



Sentiment Analysis of Pluang Applications With Naive Bayes and Support Vector Machine (SVM) Algorithm

Analisis Sentimen Terhadap Aplikasi Pluang Menggunakan Algoritma Naive Bayes dan Support Vector Machine (SVM)

**Bagas Akbar Maulana^{1*}, Muhammad Jazilul Fahmi²,
Ari Muhamad Imran³, Nutriana Hidayati⁴**

^{1,2,3,4}Fakultas Teknologi Informasi dan Komunikasi, Universitas Semarang, Indonesia

E-Mail: ¹bagasakbarmaulana123@gmail.com, ²mjazilulfahmi@gmail.com,
³arimran49@gmail.com, ⁴anna@usm.ac.id

Received Dec 12th 2023; Revised Jan 24th 2024; Accepted Feb 5th 2024
Corresponding Author: Bagas Akbar Maulana

Abstract

Sentiment analysis is the process of analyzing text to determine the sentiment expressed, such as positive, negative, or neutral. Sentiment analysis can be used for various purposes, such as understanding public opinion, measuring customer satisfaction, and detecting fraud. In this research, we apply the sentiment classification method to analyze Pluang application reviews. The Pluang application is a digital investment application that offers various investment products, such as gold, shares and mutual funds. We use two sentiment classification models, namely Naive Bayes and SVM. The research results show that the SVM model works better than the Naive Bayes model. Specifically, the SVM model has an accuracy of 99.50%, precision of 99.67%, recall of 99.33%, and F1 score of 99.50%. Meanwhile, the Naive Bayes model has an accuracy of 99.25%, precision of 99.44%, recall of 99.06%, and F1 score of 99.25%. The advantage of the SVM model compared to the Naive Bayes model is its ability to differentiate positive and negative text better. This is caused by several factors, including: Negative data has more diverse characteristics than positive data, SVM uses kernel techniques which are able to map data into a higher dimensional space, so that it can model relationships between data better. Overall, this research shows that SVM is an effective classification method for analyzing the sentiment of Pluang application reviews.

Keyword: Naive Bayes, Pluang, Sentiment Analysis, Support Vector Machine.

Abstrak

Analisis sentimen merupakan proses menganalisis teks untuk menentukan sentimen yang diungkapkan, seperti positif, negatif, atau netral. Analisis sentimen dapat digunakan untuk berbagai tujuan, seperti memahami opini publik, mengukur kepuasan pelanggan, dan mendeteksi penipuan. Dalam penelitian ini, kami menerapkan metode klasifikasi sentimen untuk menganalisis ulasan aplikasi Pluang. Aplikasi Pluang adalah aplikasi investasi digital yang menawarkan berbagai produk investasi, seperti emas, saham, dan reksa dana. Kami menggunakan dua model klasifikasi sentimen, yaitu Naive Bayes dan Support Vector Machine (SVM). Hasil penelitian menunjukkan bahwa model SVM bekerja lebih baik dibandingkan model Naive Bayes. Secara spesifik, model SVM memiliki akurasi sebesar 99,50%, presisi sebesar 99,67%, recall sebesar 99,33%, dan skor F1 sebesar 99,50%. Sedangkan model Naive Bayes memiliki akurasi sebesar 99,25%, presisi sebesar 99,44%, recall sebesar 99,06%, dan skor F1 sebesar 99,25%. Kelebihan model SVM dibandingkan model Naive Bayes adalah kemampuannya untuk membedakan teks positif dan negatif dengan lebih baik. Ini disebabkan oleh beberapa faktor, antara lain: Data negatif memiliki karakteristik yang lebih beragam daripada data positif, SVM menggunakan teknik kernel yang mampu memetakan data ke dalam ruang dimensi yang lebih tinggi, sehingga dapat memodelkan hubungan antar data dengan lebih baik. Secara keseluruhan, penelitian ini menunjukkan bahwa SVM merupakan metode klasifikasi yang efektif untuk menganalisis sentimen ulasan aplikasi Pluang.

