

STUDI KOMPARASI SVM DAN NBC TERHADAP KLASIFIKASI SENTIMEN APLIKASI DANA

Muhammad Rapi ¹⁾, Chandra Kirana ²⁾

^{1,2)} Teknik Informatika Institute Sains dan Bisnis Atma Luhur

email : 2311500044@mahasiswa.atmaluhur.ac.id¹⁾, chandra.kirana@atmaluhur.ac.id²⁾

Abstraksi

Penelitian ini mengeksplorasi efektivitas komparasi antara model *Support Vector Machine* (SVM) dan *Naïve Bayes Classifier* (NBC) dalam menganalisis sentiment ulasan pengguna aplikasi DANA pada Google Play Store. Studi ini dilakukan untuk mengisi kesenjangan penelitian sebelumnya yang belum mengeksplorasi penanganan kelas opini netral secara spesifik pada objek aplikasi finansial yang sama. Sebanyak 4.800 data tekstual diproses melalui tahapan prapemrosesan sistematis, meliputi *cleaning*, *case folding*, tokenisasi, dan *stemming* Sastrawi. Melalui pembobotan TF-IDF yang menghasilkan 3.715 fitur unik, temuan riset mengkonfirmasi dominansi SVM dengan capaian akurasi sebesar 82,29%, melampaui arsitektur NBC yang mencatatkan nilai 81,46%. Meskipun performa pada kelas positif dan negative sangat mumpuni, kedua algoritma menghadapi kendala signifikan dalam mengidentifikasi kategori netral (*f1-score* 0,00) akibat distribusi data yang tidak seimbang secara ekstrem. Hasil penelitian ini memberikan rekomendasi strategis bagi pengembang aplikasi dan akademisi dalam mengoptimalkan model klasifikasi sentiment pada sektor *fintech*.

Kata Kunci :

Analisis Sentimen, DANA, Support Vector Machine, Naïve Bayes, TF-IDF

Abstract

This study explores the comparative effectiveness of Support Vector Machine (SVM) dan Naïve Bayes Classifier (NBC) models in analyzing sentiment from DANA application user reviews on the Google Play Store. This research aims to address a gap in previous studies that have not specifically explored the handling of neutral opinion classes within the same financial application context. A total of 4,800 textual data points were processed through systematic preprocessing stages, including cleaning, case folding, tokenization, and Sastrawi stemming. Using TF-IDF weighting, which generated 3,175 unique features, the findings confirm the dominance of the SVM model with an accuracy of 82.29%, outperforming the NBC architecture which recorded a value of 81.46%. although performance on positive and negative classes is highly robust, both algorithms face significant challenges in identifying the neutral category (f1-score 0.00) due to an extremely imbalanced data distribution. These results provide strategic recommendations for applications developers and academics in optimizing sentiment classification models within the fintech sector.

Keywords :

Sentiment Analysis, DANA, Support Vector Machine, Naïve Bayes, TF-IDF

Pendahuluan

Pada era revolusi industri 4.0, penerapan teknologi modern yang terintegrasi dengan jaringan fiber dan sistem komunikasi cerdas berperan penting dalam mendukung berbagai sektor, termasuk ekonomi, konsumsi, dan produksi. Perkembangan pesat dalam bidang penyimpanan data, jaringan, serta kecerdasan buatan menjadi faktor utama yang mendorong kemajuan ini, di mana kemampuan teknologi dalam mengelola dan mentransformasi data secara efisien menjadi sangat krusial. Seiring dengan pesatnya perkembangan teknologi digital, sektor keuangan mengalami transformasi signifikan yang dikenal sebagai *Financial Technology* (Fintech). Perkembangan fintech di Indonesia mulai menunjukkan kemajuan dalam beberapa tahun

terakhir. Pada tahun 2015, dibentuk Asosiasi Fintech Indonesia (AFI) yang bertujuan menjadi mitra bisnis yang kompeten di sektor keuangan. Sejumlah penelitian membuktikan bahwa masyarakat cenderung lebih menyukai teknologi yang menawarkan layanan platform yang cepat, mudah, aman, dan bermanfaat. Salah satu wujudnya adalah hadirnya fintech yang menyediakan sistem pembayaran nontunai atau e-money[1]. Inovasi Fintech telah mengubah cara masyarakat bertransaksi, menyimpan uang, dan mengakses layanan finansial, menjadikannya bagian integral dari kehidupan sehari-hari di Indonesia.

Salah satu inovasi terkemuka di sektor Fintech Indonesia adalah aplikasi dompet digital DANA.

DANA adalah platform dompet digital Indonesia yang bertujuan memfasilitasi transaksi online dengan cepat tanpa menggunakan uang tunai[2]. Sebagai platform pembayaran digital, DANA telah diadopsi secara luas dengan total unduhan mencapai lebih dari 100 juta kali di Google Play Store dan memfasilitasi jutaan transaksi harian. Meningkatnya penetrasi penggunaan aplikasi ini berdampak pada volume umpan balik (*feedback*) dari pengguna. Ulasan yang diserahkan pengguna melalui platform Google Play Store menjadi sumber data yang sangat berharga dan spontan, mencerminkan pengalaman aktual mereka terhadap fitur, keandalan, dan layanan pelanggan aplikasi. Besarnya volume ulasan ini menciptakan *big data* tekstual yang kompleks, sehingga analisis sentimen menjadi metode yang sangat dibutuhkan untuk memetakan kepuasan pengguna secara *real-time*.

Proses identifikasi sentimen melalui teks untuk pengelompokan polaritas positif atau negatif dikenal sebagai analisis sentimen. Menurut Liu, bidang ini mempelajari beragam aspek mulai dari opini hingga emosi yang berkaitan dengan atribut produk atau peristiwa tertentu. Bidang ini merupakan ranah riset baru dalam NLP yang difungsikan untuk mendeteksi unsur subjektivitas serta melakukan ekstraksi sentimen secara otomatis [3]. Penelitian serupa pernah dilakukan pada tahun 2024 [4], namun belum mengeksplorasi efektivitas penanganan kelas opini netral secara spesifik pada objek aplikasi finansial yang sama. Penelitian ini hadir untuk mengisi kesenjangan tersebut dengan menggunakan volume data yang lebih besar, yakni 4.800 ulasan, guna memperoleh hasil komparasi kinerja model yang lebih representatif.

