplikasi Info BMKG pada Google Play Store di Indonesia," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 7, no. 4, pp. 1826–1834, 2023, [Online]. Available: <https://j-ptiik.uib.ac.id/index.php/j-ptiik/article/view/12539%0Ahttps://j-ptiik.uib.ac.id/index.php/j-ptiik/article/download/12539/5722>
- [3] F. K. Ihtada, "No TitleEAENH," *Ayan*, vol. 15, no. 1, pp. 37–48, 2024.
- [4] L. Galke *et al.*, "Are We Really Making Much Progress in Text Classification? A Comparative Review," 2022, [Online]. Available: <http://arxiv.org/abs/2204.03954>
- [5] F. Yulianto, R. Arifando, and A. A. Supianto, "Improved Eliminate Particle Swarm Optimization on Support Vector Machine for Freshwater Fish Classification," *Proc. 2019 4th Int. Conf. Sustain. Inf. Eng. Technol. SIET 2019*, pp. 38–43, 2019, doi: 10.1109/SIET48054.2019.8986119.
- [6] F. Fajri, B. Tutuko, and S. Sukemi, "Membandingkan Nilai Akurasi BERT dan DistilBERT pada Dataset Twitter," *JUSIFO (Jurnal Sist. Informasi)*, vol. 8, no. 2, pp. 71–80, 2022, doi: 10.19109/jusifo.v8i2.13885.
- [7] A. Akbar Hidayat, "Analisis Ekstraksi Fitur Pada Klasifikasi Teks Menggunakan Algoritma K-Nearest Neighbor (Studi Kasus: Berita Hoaks)," 2022.
- [8] G. Darmawan, S. Alam, and M. I. Sulistyono, "Analisis Sentimen Berdasarkan Ulasan Pengguna Aplikasi Mypertamina Pada Google Playstore Menggunakan Metode Naïve Bayes," *STORAGE – J. Ilm. Tek. dan Ilmu Komput.*, vol. 2, no. 3, pp. 100–108, 2023.
- [9] H. B. Assyafah and D. T. Yulianti, "Analisis Dataset menggunakan Sentiment Analysis (Studi Kasus Pada Tripadvisor)," *Strategi*, vol. 3, no. 2, pp. 320–331, 2021, [Online]. Available: <https://mail.strategi.it.maranatha.edu/index.php/strategi/article/view/287>
- [10] M. A. Rohman, "Teknik Ekstraksi Fitur dalam NLP - Sekolah Statistics, Data and Technical Analytics." Accessed: Dec. 01, 2024. [Online]. Available: <https://sekolahstata.com/teknik-ekstraksi-fitur-dalam-nlp/>
- [11] M. Jojoa, P. Eftekhari, B. Nowrouzi-Kia, and B. Garcia-Zapirain, "Natural language processing analysis applied to COVID-19 open-text opinions using a distilBERT model for sentiment categorization," *AI Soc.*, vol. 39, no. 3, pp. 883–890, 2024, doi: 10.1007/s00146-022-01594-w.
- [12] N. Khoirunnisaa, K. Nabila, N. Kesuma, S. Setiawan, A. Yunizar, and P. Yusuf, "KLASIFIKASI TEKS ULASAN APLIKASI NETFLIX PADA GOOGLE PLAY STORE MENGGUNAKAN ALGORITMA NAÏVE BAYES DAN SVM," *SKANIKA Sist. Komput. dan Tek. Inform.*, vol. 7, no. 1, pp. 64–73, Jan. 2024, doi: 10.36080/SKANIKA.V7I1.3138.
- [13] 啓造朝長, "済無No Title No Title," *Angew. Chemie Int. Ed.* 6(11), 951–952., vol. 2016, no. 1, pp. 1–

- 2, 1967.
- [14] A. Ilham, N. Azmi Verdikha, and A. J. Latipah, "Klasifikasi Ujaran Kebencian di Twitter Menggunakan Fitur Ekstraksi Glove dengan Support Vector Machine (SVM)," *J. Explor. IT*, vol. 5, no. 36, pp. 64–72, 2023, [Online]. Available: <https://doi.org/10.35891/explorit>
- [15] R. Q. Wang, F. R. Ouedraogo, and S. S. Pathapati, "Preprocessing," *Comput. Fluid Dyn. Appl. Water, Wastewater Stormwater Treat.*, vol. 3, no. 3, pp. 53–56, 2019, doi: 10.1061/9780784415313.ch06.

BIODATA PENULIS



Penulis Pertama

Reno Ridhoi Merupakan Mahasiswa Di Universitas Muhammadiyah Kalimantan Timur, Jurusan Teknik Informatika.



Penulis Kedua

Naufal Azmi Verdikha, S.Kom., M.Eng Merupakan dosen tetap di Universitas Muhammadiyah Kalimantan Timur



Penulis Ketiga

Fendy Yulianto, S.Kom., M.Kom. Merupakan dosen di Universitas Muhammadiyah Kalimantan Timur