

Sentiment-Based Knowledge Discovery of Wondr by BNI App Reviews Using SVM, KNN, and Naive Bayes for CRM Enhancement

Zahra Tri Zafira¹, Ken Ditha Tania^{2*}, Winda Kurnia Sari³

^{1,2,3}Sistem Informasi, Universitas Sriwijaya

trizafira99@gmail.com¹, kenya.tania@gmail.com², windakurniasari@gmail.com³

Article Info

Article history:

Received 2025-07-19

Revised 2025-08-25

Accepted 2025-09-03

Keyword:

*Sentiment Analysis,
Knowledge Discovery,
User Loyalty,
Machine Learning,
Wondr by BNI.*

ABSTRACT

The rapid development of digital banking services has necessitated a deeper understanding of user perceptions and satisfaction levels. This study analyzes sentiment from user reviews of the Wondr by BNI app using a Knowledge Discovery approach and machine learning methods. Three classification algorithms were compared: Support Vector Machine (SVM), K-Nearest Neighbor (KNN), and Naive Bayes, evaluated with accuracy, precision, recall, and f1-score. The results show that SVM and Naive Bayes achieved the best performance with F1-scores of 0.88 and 0.87, while KNN lagged behind with 0.77. An ANOVA test further confirmed that the performance differences were statistically significant ($p < 0.05$), with SVM and Naive Bayes consistently outperforming KNN. Word Cloud analysis revealed dominant positive terms such as "easy," "fast," and "transaction," alongside negative terms like "login," "difficult," and "verification." These findings highlight user appreciation for simplicity and speed, while pointing out functional issues that require attention. This research not only enriches the literature on Indonesian-language sentiment analysis in the financial sector but also provides practical insights for Customer Relationship Management (CRM), particularly in strengthening customer retention strategies and guiding UX redesign for digital banking services.



This is an open access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.

I. PENDAHULUAN

Penggunaan telepon seluler dalam berbagai aktivitas semakin hari semakin meningkat. Kecenderungan dalam memilih solusi yang praktis dan efisien mendorong individu khususnya generasi muda untuk mengandalkan teknologi dalam berbagai aspek kehidupan. Kemudahan akses internet turut memperkuat tren ini. Sebuah survei yang dilakukan pada tahun 2024 oleh Asosiasi Penyelenggara Jasa Internet Indonesia (APJII) menemukan bahwa sekitar 221,5 juta orang atau 79,5% dari total populasi Indonesia telah menggunakan internet sebagai bagian dari aktivitas sehari-hari.

Menanggapi fenomena tersebut, industri perbankan di Indonesia terus berinovasi melalui digitalisasi layanan, salah satunya dengan menghadirkan aplikasi *mobile banking*. Aplikasi ini bukan hanya memfasilitasi transaksi keuangan, tetapi juga menjadi sarana penting dalam menjaga dan meningkatkan loyalitas nasabah. Aplikasi Wondr milik BNI

misalnya, memiliki lebih dari 5 juta unduhan di Google Play Store dan 330 ribu ulasan pengguna. Aplikasi ini menyasar generasi muda melalui pendekatan gaya hidup finansial yang modern dan *digital-native*.

Ulasan yang ditulis pengguna pada platform digital seperti Google Play Store dapat menjadi sumber data yang berharga dalam mengevaluasi sejauh mana nasabah merasa puas, nyaman, serta bagaimana persepsi mereka terhadap layanan perbankan digital. Sentimen yang terekam dalam ulasan mencerminkan kualitas layanan dari sudut pandang pengguna, termasuk berbagai kelebihan dan kekurangan yang dirasakan. Dengan menganalisis sentiment ini, bank dapat memperoleh *insight* atau *knowledge discovery* untuk meningkatkan kualitas layanan dan memperkuat loyalitas nasabah [1][2][3].

Loyalitas nasabah sendiri dipengaruhi oleh beberapa faktor seperti kinerja layanan, strategi hubungan pelanggan (*customer relationship marketing*), serta bauran pemasaran

(*product, price, place & promotion*) [4]. Hal ini menegaskan bahwa strategi pelayanan yang baik dan hubungan personal yang berkelanjutan dapat meningkatkan loyalitas nasabah. Konsep ini berkembang di era modern melalui pendekatan *Electronic Customer Relationship Management* (e-CRM) yang dirancang untuk memperkuat efektivitas dan efisiensi hubungan antara bank dan nasabah. Penerapan e-CRM memungkinkan personalisasi layanan berbasis data, mempercepat respons terhadap kebutuhan pengguna, dan mendukung pengambilan keputusan strategis berbasis analitik [5].

Dalam kerangka teori *behavioral intention* dan CRM, sikap dan persepsi positif dapat mencerminkan loyalitas pengguna yang dimana hal tersebut dapat mendorong mereka untuk menggunakan layanan secara terus menerus (*continuance intention*) dan bahkan merekomendasikannya kepada orang lain (*advocacy*). Sebaliknya, sentiment negatif bisa menjadi tanda adanya resiko pengguna berpindah ke aplikasi lain. Karena itu, analisis sentimen tidak hanya berguna untuk menilai aspek teknis lainnya, tetapi juga penting sebagai penghubung antara persepsi pengguna dan tingkat loyalitas mereka.

Analisis sentimen juga dapat mencerminkan dimensi penting dalam kualitas layanan digital seperti efisiensi, kepuasan pengguna, kemudahan pembelajaran (*learnability*), daya ingat (*memorability*), serta tingkat kesalahan (*error rate*) [6]. Dimensi tersebut sangat relevan dalam membangun pengalaman pengguna yang positif, sehingga memungkinkan pihak perbankan melakukan perbaikan layanan secara berkelanjutan dan membuat nasabah merasa nyaman dan tetap setia. Penelitian [7] mengungkapkan bahwa pemanfaatan *Big Data Analytical Intelligence* (BDAI) dalam sistem *Customer Relationship Management* (CRM) memungkinkan organisasi untuk memberikan layanan yang lebih personal dan responsif, sekaligus mempererat hubungan jangka panjang antara bank dan nasabah.