Dalam praktiknya, pemilihan metode *Support Vector Machine* (SVM) dan *Naïve Bayes Classifier* (NBC) didasarkan pada efektivitas komputasionalnya dalam menangani data tekstual dengan dimensi tinggi. Berbeda dengan model *Deep learning* seperti BERT atau LSTM yang memerlukan sumber daya komputasi intensif, SVM dan NBC menawarkan pendekatan yang lebih efisien dan ringan untuk klasifikasi teks pada skala data menengah. SVM dipilih karena ketangguhannya dalam mencari margin optimal, sementara NBC dipilih sebagai *baseline* klasifikasi probabilistik yang cepat dalam menangani fitur yang tereduksi oleh metode TF-IDF.

Berdasarkan latar belakang tersebut, penelitian ini merumuskan permasalahan sebagai berikut:

1. Bagaimana performa komparasi algoritma *Support Vector Machine* (SVM) dan *Naïve Bayes Classifier* (NBC) dalam melakukan klasifikasi sentimen pada ulasan aplikasi DANA?
2. Sejauh mana pengaruh ketidakseimbangan data (*data imbalanced*) terhadap Tingkat akurasi dan *f1-score* kedua model tersebut pada kelas sentiment netral?

Tujuan utamanya adalah mengevaluasi algoritma mana yang paling akurat dan relevan dalam menangani data ulasan yang sudah melalui tahapan *preprocessing* seperti *Stemming* Sastrawi dan *Feature Extraction* TF-IDF. Penelitian ini diharapkan dapat memberikan rekomendasi yang valid bagi Perusahaan DANA maupun akademisi mengenai model terbaik untuk analisis sentimen dalam konteks fintech di Indonesia.

Tinjauan Pustaka

Aplikasi DANA adalah salah satu platform dompet digital terkemuka di Indonesia, yang menyediakan berbagai layanan pembayaran dan transaksi finansial. Tingginya tingkat adopsi DANA di masyarakat membuat aplikasi ini menjadi objek penting untuk analisis sentimen, karena umpan balik pengguna secara langsung merefleksikan kualitas layanan dan fitur yang disediakan. Ulasan Pengguna (User Reviews) dari platform seperti Google Play Store berfungsi sebagai sumber data primer dalam penelitian ini. Berbeda dengan media sosial seperti Twitter yang berisi opini terhadap berbagai topik, ulasan aplikasi cenderung fokus dan spesifik pada pengalaman interaksi pengguna terhadap produk tersebut. Data ulasan ini bersifat terstruktur dan mudah diakses, menjadikannya data yang ideal untuk pemodelan klasifikasi sentimen.

Sebagai salah satu teknik pembelajaran mesin yang banyak diminati, SVM menawarkan keunggulan melalui landasan matematika yang kuat untuk menghasilkan model klasifikasi yang andal. Dibandingkan dengan algoritma klasifikasi lainnya, SVM terbukti mampu memberikan hasil yang lebih akurat dalam berbagai skenario implementasi. Keandalan performanya dalam memproses bahasa alami juga telah divalidasi oleh banyak peneliti, yang menempatkan SVM sebagai salah satu metode dengan akurasi terbaik untuk klasifikasi teks [5].

Naïve Bayes merupakan algoritma klasifikasi probabilistik yang bekerja dengan menerapkan Teorema Bayes serta memanfaatkan informasi frekuensi data yang tersedia dalam basis data. Mekanisme utama dalam proses klasifikasi ini adalah penentuan label kategori bagi sebuah sampel melalui analisis terhadap serangkaian fitur atau karakteristik khusus yang menyusunnya. Sebagai gambaran, penggunaan data terkait usia, kategori gender, serta tingkat penghasilan dapat diolah untuk memproyeksikan probabilitas seseorang dalam meraih posisi pekerjaan dengan kompensasi gaji yang tinggi [6].

Skema pembobotan fitur dalam penelitian ini memanfaatkan metode *Term Frequency-Inverse Document Frequency* (TF-IDF) untuk mengevaluasi signifikansi suatu kata berdasarkan perbandingan kemunculannya dalam sebuah dokumen dengan distribusinya di seluruh dataset. Pendekatan ini

mengombinasikan dua elemen fundamental, yakni *Term Frequency* (TF) yang mengukur intensitas kemunculan kata secara lokal dan *Inverse Document Frequency* (IDF) yang berfungsi menyeimbangkan bobot kata tersebut terhadap keseluruhan korpus. Dalam mekanisme ini, istilah yang sering muncul pada satu ulasan akan memperoleh bobot TF yang lebih besar, namun nilai IDF akan meningkat secara signifikan apabila istilah tersebut bersifat langka atau jarang ditemukan di dokumen lain. Sebaliknya, fungsi IDF adalah mereduksi kepentingan kata-kata yang terlalu umum muncul di banyak dokumen agar tidak mendominasi hasil klasifikasi. Melalui proses ini, data teks ditransformasikan menjadi representasi numerik yang dapat diproses secara komputasi. Implementasinya dimulai dengan menghitung frekuensi setiap istilah pada dataset setelah seluruh kata direduksi ke bentuk dasarnya melalui tahap pra-proses[7].

Sebelum dilakukan klasifikasi, data teks perlu melalui tahap pra-proses seperti:

1. **Tokenization:** Proses yang digunakan untuk memecah sebuah kalimat menjadi token atau kata-kata yang terpisah [8].
2. **Case Folding:** tahapan dalam tekt preprocessing yang berfungsi untuk menyeragamkan bentuk karakter pada data[9].
3. **Stopword Removal:** di proses ini, menghapus/membuang kata-kata yang tidak penting, seperti yang, dan, di, adalah dll [10].
4. **Stemming:** Proses untuk memperoleh bentuk dasar sebuah kata dengan menghapus awalan, akhiran, sisipan, serta konfiks (gabungan awalan dan akhiran)[11]. Proses ini menggunakan *library* Sastrawi yang spesifik untuk Bahasa Indonesia.
5. **Cleaning:** Penghapusan karakter yang tidak relevan, seperti tautan (URL), tanda baca, angka, dan karakter non-alfanumerik.

