

Customer Review Classification on the MitraShopee Application Using Support Vector Machine and Naïve Bayes Methods

[Klasifikasi Ulasan Pelanggan pada Aplikasi MitraShopee Menggunakan Metode Support Vector Machine dan Naïve Bayes]

Adika Surya Perdana¹⁾, Istian Kriya Almanfaluti²⁾, Alshaf Pebrianggara³⁾, Mochammad Rizal Yulianto⁴⁾

¹⁾Program Studi Bisnis Digital, Universitas Muhammadiyah Sidoarjo, Indonesia

²⁾Program Studi Bisnis Digital, Universitas Muhammadiyah Sidoarjo, Indonesia

³⁾Program Studi Manajemen, Universitas Muhammadiyah Sidoarjo, Indonesia

⁴⁾Program Studi Bisnis Digital, Universitas Muhammadiyah Sidoarjo, Indonesia

*Email Penulis Korespondensi: istian.alman@umsida.ac.id

Abstract. This study aims to classify customer reviews of the MitraShopee application into three categories: criticism, suggestions, and questions using Naïve Bayes and Support Vector Machine (SVM) algorithms. A total of 6,000 reviews were collected from Google Play Store and processed through data cleaning stages such as case folding, tokenizing, stopword removal, and normalization using Python in Google Colab. The cleaned data was then analyzed with TF-IDF and labeled using a rule-based approach. Evaluation results showed that the SVM algorithm performed better with an accuracy of 81%, while Naïve Bayes achieved 56% accuracy. SVM demonstrated a balanced performance in terms of precision and recall, particularly in detecting suggestion and question reviews. Meanwhile, Naïve Bayes tended to be biased toward a single class and struggled to accurately identify criticisms. This research highlights the importance of algorithm selection and thorough preprocessing in text classification. The results are expected to support app developers in quickly and accurately understanding user needs through an automated review classification system.

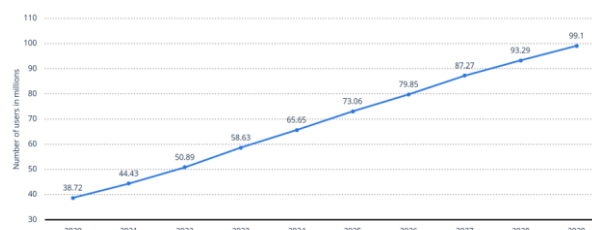
Keywords - Classification, Reviews, Customers, Support Vector Machine, Naïve Bayes

Abstrak. Penelitian ini bertujuan untuk menghasilkan ulasan pelanggan aplikasi MitraShopee ke dalam tiga kategori, yaitu, kritik, saran, dan pertanyaan, menggunakan algoritma Naïve Bayes dan Support Vector Machine (SVM). Sebanyak 6.000 ulasan diambil dari Google Play Store dan diproses melalui tahapan pembersihan data seperti case folding, tokenizing, stopword removal, dan normalisasi menggunakan Python di Google Colab. Data yang telah dibersihkan kemudian dianalisis dengan TF-IDF dan dilabeli berdasarkan pendekatan berbasis aturan. Hasil evaluasi menunjukkan bahwa algoritma SVM memiliki performa lebih unggul dengan akurasi sebesar 81%, sedangkan Naïve Bayes mencatatkan akurasi 56%. SVM menunjukkan keseimbangan yang baik antara precision dan recall, khususnya dalam mendeteksi ulasan saran dan pertanyaan. Sementara itu, Naïve Bayes cenderung bias terhadap satu kelas dan kurang mampu mengenali kritik dengan baik. Penelitian ini menunjukkan pentingnya pemilihan algoritma dan tahap preprocessing dalam klasifikasi teks. Hasil yang diperoleh diharapkan dapat membantu pengembang aplikasi dalam memahami kebutuhan pengguna secara lebih cepat dan tepat sasaran melalui sistem klasifikasi otomatis.

Kata Kunci - Klasifikasi, Ulasan, Pelanggan, Support Vector Machine, Naïve Bayes

I. PENDAHULUAN

E-commerce adalah proses jual beli yang berlangsung secara online melalui berbagai platform elektronik, yang mempertemukan konsumen dengan penjual berinteraksi dan melakukan bisnis tanpa harus bertemu secara fisik. Sumber pada KEMENKOPUKM Smesco mencatat terdapat ada sekitar 16.4 juta pelaku UMKM yang telah bergabung di platform digital, menunjukkan pertumbuhan lebih dari 100% sejak awal pandemi. Banyak pemilik usaha baik skala besar maupun kecil yang memindahkan operasi dalam bisnisnya dari konvensional menjadi digital, terutama pada e-commerce [1].



Gambar 1. Data peningkatan e-commerce di Indonesia

Jumlah pengguna di Indonesia jumlah pengguna *e-commerce* terus meningkat sejak tahun 2020. Pada tahun 2023, tercatat ada sekitar 38.63 juta pengguna, dan diperkirakan angka ini akan terus bertumbuh hingga mencapai 99.1 juta pengguna pada tahun 2029. *E-Commerce* menyediakan jangkauan pasar yang sangat luas tanpa batasan dan memberikan kemudahan transaksi bagi penjual maupun pembeli. Pembeli tanpa harus keluar rumah untuk membeli suatu barang yang dibutuhkan bahkan akses untuk mencari produk dan membandingkan harga dapat dilakukan dengan lebih mudah.



Gambar 2. Jumlah data penipuan transaksi online

Pada Gambar 2. menyajikan data laporan penipuan transaksi online di Indonesia periode tahun 2017 hingga 2022. Terlihat bahwa penipuan transaksi online umum merupakan jenis penipuan yang paling banyak dilaporkan, dengan angka mencapai 405.000 kasus. Dibandungkan dengan penipuan investasi online sebanyak 19.000 kasus dan penipuan jual beli online 12.000 kasus, penipuan transaksi online menunjukkan angka yang jauh lebih tinggi, menandakan perlunya peningkatan kewaspadaan dan upaya pencegahan yang efektif. Dengan banyaknya *e-commerce* saat ini masyarakat bingung dalam menentukan *e-commerce* yang benar dipercaya. Sehingga, perlu adanya penilaian yang dapat mengidentifikasi tingkat popularitas *e-commerce* di Indonesia mengacu pada klasifikasi ulasan dalam Google Play Store yang disebut dalam teks berbentuk pandangan kritik, saran, dan pertanyaan [2].