Seiring dengan kemajuan dalam teknik analitik, penggunaan *machine learning* untuk klasifikasi sentimen kini semakin luas. digunakan dalam klasifikasi sentiment. Berbagai algoritma seperti Naïve Bayes, *K-Nearest Neighbor* (KNN), dan *Support Vector Machine* (SVM), telah banyak dimanfaatkan untuk menganalisis ulasan dari pengguna. Beberapa studi menunjukkan bahwa SVM secara konsisten menghasilkan akurasi yang tinggi dalam mengklasifikasikan data teks. Misalnya penelitian [8], melaporkan akurasi SVM sebesar 92% dalam menganalisis ulasan aplikasi BRImo. Sementara [9], menemukan bahwa SVM mengungguli Naïve Bayes dalam mengklasifikasikan ulasan aplikasi SeaBank, dengan akurasi 93,99%.

Lebih lanjut, [10] menggunakan pendekatan berbasis *cloud* dengan algoritma Stanford CoreNLP dan Naïve Bayes untuk mengklasifikasikan sentimen pada data Twitter bank di Maroko. Hasilnya menunjukkan bahwa CoreNLP menghasilkan presisi 80,7% pada ulasan berbahasa Inggris. Penelitian oleh [11] pada aplikasi NeoBank juga menunjukkan efektivitas SVM dalam klasifikasi data ulasan dengan akurasi terbaik 82,33%, serta F1- score 86% untuk

sentimen positif. Pendekatan yang lebih kompleks dilakukan oleh [12], yang menggunakan model ELECTRA berbasis substitusi sinonim dan antonim untuk meningkatkan akurasi klasifikasi sentimen dalam konteks perbankan, mencapai akurasi 93,47% dan mengungguli model lain seperti FinBERT.

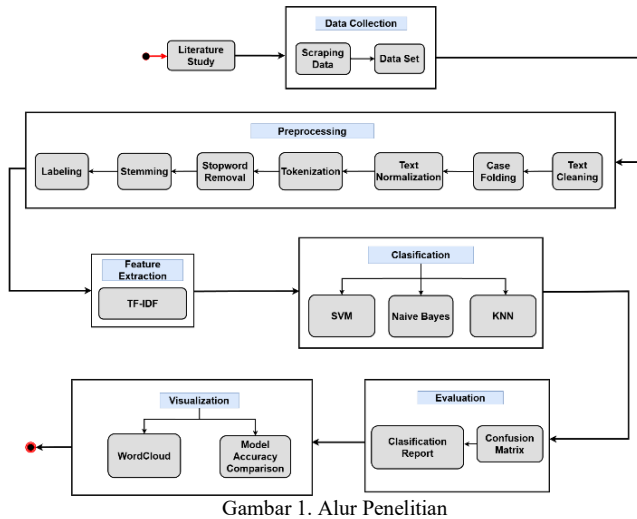
Sementara itu, [13] menerapkan optimasi robust pada CNN dan RNN untuk mengatasi ketidakseimbangan data ulasan nasabah bank di Thailand. Model CNN berhasil meningkatkan akurasi dari 83,2% menjadi 87,5%. Dalam studi oleh [14], SVM dan *Random Forest* dibandingkan dengan tiga algoritma lain dimana SVM mencatat akurasi hingga 91% pada klasifikasi sentimen dan 84% pada klasifikasi emosi. Analisis lebih lanjut oleh [15] menunjukkan bahwa *Decision Tree* menghasilkan akurasi prediksi risiko kebangkrutan bank sebesar 95% dengan tingkat kesalahan *false positive* dan *false negative* yang sangat rendah (2,94%), saat dikombinasikan dengan analisis sentimen laporan keuangan.

Berbagai studi terdahulu menunjukkan bahwa algoritma SVM seringkali memberikan performa terbaik dalam klasifikasi ulasan aplikasi *mobile banking*. Namun penelitian-penelitian tersebut belum secara spesifik menyoroti aplikasi Wondr by BNI dan belum ada kajian yang mendalam tentang bagaimana persepsi pengguna berhubungan dengan loyalitas nasabah. Penelitian sebelumnya lebih fokus pada kajian lain dengan metode berbeda, sehingga pemahaman tentang sentimen pengguna Wondr dan dampaknya terhadap loyalitas masih kurang. Dalam konteks inilah, proses *knowledge discovery* menjadi sangat penting karena memungkinkan penggalian informasi yang berharga dari data ulasan pengguna secara sistematis [16], guna menghasilkan pengetahuan yang relevan dan bermanfaat bagi penguatan strategi bisnis, khususnya dalam pengembangan e-CRM.

Tujuan penelitian ini adalah untuk mengisi celah tersebut dengan menganalisis ulasan pengguna aplikasi Wondr yang menggunakan tiga algoritma *machine learning*: Naive Bayes, *K- Nearest Neighbor* (KNN), dan *Support Vector Machine* (SVM). Algoritma-algoritma ini digunakan untuk mengklasifikasikan sentimen dalam kategori positif dan negatif serta menelaah implikasinya terhadap loyalitas nasabah. Dengan demikian, penelitian ini tidak hanya menilai kinerja model klasifikasi, tetapi juga mengeksplorasi bagaimana pola sentiment dapat menjadi indikator perilaku loyalitas pengguna. Hasil analisis diharapkan memberikan insight strategis yang aplikatif bagi BNI maupun institusi perbankan lainnya, khususnya dalam merancang strategi *customer retention*, komunikasi personalisasi, dan perbaikan pengalaman pengguna (UX *redesign*) melalui pendekatan CRM berbasis data.

II. METODE

Studi ini dilakukan dengan pendekatan kuantitatif dan analisis sentimen yang didukung oleh teknik *machine learning*. Proses penelitian disusun secara terstruktur dan mengikuti alur kerja yang sistematis dari tahap awal hingga tahap akhir. Alur kerja tersebut menggambarkan keseluruhan tahapan yang dilakukan oleh peneliti, sebagaimana ditampilkan pada Gambar 1 berikut.



Gambar 1. Alur Penelitian

A. Literature Study

Tahap studi literatur dilakukan dengan mengumpulkan data dari berbagai sumber, seperti jurnal dan artikel, yang nantinya digunakan sebagai landasan untuk mendukung jalannya penelitian. Dalam studi ini, penulis mencari data dan bahan literatur dari jurnal atau artikel yang berhubungan dengan analisis sentimen, *Customer Relationship Management* (CRM), *mobile banking*, klasifikasi Naive Bayes, SVM, dan KNN.