Untuk memvalidasi performa serta akurasi dari model klasifikasi yang telah dirancang, diperlukan suatu fase evaluasi yang sistematis dan mendalam. Instrumen *Confusion Matrix* menjadi pilihan utama dalam proses ini karena kemampuannya dalam menyajikan ringkasan perbandingan antara label prediksi dengan nilai riil dari data uji. Melalui matriks ini, kualitas klasifikasi diurai ke dalam empat dimensi dasar, yakni *True Positive* (TP) dan *True Negative* (TN) untuk merepresentasikan prediksi yang tepat, serta *False Positive* (FP) dan *False Negative* (FN) untuk menandai adanya kesalahan identifikasi kelas. Kuantifikasi dari keempat elemen tersebut nantinya diintegrasikan ke dalam berbagai metrik performansi standar, seperti tingkat akurasi, presisi, *recall*, dan *f1-score*, guna memastikan objektivitas hasil pengujian secara menyeluruh.[12].

Konteks penelitian ini diperkuat oleh sejumlah studi terdahulu yang telah menyediakan landasan kuat bagi perbandingan algoritma klasifikasi.

Eksplorasi mengenai efektivitas model klasifikasi dilakukan dalam studi [13] yang mengadu performa antara algoritma *Naïve Bayes* dan *Support Vector Machine* (SVM) untuk memetakan opini publik terkait pembangunan Ibu Kota Nusantara berdasarkan data yang dihimpun dari platform Twitter. Temuan riset tersebut mengonfirmasi keunggulan SVM yang secara konsisten mampu menghasilkan tingkat akurasi analisis sentimen sebesar 94%, melampaui capaian algoritma *Naïve Bayes* yang mencatatkan angka 91%. Perbedaan margin akurasi ini memberikan indikasi bahwa dalam lingkup pengolahan data opini masyarakat di media sosial mengenai IKN, penggunaan pendekatan SVM menawarkan solusi yang lebih presisi dan reliabel dibandingkan metode klasifikasi konvensional lainnya.

Studi oleh [14] melakukan penelitian untuk mengetahui pandangan Masyarakat terhadap *paylater*. Metode yang digunakan dalam penelitian tersebut menerapkan analisis sentimen menggunakan metode *Naïve Bayes* dan *TextBlob* dari Pustaka *TextBlob* dengan Bahasa pemrograman Python. Pada pengujian model dengan *confusion matrix* didapatkan bahwa algoritma *Naïve Bayes* lebih akurat sebesar 91% dibandingkan dengan *TextBlob* yang hanya 61%.

Dalam studinya [3], mengeksplorasi penggunaan algoritma SVM untuk memetakan opini konsumen terhadap merek perawatan kulit *Skintif* berdasarkan data yang dikumpulkan dari Youtube dan Instagram. Implementasi model dengan fungsi kernel linear menunjukkan performa yang sangat mumpuni dalam membedakan polaritas sentimen positif dan negatif, di mana tingkat akurasi klasifikasi yang dihasilkan mencapai angka 94%.

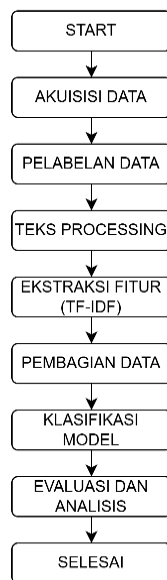
Sementara itu, pada penelitian [4] menganalisis sentimen aplikasi Dana menggunakan *Naïve Bayes* dan *Support Machine Learning*. Didapatkan Kesimpulan bahwa mayoritas pengguna mengkespresikan sentimen negatif, yang mencap 84,7%. Hal ini menandakan adanya ketidakpuasan yang cukup signifikan terhadap layanan tersebut.

Dalam mengklasifikasikan berita palsu di Indonesia, penelitian [15] memanfaatkan empat variasi model NBC dan empat jenis kernel SVM dengan rasio data 80:20. Hasilnya, *Complement Naïve Bayes* menjadi model NBC paling optimal dengan akurasi 93,1%. Di sisi lain, penggunaan kernel *Sigmoid* pada algoritma SVM menunjukkan hasil yang lebih dominan, mencapai akurasi 96,5% dan *recall* 100%. Capaian

ini menegaskan keunggulan SVM dalam menangani data teks berita dibandingkan pendekatan NBC pada studi tersebut.

Metode Penelitian

Metodologi yang diterapkan dalam studi ini disusun secara sistematis guna menjawab tantangan dalam membandingkan performansi algoritma klasifikasi. Rangkaian penelitian dimulai dengan akuisisi data umpan balik pengguna aplikasi DANA yang bersumber dari Google Play Store, yang kemudian dilanjutkan dengan prosedur praproses data yang ketat, termasuk implementasi teknik *stemming* Sastrawi untuk menghasilkan kumpulan data teks yang bersih dan terstandarisasi. Data yang telah diproses tersebut selanjutnya ditransformasikan menjadi representasi fitur numerik melalui metode pembobotan *Term Frequency-Inverse Document Frequency* (TF-IDF). Kumpulan fitur ini kemudian didistribusikan ke dalam subset data latih (*training*) serta data uji (*testing*) untuk melatih dua arsitektur model utama, yaitu *Support Vector Machine* (SVM) dan *Naïve Bayes Classifier* (NBC). Tahap final melibatkan evaluasi mendalam terhadap kinerja model dengan memanfaatkan metrik akurasi, presisi, serta *recall* guna mengidentifikasi algoritma yang memberikan hasil klasifikasi sentimen paling efektif. Tahapan penelitian ini disusun secara sistematis untuk mencapai tujuan klasifikasi sentimen ulasan dompet digital DANA, sebagaimana ditunjukkan pada alur kerja penelitian pada Gambar 1.