Kata Kunci: Analisis Sentimen, Naive Bayes, Pluang, Support Vector Machine.

1. PENDAHULUAN

Kemajuan teknologi yang pesat mengubah cara orang berinvestasi. Hampir semua perusahaan berlomba-lomba menciptakan platform digital agar dapat terus memberikan pelayanan kepada konsumen,

terutama dalam hal platform investasi. Seiring dengan meningkatnya minat dalam dunia investasi, pilihan instrumen investasi juga semakin meluas dan beragam. Aplikasi Pluang hadir sebagai solusi platform investasi multi-aset yang memberikan respons terhadap kebutuhan investor muda yang penuh optimisme terhadap pertumbuhan di berbagai kelas aset. Inovasi keuangan digital PT Bumi Santosa Cemerlang mengembangkan aplikasi investasi digital bernama Pluang, yang memungkinkan masyarakat untuk dengan mudah berinvestasi dalam emas. Harga emas satu gram senilai Rp. 8.638,-. Jika ingin mendapatkan keuntungan dari investasi di aplikasi pluang, Anda harus membeli emas di harga sedang turun dan menjualnya kembali di harga sedang naik[1]. Inovasi keuangan digital yang diberikan Pluang tidak hanya menawarkan emas saja, ada beragam produk investasi seperti saham, indeks AS, aset kripto, dan reksa dana. Analisis sentimen adalah tahapan di mana data teks dianalisis secara otomatis dengan tujuan untuk memahami dan mengekstrak sentimen yang ada didalam sebuah ulasan[2]. Dengan menggunakan analisis sentimen peneliti ini dapat menilai polaritas opini yang dikumpulkan dan dapat memprediksi apakah sentimen itu negatif atau Positif. Analisis sentimen merupakan sebuah penyelesaian yang memungkinkan penelitian ini untuk dapat mengukur tingkat kepuasan pengguna. Melalui metode analisis sentimen, penelitian ini dapat menyimpulkan informasi dari data yang bersifat acak [3]. Upaya menganalisa teks digital terhadap nada emosional pesan positif atau negatif dapat dilakukan dengan analisa sentimen dengan berbagai metode diantaranya metode Naive Bayes (NB), Decision Tree, K-Nearest Neighbor (K-NN), Neural Networks, dan Support Vector Machine (SVM) [4]. Penelitian ini mengambil pendekatan Algoritma Naive Bayes dan SVM dalam menentukan terhadap data teks pada pengguna aplikasi Pluang.