B. Data Collection

Data dalam penelitian ini diperoleh dari ulasan pengguna aplikasi Wondr by BNI yang tersedia di platform Google Play Store. Proses pengumpulan data dilakukan dengan teknik *scraping* menggunakan bahasa pemrograman Python di platform Google Colaboratory. Data awal yang discraping berjumlah 5.000 ulasan, kemudian data yang bersifat duplikat dihapus sehingga menyisakan 3.191 data ulasan, setelah itu data dengan nilai kosong pada kolom juga dihapus, sehingga jumlah ulasan yang digunakan untuk analisis sentiment menjadi 3.132. Data yang dikumpulkan mencakup *Review ID*, *username*, *rating* (penilaian 1-5), *review text*, dan *date*. Sampel hasil *scraping* yang telah disimpan dalam format CSV dapat dilihat pada Tabel 1.

TABEL I
SAMPEL HASIL SCRAPING DATA

Review ID	Username	Rating	Review Text	Date
5663b04d-ae65-49b4-955d-9ff8ee984b87	Pengguna Google	1	kurang informatif terutama soal histori transaksi semoga update berikutnya.....	6/3/2025 6:50:04 AM
02502d56-c484-40a8-afde-571b21b7b0	Pengguna Google	5	sangat membantu	6/3/2025 6:50:04 AM
37df4ef5-9790-47be-8d45-2a327cf78afa	Pengguna Google	5	mudah dan aman	6/3/2025 6:50:04 AM
...
1512109e-cba0-4239-b091-00902e96e1d1	This is the way Official	5	Pelayanan ok memudahkan	6/3/2025 6:50:04 AM

C. Preprocessing

Preprocessing merupakan tahap paling penting dalam analisis sentimen, karena berpengaruh pada performa klasifikasi [17]. Proses ini dilakukan untuk membersihkan dan menyelaraskan data teks sehingga lebih mudah diproses dan dianalisis oleh model *machine learning*. Langkah pertama yang dilakukan dalam *preprocessing* adalah *text cleaning*, yaitu proses pembersihan data teks dari komponen yang dianggap *noise* seperti URL, tag HTML, emoji, simbol, dan angka [18]. Setelah itu dilakukan *case folding* dengan mengubah semua kata dari huruf a hingga z menjadi huruf kecil untuk menjaga konsistensi penulisan [19]. Proses dilanjutkan dengan normalisasi kata menggunakan kamus kata tidak baku, sehingga kata-kata slang atau informal diganti ke bentuk bakunya.

Tahap berikutnya adalah *tokenization*, yaitu memecah teks ulasan menjadi unit-unit kecil. Pada tahap ini tanda baca, karakter khusus, atau simbol yang tidak relevan juga dihilangkan agar hanya tersisa kata-kata bermakna [20]. Setelah teks dipecah menjadi token, dilakukan *stopword removal* untuk membuang kata-kata umum yang frekuensinya tinggi namun tidak membawa makna sentiment, seperti “di”, “yang”, “dan”, “untuk”, atau “dengan” [21]. Kemudian dilakukan *stemming* dengan menggunakan algoritma bahasa Indonesia (misalnya Sastrawi), yang mengubah kata berimbuhan menjadi bentuk dasarnya. Tahap ini penting untuk menyatukan variasi kata, misalnya kata “membantu”, “dibantu”, dan “bantuan” semuanya diubah menjadi kata dasar “bantu” [22]. Dan tahap terakhir adalah *labeling*, yaitu

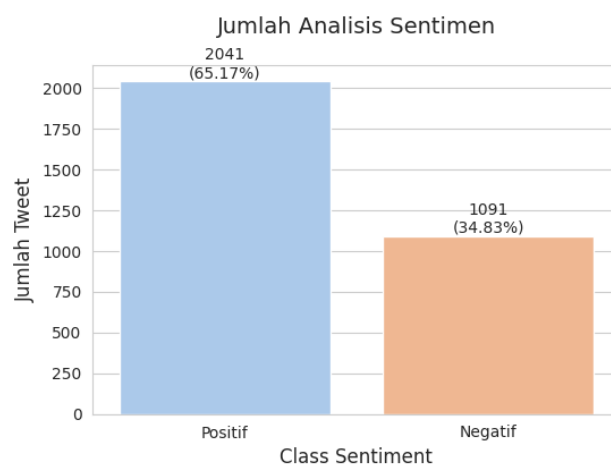
pemberian kategori pada setiap ulasan agar bisa dipahami model klasifikasi [23]. Dalam penelitian ini, rating digunakan sebagai acuan: ulasan dengan rating ≤ 3 diberi label “Negatif”, sedangkan ulasan dengan rating > 3 diberi label “Positif”. Hasil implementasi *preprocessing* dapat dilihat pada Tabel 2 dan 3 berikut.

TABEL II
HASIL PREPROCESSING

Tahap Preprocessing	Hasil
Sebelum Preprocessing	aplikasi nya bagus, tapi kalau sudah keluar mau masuk lagi susah minta ampun.. wajah yg punya dia gak kenal 😞😞 Makin lama makin jelas aplikasi ini sulit mengenali wajah tapi passwordnya pakai wajah, are you serius yg punya ?
Cleaning	aplikasi nya bagus tapi kalau sudah keluar mau masuk lagi susah minta ampun wajah yg punya dia gak kenal Makin lama makin jelas aplikasi ini sulit mengenali wajah tapi passwordnya pakai wajah are you serius yg punya
Case Folding	aplikasi nya bagus tapi kalau sudah keluar mau masuk lagi susah minta ampun wajah yg punya dia gak kenal makin lama makin jelas aplikasi ini sulit mengenali wajah tapi passwordnya pakai wajah are you serius yg punya
Normalization	aplikasi ya bagus tapi kalau sudah keluar mau masuk lagi susah meminta ampun wajah yang punya dia tidak kenal makin lama makin jelas aplikasi ini sulit mengenali wajah tapi passwordnya pakai wajah are you serius yang punya
Tokenization	['aplikasi', 'ya', 'bagus', 'tapi', 'kalau', 'sudah', 'keluar', 'mau', 'masuk', 'lagi', 'susah', 'meminta', 'ampun', 'wajah', 'yang', 'punya', 'dia', 'tidak', 'kenal', 'makin', 'lama', 'makin', 'jelas', 'aplikasi', 'ini', 'sulit', 'mengenali', 'wajah', 'tapi', 'passwordnya', 'pakai', 'wajah', 'are', 'you', 'serius', 'yang', 'punya']
Stopword Removal	['aplikasi', 'ya', 'bagus', 'masuk', 'susah', 'ampun', 'wajah', 'kenal', 'aplikasi', 'sulit', 'mengenali', 'wajah', 'passwordnya', 'pakai', 'wajah', 'are', 'you', 'serius']
Steaming	aplikasi ya bagus masuk susah ampun wajah kenal aplikasi sulit nali wajah passwordnya pakai wajah are you serius