Gambar 1. Diagram Alir Metologi Penelitian

Mengadopsi kerangka kerja kuantitatif berbasis pembelajaran mesin, studi ini berupaya membedah sentimen dari opini publik terhadap platform dompet digital DANA. Orientasi utama dari metodologi yang diterapkan adalah untuk melakukan komparasi performansi antara arsitektur *Support Vector*

Machine (SVM) dan *Naïve Bayes Classifier* (NBC). Evaluasi efektivitas ini dilakukan terhadap sekumpulan data tekstual ulasan yang sebelumnya telah ditransformasikan menjadi representasi numerik melalui teknik ekstraksi fitur *Term Frequency-Inverse Document Frequency* (TF-IDF).

1. Sumber Data dan Akuisisi Data

Himpunan data primer dalam studi ini diekstraksi dari sekumpulan ulasan konsumen aplikasi DANA yang tersedia pada portal digital Google Play Store. Pemilihan sumber data ini didasarkan pada ketersediaan opini aktual pengguna yang memberikan gambaran nyata mengenai pengalaman mereka dalam berinteraksi dengan layanan dompet digital tersebut

- a. Akuisisi Data: Prosedur pengumpulan informasi dalam studi ini dijalankan melalui mekanisme otomatisasi dengan memanfaatkan antarmuka pemrograman aplikasi pihak ketiga, yaitu *Google Play Scraper API*. Pendekatan tersebut diadopsi mengingat keunggulannya dalam menyediakan akses terhadap data umpan balik pengguna yang bersifat organik sekaligus memiliki format yang terorganisir secara sistematis.
- b. Volume Data: Total data mentah yang berhasil dikumpulkan adalah ≈ 5000 ulasan. Setelah proses pembersihan awal (menghapus ulasan kosong), jumlah sampel data yang digunakan untuk pemodelan adalah ≈ 4.800 ulasan.
- c. Karakteristik Data: Setiap ulasan mencakup kolom konten teks (*content*) dan skor rating bintang (*score*) yang digunakan untuk proses pelabelan.

2. Pelabelan Data (*Data Labeling*)

Data ulasan diambil dari Google Play Store dengan pelabelan otomatis berbasis *rating* (1-2: Negatif, 3: Netral, 4-5: Positif). Untuk memastikan kualitas data dari bias *rating*, dilakukan validasi manual (*cross-check*) terhadap 10% sampel dataset secara acak. Hasil validasi manual ini menunjukkan Tingkat keselarasan yang tinggi antara label antara label otomatis dan sentimen tekstual, sehingga dataset dianggap valid untuk proses klasifikasi

Tabel 1. Pelabelan Data (*Data Labeling*)

Rating Bintang (Skor)	Sentimen	Jumlah Sampel (Awal)
4 dan 5	Positif	3.584
1 dan 2	Negatif	1.182
3	Netral	234

3. Tahap Praproses Data (*Preprocessing*)

Tahap praproses (*preprocessing*) dilakukan untuk membersihkan data mentah agar siap diproses pada tahap klasifikasi. Ilustrasi transformasi data pada setiap tahapan praproses dapat dilihat pada Tabel 2.

Tabel 2. Contoh Tahapan *Preprocessing*

No	Tahap	Operasi	Contoh hasil Ulasan
1	Data Mentah	(<i>original Review</i>)	Aplikasi Dana-nya sering <i>ERROR</i> , padahal udah di- <i>update</i> , mohon diperbaiki ya min
2	Case Folding	Mengubah semua teks menjadi huruf kecil (<i>lowercase</i>)	aplikasi dana-nya sering <i>error</i> , padahal udah di <i>update</i> , mohon diperbaiki ya min
3	Cleaning	Menghapus tanda baca, angka, dan karakter non-alfanumerik	aplikasi dananya sering <i>error</i> padahal udah di <i>update</i> mohon diperbaiki ya min
4	Tokenization	Memecah teks menjadi unit-unit kata (<i>token</i>)	['aplikasi', 'dana', 'nya', 'padahal', 'udah', 'di', 'update', 'mohon', 'diperbaiki', 'ya', 'min']
5	Stopword Removal	Menghapus kata umum yang tidak memenuhi sentimen, (sepertinya,sering,udah,di,ya,min)	['aplikasi', 'dana', 'error', 'padahal', 'update', 'mohon', 'diperbaiki']
6	Stemming	Mengubah kata berimbuhan menjadi kata dasar (menggunakan Sastrawi)	['aplikasi', 'dana', 'error', 'padahal', 'update', 'mohon', 'perbaiki']

4. Ekstraksi Fitur (Feature Extraction) dan Pembagian Data

Setelah data teks bersih dihasilkan, data diubah menjadi representasi numerik menggunakan metode TF-IDF.

- a. Feature Extraction (TF-IDF): Digunakan untuk menghitung bobot kata yang mencerminkan kepentingan kata tersebut dalam seluruh korpus. Dalam penelitian ini, model TF-IDF membatasi fitur (kata unik) maksimal sebanyak ≈ 3.715 fitur teratas, menghasilkan dimensi data akhir (sampel, fitur) sebesar ($\approx 4.800, 3.715$).
- b. Pembagian Data dan Prosedur Validasi: Untuk kebutuhan pemodelan, data berformat vector hasil pembobotan TF-IDF dipisahkan menjadi dua kelompok data, yaitu data latih dan data uji. Penelitian ini menerapkan metode *single split* melalui fungsi *train_test_split* dengan rasio 80% untuk pelatihan (3.840 sampel) dan 20% untuk pengujian (960 sampel) guna mendapatkan hasil klasifikasi yang akurat dan kredibel

Ditemukan bahwa data ulasan memiliki kondisi *Imbalanced data*, dimana jumlah sampel pada kelas Netral (234) jauh lebih kecil dibandingkan kelas Positif (3.584). Untuk mengatasi potensi bias secara algoritmik, proses pembagian data menggunakan parameter *stratify = y*. Pendekatan ini memastikan rasio sentimen di data latih dan data uji tetap konsisten, sehingga model dapat mempelajari karakteristik kelas minoritas (Netral) secara lebih proporsional tanpa harus melakukan manipulasi data sintesis (*oversampling*).