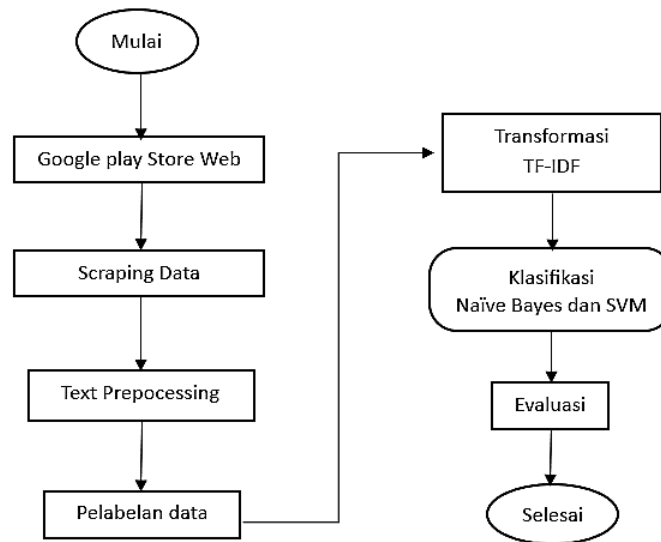
Berdasarkan penelitian-penelitian terdahulu, metode analisis sentimen menggunakan *Naive Bayes* dan SVM telah terbukti memberikan hasil evaluasi yang optimal. Oleh karena itu, penelitian ini diharapkan dapat memberikan manfaat sebagai saran dan masukan untuk mendukung peningkatan aplikasi Pluang. Pada penelitian terdahulu membahas perbandingan *naive bayes* dan SVM dalam melakukan analisis sentimen terhadap 1007 dataset yang diambil dari ulasan aplikasi *Zoom Cloud Meetings* di Play Store. SVM memberikan nilai akurasi dan acuan yang sedikit lebih baik daripada *Naive Bayes* [6]. Perbedaan penelitian terdahulu yaitu tidak menggunakan wordcloud, dimana wordcloud memiliki kelebihan untuk dapat mengetahui sentiment positif dan negative yang sering diberikan oleh customer[7]. Sedangkan pada penelitian terdahulu lainnya menemukan bahwa algoritma SVM memiliki akurasi 90,01 persen, *Naive Bayes* 79,20 persen, dan K-NN 62,10 persen. Hasilnya menunjukkan bahwa algoritma SVM memiliki tingkat akurasi yang lebih tinggi daripada algoritma Naive Bayes dan KNN [8]. Perbedaan penelitian ini yaitu menggunakan *K-Cross Validation*, yang memiliki kelebihan dapat melihat model yang memiliki akurasi terbaik karena data secara random dibagi menjadi K-partisi sehingga dapat diketahui komposisi model terbaik[9]. Kemudian untuk penelitian terdahulu lainnya membahas analisis sentimen untuk ulasan aplikasi E-wallet. Peneliti menemukan peningkatan akurasi yang cukup besar saat menggunakan fitur *Optimization Swarm Particle* [10]. Perbedaan penelitian ini yaitu peneliti menggunakan *Feature Selection*, yang memiliki kelebihan untuk dapat mengoptimalkan kinerja dari *classifier*, yang dapat meningkatkan akurasi dalam hasil penelitian[11].

Penelitian ini bertujuan untuk mengevaluasi kepuasan pengguna, memberikan pandangan mendalam tentang aspek-aspek aplikasi yang mempengaruhi kepuasan, memahami tren sentimen dari waktu ke waktu, dan memberikan dasar informasi untuk pengambilan keputusan strategis. Urgensinya terletak pada kemampuan memberikan wawasan mendalam terhadap aplikasi Pluang. Evaluasi ini dapat mencakup pengujian data pelatihan dan pengujian menggunakan metrik seperti *akurasi*, *presisi*, *recall*, dan *f1-score*. [5]. membandingkan kinerja klasifikasi antara algoritma *Naive Bayes* dan SVM dengan fokus pada akurasi, presisi, dan *recall*. Rujukan pada penelitian terdahulu Peneliti memberikan saran agar penelitian mendatang dapat meningkatkan akurasi dengan menambahkan jumlah data penelitian. Selain itu, penelitian dapat diperkaya dengan penambahan parameter pada pengujian SVM, seperti parameter literasi, gamma, dan lambda pada kernel[12]. Dengan pendekatan ini, penelitian selanjutnya diharapkan dapat lebih mendalam, menghasilkan hasil yang lebih akurat, dan lebih relevan dalam konteks penelitian yang bersangkutan. Harapannya, penelitian ini akan berperan sebagai indikator berharga dalam menilai sejauh mana analisis sentimen yang diimplementasikan pada aplikasi Pluang dapat menjadi dasar evaluasi yang lebih baik untuk penelitian berikutnya.

2. METODE PENELITIAN

Metode yang digunakan dalam penelitian ini adalah kuantitatif. Metode kuantitatif digunakan sebagai metode analisis untuk memahami persepsi pengguna terhadap aplikasi Pluang. Dalam konteks ini, penelitian ini menggunakan algoritma klasifikasi seperti Support Vector Machine (SVM) dan *Naive Bayes* untuk menganalisis dan mengukur sentimen dalam ulasan pengguna terhadap aplikasi Pluang. Pendekatan ini memanfaatkan teknik statistik dan komputasi untuk mengelompokkan ulasan pengguna ke dalam kategori sentimen positif dan negatif. Dengan menerapkan SVM dan *Naive Bayes*, penelitian ini bertujuan untuk mengukur dan mengevaluasi keakuratan dan kinerja model dalam menafsirkan sentimen pengguna terhadap

aplikasi Pluang, sehingga memberikan wawasan kesadaran pengguna yang lebih dalam terhadap platform investasi ini. Penelitian ini dijelaskan secara berurutan pada Gambar 1.