TABEL III
HASIL LABELING

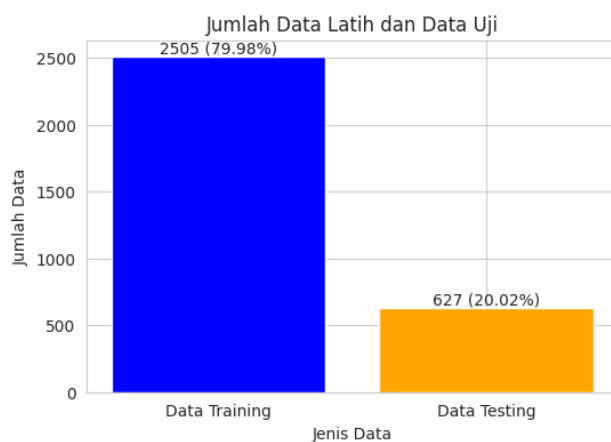
Hasil Steaming	Rating	Label
aplikasi ya bagus masuk susah ampun wajah kenal aplikasi sulit nali wajah passwordnya pakai wajah are you serius	3	Negatif



Gambar 2. Jumlah Analisis Sentimen

Berdasarkan hasil labeling, diperoleh 2.041 ulasan positif atau sekitar 65,17% dari seluruh data ulasan dan 1.091 ulasan negatif atau sekitar 34,83%, sebagaimana ditunjukkan pada Gambar 2. Meskipun aplikasi Wondr ini masih terbilang baru tetapi distribusi tersebut menunjukkan bahwa mayoritas pengguna aplikasi ini memberikan penilaian yang positif, yang menandakan bahwa pengguna mengapresiasi kemudahan dan manfaat layanan yang diberikan oleh aplikasi ini. Namun proporsi ulasan negatif yang cukup besar yaitu hampir sepertiga dari total data, menjadi hal yang perlu diperhatikan oleh pihak BNI terkait permasalahan yang dirasakan oleh pengguna. Ketidakeimbangan distribusi kelas ini menjadi pertimbangan penting dalam proses klasifikasi, karena model *machine learning* akan lebih mudah mengenali pola pada kelas dominan (positif) dibanding dengan kelas minoritas (negatif). Dari perspektif CRM, dominasi ulasan positif dapat dimanfaatkan dalam strategi *customer retention* dengan memperkuat aspek layanan yang sudah diapresiasi, sementara ulasan negatif berperan penting sebagai masukan bagi tim pengembangan untuk melakukan UX *redesign* pada fitur yang sering menimbulkan keluhan, seperti *login* atau verifikasi.

D. Feature Extraction



Gambar 3. Pembagian Data Latih Dan Data Uji

Sebelum dilakukan *feature extraction*, dataset terlebih dahulu dibagi menjadi data latih dan data uji untuk memastikan model dapat dilatih dan dievaluasi secara objektif seperti terlihat pada Gambar 3, sebanyak 2.505 data (79,98%) digunakan sebagai data *training*, sedangkan 627 data (20,02%) dialokasikan sebagai data *testing*. Proporsi pembagian 80:20 ini dipilih karena umum digunakan dalam penelitian klasifikasi teks, Dimana Sebagian besar data digunakan untuk melatih model agar dapat mengenali pola, sementara Sebagian lainnya disisihkan untuk menguji performa model terhadap data baru yang belum pernah dilihat sebelumnya.

Setelah pembagian data proses ekstraksi fitur baru bisa dilakukan. Proses ini digunakan untuk mengubah data teks menjadi bentuk numerik sehingga dapat diolah secara matematis [24]. Metode TF-IDF (*Term Frequency - Inverse Document Frequency*) ini menggabungkan dua ide: seberapa sering kata muncul dalam sebuah dokumen dan seberapa jarang kata tersebut muncul dalam kumpulan dokumen [25]. Fitur ini memiliki keunggulan dalam implementasinya yang mudah dan memiliki nilai akurasi yang lebih besar dibanding fitur ekstraksi lainnya [26]. Perhitungan manual TF-IDF dapat dilakukan menggunakan persamaan berikut.

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

$$IDF(t) = \log \left(\frac{N}{df(t)} \right) \quad (2)$$

Keterangan :

N = Jumlah Total Dokumen

$df(t)$ = Jumlah Dokumen yang Mengandung Kata t

$$TF-IDF(t,d) = TF(t,d) \times IDF(t) \quad (3)$$

TABEL IV
HASIL LABELING

Kata	Bobot TF-IDF
sip	0.655186
gampang	0.544538
praktis	0.523649
pada	0.000000
padah	0.000000
padahal	0.000000
pagi	0.000000
paham	0.000000
pajak	0.000000
pandai	0.000000

E. Clasification

Pada tahap klasifikasi, model *machine learning* dilatih untuk mempelajari pola-pola dalam data fitur numerik (hasil dari TF- IDF) dan mengaitkannya dengan label sentimen yang sudah ditentukan (Positif atau Negatif, berdasarkan rating ulasan). Berikut tiga algoritma klasifikasi yang digunakan:

1) *Support Vector Machine (SVM)*: SVM adalah algoritma klasifikasi berbasis *margin* yang berfungsi untuk menemukan *hyperplane* terbaik guna memisahkan dua kelas

data [27]. Untuk data yang tidak dapat dipisahkan secara linier, algoritma ini bekerja dengan memaksimalkan jarak atau *margin* antara masing-masing kelas data, SVM dapat menggunakan fungsi *kernel* agar tetap mampu melakukan klasifikasi. Model ini dikenal efektif dalam menangani data bedimensi tinggi maupun data yang bersifat *nonlinier*, sehingga sering digunakan dalam analisis sentimen [28]. Secara umum prediksi kelas dilakukan menggunakan persamaan 4 berikut.

$$y(x) = \text{sign}(w \cdot x + b) \quad (4)$$

Keterangan :

w = Vektor Bobot

x = Vektor Fitur

b = Bias

2) *K-Nearest Neighbor (KNN)*: Algoritma ini mengklasifikasikan data berdasarkan seberapa mirip data lain di sekitarnya [29]. Sebuah data baru diklasifikasikan kedalam kelas mayoritas dari k data latih terdekat dari ruang fitur. Jarak antar data dihitung menggunakan metrik tertentu seperti *Euclidean* atau *Cosine similarity*. Model ini terbukti mampu memberikan akurasi yang tinggi dalam klasifikasi sentimen [30]. Secara umum jarak *Euclidean* antara dua vektor x dan y dihitung menggunakan persamaan 5 berikut.

$$\text{Jarak}(x, y) = \sum_{i=1}^n (x_i - y_i)^2 \quad (5)$$

Setelah dihitung, algoritma menentukan kelas berdasarkan frekuensi kemunculan kelas dari k tetangga terdekat.