5. Klasifikasi dan Evaluasi Model

Tahap ini merupakan inti dari proses penelitian, di mana data ulasan yang telah melalui tahap pra-proses dan ekstraksi fitur TF-IDF akan diproses menggunakan dua algoritma klasifikasi untuk dibandingkan kinerjanya.

- a. Support Vector Machine (SVM)

Sebagai teknik pembelajaran mesin, SVM beroperasi dengan mengidentifikasi bidang pemisah atau *hyperplane* terbaik guna membedakan kategori data di dalam ruang fitur berdimensi tinggi. Pemilihan *Kernel Linear* dalam studi ini didasari oleh efisiensinya dalam mengolah dataset tekstual yang cenderung bersifat *sparse* serta memiliki kompleksitas fitur yang besar sebagai hasil dari pembobotan TF-IDF. Adapun formulasi matematis yang merepresentasikan fungsi keputusan pada klasifikasi linear SVM dituangkan dalam Persamaan (1):

$$f(x) = \text{sign}(w \cdot x + b) \quad (1)$$

Keterangan :

- x : Vektor fitur masukan (bobot TF-IDF)
- w : Vektor bobot (*weight vector*) yang menentukan orientasi *hyperplane*.
- b : *Bias* yang menentukan posisi bidang relative terhadap titik pusat
- Sign: Fungsi untuk menentukan label kelas (Positif, Negatif, atau Netral) berdasarkan nilai kalkulasi.

b. Naïve Bayes Classifier (NBC)

Klasifikasi berbasis probabilitas dijalankan oleh algoritma Naïve Bayes Classifier dengan menerapkan prinsip Teorema Bayes serta asumsi independensi fitur yang kuat. Multinomial Naïve Bayes menjadi varian model yang dipilih dalam studi ini karena secara teoritis sangat optimal untuk mengolah representasi data teks dalam format pembobotan TF-IDF. Sebagai dasar penentuan prediksi kelas, nilai peluang tertinggi diidentifikasi dengan merujuk pada formulasi dalam Persamaan (2):

$$P(c|d) = \frac{P(c) \prod_{i=1}^n P(ti|c)^{fti,d}}{P(d)} \quad (2)$$

Keterangan:

- $P(c|d)$: Probabilitas posterior kelas c . berdasarkan dokumen ulasan d
- $P(c)$: Probabilitas prior dari kelas c .
- $P(ti|c)$: Probabilitas kemunculan kata ti pada kelas c .
- fti, d : Frekuensi atau bobot kata ti dalam dokumen d .

Untuk menjamain objektivitas dalam studi komparasi ini, kedua algoritma diimplementasikan menggunakan parameter standar (*default*) dari pustaka scikit-learn. Model *Support Vector Machine* (SVM) dikonfigurasi menggunakan kernel RBF (*Radial Basis Function*), nilai parameter penalty $C = 1.0$, dan $\gamma = \text{'scale'}$. Sementara itu, model *Multinomial Naïve Bayes* menggunakan nilai parameter *smoothing* $\alpha = 1.0$ (*Laplace smoothing*) guna mengantisipasi nilai probabilitas nol pada fitur kata yang tidak muncul dalam data latih

- c. Kinerja dari model SVM dan NBC dievaluasi secara mendalam menggunakan instrumen *Confusion Matrix*. Matriks ini memetakan hasil prediksi sistem terhadap label aktual dari data uji. Mengingat penelitian ini menggunakan tiga kategori sentimen (Positif, Negatif, dan Netral), maka evaluasi didasarkan pada metrik berikut:

- 1) Akurasi : Mengukur sejauh mana model mampu mengklasifikasi seluruh data dengan benar.

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

- 2) Presisi : Mengukur Tingkat keakuratan antara hasil prediksi positif dengan data yang sebenarnya positif.

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

- 3) Recall : Mengukur kemampuan model dalam menemukan Kembali informasi atau label yang benar dari seluruh dataset.

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

Keterangan komponen : *True Positive* (TP), *True Negative* (TN), *False Positive* (FP), dan *False Negative* (FN).

Seluruh proses eksperimen dikembangkan menggunakan Bahasa pemrograman Python versi 3.9. pembangunan arsitektur model dan pemrosesan data dilakukan dengan memanfaatkan pustaka scikit-learn (versi 1.2.2). yang dipilih karena ketersediaan fungsi klasifikasi yang terstandarisasi untuk memastikan reproducibility penelitian.

Hasil dan Pembahasan

Bagian ini memaparkan temuan ilmiah yang diperoleh dari serangkaian pengujian model klasifikasi sentimen terhadap ulasan pengguna aplikasi DANA yang bersumber dari Google Play Store. Fokus utama dari pembahasan ini adalah mengevaluasi kinerja algoritma *Support Vector Machine* (SVM) dan *Naïve Bayes Classifier* (NBC) setelah melalui tahapan *preprocessing* yang ketat serta ekstraksi fitur menggunakan metode TF-IDF. Penilaian dilakukan secara komprehensif dengan membandingkan nilai akurasi, presisi, dan *recall* guna menentukan efektivitas model dalam menangani data teks bahasa Indonesia yang memiliki karakteristik kompleks. Hasil penelitian ini juga diarahkan untuk memberikan penjelasan logis terkait distribusi sentimen positif, negatif, dan netral serta

bagaimana ketidakseimbangan data memengaruhi hasil akhir pemodelan. Seluruh temuan disajikan secara sistematis dalam bentuk tabel dan visualisasi agar memudahkan interpretasi terhadap perbandingan kinerja kedua algoritma yang diuji.

Komparasi Performa Model

Perbandingan performansi antara arsitektur *Support Vector Machine* (SVM) dan *Naïve Bayes Classifier* (NBC) berhasil didapatkan usai merampungkan fase pembelajaran serta validasi pada 20% porsi dataset uji. Penyajian data secara komprehensif terkait metrik evaluasi yang mencakup parameter akurasi, presisi, serta *recall* dapat diamati pada Tabel 3 sebagai bentuk perangkuman hasil pengujian yang telah dilakukan.