Gambar 1. Metode Penelitian

2.1. Scraping Data

Langkah pertama yang diambil untuk menganalisis sentimen aplikasi Pluang adalah eliminasi. *Scraping* adalah teknik yang mengubah data web tidak terstruktur menjadi data terstruktur yang dapat disimpan dan dianalisis dalam database atau spreadsheet. Proses web *Scraping* dibagi menjadi 3 tahapan yaitu scanning, ekstraksi dan konversi. Langkah *Fetch* dilakukan dengan mengakses melalui protokol *HTTP*, untuk mengirim dan menerima permintaan dari web server. kemudian dilakukan ekstraksi melalui halaman *HTML* dengan cara parsing data. Setelah data terkumpul, dilakukan transformasi data untuk memperoleh data terstruktur [13]. Dataset yang dihasilkan kemudian disimpan dalam format CSV. Dataset ini merepresentasikan ciri-ciri data berupa teks berupa kalimat atau kata sebanyak pendapat, opini atau sentimen, yang kemudian diolah dalam beberapa langkah hingga diperoleh nilai atau angka jumlah teks.

2.2. Dataset

Data yang digunakan adalah ulasan dari pengguna aplikasi Pluang di *Google Playstore*. *Web scraping* dilakukan pada 1 Januari 2024 dengan *Google Colaboratory*. Data hasil web *Scraping* diperoleh sebanyak 44,5 ribu ulasan yang kemudian dieliminasi menjadi sebanyak 2000 ulasan untuk Penelitian dalam jangka waktu real time. Untuk mengetahui kinerja aplikasi Pluang dalam hasil positif dan negatif, kemudian dilakukan filtering dan cleaning data.

Tabel 1. Dataset

Index	at	content	score
0	12/01/2024 12:12	aman, mudah, terpercaya, tentunya profit	5
1	11/01/2024 11:14	aplikasi sangat bagus dan membantu ke uangan	5
3	10/01/2024 05:32	Terimakasih karena sekarang semakin banyak pilihan coinnya. Tinggal update stacking coin selain btc eth.	5
1998	23/06/2019 11:47	akun sudah divalidasi...tapi email belum terverifikasi...sudah nyoba berkali2 tetep aja belum terkirim. gimana ini min	3
1999	15/11/2018 11:28	Ini aplikasi penipuan uang saya 6jt 2rts lebih di akun saya. malah akun saya di blok.dan saya mengalami ke rugian uang hampir 10jt. Dari ini aplikas. masa mau ganti no yg terdaftar gak bisa.dasar aplikasi penipu.u	1

2.3. Pelabelan Data

Proses pelabelan data dalam analisis sentimen merupakan langkah penting dalam menentukan kategori sentimen dari informasi yang tersedia. Dalam kasus ini, menentukan aturan yang memberikan label -1 negatif pada data dengan nilai kurang dari 3, dan label positif 1 pada data dengan nilai 4 dan 5. Langkah ini untuk menandai atau mengelompokkan setiap ulasan berdasarkan tingkat sentimen yang disertakan dalam skor ulasan. Tujuan dari proses ini adalah untuk mempersiapkan data untuk interpretasi oleh model analisis sentimen

selanjutnya untuk memahami opini dan pandangan yang terkandung dalam ulasan pengguna terhadap aplikasi tersebut.