3) *Naïve Bayes*: Algoritma klasifikasi menggunakan teorema probabilitas, dengan asumsi bahwa setiap fitur saling independen. Algoritma ini cukup populer dalam analisis sentimen karena sifatnya yang sederhana, cepat, dan mampu menghasilkan akurasi yang cukup baik [31][32]. Berikut rumus dasar untuk menghitung probabilitas suatu kelas C berdasarkan fitur (kata).

$$P(C|w_1, w_2, \dots, w_n) \propto P(C) \cdot \prod_{i=1}^n P(w_i|C) \quad (6)$$

Kelas yang memiliki nilai probabilitas tertinggi dipilih sebagai hasil klasifikasi. Estimasi probabilitas dilakukan dari data latih, dan untuk menghindari nilai nol digunakan teknik *smoothing* seperti Laplace *smoothing*.

F. Evaluation

Untuk mengetahui seberapa baik model klasifikasi memprediksi sentiment pada data uji (data yang belum pernah dilihat), tahap evaluasi sangat penting [33]. Tabel *confusion matrix* menunjukkan jumlah prediksi yang benar baik untuk kelas positif maupun negatif serta jumlah kesalahan prediksi yaitu (*False Positives*) dan (*False Negatives*) untuk masing-masing kelas sentimen [34]. Berdasarkan *confusion matrix*, metrik kinerja seperti *Accuracy*, *Precision*, *Recall*, dan *F1-score* dapat dihitung secara manual melalui persamaan berikut.

$$\text{Accuracy} = \frac{TP+TN}{TP+FP+TN+FN} \quad (7)$$

Accuracy adalah metrik yang mengukur kemampuan model untuk mengklasifikasikan data dengan benar secara keseluruhan. Ini adalah rasio dari jumlah prediksi yang benar terhadap jumlah data keseluruhan.

$$Precision = \frac{TP}{TP+FP} \quad (8)$$

Persentase prediksi positif yang benar dari keseluruhan prediksi positif yang dibuat oleh model disebut *Precision*. Metrik ini penting untuk situasi dimana kesalahan prediksi positif memiliki konsekuensi yang signifikan.

$$Recall = \frac{TP}{TP+FN} \quad (9)$$

Recall atau *sensitivity* mengukur kemampuan model dalam mendeteksi seluruh data positif yang benar. Metrik ini menjadi penting ketika tujuan utama adalah meminimalkan kesalahan klasifikasi pada kelas positif.

$$F1-Score = \frac{2}{\frac{1}{Recall} + \frac{1}{Precision}} \quad (10)$$

F1-score merupakan rata-rata dari *precision* dan *recall*. Metrik ini berguna ketika terdapat ketidakseimbangan kelas dalam data, dan dibutuhkan kompromi antara *precision* dan *recall*.

G. Visualization

Tahap terakhir dalam alur penelitian ini yaitu visualisasi, yang bertujuan untuk menyajikan hasil *processing* dan kinerja model dalam bentuk grafis agar mudah dipahami dan dianalisis oleh pembaca [35]. Pada tahap ini, penulis memvisualisasikan perbandingan akurasi dari ketiga model dalam bentuk *bar chart*, yang memberikan gambaran secara langsung mengenai model yang paling optimal dalam mengklasifikasi sentiment pengguna. Selain itu, visualisasi sentimen positif dan negatif disajikan dalam bentuk *WordCloud* guna memudahkan identifikasi kata yang paling sering muncul dalam masing-masing kategori ulasan. Oleh karena itu visualisasi ini tidak hanya memperkuat pemahaman terkait hasil analisis data, tetapi juga mengungkap informasi tersembunyi yang bermanfaat bagi pertumbuhan ekonomi dan pengelolaan sistem perbankan khususnya dalam *Customer Relationship Managemengt* (CRM), visualisasi sentiment yang dihasilkan membantu industri perbankan dalam memahami persepsi pengguna terhadap fitur dan layanan, yang menjadi indikator penting dalam mengukur serta meningkatkan loyalitas nasabah. Pola-pola sentiment inilah yang menjadi *Knowledge Discovery* sebagai dasar penyusunan strategi perbaikan layanan berbasis data [36].

III. HASIL DAN PEMBAHASAN

A. Hasil Klasifikasi Model SVM

SVM Confusion Matrix

Actual	Negatif	184	34
	Positif	44	365
		Negatif	Positif
		Predicted	

Gambar 4. Confusion Matrix Model SVM

Hasil *confusion matrix Support Vector Machine* (SVM) berukuran 2 x 2 menunjukkan bahwa model menemukan 184 data negatif yang benar (*True Negative*) dan 365 data positif yang benar. Namun, terdapat klasifikasi tambahan yaitu 34 data negatif yang keliru dianggap sebagai positif (*False Positive*) dan 44 data positif yang salah dianggap sebagai negative (*False Negative*). Hasil ini membuktikan bahwa algoritma SVM memiliki akurasi klasifikasi yang cukup baik dengan jumlah prediksi benar yang jauh lebih besar dibandingkan dengan jumlah kesalahan. Berikut *classification report* yang dihasilkan berdasarkan *confusion matrix SVM*.