Berdasarkan hasil survei “Pola Perilaku Masyarakat saat Belanja di Marketplace” dalam survei yang dilakukan oleh GoodStats pada tahun 2022, responden menempatkan Shopee sebagai pilihan *marketplace* kedua mereka. Dengan meningkatnya pengguna pada aplikasi Shopee ini, akhirnya Shopee berinovasi membuat aplikasi baru yaitu MitraShopee. MitraShopee sendiri merupakan platform yang menawarkan jual beli berbagai produk digital yang tersedia untuk diunduh di Google Play Store. Pada platform Google Play Store terdapat berbagai fitur-fitur yang bisa menunjang bagi pengguna seperti peringkat dan ulasan [3].

Memberikan ulasan *online* setelah pembelian menjadi salah satu tahapan penting dalam perjalanan pelanggan. Dengan adanya ulasan, pengguna menjadi lebih mudah dalam memperoleh informasi tentang suatu produk atau layanan, pengguna dapat dengan mudah mempercayai rekomendasi yang telah digunakan oleh pengguna sebelumnya [4]. Belum ditemukan pendekatan yang sistematis dan benar untuk mengidentifikasi apakah sebuah ulasan termasuk dalam kategori kritik, saran, dan pertanyaan. Mengkategorikan ulasan sangat penting bagi pengembang dalam mengambil data dari para ulasan pengguna pada aplikasi guna menutupi keresahan pengguna dalam aplikasi.

Pada Penelitian sebelumnya masih belum ada yang secara spesifik mengkategorikan ulasan pelanggan aplikasi *e-commerce* ke dalam bentuk kritik, saran, dan pertanyaan. Padahal pengelompokkan ini penting untuk membantu pengembang memahami kebutuhan dan keluhan pengguna secara lebih mendalam [5]. Sebagian besar penelitian terdahulu masih berfokus pada analisis sentiment dua kelas, yakni positif dan negatif, tanpa menelusuri lebih jauh bentuk konkret dari ulasan tersebut. Penelitian mengenai klasifikasi ulasan pada aplikasi MitraShopee masih sangat terbatas, sehingga diperlukan pendekatan sistematis untuk mengeksplorasi ulasan pada pengguna secara lebih komprehensif. Dalam konteks ini, belum ada perbandingan yang mendalam terkait performa algoritma Naive Bayes dan *Support Vector Machine* (SVM) dalam mengelompokkan multi-kelas khusus untuk ulasan aplikasi MitraShopee. Penelitian ini diharapkan dapat menjawab kekosongan tersebut melalui penerapan dua metode klasifikasi untuk mengolah ulasan berbasis data dari Google Play Store secara lebih efektif dan efisien.

Metode analisis sentimen yang digunakan dalam memahami kondisi emosional individu apakah itu positif, atau negatif [6]. Analisis sentimen diterapkan pada aplikasi Shopee, tempat dimana sistem memperoleh dataset ulasan dari website Google Play Store system selanjutnya akan mengelompokkan ulasan untuk menilai apakah terdapat sentimen negatif atau positif. Berdasarkan klasifikasi tersebut, pengguna bisa mendapatkan informasi dan mengevaluasi aplikasi Shopee, yang membantu mereka dalam membuat keputusan. Analisis sentimen dalam komentar dan ulasan pada Google Play Store dapat dilakukan menggunakan dua pendekatan populer yaitu *Support Vector Machine* (SVM) dan Naive Bayes. Ulasan pelanggan pada Google Play Store menjadi penting dalam membantu user memilih platform yang terpercaya. Klasifikasi ulasan pelanggan pada aplikasi MitraShopee dengan penerapan metode *Support Vector Machine* dan Naive Bayes menunjukkan potensi besar dalam meningkatkan pengalaman pengguna dan memperbaiki layanan *e-commerce*. Proses pengujian dilakukan dengan membagi data diperuntukkan untuk melatih model, sementara 20% sisanya dimanfaatkan untuk menguji sejauh mana kinerja model yang telah dilatih. Menurut temuan penelitian, algoritma Naive Bayes menunjukkan akurasi sebesar 85%, sementara algoritma *Support Vector Machine*

mendapatkan akurasi 81% dalam analisis sentimen. Dalam analisis sentimen, Naïve Bayes terbukti unggul jika dibandingkan dengan metode SVM [1]. Penelitian tentang analisis sentimen sangat diminati karena topik ini memiliki relevansi dan daya tarik yang signifikan. Di bidang ini, telah dilakukan penelitian tentang sentimen masyarakat terhadap PayLater, yang memanfaatkan algoritma Naïve Bayes Classifier. Pada penelitiannya diperoleh nilai akurasi dengan menggunakan model *confusion matrix* menghasilkan tingkat akurasi sebesar 95%. Penelitian selanjutnya tentang studi klasifikasi *teks review* Tokopedia menggunakan SVM menghasilkan tingkat akurasi 75% [7].

II. LITERATURE REVIEW

E-Commerce

E-commerce atau perdagangan elektronik, yang merupakan fenomena telah mengubah cara transaksi bisnis di seluruh dunia. *E-commerce* meliputi proses pembelian, penjualan, dan pertukaran barang serta jasa melalui platform digital, jaringan computer, terutama internet. *E-commerce* melibatkan transaksi yang melibatkan pembeli dan penjual dengan memanfaatkan teknologi jaringan untuk meningkatkan efisiensi dan efektivitas dalam berbisnis. *E-commerce* juga mencakup berbagai model transaksi untuk memperluas cakupan pasar dan meningkatkan aksesibilitas bagi konsumen. Perkembangan teknologi informasi, khususnya internet, telah mendorong pertumbuhan *e-commerce* secara signifikan, memberikan kemudahan untuk transaksi yang lebih cepat dan menghemat biaya [8].

Klasifikasi Ulasan

Klasifikasi ulasan pelanggan dalam analisis sentiment merupakan pendekatan yang krusial untuk memahami opini dan pengalaman konsumen terhadap produk di *platform e-commerce*. Proses ini melibatkan pengkategorian ulasan menjadi beberapa kelas, seperti positif, negatif, dan netral, memungkinkan pemilik bisnis untuk mengevaluasi persepsi pelanggan secara lebih efektif. Klasifikasi ulasan tidak hanya membantu dalam memahami reaksi pelanggan tetapi juga berkontribusi pada pengambilan keputusan strategis yang lebih baik bagi perusahaan dalam meningkatkan produk dan layanannya [9].