Tabel 3. Hasil Komparasi Kinerja Algoritma

Model	Akurasi	Presisi	Recall
SVM	82,29%	78,20%	82,29%
Naïve Bayes	81,46%	77,22%	81,46%

Dominansi pada seluruh parameter pengujian diperlihatkan oleh algoritma SVM terhadap Naïve Bayes, sebagaimana terpotret dalam rangkuman data pada Tabel 3. Capaian akurasi SVM yang menyentuh angka 82,29% unggul tipis dengan selisih 0,83% di atas perolehan Naïve Bayes yang berada pada level 81,46%. Margin yang sangat sempit tersebut mengindikasikan bahwa kedua arsitektur model memiliki kompetensi yang relatif seimbang dalam memproses sentimen ulasan DANA, sekalipun SVM terbukti lebih andal dalam mendefinisikan *hyperplane* pada ruang fitur TF-IDF yang memiliki dimensi tinggi.

Jika dibandingkan dengan penelitian terdahulu yang dilakukan oleh [4] pada objek aplikasi DANA, model SVM yang dikembangkan dalam penelitian ini mencapai akurasi sebesar 82,29%. Peningkatan akurasi ini dicapai melalui penggunaan dataset yang lebih besar (4.800 ulasan) dibandingkan penelitian [4], yang membuktikan bahwa penambahan volume data tesktual secara signifikan meningkatkan kemampuan model dalam mempelajari variasi pola sentimen pada ulasan aplikasi finansial

Kemampuan sistem dalam mengidentifikasi kategori sentimen secara spesifik tidak dapat sepenuhnya tergambarkan hanya melalui metrik akurasi global, kendati nilai tersebut telah menyajikan gambaran performa secara makro. Atas dasar tersebut, pendekatan evaluasi tambahan ditempuh melalui peninjauan parameter *f1-score*, *recall*, dan presisi pada tiap-tiap label demi mengidentifikasi adanya hambatan teknis atau kecenderungan bias pada kelompok data tertentu. Melalui Tabel 4 di bawah ini, detail pencapaian klasifikasi untuk setiap kelompok sentimen dapat dicermati secara mendalam.

Tabel 4. Detail Performa Klasifikasi Per Kategori Sentimen

Algoritma	Kelas Sentimen	Precisi	Recall	F1-Score	Support
SVM	Negatif	0,69	0,69	0,69	235
	Netral	0,00	0,00	0,00	45
	Positif	0,87	0,93	0,89	680
Naïve Bayes	Negatif	0,77	0,55	0,64	235
	Netral	0,00	0,00	0,00	45
	Positif	0,82	0,96	0,89	680

Berdasarkan hasil detail performa pada tabel 4, dapat dilakukan penajaman analisis terhadap kemampuan model dalam mengklasifikasikan setiap kategori sentimen sebagai berikut:

a. Dominasi Kelas Positif

Kedua model menunjukkan performa paling optimal pada kategori Positif, dimana SVM mencapai nilai *recall* 0,93 dan Naïve Bayes mencapai 0,96. Hal ini mengidentifikasi bahwa model sangat sensitif dalam mengenali ulasan apresiatif pengguna, yang didukung oleh ketersediaan data latih yang melimpah pada kelas ini (3.584 sampel).

b. Stabilitas pada Kelas Negatif

Capaian *f1-score* sebesar 0,69 menegaskan dominasi signifikan SVM dalam memproses sentimen negatif jika dibandingkan dengan perolehan Naïve Bayes yang hanya berada pada angka 0,64. Bias klasifikasi terdeteksi pada model Naïve Bayes akibat rendahnya nilai *recall* (0,55), yang mengakibatkan ulasan berkategori negatif sering kali salah diprediksi sebagai label positif oleh sistem tersebut.

c. Kegagalan Klasifikasi Kelas Netral

Temuan paling kritis adalah perolehan nilai *precision*, *recall* dan *f1-Score* sebesar 0,00 untuk kategori Netral pada kedua algoritma. Angka 0 pada kolom prediksi Netral di *Confusion Matrix* membuktikan bahwa model gagal mengidentifikasi satu pun alasan netral dengan benar.

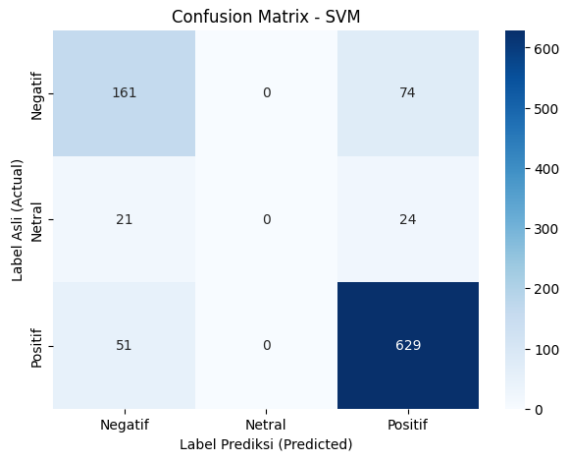
d. Dampak Ketidakseimbangan Data (*Imbalanced data*)

Kegagalan pada kelas Netral merupakan konsekuensi langsung dari kondisi *Imbalanced data* yang ekstrem, di mana jumlah sampel Netral (234 data) sangat minoritas dibandingkan kelas Positif. Keterbatasan variasi fitur pada data Netral menyebabkan model cenderung menggeneralisasi ulasan tersebut ke dalam kelas mayoritas, sehingga menurunkan kemampuan diskriminasi model secara keseluruhan.

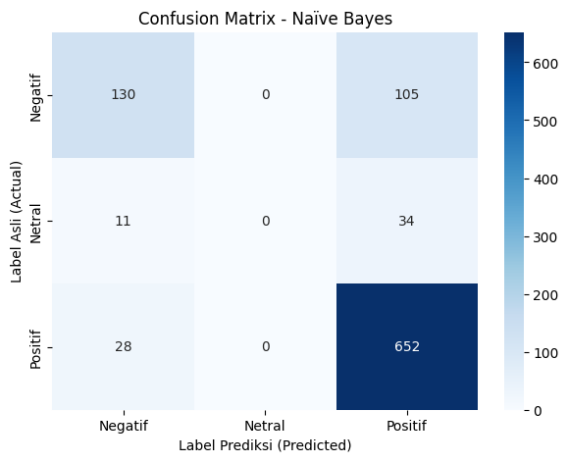
Visualisasi Confusion Matrix

Visualisasi *Confusion Matrix* dalam bentuk *heatmap* digunakan untuk memberikan gambaran yang lebih jelas mengenai distribusi prediksi yang dilakukan oleh masing-masing model terhadap label aktual.