Classification Report for SVM:

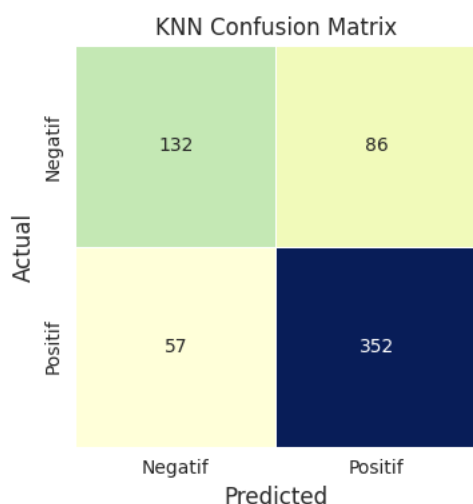
	precision	recall	f1-score	support
Negatif	0.807	0.844	0.825	218.000
Positif	0.915	0.892	0.903	409.000
accuracy	0.876	0.876	0.876	0.876
macro avg	0.861	0.868	0.864	627.000
weighted avg	0.877	0.876	0.876	627.000

Gambar 5. Clasification Report Model SVM

Secara keseluruhan, akurasi model SVM mencapai 87,6%, menandakan bahwa model ini cukup andal dalam mengklasifikasikan data. Kinerja terbaik ditunjukkan pada kelas positif dengan *precision* 0,915, *recall* 0,892, dan *f1-score* 0,903, sementara itu pada kelas negatif nilainya sedikit lebih rendah dengan *f1-score* 0,825, menandakan masih adanya ulasan negatif yang salah diklasifikasi. Nilai *macro average* (0,864) dan *weighted average* (0,876) menunjukkan

bahwa model tetap seimbang meskipun data tidak seimbang antara ulasan positif dan negatif. Temuan ini menegaskan bahwa SVM mampu memberikan representasi yang lebih baik dalam memprediksi sentimen, khususnya pada kelas dominan. Dengan demikian, model ini tidak hanya memfasilitasi pemetaan pola kepuasan pelanggan, tetapi juga menjadi dasar bagi strategi pengembangan layanan dan perbaikan fitur aplikasi.

B. Hasil Klasifikasi Model KNN



Gambar 6. Confusion Matrix Model KNN

Berdasarkan *confusion matrix* untuk model KNN di atas, diketahui bahwa model berhasil mengklasifikasikan 352 data positif dan 132 data negatif secara benar (*True Positive* dan *True Negative*). Namun, terdapat kesalahan klasifikasi lainnya, yaitu 57 data positif yang keliru diprediksi sebagai negatif (*False Negative*) dan 86 data negatif yang salah diprediksi sebagai Positif (*False Positive*). Jika dibandingkan dengan model SVM sebelumnya, model KNN menghasilkan jumlah kesalahan klasifikasi yang lebih tinggi. Hal ini menunjukkan bahwa akurasi dan ketepatan klasifikasi KNN lebih rendah dibandingkan SVM dalam konteks pemrosesan data sentimen.

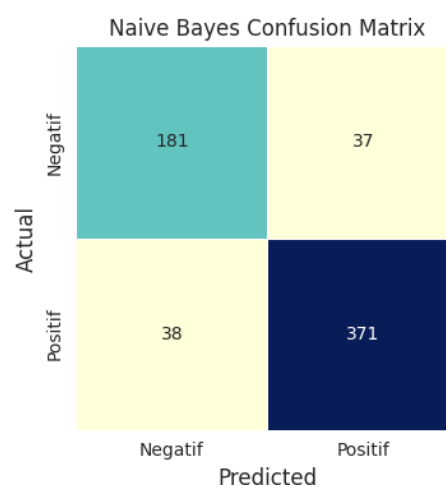
Classification Report for KNN:				
	precision	recall	f1-score	support
Negatif	0.698	0.606	0.649	218.000
Positif	0.804	0.861	0.831	409.000
accuracy	0.772	0.772	0.772	0.772
macro avg	0.751	0.733	0.740	627.000
weighted avg	0.767	0.772	0.768	627.000

Gambar 7. Clasification Report Model KNN

Classification report model KNN, memiliki performa yang relatif rendah dibandingkan dengan model lainnya, dengan akurasi keseluruhan hanya 77,2%, hal ini

menunjukkan kemampuan model dalam mengklasifikasi sebagian besar data uji dengan tepat. Kinerja pada kelas positif juga cukup baik dengan *precision* sebesar 0.804, *recall* 0.861, dan *f1-score* 0.831, yang berarti bahwa model jarang melakukan kesalahan klasifikasi terhadap data positif. Namun kelemahan utama terlihat pada kelas negatif, dimana *recall* hanya mencapai 0,606 dengan *f1-score* 0,649, menandakan bahwa banyak ulasan negatif tidak berhasil teridentifikasi dengan baik. Nilai *macro average* (0,740) dan *weighted average* (0,768) juga menunjukkan ketidakseimbangan performa antar kelas. Hal ini mengindikasikan bahwa KNN kurang efektif jika dijadikan model utama untuk mendukung strategi CRM yang membutuhkan sensitivitas tinggi terhadap keluhan pengguna.

C. Hasil Klasifikasi Model Naïve Bayes



Gambar 8. Confusion Matrix Naïve Bayes

Confusion Matrix pada model Naive Bayes, diperoleh hasil klasifikasi yang cukup baik, dengan 371 data positif diklasifikasikan secara benar (*True Positive*) dan 181 data negatif yang juga terklasifikasi dengan tepat (*True Negative*). Kesalahan klasifikasi yang terjadi relatif rendah, yaitu 38 data positif yang diprediksi sebagai negatif (*False Negative*) dan data negatif yang diprediksi sebagai positif (*False Positive*). Jika dibandingkan dengan model sebelumnya, yaitu KNN dan SVM, model Naive Bayes menunjukkan performa yang mendekati model SVM, bahkan dalam beberapa metrik seperti *recall* untuk kelas positif, Naive Bayes kemungkinan unggul karena jumlah *False Negative* yang lebih kecil.

< 0,05), di mana SVM dan Naive Bayes tidak berbeda satu sama lain, tetapi keduanya lebih baik secara konsistensi dibandingkan KNN. Temuan ini memberikan implikasi langsung pada strategi *Customer Relationship Management* (CRM), di mana ulasan positif dapat dimanfaatkan untuk memperkuat *customer retention* dengan menonjolkan keunggulan aplikasi pada kemudahan dan kecepatan layanan, sementara ulasan negatif penting dijadikan dasar perbaikan fitur melalui *UX redesign*. Dengan demikian, analisis sentimen ini tidak hanya mengungkap persepsi pengguna, tetapi juga memberi arahan praktis yang bisa diterapkan BNI dalam meningkatkan kualitas aplikasi sekaligus menjaga loyalitas nasabah di tengah persaingan perbankan digital.