Naïve Bayes

Sebagai algoritma klasifikasi probalistik, Naïve Bayes efektif dan sering digunakan dalam data mining, terutama untuk mengelompokkan teks, analisis sentiment dan deteksi spam. Naïve Bayes bekerja berdasarkan dengan asumsi bahwa setiap fitur dalam dataset bersifat independen terhadap fitur lainnya. Meskipun asumsi ini jarang sepenuhnya terpenuhi dalam praktik, pendekatan ini tetap efektif dalam berbagai kasus klasifikasi. Keunggulan Naïve Bayes terletak pada kecepatan komputasi dan efisiensi dalam menangani dataset besar dengan jumlah fitur yang tinggi. Naïve Bayes dapat diterapkan dalam berbagai bidang, seperti prediksi penyakit, klasifikasi dokumen, dan sistem rekomendasi [10].

Support Vector Machine (SVM)

Support Vector Machine merupakan metode pembagian mesin yang digunakan untuk klasifikasi dan regresi dengan pendekatan berbasis margin maksimum. SVM bekerja melalui pencairan *hyperplane* optimal untuk dapat dengan tujuan memisahkan kelas dengan margin yang maksimal, sehingga meningkatkan generalisasi sistem. Keunggulan utamanya terdapat dalam hal kemampuannya untuk memproses data berdimensi besar dengan fleksibilitasnya pada menangani kasus nonlinear melalui penggunaan fungsi kernel. Dengan optimisasi parameter yang tepat dan pemilihan kernel yang sesuai, *Support Vector Machine* tersebut sebagai salah satu algoritma yang populer digunakan pada penelitian dan aplikasi industri [11].

Term Frequency-Inverse Document (TF-IDF)

TF-IDF adalah sebuah teknik pembobotan yang banyak dimanfaatkan dalam pemrosesan teks dan klasifikasi dokumen. TF-IDF mengintegrasikan dua bagian penting, yaitu *Term Frequency* yang mengukur berapa sering kata muncul dalam sebuah dokumen, serta *Inverse Document Frequency*, yang mengukur frekuensi kemunculan kata dalam sebuah koleksi dokumen. Kedua komponen ini bertujuan untuk memberikan nilai lebih pada kata-kata yang sering muncul dalam satu dokumen, namun jarang muncul di dokumen lainnya. Itulah yang dilakukan oleh TF-IDF sangat efektif untuk menghilangkan kata umum yang sering ditemukan di hampir semua file, tetapi tidak memiliki nilai dalam klasifikasi [10].

III. METODE

Dalam penelitian ini menggunakan 6.000 kumpulan data yang mencakup ulasan yang bersumber dari aplikasi *e-commerce* MitraShopee yang dihosting di platform Google Play Store, yang kemudian dikategorikan menjadi tiga kelompok berbeda yaitu kritik, saran, dan pertanyaan. Pemilihan jumlah data menggunakan metode *Time Series* diambil dari rentan waktu 09 Februari 2024 sampai dengan 14 April 2025 [12]. Pengambilan data didasarkan pada pertimbangan metodologis dan acuan dari studi-studi terdahulu dalam bidang *Natural Language Processing* (NLP) dan machine learning. Dalam literatur lain disebutkan bahwa dataset dengan jumlah lebih dari 1.000 hingga 5.000 sampel sudah mencukupi untuk membangun model klasifikasi teks yang valid, dengan catatan data telah melalui tahap pembersihan dan pelabelan yang baik [13]. Dalam kelas kritik mencakup ulasan yang mengekspresikan ketidakpuasan atau masalah yang dihadapi pengguna saat menggunakan aplikasi. Ini bisa berhubungan dengan performa aplikasi, fitur yang tidak berfungsi, atau pengalaman pengguna yang kurang memuaskan. Kelas Saran berisi masukan konstruktif dari pengguna tentang bagaimana aplikasi dapat ditingkatkan. Pengguna memberi ide atau rekomendasi yang dapat membantu pengembang dalam meningkatkan kualitas aplikasi [14]. Sedangkan, kelas pertanyaan mencakup ulasan yang menunjukkan keingintahuan pengguna tentang fitur atau aspek tertentu dari aplikasi. Ini sering kali berkaitan dengan cara penggunaan aplikasi atau klarifikasi mengenai kebijakan tertentu. Tahap alur penelitian digambarkan pada gambar 3.



Gambar 3. Tahapan penelitian

3.1 Pengumpulan Data (*Scrapping*)

Proses pengumpulan data Penelitian menggunakan 6.000 kumpulan data yang terdiri dari data dari kelas kritik, saran, pertanyaan, lainnya, dan data untuk data testing yang diperoleh dari ulasan aplikasi MitraShopee dimana dapat diakses di Google Play Store melalui *Google Play Scrapping*. Langkah berikut melibatkan *pre-processing*. Setelah menyelesaikan pembersihan kumpulan data yang tidak dimurnikan, fase berikutnya memerlukan pelabelan untuk menggambarkan kategori kritik, saran, dan pertanyaan. Setelah ini, prosedur pembobotan *Term Frequency-Inverse Document Frequency* dijalankan dalam menetapkan signifikan disetiap istilah yang ada pada dokumen. Teknik pengelompokan diterapkan memanfaatkan algoritma SVM dan Naïve Bayes.

Tabel 1. Contoh hasil pengumpulan data

ID	UserName	Score	At	Comment
1	Penggunaan Google	5	11/17/2022 1:43	Cukup bagus apk nya bpjs
2	Penggunaan Google	5	11/17/2022 1:42	Sangat cepat dan mudah
3	Penggunaan Google	1	11/17/2022 1:42	Bintang setengan ada ga?
...
279876	Penggunaan Google	5	2/20/2018 7:32	Mantap gannn

3.2 *Pre-processing* Data

Di bagian *pre-processing*, kalimat harus disesuaikan untuk menjamin perkembangan data yang lancar ke fase berikutnya. Proses pengolahan teks melibatkan *case folding*, *tokenizing*, *stopwords* serta *normalization*.

3.2.1 Case Folding

Case Folding merupakan sebuah metode untuk menyesuaikan huruf besar menjadi huruf kecil atau dalam bentuk yang konsisten. Proses ini memudahkan pencarian, mengingat tidak semua dokumen teks mengikuti aturan penggunaan huruf kapital yang konsisten [15].

Tabel 2. Contoh hasil pengumpulan data

Sebelum	Sesudah
Cukup bagus apk nya bpjs	cukup bagus apk nya bpjs
Sangat cepat dan mudah	sangat cepat dan mudah
Bintang setengah ada ga?	bintang setengah ada ga?
Cucokk mantep	cucokk mantep
Mantap gannn	mantap gannn

3.2.2 Tokenizing

Proses ini memecah kalimat menjadi kata-kata dengan membaginya berdasarkan spasi antar kata [15].