Tabel 3. Pelabelan Data

Index	at	content	score	sentiment
0	12/01/2024 12:12	aman, mudah, terpercaya, tentunya profit	5	1
1	11/01/2024 11:14	aplikasi sangat bagus dan membantu ke uangan	5	1
3	10/01/2024 05:32	Terimakasih karena sekarang semakin banyak pilihan coinnnya. Tinggal update stacking coin selain btc eth.	5	1
1998	23/06/2019 11:47	akun sudah divalidasi...tapi email belum terverifikasi...sudah nyoba berkali2 tetep aja belum terkirim. gimana ini min	3	0
1999	15/11/2018 11:28	Ini aplikasi penipuan uang saya 6jt 2rts lebih di akun saya. malah akun saya di blok.dan saya mengalami ke rugian uang hampir 10jt. Dari ini aplikasi. masa mau ganti no yg terdaftar gak bisa. dasar aplikasi penipu.u	1	-1

2.4. Text Preprocessing

Tahap *text preprocessing* digunakan untuk memperoleh data yang sesuai dengan proses dari sistem yang akan dibangun, dimana sudah ada pengurangan kosa kata, menghilangkan noise, membuat data lebih terstruktur sehingga dihapuskan dapat mempermudah dan mempercepat proses [14]. Langkah pertama dalam pengolahan kata adalah *Case folding* mengubah huruf besar menjadi kecil untuk menyamakan atau menyeragamkan karakter pada data teks [15]. Kemudian, *stop word* dihilangkan, yaitu kata-kata umum yang tidak memberikan kontribusi signifikan terhadap pemahaman konten. Teks tersebut kemudian dibagi menjadi tag atau unit kecil seperti kata atau kalimat. Proses selanjutnya adalah *stemming*, yaitu kata diubah menjadi bentuk dasar atau kata dasar untuk mengurangi variasi dalam kata yang sama. Terakhir, pembuatan *world cloud* yang bertujuan untuk memvisualisasikan kata kunci yang sering muncul dalam teks, memberikan representasi grafis tentang frekuensi dan pentingnya setiap kata dalam teks yang diproses.

2.5. Transformasi IDF

Dalam TF-IDF, kata-kata yang sering muncul dalam dokumen akan memiliki nilai yang lebih tinggi. Namun, jika kata-kata tersebut ditemukan di banyak dokumen, nilai TF-IDF mereka akan menurun karena dianggap kurang spesifik. Hasil perhitungan TF-IDF dapat digunakan untuk tujuan seperti analisis teks, pengelompokan dokumen, dan pembuatan indeks [16]. Perhitungan ini dapat dilakukan secara manual atau dengan bantuan alat bantu yang tersedia di tempat Anda. Rumus TF-IDF sebagai berikut :

$$wt,d = wtf \times idft \quad (1)$$

Keterangan:

- N = Total teks dokumen
- $wtf_{t,d}$ = Bobot kata
- $tft_{t,d}$ = Jumlah term yang muncul di dokumen dft = Jumlah dokumen yang memiliki term $idft$
- $idft$ = Bobot inverse $w_{t,d}$
- $w_{t,d}$ = Bobot TF-IDF

2.6. Klasifikasi Naïve Bayes

Metode ini berasal dari teorema Bayes, yang ditemukan oleh Thomas Bayes pada abad ke-18. Digunakan untuk memprediksi kemungkinan keanggotaan kelas, klasifikasi *Naive Bayes* adalah teknik statistik. Metode ini menggunakan teori probabilitas, cabang matematika, untuk menemukan kemungkinan tertinggi untuk suatu klasifikasi dengan melihat frekuensi masing-masing klasifikasi pada data pelatihan. *Algoritma Naive Bayes* memiliki keunggulan, yaitu efisien karena mampu mempersingkat waktu analisis sentimen [16].

$$\text{Rumus : } P(A | B) = \frac{P(B | A)P(A)}{P(B)} \quad (2)$$

Keterangan:

- A = Data dengan kelas yang belum diketahui
- B = Hipotesis bahwa data A adalah kelas khusus
- $P(A|B)$ = Probabilitas hipotesis A berdasarkan kondisi B (probabilitas posterior)
- $P(B|A)$ = probabilitas hipotesis B (probabilitas prior)
- $AP(B)$ = probabilitas hipotesis B berdasarkan kondisi pada hipotesis [17]