DAFTAR PUSTAKA

- [1] R. A. Akob and Z. Sukarno, "Pengaruh Kualitas Layanan Mobile Banking terhadap Kepuasan dan Loyalitas Nasabah Bank BUMN di Makassar," *Jurnal Maksipreneur: Manajemen, Koperasi, dan Entrepreneurship*, vol. 11, no. 2, pp. 269–283, Feb. 2022, doi: 10.30588/jmp.v11i2.889.
- [2] N. A. Rahman, S. D. Idrus, and N. L. Adam, "Classification of customer feedbacks using sentiment analysis towards mobile banking applications," *IAES International Journal of Artificial Intelligence*, vol. 11, no. 4, pp. 1579–1587, Dec. 2022, doi: 10.11591/ijai.v11i4.pp1579-1587.
- [3] T. Winasih and L. Hakim, "Peran Kepuasan Sebagai Variabel Mediasi Kualitas Layanan dan Kepercayaan Nasabah terhadap Loyalitas Nasabah Bank Syariah," *Iqtishadia : Jurnal Ekonomi dan Perbankan Syariah*, vol. 8, no. 2, pp. 206–219, Dec. 2021, doi: 10.1905/iqtishadia.v8i2.4152.
- [4] S. Muliadi, "Determinan Loyalitas Nasabah Bank NTB Syariah," *IQTISHADUNA: Jurnal Ilmiah Ekonomi Kita*, vol. 11, no. 2, pp. 184–199, Dec. 2022, doi: 10.46367/iqtishaduna.v11i2.825.
- [5] A. Govindaraj and R. A. Muthalif, "Analyzing The Implementation And Impact Of Electronic Customer Relationship Management (E-Crm) In State Bank Of India: A Case Study On Enhancing Customer Engagement And Service Efficiency," *Proceedings on Engineering Sciences*, vol. 6, no. 3, pp. 989–994, 2024, doi: 10.24874/PES06.03.011.
- [6] H. Basri, M. B. S. Junianto, and I. Kusyudi, "Enhancing Usability Testing Through Sentiment Analysis: A Comparative Study Using SVM, Naive Bayes, Decision Trees and Random Forest," *Jurnal Teknologi Sistem Informasi dan Aplikasi*, vol. 7, no. 4, pp. 1603–1610, Oct. 2024, doi: 10.32493/jtsi.v7i4.45117.
- [7] C. Zhang, X. Wang, A. P. Cui, and S. Han, "Linking big data analytical intelligence to customer relationship management performance," *Industrial Marketing Management*, vol. 91, pp. 483–494, Nov. 2020, doi: 10.1016/j.indmarman.2020.10.012.
- [8] M. D. Bimantara and I. Zufria, "Text Mining Sentiment Analysis on Mobile Banking Application Reviews using TF-IDF Method with Natural Language Processing Approach," *JINAV: Journal of Information and Visualization*, vol. 5, no. 1, pp. 115–123, Jul. 2024, doi: 10.35877/454ri.jinav2772.
- [9] Y. Christian, T. Wibowo, and M. Lyawati, "Sentiment Analysis by Using Naive Bayes Classification and Support Vector Machine, Study Case Sea Bank," *Sinkron*, vol. 9, no. 1, pp. 258–275, Jan. 2024, doi: 10.33395/sinkron.v9i1.13141.
- [10] A. Riadsolh, I. Lasri, and M. Elbelkacemi, "Cloud-based sentiment analysis for measuring customer satisfaction in the moroccan banking sector using naive bayes and stanford nlp," *Journal of Automation, Mobile Robotics and Intelligent Systems*, vol. 14, no. 4, pp. 64–71, 2020, doi: 10.14313/JAMRIS/4-2020/47.
- [11] Kusnawi, M. Rahardi, and V. D. Pandiangan, "International Journal On Informatics Visualization journal homepage: www.joiv.org/index.php/joiv International Journal On Informatics Visualization Sentiment Analysis of Neobank Digital Banking Using Support Vector Machine Algorithm in Indonesia," *JOIV: International Journal On Informatics Visualization*, vol. 7, no. 2, pp. 377–383, 2023, Accessed: Jul. 08, 2025. [Online]. Available: <https://dx.doi.org/10.30630/joiv.7.2.1652>
- [12] A. Mohanty and R. C. Cherukuri, "Sentiment Analysis on Banking Feedback and News Data using Synonyms and Antonyms," *IJACSA International Journal of Advanced Computer Science and Applications*, vol. 14, no. 12, pp. 934–944, 2023, doi: 10.14569/IJACSA.2023.0141294.
- [13] B. Plubin, W. Bunyatisai, S. Plubin, and K. Jiamwattanapong, "Robust Optimization Base Deep Learning Model for Thai Banking Reviews Sentiment Analysis with Imbalanced Data," *Pak J Life Soc Sci*, vol. 22, no. 2, pp. 3330–3350, 2024, doi: 10.57239/PJLSS-2024-22.2.00245.
- [14] Edwina and T. Mauritsius, "Data-Driven Insights for Mobile Banking App Improvement: A Sentiment Analysis and Topic Modelling Approach for SimobiPlus User Reviews," *International Journal of Engineering Trends and Technology*, vol. 72, no. 6, pp. 347–360, Jun. 2024, doi: 10.14445/22315381/IJETT-V72I6P132.
- [15] D. P. de Jesus and C. da N. Besarria, "Machine learning and sentiment analysis: Projecting bank insolvency risk," *Research in Economics*, vol. 77, no. 2, pp. 226–238, Jun. 2023, doi: 10.1016/j.rie.2023.03.001.
- [16] N. A. Sofiah, K. D. Tania, A. Meiriza, and A. Wedhasmara, "A Comparative Assessment SARIMA and LSTM Models for the Gurugram Air Quality Index's Knowledge Discovery," in *2024 International Conference on Electrical Engineering and Computer Science (ICECOS)*, IEEE, Sep. 2024, pp. 26–31. doi: 10.1109/ICECOS63900.2024.10791243.
- [17] S. Khairunnisa, Adiwijaya, and S. Al Faraby, "Pengaruh Text Preprocessing terhadap Analisis Sentimen Komentar Masyarakat pada Media Sosial Twitter (Studi Kasus Pandemi COVID-19)," *Jurnal Media Informatika Budidarma*, vol. 5, no. 2, p. 406, Apr. 2021, doi: 10.30865/mib.v5i2.2835.
- [18] H. C. Husada and A. S. Paramita, "Analisis Sentimen Pada Maskapai Penerbangan di Platform Twitter Menggunakan Algoritma Support Vector Machine (SVM)," *Teknika*, vol. 10, no. 1, pp. 18–26, Feb. 2021, doi: 10.34148/teknika.v10i1.311.
- [19] Herwinskyah and A. Witanti, "Analisis Sentimen Masyarakat Terhadap Vaksinasi Covid-19 Pada Media Sosial Twitter Menggunakan Algoritma Support Vector Machine (SVM)," *Jurnal Sistem Informasi dan Informatika (Simika)*, vol. 5, no. 1, pp. 59–67, 2022, doi: 10.47080/simika.v5i1.1411.
- [20] B. Laurensz and E. Sediyo, "Analisis Sentimen Masyarakat terhadap Tindakan Vaksinasi dalam Upaya Mengatasi Pandemi Covid-19," *Jurnal Nasional Teknik Elektro dan Teknologi Informasi*, vol. 10, no. 2, 2021, doi: 10.22146/jnteti.v10i2.1421.
- [21] H. Utami, "Analisis Sentimen dari Aplikasi Shopee Indonesia Menggunakan Metode Recurrent Neural Network," *Indonesian Journal of Applied Statistics*, vol. 5, no. 1, p. 31, May 2022, doi: 10.13057/ijas.v5i1.56825.
- [22] S. Juanita, "Analisis Sentimen Persepsi Masyarakat Terhadap Pemilu 2019 Pada Media Sosial Twitter Menggunakan Naive Bayes," *Jurnal Media Informatika Budidarma*, vol. 4, no. 3, pp. 552–558, Jul. 2020, doi: 10.30865/mib.v4i3.2140.
- [23] G. K. Putri, H. Sujaini, and D. I. Ulumi, "Perbandingan Algoritma Support Vector Machine (SVM) dan Naive Bayes pada Analisis Sentimen Bahasa Jawa dan Sunda," *JUSTIN: Jurnal Sistem dan Teknologi Informasi*, vol. 13, no. 2, pp. 299–306, 2025, doi: 10.26418/justin.v13i2.88285.
- [24] A. S. Dharma and Y. G. R. Saragih, "Comparison of Feature Extraction Methods on Sentiment Analysis in Hotel Reviews," *Sinkron: Jurnal dan Penelitian Teknik Informatika*, vol. 6, no. 4, pp. 2349–2354, Oct. 2022, doi: 10.33395/sinkron.v7i4.11706.
- [25] M. Lestandy, Abdurrahim, and L. Syafa'ah, "Analisis Sentimen Tweet Vaksin COVID-19 Menggunakan Recurrent Neural Network dan Naive Bayes," *Jurnal RESTI*, vol. 5, no. 4, pp. 802–808, Aug. 2021, doi: 10.29207/resti.v5i4.3308.