Tabel 3. Contoh data yang telah melalui proses tokenisasi

Sebelum	Sesudah
cukup bagus apk nya bpjs	[cukup, bagus, apk, nya, bpjs]
sangat cepat dan mudah	[sangat, cepat, dan, mudah]
bintang setengah ada ga?	[bintang, setengah, ada, ga]
cucokk mantep	[cucokk, mantep]
mantap gannn	[mantap, gannn]

3.2.3 Stopword Removal

Stopword removal, juga dikenal sebagai *filtering*, merupakan langkah pemilihan kata-kata utama yang muncul dari hasil tokenisasi. Pada Langkah ini, dipilih kata-kata yang mewakili isi dokumen [16].

Tabel 4. Contoh hasil *stopwords*

Sebelum	Sesudah
[cukup, bagus, apk, nya, bpjs]	[bagus, apk, bpjs]
[sangat, cepat, dan, mudah]	[cepat, mudah]
[bintang, setengah, ada, ga]	[bintang, ga]
[cucokk, mantep]	[cucok, mantep]
[mantap, gannn]	[mantap]

3.2.4 Normalization

Normalisasi dilakukan dengan cara menghapus kata-kata yang memiliki imbuhan atau tambahan, serta memperbaiki kata yang salah ejaannya menjadi yang benar [16].

Tabel 5. Contoh proses normalisasi

Sebelum	Sesudah
[bagus, apk, bpjs]	bagus aplikasi bpjs
[cepat, mudah]	cepat mudah
[bintang, ga]	bintang ga
[cucok, mantep]	cucok mantep
[mantap]	mantap

3.3 Pembobotan atau Pelabelan Kata

3.3.1 Term Frequency (TF)

Proses ini menghitung frekuensi kemunculan kata dalam kumpulan data [17]. Frekuensi data dapat dihitung menggunakan rumus:

$$TF(t, d) = \frac{\text{jumlah kemunculan data dalam dokumen}}{\text{total kata dalam dokumen}} \quad (1)$$

Keterangan :

t = term (kata tertentu yang sedang diproses frekuensinya)

d = dokumen (dokumen dimana kata itu muncul)

3.3.2 Inverse Document Frequency (IDF)

Proses ini menghitung jumlah dokumen yang mengandung istilah kunci tertentu pada dataset [18]. Jumlah dokumen yang berisi kata kunci dapat dihitung dengan rumus berikut:

$$IDF(t, d) = \left(\frac{N}{df} + 1\right) \quad (2)$$

Keterangan :

N = Jumlah semua dokumen

Df = Jumlah kata pada dokumen

3.3.3 Confusion Matrix

Confusion matrix merupakan sebuah matriks 3x3 yang bertujuan untuk mengevaluasi kinerja model dengan membandingkan prediksi model terhadap kata actual. *Confusion matrix* menghasilkan nilai akurasi, presisi, recall dan F1-Score dihitung dan ditampilkan dalam format presentase untuk menunjukkan performa model [17]. Tabel berikut menunjukkan *confusion matrix*:

Tabel 6. Confusion Matrix

Data Aktual	Data Prediksi		
	Kritik	Saran	Pertanyaan
Kritik	True Kritik	True Saran	False Pertanyaan
Saran	False Kritik	True Saran	False Pertanyaan
Pertanyaan	False Kritik	False Saran	True Pertanyaan

$$Accuracy = \frac{TK+TS+TP}{TK+FP+FS+FK+TP+TS} \quad (3)$$

Keterangan :

TK = True Kritik

FK = False Kritik

FP = False Pertanyaan

TP = True Pertanyaan

FS = False Sara

TS = True Saran

3.4 Klasifikasi Support Vector Machine (SVM)

Support Vector Machine merupakan algoritma klasifikasi dimana sangat populer dalam keperluan pengelompokan dan analisis data. *Support Vector Machine* digunakan untuk analisis data dan pengenalan pola. Salah satu dasar *Support Vector Machine*, *linear classifier*, hanya berfokus pada dua kelas, namun umumnya masalah yang ditemukan di dunia nyata bersifat *multiclass*. Kemudian dikembangkan untuk menangani masalah non-linear yang diselesaikan menggunakan trik kernel dan memetakan data ke ruang yang lebih besar. Algoritma SVM menggunakan hyperplane untuk membedakan antara senitmen positif dan negatif [19]. Metode perhitungan *Support Vector Machine*:

$$(w.x_i) + b = 0 \quad (4)$$

Data x_i , yang termasuk kelas -1 dapat dirumuskan sebagai berikut [20]

$$(w.x_i + b) \leq -1, y_i = -1 \quad (5)$$

Rumus untuk data x_i pada kelas +1 adalah sebagai berikut [21]

$$(w.x_i + b) \geq 1, y_i = 1 \quad (6)$$

3.5 Naïve Bayes

Algoritma yang diterapkan pada proses pengelompokan ini merupakan Naïve Bayes, tepatnya varian Naïve Bayes Multinomial, yang memang banyak digunakan dalam pemrosesan teks dan pengelompokan data berbasis kata. Naïve Bayes bekerja berdasarkan prinsip probabilitas, di mana Naïve Bayes memprediksi kelas dari sebuah data berdasarkan kemungkinan (probabilitas) suatu kata muncul dalam setiap kelas yang telah ditentukan. Cara perhitungan dengan Naïve Bayes [17]:

$$P(H|X) = \frac{P(H|X) P(H)}{P(X)!} \quad (7)$$

Keterangan :

X = Data kelas tidak diketahui

H = Data Hipotesis X adalah kelas yang terpisah

$P(H|X)$ = Probabilitas hipotesis kondisional H

$P(H)$ = Probabilitas hipotesis H

$P(X|H)$ = Probabilitas X

IV. HASIL DAN PEMBAHASAN

Untuk mendukung penelitian ini, sebanyak 6.000 ulasan dari user aplikasi MitraShopee mengambil langsung dari platform Google Play Store menggunakan Google Play Scraper digunakan sebagai sumber data utama. Seluruh data kemudian diproses dan dianalisis menggunakan bahasa pemrograman Python. Proses pengolahan ini dilakukan di Google Colab karena lebih praktis dan fleksibel, terutama untuk kebutuhan analisis berbasis *machine learning*.

4.1 Pengumpulan Data

Total ada sekitar 6.000 ulasan yang berhasil dikumpulkan dan digunakan sebagai bahan utama untuk dianalisis. Ulasan-ulasan ini mencerminkan berbagai pengalaman, keluhan, pertanyaan, hingga saran yang diberikan langsung oleh para pengguna aplikasi.

Tabel 7. Isi dataset

No	Rating	Review Text
0	5	bagus, harga kompetitif semoga...
1	5	saangat baik dan bagus
2	5	bisa lebih cepat waktu transaksi...
3	5	mantul
4	5	kemudahan dan kenyamanan...