Melalui visualisasi ini, dapat diidentifikasi kategori sentimen mana yang berhasil diklasifikasikan dengan tepat dan kategori mana yang mengalami salah klasifikasi (*misclassification*).



Gambar 2. Matrix Evaluasi Kinerja untuk model SVM



Gambar 3. Matrix Evaluasi Kinerja untuk model Algoritma Naïve Bayes

Berdasarkan visualisasi pada Gambar 2 dan Gambar 3, dapat dilakukan analisis mendalam mengenai performa kedua model dalam mengklasifikasikan sentimen sebagai berikut:

a. Analisis Model SVM

Pada matriks SVM, terlihat bahwa model berhasil memprediksi 629 data dengan benar pada kategori Positif dan 161 data dengan benar pada kategori Negatif. Namun, terdapat 74 data negatif yang salah diprediksi sebagai positif, serta 51 data positif yang salah diprediksi sebagai negatif. Hal ini menunjukkan bahwa SVM memiliki tingkat akurasi yang cukup stabil dalam mengenali ulasan yang bersifat ekstrem (sangat positif atau sangat negatif).

b. Analisis Model Naïve Bayes

Pada matriks Naïve Bayes, model menunjukkan kecenderungan yang lebih agresif dalam memprediksi kelas mayoritas (Positif), dengan berhasil menebak 652 data positif secara benar. Namun, performanya pada kelas Negatif lebih

rendah dibandingkan SVM, di mana terdapat 105 data negatif yang justru salah diklasifikasikan ke dalam kategori positif. Hal ini menjelaskan mengapa nilai *recall* untuk kelas negatif pada Naïve Bayes lebih rendah.

c. Kegagalan pada kelas Netral

Hal yang paling menonjol dari kedua gambar tersebut adalah kosongnya angka pada kolom prediksi Netral. Baik SVM maupun Naïve Bayes tidak berhasil memprediksi satu pun data ke dalam kategori Netral (terlihat dari angka 0 pada kolom tengah). Seluruh data asli Netral (sebanyak 45 data) justru terdistribusi ke dalam prediksi Negatif atau Positif. Fenomena ini membuktikan adanya kendala *imbalanced data*, di mana jumlah sampel Netral yang sangat sedikit membuat model tidak mampu mempelajari karakteristik sentimen tersebut secara maksimal.

Analisis lebih mendalam dilakukan melalui *error analysis* terhadap *Confusion Matrix* untuk memahami batasan model. Ditemukan bahwa kesalahan klasifikasi paling sering terjadi pada ulasan yang bersifat ambigu atau memiliki kalimat yang sangat pendek (misal: 'lumayan', 'ok'). Model kesulitan membedakan kelas netral karena sering kali ulasan tersebut memiliki kedekatan semantik dengan kata-kata yang digunakan dalam kelas positif maupun negatif. Hal ini mengonfirmasi bahwa rendahnya performa pada kelas netral dipengaruhi oleh kurangnya fitur pembeda yang kuat dalam teks ulasan yang singkat.

Kesimpulan dan Saran

Beberapa poin simpulan berhasil dirumuskan setelah melalui rangkaian analisis sentimen terhadap opini pengguna aplikasi DANA menggunakan pendekatan *Support Vector Machine* (SVM) dan *Naïve Bayes Classifier* (NBC):

1. Penyelenggaraan model klasifikasi sentimen telah terlaksana secara efektif dalam memproses ulasan konsumen dengan hasil performansi yang memuaskan. Keunggulan kompetitif diperlihatkan oleh algoritma SVM yang mencatatkan skor akurasi sebesar 82,29%, melampaui capaian akurasi dari metode Naïve Bayes yang berada pada angka 81,46%.
2. Penyederhanaan fitur teks secara signifikan sangat bergantung pada efektivitas tahapan prapemrosesan yang mencakup prosedur *stemming*, *stopword removal*, *tokenization*, *cleaning*, serta *case folding*. Keberhasilan rangkaian metode ini tercermin dari tingginya kemampuan model dalam mengidentifikasi sentimen positif, yang dibuktikan dengan raihan nilai *recall* pada rentang 0,93 hingga 0,96.
3. Terdapat kendala signifikan pada klasifikasi sentimen Netral yang mendapatkan nilai *f1-score* 0,00 pada kedua model. Hal ini disebabkan ketidakseimbangan data (*imbalanced data*) yang ekstrem, di mana

jumlah sentimen positif (3.584 data) jauh mendominasi dibandingkan sentimen netral (234 data).

4. Secara keseluruhan, algoritma SVM terbukti lebih stabil dibandingkan Naïve Bayes dalam menangani data ulasan ini, ditunjukkan dengan performa kelas negatif yang lebih seimbang antara *precision* dan *recall*.

Beberapa rekomendasi dikemukakan sebagai acuan bagi pengembangan studi mendatang dengan merujuk pada temuan yang telah didapatkan:

1. Implementasi teknik penyeimbangan dataset (*data balancing*), misalnya melalui metode *Random Over-sampling* atau SMOTE (*Synthetic Minority Over-sampling Technique*), sangat disarankan bagi riset berikutnya guna menanggulangi kendala pada kategori minoritas (Netral). Prosedur ini bertujuan untuk meningkatkan sensitivitas serta kemampuan model dalam mengidentifikasi sentimen tersebut secara lebih akurat.
2. Penggunaan metode ekstraksi fitur yang lebih modern seperti *Word2Vec* atau *FastText* disarankan untuk melengkapi metode TF-IDF guna menangkap konteks makna kata yang lebih mendalam pada ulasan pengguna.
3. Perlu dilakukan penambahan dataset dari periode waktu yang berbeda atau dari platform lain agar model yang dihasilkan memiliki generalisasi yang lebih luas terhadap berbagai jenis keluhan atau pujian pengguna.