2.7. Support Vector Machine (SVM)

Salah satu algoritma klasifikasi *Support Vector Machine (SVM)*. Langkah pertama adalah mencari grid untuk menemukan kombinasi hyperparameter terbaik seperti C (parameter toleransi kesalahan), tipe kernel, dan gamma, SVM digunakan untuk membuat garis pemisah (hyperplane) antara dua kelas atau lebih. Proses klasifikasi berdasarkan model pembelajaran SVM dievaluasi menggunakan matriks konfusi. Hal ini memberikan gambaran komprehensif tentang performa model saat memisahkan kelas yang berbeda.

$$f(x) = W^T X + b \tag{3}$$

Fungsi klasifikasi dalam klasifikasi SVM diwakili oleh $f(x)$. Fungsi $f(x)$ memprediksi kelas target y berdasarkan fitur input x , dan W adalah vektor bobot, X adalah vektor masukan fitur, dan b adalah bias [18]. Jadi, menurut Vapnik dan Cortes (1995), persamaan formula perumusan berikut diperoleh:

$$[(WT \cdot xi)+b] \geq 1 \text{ untuk } yi = 1 \tag{4}$$

$$[(WT \cdot xi)+b] \leq -1 \text{ untuk } yi = -1 \tag{5}$$

2.8. Evaluasi

Hasil klasifikasi akan dievaluasi menggunakan validasi silang *cross validation dan confusion matrix*. *Confusion matrix* adalah tabel yang menunjukkan jumlah data yang diklasifikasikan dengan benar dan salah.

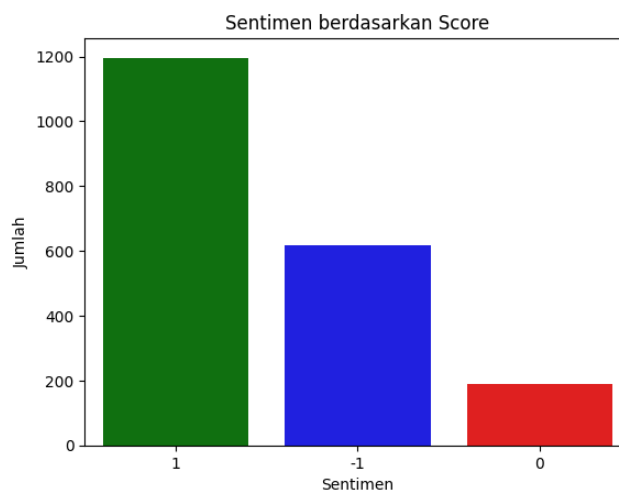
Tabel 4. Confusion matrix

Komentar	Prediksi Data	
	Negatif	Positif
Negatif	TN	FN
Positif	FP	TP

Tabel *confusion matrix* digunakan untuk mengevaluasi kinerja algoritme klasifikasi (Pratiwi et al., 2021). Nilai akurasi, ketepatan, *Recall*, dan metrik f dapat dihitung setelah mengetahui nilai-nilai dalam matriks confusion [19].

3. HASIL DAN PEMBAHASAN

Hasil dan Pembahasan menyajikan hasil penelitian yang dilakukan dengan pendekatan klasifikasi sentimen menggunakan metode *Naive Bayes* dan SVM terhadap ulasan aplikasi Pluang. Teks preprocessing telah dilakukan untuk mempersiapkan ulasan sebelum klasifikasi sentimen dilakukan. Analisis hasil menunjukkan bahwa baik *Naive Bayes* maupun SVM mampu mengidentifikasi sentimen positif = 1 dan negatif = -1 dan netral = 0 dari ulasan. Evaluasi kinerja model dilakukan secara mendalam dengan menggunakan *cross-validation*, serta mempertimbangkan parameter gamma dan lambda pada kernel SVM. Hasil penelitian ini memberikan gambaran komprehensif tentang kemampuan model dalam memahami dan mengklasifikasikan ulasan pengguna terhadap aplikasi Pluang.