Pengambilan data yang terbaru, agar informasi yang didapat bersifat beragam dan tidak bias terhadap waktu atau tren tertentu. Semua data yang sukses dikumpulkan kemudian disimpan sebagai file CSV, sehingga lebih mudah untuk diproses lebih lanjut dalam tahap analisis.

4.2 Pre-processing Data

Proses *Pre-processing* sebagai langkah pembersihan dan menyiapkan data supaya lebih dapat langsung dimanfaatkan pada pemodelan machine learning. Tahapan ini penting karena data teks mentah biasanya masih mengandung banyak unsur yang tidak relevan atau menyulitkan saat analisis dilakukan.

4.2.1 Case Folding

Semua karakter huruf dalam teks ulasan dikonversi ke huruf kecil (*lowercase*), dalam menyamakan bentuk kata, sehingga kata dianggap sama dan tidak diproses dua entitas berbeda.

Tabel 8. Hasil Case Folding

No	Before	After
0	bagus harga kompetitif semoga...	[bagus, harga, kompetitif, ...]
1	Saangat baik dan bagus	[saangat, baik, dan, bagus]
2	Bisa lebih cepat waktu...	[bisa, lebih, cepat, waktu, ...]
3	mantul	[mantul]
4	kemudahan dan kenyamanan...	[kemudahan, dan, kenyamanan...]

Pada tahap *case folding*, seluruh huruf dalam teks ulasan diubah menjadi huruf kecil. Dengan tujuan agar sistem bisa membaca kata secara konsisten, misalnya “Bagus” dan “bagus” dianggap sama dan tidak dihitung sebagai dua kata berbeda. Ini penting untuk menghindari perhitungan ganda akibat perbedaan penulisan huruf kapital, sehingga analisis data menjadi lebih akurat dan konsisten.

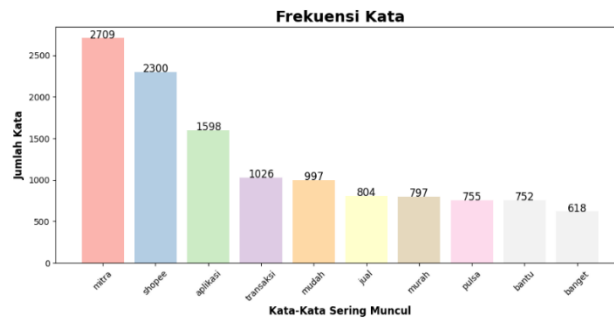
4.2.2 Tokenizing

Proses memecah teks panjang menjadi potongan-potongan kata atau token. Tahap ini untuk mempersiapkan data agar bisa dianalisis per kata.

Tabel 9. Tokenizing

No	Before	After
0	bagus harga kompetitif semoga...	[bagus, harga, kompetitif, ...]
1	Saangat baik dan bagus	[saangat, baik, dan, bagus]
2	Bisa lebih cepat waktu...	[bisa, lebih, cepat, waktu, ...]
3	mantul	[mantul]
4	kemudahan dan kenyamanan...	[kemudahan, dan, kenyamanan...]

Proses tokenisasi memudahkan sistem untuk menganalisis dan mengelola data dalam bentuk unit kata yang lebih terstruktur dan dapat diproses lebih lanjut.



Gambar 5. Frekuensi kata

Beberapa kata lain yang cukup mencolok seperti “transaksi”, “mudah”, “bagus”, “untung”, “murah”, dan “cepat” mengindikasikan bahwa banyak pengguna merasa puas dengan kemudahan dan keuntungan dalam menggunakan MitraShopee, terutama dari segi transaksi dan harga. Kata-kata positif seperti “mantap”, “alhamdulillah”, dan “puas” juga muncul cukup sering, yang bisa mencerminkan tingkat kepuasan pengguna terhadap layanan yang diberikan.

Kemunculan kata seperti “username”, “masuk”, atau “saldo” juga bisa menjadikannya masalah teknis atau kendala *login* yang cukup sering dibicarakan, sehingga dapat menjadi perhatian khusus untuk pengembangan aplikasi kedepannya. *Wordcloud* memberikan gambaran cepat dan *intuited* mengenai topik yang paling sering dibahas pengguna serta nada umum dari ulasan mereka apakah cenderung positif, netral, atau mengandung keluhan.

4.4 Pelabelan Data

Tabel 12. Hasil pelabelan data

No	Review Text	Label
0	saya kasih bintang	Lainnya
1	mohon struk untuk admin bank dan admin loket ...	Saran
2	ni yang nawarin gk jelas langsung ilang biking gw...	Kritik
3	mengapa saldo mitra di tutup...	Pertanyaan
...
6000	sangat rekomendasi	Lainnya

Pada tahap selanjutnya adalah merupakan pelabelan data terhadap ulasan yang telah dikumpulkan dari Google Play Store pada aplikasi MitraShopee. Dari total 6.000 data, setiap ulasan dianalisis secara otomatis menggunakan Google Colab. Dengan memanfaatkan pendekatan berbasis aturan (*rule-based-classification*), setiap ulasan yang telah dibersihkan dianalisis menggunakan kata kunci tertentu yang telah didefinisikan sebelumnya untuk mengidentifikasi konteks dari isi ulasan dan diklasifikasikan ke dalam tiga kategori utama, yaitu: kritik, saran, dan pertanyaan. Label lainnya diberikan pada ulasan yang bersifat umum, pujian, atau tidak mengandung elemen dari ketiga kategori utama.

Dalam membedakan antara kategori kritik, saran, dan pertanyaan pada ulasan pengguna, terdapat beberapa pendekatan yang digunakan baik secara *linguistic* maupun teknis. Kritik umumnya ditandai dengan ekspresi ketidakpuasan yang mengandung kata-kata bernada negatif seperti “gagal”, “lemot”, atau “jelek”. Label saran sering kali muncul dalam bentuk kalimat usulan atau harapan dengan kata-kata seperti “sebaiknya”, “lebih baik”, “mohon”, atau “diharapkan”, yang terlihat netral hingga positif. Pada label pertanyaan memiliki struktur yang lebih eksplisit, biasanya dimulai dengan kata tanya seperti “apa”, “bagaimana”, atau “kenapa”. Label lainnya digunakan untuk mengelompokkan ulasan yang tidak termasuk ke dalam kategori kritik, saran, atau pertanyaan. Label ini mencakup komentar seperti “terima kasih” atau “bagus”, pernyataan yang singkat atau ambigu, hingga komentar yang tidak relevan. Dalam pendekatan *Natural Language Processing* (NLP) model klasifikasi dapat mempelajari pola-pola ini melalui representasi teks seperti TF-IDF, frekuensi kata kunci, jenis kata yang digunakan, serta nada atau emosi yang terkandung dalam ulasan [22].