- [26] S. M. P. Tyas, B. S. Rintyarna, and W. Suharso, "The Impact of Feature Extraction to Naïve Bayes Based Sentiment Analysis on Review Dataset of Indihome Services," *Digital Zone: Jurnal Teknologi Informasi dan Komunikasi*, vol. 13, no. 1, pp. 1–10, Apr. 2022, doi: 10.31849/digitalzone.v13i1.9158.
- [27] P. Arsi and R. Waluyo, "Analisis Sentimen Wacana Pemindahan Ibu Kota Indonesia Menggunakan Algoritma Support Vector Machine (SVM)," *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIIK)*, vol. 8, no. 1, pp. 147–156, 2021, doi: 10.25126/jtiik.202183944.
- [28] N. Fitriyah, B. Warsito, and D. A. I. Maruddani, "Analisis Sentimen Gojek Pada Media Sosial Twitter Dengan Klasifikasi Support Vector Machine (SVM)," *Jurnal Gaussian*, vol. 9, no. 3, pp. 376–390, 2020, doi: 10.14710/j-gauss.9.3.376-390.
- [29] A. R. Isnain, J. Supriyanto, and M. P. Kharisma, "Implementation of K-Nearest Neighbor (K-NN) Algorithm For Public Sentiment Analysis of Online Learning," *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, vol. 15, no. 2, p. 121, Apr. 2021, doi: 10.22146/ijccs.65176.
- [30] F. M. J. M. Shamrat *et al.*, "Sentiment analysis on twitter tweets about COVID-19 vaccines using NLP and supervised KNN classification algorithm," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 23, no. 1, pp. 463–470, Jul. 2021, doi: 10.11591/ijeecs.v23.i1.pp463-470.
- [31] H. A. Santoso, E. H. Rachmawanto, A. Nugraha, A. A. Nugroho, D. R. I. M. Setiadi, and R. S. Basuki, "Hoax classification and sentiment analysis of Indonesian news using Naive Bayes optimization," *Telkomnika (Telecommunication Computing Electronics and Control)*, vol. 18, no. 2, pp. 799–806, Apr. 2020, doi: 10.12928/TELKOMNIKA.V18I2.14744.
- [32] A. R. Isnain, N. S. Marga, and D. Alita, "Sentiment Analysis Of Government Policy On Corona Case Using Naive Bayes Algorithm," *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, vol. 15, no. 1, p. 55, Jan. 2021, doi: 10.22146/ijccs.60718.
- [33] M. F. Fakhrezi, A. F. Rochim, and D. M. K. Nugraheni, "Comparison of Sentiment Analysis Methods Based on Accuracy Value Case Study: Twitter Mentions of Academic Article," *Jurnal RESTI*, vol. 7, no. 1, pp. 161–167, Feb. 2023, doi: 10.29207/resti.v7i1.4767.
- [34] E. D. Madyatmadja, Shinta, D. Susanti, F. Anggreani, and D. J. M. Sembiring, "Sentiment Analysis on User Reviews of Mutual Fund Applications," *Journal of Computer Science*, vol. 18, no. 10, pp. 885–895, 2022, doi: 10.3844/jccsp.2022.885.895.
- [35] D. A. Agustina, S. Subanti, and E. Zukhronah, "Implementasi Text Mining Pada Analisis Sentimen Pengguna Twitter Terhadap Marketplace di Indonesia Menggunakan Algoritma Support Vector Machine," *IJAS: Indonesian Journal of Applied Statistic*, vol. 3, no. 2, pp. 109–122, 2020, doi: 10.13057/ijas.v3i2.44337.
- [36] V. Novalia, K. Ditha Tania, A. Meiriza, and A. Wedhasmara, "Knowledge Discovery of Application Review Using Word Embedding's Comparison with CNN-LSTM Model on Sentiment Analysis," in *2024 International Conference on Electrical Engineering and Computer Science (ICECOS)*, IEEE, Sep. 2024, pp. 234–238. doi: 10.1109/ICECOS63900.2024.10791113.