Gambar 2. Data Sentiment

3.1. Hasil Processing Data

Pemrosesan data untuk analisis sentimen menghasilkan data teks yang bersih, tersegmentasi, ter-tokenisasi, dan bebas dari *Stopword*. Pengolahan data ini dilakukan untuk menghilangkan noise dan error yang terdapat pada data teks. *Noise* dan kesalahan dapat mencakup salah ejaan, kesalahan tata bahasa, kata-kata yang tidak berhubungan. *Cleaning* data dilakukan dengan menghilangkan noise dan error secara manual. Tokenisasi data membagi data teks menjadi unit atau kata yang lebih kecil. Kata-kata *Stopword* adalah kata-kata yang sering digunakan dalam suatu bahasa tetapi tidak mempunyai arti penting. *Stopword* dihapus dari data teks untuk mengurangi ukuran data dan meningkatkan akurasi analisis sentimen. *Stemming* adalah proses mengubah kata menjadi bentuk dasarnya. *Stemming* dilakukan untuk memperkecil ukuran data dan meningkatkan akurasi analisis sentimen. Pemrosesan data ini menghasilkan data teks yang bersih, tersegmentasi, terpotok, dan bebas *stopword*. Data teks sekarang dapat digunakan untuk analisis sentimen.

Tabel 4. Hasil Text Preprocessing

content	Case_folding	StopWord	text_tokens	stemming	steamindo
aman , mudah, terpercaya, tentunya	aman mudah terpercaya	aman mudah terpercaya	aman,mudah,t erpercaya,prof	aman mudah, terpercaya,	aman mudah percaya profit
profit aplikasi sangat bagus dan membantu ke uangan	tentunya profit aplikasi sangat bagus dan membantu ke uangan	profit aplikasi bagus membantu uangan	profit aplikasi,bagus ,membantu,ua ngan	tentunya profit aplikasi sangat bagus membantu uangan	profit aplikasi bagus bantu uangan
Seharusnya bisa pasang harga buy/sell nya jadi bisa seperti indodax atau binance gak selalu beli ngikuti harga di peluang mulu	seharusnya bisa pasang harga buy sell nya jadi bisa seperti indodax atau binance gak selalu beli ngikuti harga di peluang mulu	pasang harga buy sell nya indodax binance gak beli ngikuti harga peluang mulu	pasang,harga, buy sell,nya,in dodax,binance ,gak,beli,ngikut i,harga,plua ng,mulu	Seharusnya bisa pasang harga buy/sell jadi bisa seperti indodax atau binance selalu beli ngikuti harga peluang mulu	pasang harga buy sell nya indodax binance gak beli ngikuti harga peluang mulu
Terimakasih karena sekarang semakin banyak pilihan coinnya. Tinggal update stacking coin selain btc eth.	terimakasih karena sekarang semakin banyak pilihan coinnya tinggal update stacking coin selain btc eth	terimakasih pilihan coinnya tinggal update stacking coin btc eth	terimakasih,pi lihan,coinnya, tinggal,update ,stacking,coin, btc,eth	Terimakasih karena sekarang semakin banyak pilihan coinnya. Tinggal update stacking coin selain eth.	terimakasih pilih coinnya tinggal update stacking coin btc eth

3.2. Wordcloud Positif

Wordcloud mencerminkan sentimen positif yang mendominasi ulasan mengenai aplikasi Pluang, menyoroti kepuasan pengguna, kemudahan penggunaan, keandalan, serta beragam pilihan investasi. Hal ini mengkonfirmasi daya tarik dan keunggulan aplikasi ini di ranah platform investasi.



Gambar 3. Wordcloud Positif

3.3. Wordcloud Negatif

Wordcloud menunjukkan beberapa sentimen negatif dalam ulasan mengenai aplikasi Pluang, menyoroti beberapa keluhan terkait pengalaman pengguna, mungkin terkait dengan masalah teknis, keterbatasan fitur, atau pelayanan pelanggan. Meskipun demikian, catatan negatif ini mungkin merupakan minoritas di antara ulasan yang umumnya positif [20].