Dari hasil pelabelan data ulasan terdapat 254 ulasan yang tergolong kritik, 372 saran, 728 pertanyaan, dan 3.711 label lainnya. Terlihat bahwa mayoritas ulasan pengguna masih termasuk dalam kategori lainnya, yang berarti banyak pengguna lebih cenderung memberikan *testimoni* umum atau pujian tanpa menyampaikan masukan spesifik. Namun demikian, ditemukan juga sejumlah ulasan yang memuat kritik dan saran konstruktif, serta beberapa pertanyaan yang menandakan adanya kebutuhan akan informasi lebih lanjut dari pihak pengembang [23].

4.5 Evaluasi Model

Tabel 13. Evaluasi model SVM

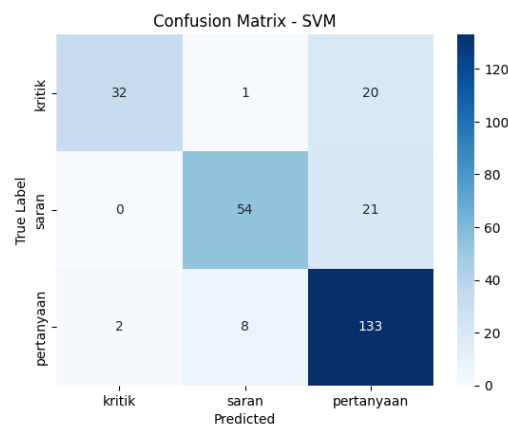
	Precision	Recall	F1-Score	Support
Kritik	0.94	0.60	0.74	53
Saran	0.76	0.93	0.84	143
Pertanyaan	0.86	0.72	0.78	75
Accuracy			0.81	271
Macro avg	0.85	0.75	0.79	271
Weighted avg	0.82	0.81	0.80	271

Pembagian data dilakukan menggunakan pendekatan *Hold-Out Validation* dengan proporsi 80% untuk data latih dan 20% untuk data uji [24]. Dari total 1.354 ulasan yang telah diklasifikasikan ke dalam kategori kritik, saran, dan pertanyaan, sebanyak 271 data dialokasikan sebagai data uji. Pemilihan proporsi ini didasarkan pada praktik umum dalam bidang *machine learning*, di mana 20% data uji dianggap cukup *representative* untuk mengukur performa model secara objektif tanpa mengurangi jumlah data latih yang diperlukan untuk membangun model secara optimal.

Berdasarkan pengelompokan menggunakan SVM menghasilkan akurasi sebesar 80.8%, yang berarti bahwa dari seluruh data uji yang dipakai (sebanyak 271 ulasan), sekitar 81% dapat diklasifikasikan secara tepat ke dalam tiga kategori yang telah ditentukan, yaitu kritik, saran, dan pertanyaan. Secara lebih rinci, performa model dapat dijelaskan melalui metrik evaluasi.

Kategori kritik memiliki nilai *precision* sebesar 0.94, yang mengindikasikan bahwa model memiliki kinerja baik dalam mengenali ulasan yang sebenarnya tergolong kritik. Namun, nilai *recall* untuk kategori ini berada di angka 0.60, yang artinya masih ada sejumlah kritik yang tidak terdeteksi dengan baik oleh model [25]. Kategori saran menunjukkan performa paling stabil, dengan nilai *precision* 0.76 dan *recall* 0.93. Ini berarti Sebagian besar ulasan berupa saran berhasil diklasifikasikan dengan baik. Nilai *F1-Score* yang tinggi, yakni 0.84, memperkuat bahwa model cukup konsisten dalam mengenali saran dari pengguna. Kategori pertanyaan memperoleh nilai *precision* 0.86 dan *recall* 0.72 ini mengindikasikan bahwa model dapat mengidentifikasi pertanyaan dengan cukup baik, meskipun ada beberapa pertanyaan yang masih keliru diklasifikasikan [26].

Performa model Support Vector Machine (SVM) dapat dikatakan cukup optimal dalam memahami dan mengklasifikasikan bentuk komunikasi pengguna dalam ulasan aplikasi MitraShopee. Hasil ini juga menunjukkan bahwa pendekatan klasifikasi berbasis *machine learning* mampu membantu dalam menyaring masukan dari pengguna secara otomatis dan efisien.



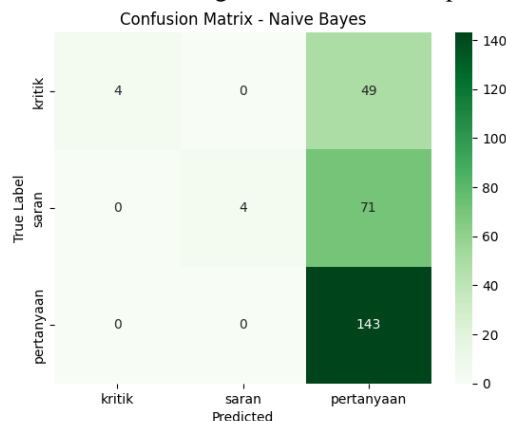
Gambar 6. Confusion Matrix SVM

Pada gambar di atas terlihat bahwa model SVM cukup mampu membedakan antara kategori ulasan. Untuk kelas kritik, dari total 53 data. Sebanyak 32 data berhasil diklasifikasikan dengan benar, sementara 20 data salah prediksi sebagai pertanyaan, dan 1 data salah dikira saran. Ini menunjukkan adanya sedikit kebingungan model dalam membedakan kritik dan pertanyaan, meski sebagian besar prediksi sudah tepat. Untuk kelas saran, performa model sangat baik dengan 54 data terklasifikasi benar, dan hanya 21 yang keliru diprediksi sebagai pertanyaan [27]. Tidak ada data saran yang salah dikira kritik, yang menunjukkan konsistensi prediksi pada kelas ini. Pada kelas pertanyaan, dari total 143 data, model berhasil memprediksi dengan sangat baik sebanyak 133 data terklasifikasi benar, hanya 10 sisanya tersebar salah ke kelas lain.