Daftar Pustaka

- [1] S. U. Khoiriyah, M. Halim, and A. S. Zulkarnaeni, "Pengaruh persepsi manfaat, persepsi kemudahan, dan persepsi keamanan terhadap minat menggunakan financial technology pada aplikasi dana," pp. 1–15, [Online]. Available: https://repository.unmuhjember.ac.id/14565/10/10_ARTIKEL.pdf
- [2] F. Aini, F. Muttakin, T. K. Ahsyar, and E. Saputra, "Analisis Kepuasan Pengguna Aplikasi DANA Menggunakan Metode TAM dan EUCS," ©*Asosiasi Prakarsa Indones. Cerdas*, vol. 06, no. 01, pp. 65–76, 2023, [Online]. Available: <https://apic.id/jurnal/index.php/jsc/article/view/288>
- [3] R. A. Harnelia, Saputra, "ANALISIS SENTIMEN REVIEW SKINCARE SKINTIFIC DENGAN ALGORITMA SUPPORT VECTOR MACHINE," *JITET (Jurnal Inform. dan Tek. Elektro Ter.)*, vol. 12, no. 2, pp. 994–1002, 2024, [Online]. Available: <https://journal.eng.unila.ac.id/index.php/jitet/article/view/4095>
- [4] G. G. Warow and H. Pandia, "Analisis Sentimen Aplikasi Dana Menggunakan Naïve Bayes Classifier dan Support Vector Machine," *Jutisi J. Ilm. Tek. Inform. dan Sist. Inf.*, vol. 13, no. 1, pp. 608–621, 2024, [Online]. Available: <https://ojs.stmik-banjarbaru.ac.id/index.php/jutisi/article/view/189>
- [5] I. S. K. Idris, Y. A. Mustofa, and irvan A. Salihi, "Analisis Sentimen Terhadap Penggunaan Aplikasi Shopee Menggunakan Algoritma Support Vector Machine (SVM)," *Jambura J. Electr. Electron. Eng.*, vol. 5, no. 1, pp. 32–35, 2023, [Online]. Available: <https://ejournal.ung.ac.id/index.php/jjee/article/view/16830>
- [6] S. D. Prasetyo, S. S. Hilabi, and F. Nurapriani, "Analisis Sentimen Relokasi Ibukota Nusantara Menggunakan Algoritma Naïve Bayes dan KNN," *J. KomtekInfo*, vol. 10, no. 1, pp. 1–7, 2023, doi: 10.35134/komtekinfo.v10i1.330.
- [7] G. Yunanda, D. Nurjanah, and S. Meliana, "Recommendation System from Microsoft News Data using TF-IDF and Cosine Similarity Methods," *Build. Informatics, Technol. Sci.*, vol. 4, no. 1, pp. 277–284, 2022, doi: 10.47065/bits.v4i1.1670.
- [8] F. D. Ananda and Y. Pristyanto, "Analisis Sentimen Pengguna Twitter Terhadap Layanan Internet Provider Menggunakan Algoritma Support Vector Machine Sentiment Analysis of Twitter Users on Internet Service Providers Using Support Vector Machine Algorithm," *Matrik J. Manajemen, Tek. Inform. dan Rekayasa Komput.*, vol. 20, no. 2, pp. 407–416, 2021, doi: 10.30812/matrik.v20i2.1130.
- [9] T. P. Lestari, "Analisis Text Mining pada Sosial Media Twitter Menggunakan Metode Support Vector Machine (SVM) dan Social Network Analysis (SNA)," *J. Inform. Ekon. bisnis*, vol. 4, no. 3, pp. 65–71, 2022, doi: 10.37034/infab.v4i3.146.
- [10] M. Algoritma and N. Bayes, "Analisis Sentimen Pengguna Twitter di Indonesia Terhadap ChatGPT Menggunakan Algoritma C4.5 dan Naïve Bayes Yuma," *J. Sains dan Teknol.*, vol. 5, no. 1, pp. 115–122, 2023, [Online]. Available: <https://ejournal.sisfokomtek.org/index.php/sainte/article/view/1368>
- [11] R. R. Putri and N. Cahyono, "PUBLIK PEMERINTAH DKI JAKARTA DENGAN ALGORITMA," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 8, no. 2, pp. 2363–2371, 2024, [Online]. Available: <https://ejournal.itn.ac.id/index.php/jati/article/view/9472>
- [12] K. M. Yunita Rani, "Perbandingan Algoritma SVM Dan Naïve Bayes Pada Analisis Sentimen Kebijakan Penghapusan Kewajiban Skripsi," *Indones. J. Comput. Sci.*, vol. 12, no. 5, pp. 2879–2890, 2023, [Online]. Available: <http://ijcs.net/ijcs/index.php/ijcs/article/view/3415>
- [13] A. Supian, B. T. Revaldo, N. Marhadi, L. Efrizoni, and Rahmaddeni, "Perbandingan Kinerja Naïve Bayes dan SVM pada Analisis Sentimen Twitter Ibukota Nusantara," *J. Ilm. Inform.*, vol. 12, no. 1, pp. 15–21, 2024, [Online]. Available: <https://pdfs.semanticscholar.org/aa0e/03bde976e1e3986680fd7b88eb814f781547.pdf>
- [14] A. Safira and F. N. Hasan, "ANALISIS SENTIMEN MASYARAKAT TERHADAP PAYLATER MENGGUNAKAN METODE NAIVE BAYES CLASSIFIER," *Zo. J. Sist. Inf.*, vol. 5, no. 1, pp. 59–70, 2023, [Online]. Available: <https://journal.unilak.ac.id/index.php/zn/article/view/3>

- ew/12856
- [15] R. R. Sani, Y. A. Pratiwi, S. Winarno, E. D. Udayanti, and F. Al Zami, "Analisis Perbandingan Algoritma Naive Bayes Classifier dan Support Vector Machine untuk Klasifikasi Hoax pada Berita Online Indonesia," vol. 13, no. 2, pp. 85–98, 2022, [Online]. Available: <https://ejournal.undip.ac.id/index.php/jmasif/article/view/47983>