Tabel 14. Evaluasi model Naïve Bayes

	Precision	Recall	F1-Score	Support
Kritik	1.00	0.08	0.14	53
Saran	0.54	1.00	0.70	143
Pertanyaan	1.00	0.05	0.10	75
Accuracy			0.56	271
Macro avg	0.85	0.38	0.32	271
Weighted avg	0.76	0.56	0.43	271

Pada tahap evaluasi model klasifikasi melalui penerapan algoritma Naïve Bayes, akurasi yang diperoleh mencapai 55.72% dari total 271 data uji. Hasil ini mengindikasikan bahwa lebih dari setengah data berhasil diklasifikasikan dengan benar oleh model, namun masih terdapat kelemahan signifikan dalam mengenali sebagian kelas. Berdasarkan kategori label, model menunjukkan performa terbaik pada kelas saran. Ini berarti model sangat efektif dalam mengenali hampir semua data yang tergolong sebagai saran. Namun, sebaliknya performa model untuk kategori kritik dan pertanyaan menunjukkan hasil yang kurang memuaskan. Meskipun *precision* untuk kedua kategori ini tinggi (1.00), hal ini menunjukkan bahwa model hanya memprediksi sangat sedikit data sebagai kritik dan pertanyaan, namun prediksi tersebut benar. Model tidak sensitif dalam mengenali data kritik dan pertanyaan secara menyeluruh [28].

**Gambar 7.** Confusion matrix Naïve Bayes

Pada gambar di atas terlihat bahwa model ini kurang mampu membedakan antara kategori ulasan. Untuk kelas kritik, hanya 4 data yang benar, dan 49 sisanya justru salah dikira sebagai pertanyaan. Pada kelas saran juga hanya 4 yang benar, sedangkan 71 juga diprediksi salah ke kelas pertanyaan. Meskipun terlihat seolah-olah akurat di kelas pertanyaan (143 benar), tapi ini bukan karena modelnya benar. Akurasinya hanya 56%, dan *recall* untuk kritik dan saran sangat rendah. Performa Naïve Bayes kurang cocok untuk klasifikasi ulasan MitraShopee yang kelasnya tidak seimbang [29].

V. SIMPULAN

Penggunaan algoritma klasifikasi SVM dan Naïve Bayes telah berhasil diimplementasikan dalam mengelompokkan ulasan pengguna aplikasi MitraShopee ke dalam tiga kategori utama, yaitu kritik, saran, dan pertanyaan. Pada penelitian ini menggunakan data training 271 data ulasan, dengan tujuan sebagai upaya mengetahui model yang paling efektif dalam mengolah dan memahami masukan dari pengguna secara otomatis. Berdasarkan hasil evaluasi algoritma SVM mencatatkan performa terbaik dengan capaian akurasi sebesar 81%, *precision* 94% pada kelas kritik, *recall* tertinggi 93% pada kelas saran. Hal ini menunjukkan bahwa SVM cukup seimbang dalam mendeteksi tiap kategori dan mampu mengurangi kesalahan klasifikasi secara signifikan. Algoritma Naïve Bayes hanya mencatatkan akurasi sebesar 56%, dengan *precision* tinggi tetapi *recall* yang sangat rendah, terutama pada kelas kritik dan pertanyaan. Klasifikasi ini tampak bias terhadap satu kategori, yaitu pertanyaan, yang menyebabkan penurunan kualitas prediksi secara keseluruhan.

Dapat disimpulkan bahwa pemilihan algoritma dan proses preprocessing data sangat krusial dalam klasifikasi teks. Setiap model memiliki sensitivitas yang berbeda terhadap karakteristik data, sehingga penyesuaian metode pra-proses seperti tokenisasi, pembersihan teks, dan representasi fitur menjadi hal yang tidak boleh diabaikan. Diharapkan hasil penelitian ini dapat menjadi landasan untuk pengembangan sistem klasifikasi otomatis kedepannya, khususnya untuk menganalisis masukan pengguna dalam platform digital seperti MitraShopee secara lebih cepat dan tepat sasaran.

UCAPAN TERIMA KASIH

Ucapan terimakasih penulis sampaikan kepada pengguna aplikasi MitraShopee yang telah membantu dan memberikan ulasan pada Google Play Store dan membantu dalam penyelesaian artikel ini. Peneliti juga sampaikan Ucapan terimakasih kepada Universitas Muhammadiyah Sidoarjo, Fakultas Bisnis, Hukum, dan Ilmu Sosial (FBHIS), serta Program Studi Bisnis Digital.

REFERENSI

- [1] Saifudin, and M. Hadziq Affan, "Analisa Pemasaran Platform P2p Lending Syariah Pada Pt Alami Fintek Sharia," *Jurnal Ekonomi dan Perbankan Syariah*, vol. 8, no. 1, pp. 436–448, 2023, doi: 10.30651/jms.v8i1.17706.
- [2] Febby Adelia Irawan, Aldy Rialdy Atmadja, and Agung Wahana, "Analisis Sentimen Ulasan Aplikasi Bank Digital Menggunakan Algoritma Naive Bayes," *Journal of Computer Science and Information Technology*, vol. 4, no. 2, pp. 60–68, 2024, doi: 10.31949/infotech.v10i2.11801
- [3] Mahda Nurhamidah, "Pengaruh Complaint Handling Bank Syariah Indonesia Terhadap Kepuasan dan Loyalitas Nasabah Ex BNI Syariah dan Ex BRI Syariah," July 2024, repository: <https://repository.uinjkt.ac.id/dspace/handle/123456789/62339>
- [4] Amalia Elma Sari, Sri Widowati, and Kemas Muslim Lhaksmana, "Klasifikasi Ulasan Pengguna Aplikasi Mandiri Online di Google Play Store dengan Menggunakan Metode Information Gain dan Naive Bayes Classifier," *e-Proceeding of Engineering*, vol. 6, no. 2, pp. 9143–9157, 2019.
- [5] Z. S. Li, N. N. Arony, K. Devathanan, M. Sihag, N. Ernst, and D. Damian, "Unveiling the Life Cycle of User Feedback: Best Practices from Software Practitioners," Sep. 2023, [Online]. Available: <http://arxiv.org/abs/2309.07345>
- [6] M. Qamal and W. Fuadi, "Analisis Sentimen Toko Online Menggunakan Algoritma Naive Bayes Classifier," pp. 641–650.
- [7] Ipan Saepul Milal, M. Hasanudin, M. Aliffiallathifa Nur Azhari, Rifki Aditya Nugraha, Nova Agustina, and Sri Erina Damayanti, "Klasifikasi Teks Review Pada E-Commerce Tokopedia Menggunakan Algoritma SVM," *Jurnal Ilmiah Nasional Riset Aplikasi dan Teknik Informatika* vol. 05, no. 01, 2023, doi: 10.53580/naratif.v5i1.191.
- [8] Albert Lodewyk Sentosa Siahaan, dkk, "E-Commerce," ISBN: 978–623–151–879–8, no. 225/JTE/2021, November 2023.
- [9] Prasyadi Wibawa Rahayu, dkk, "Buku Ajar Data Mining," ISBN: 978–623–8483–96–9, no. 006/JBI/2023, Januari 2024.
- [10] Detty Purnamasari, dkk, "Pengantar Metode Analisis Sentimen," 2023.
- [11] Laurensia Simanihuruk, and Hari Suparwito, "Long Short-Term Memory and Bidirectional Long Short-Term Memory Algorithms for Sentiment Analysis of Skintific Product Reviews," *ITM Web Conf.*, vol. 71, 2025, doi: 10.1051/itmconf/20257101016.
- [12] Pooja Saigal, "Support Vector Machines Evolution and Applications," ISBN: 978-1-53618-865-3, 2021.
- [13] Poomrape Poomka, Nittaya Kerdprasop, and Kittisak Kerdprasop, "Machine Learning Versus Deep Learning Performances on the Sentiment Analysis of Product Reviews," *International Journal of Machine Learning and Computing*, vol. 11, no. 02, pp 103-109, 2021, doi: 10.18178/ijmlc.2021.11.2.1021
- [14] Alshaf Pebrianggara, Istian Almanfaluti and Wahyu Karobby, "Design Thinking For Business," ISBN: 78-623-464-077-9, no. 218/JTI/2019, September 2023
- [15] Irna Putri Rahayu, Ahmad Fauzi, and Jamaludin Indra, "Analisis Sentimen Terhadap Program Kampus Merdeka Menggunakan Naive Bayes Dan Support Vector Machine," *Jurnal Sistem Komputer dan Informatika (JSON)*, vol. 4, no. 2, p. 296, Dec. 2022, doi: 10.30865/json.v4i2.5381.
- [16] Ayu Sri Rahayu, Ahmad Fauzi, and Rahmat, "Komparasi Algoritma Naive Bayes Dan Support Vector Machine (SVM) Pada Analisis Sentimen Spotify," *Jurnal Sistem Komputer dan Informatika (JSON)*, vol. 4, no. 2, p. 349, Dec. 2022, doi: 10.30865/json.v4i2.5398.
- [17] Nurhaliza Agustina. C.A, Desy Herlina Citra, dkk, "Implementasi Algoritma Naive Bayes untuk Analisis Sentimen Ulasan Shopee pada Google Play Store," *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 2, pp. 47–54, 2022, doi: 10.57152/malcom.v2i1.195.
- [18] Ahyawa Aulia Khoirin Nisa, "Pengaruh Promosi, Harga, Kualitas Layanan, Kualitas Produk Terhadap Kepuasan Pelanggan Shopee: Sebuah Kajian Konseptual," *Jurnal Riset Manajemen*, vol. 1, no. 3, pp. 315–328, Sep. 2023, doi: 10.54066/jurma.v1i3.899.

- [19] Elisa Febriyani, and Herny Februriyanti, “Analisis Sentimen Terhadap Program Kampus Merdeka Menggunakan Algoritma Naive Bayes Classifier Di Twitter,” *Jurnal Tekno Kompak*, vol. 17, no. 1, pp. 25–38, 2023, doi: 10.33365/jtk.v17i1.2061
- [20] Wachid Darmawan, Muhammad Faizal Kurniawan, Wahyu Setianto, and Wim Hapsoro, “Analisis Sentimen Penerapan Kurikulum Merdeka Pada Pengguna Twitter Menggunakan Metode K-Nearest Neighbor Dengan Forward Selection,” *Jurnal Smart Comp*, vol. 12, no. 1, pp. 245–253, 2023, doi: 10.30591/smartcomp.v12i1.4634
- [21] Hennie Tuhuteru, “Analisis Sentimen Masyarakat Terhadap Pembatasan Sosial Berskala Besar Menggunakan Algoritma Support Vector Machine,” *Journal Information System Development*, vol. 5, no. 2, pp. 7–13, 2020.
- [22] Muhammad Ibnu Alfarizi, Lailis Syaafaah, and Merinda Lestandy, “*Emotional Text Classification Using TF-IDF (Term Frequency-Inverse Document Frequency) And LSTM (Long Short-Term Memory)*,” *JUITA : Jurnal Informatika*, vol. 10, no. 2, pp. 225-232, 2022, doi: 10.30595/juita.v10i2.13262
- [23] Muhammad Mujahid, Erol Kina, dkk, “*Data Oversampling and Imbalanced Datasets: As Investigation Of Performance for Machine Learning and Feature Engineering*,” *Journal of Big Data*, vol. 11. no. 87, 2024, doi: 10.1186/s40537-024-00943-4
- [24] Zahra Bami, Ali Behnampour, and Hassan Doosti, “*A New Flexible Train-Test Split Algorithm, an Approach For Choosing Among the Hold-out, K-fold Cross-validation, and Hold-out Iteration*,” vol. 01, 2025, doi: 10.48850/arXiv.2501.06492
- [25] I Gede Bintang Arya Budaya, and I Ketut Putu Suniantara, “*Comparison of Sentiment Analysis Algorithms with SMOTE Oversampling and TF-IDF Implementation on Google Reviews for Public Health Centers*,” *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 4, iss. 3, pp. 1077-1086, 2024, doi: 10.57152/Malcom.v4i3.1459
- [26] Amjad Iqbal, Rashid Amin, dkk, “*Sentiment Analysis of Consumer Reviews Using Deep Learning*,” *MDPI Journal*, vol. 14, iss. 17, 2022, doi: 10.3390/su141710844
- [27] Biyang Guo, Songqiao Han, dkk, “*Label Confusion Learning to Enhance Text Classification Model*,” *Association for the Advancement of Artificial Intelligence*, pp. 12929-12936, 2020, doi: 10.48550/arXiv.2012.04987
- [28] Satria Budi, “Implementasi Metode Naïve Bayes Untuk Klasifikasi Ulasan Pada Aplikasi Telegram,” *Journal of Information Technology and Computer Science (INTECOMS)*, vol. 6, no. 2, pp. 1282-1288, 2023, doi: 10.31539/intercoms.v6i2.8284
- [29] Ni Made Dina Aprilianti, Jeanitha Gein, dkk, “Analisis Perbandingan Algoritma KNN, Gaussian Naïve Bayes, Random Forest Untuk Data Tidak Seimbang dan Data yang Diseimbangkan Dengan Metode Tomek Link Undersampling pada Dataset LCMS Tanaman Keladi Tikus,” *Prosiding Sains Nasional dan Teknologi*, vol. 13, no. 1, pp. 156-160, 2023, doi: 10.36499/psnst.v13i1.9002

Conflict of Interest Statement:

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.