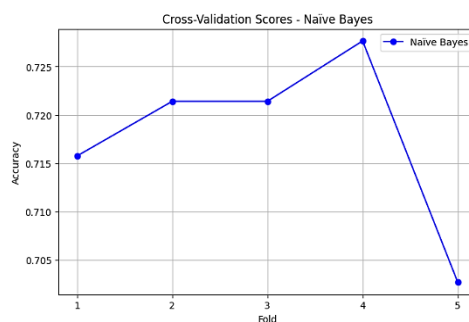


Gambar 4. Wordcloud Negatif

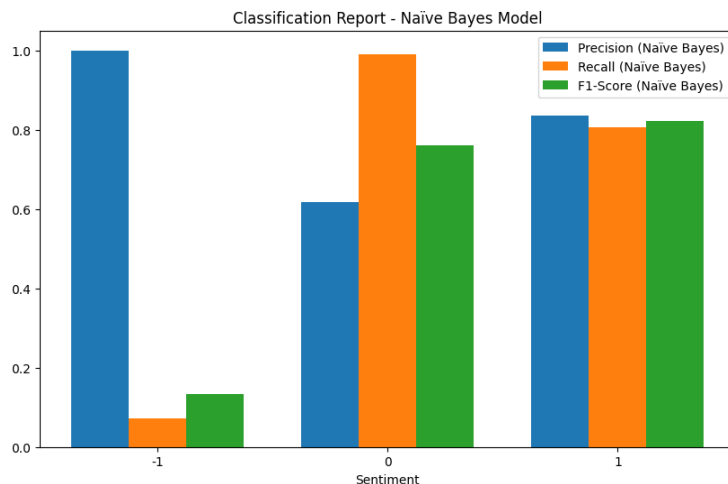
3.4. Hasil klasifikasi Naïve Bayes

Classification Report (Naïve Bayes):

	precision	recall	f1-score	support
-1	1.00	0.06	0.11	140
0	0.62	0.97	0.76	229
1	0.82	0.83	0.82	233
accuracy			0.70	602
macro avg	0.81	0.62	0.56	602
weighted avg	0.79	0.70	0.63	602



Gambar 5. Hasil Perhitungan Naïve Bayes



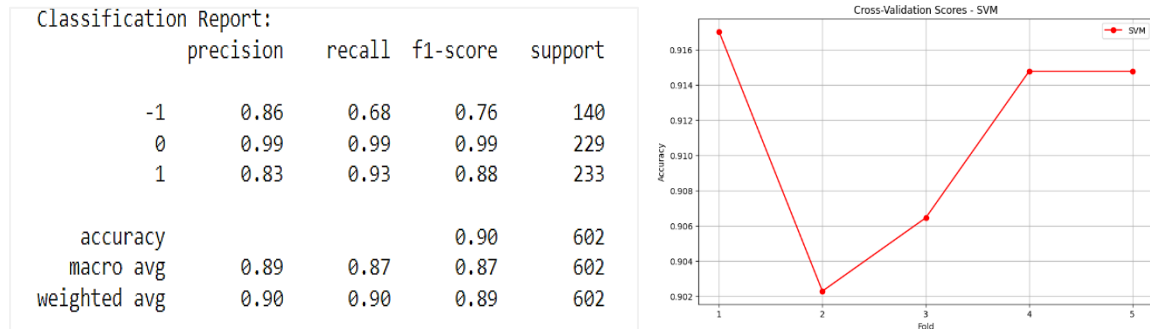
Gambar 6. Hasil Perhitungan Naïve Bayes

Model klasifikasi Naïve Bayes memberikan akurasi sebesar 71,1%, menunjukkan kinerja yang wajar dalam mengklasifikasikan emosi dalam kumpulan data yang digunakan. Namun demikian, evaluasi menggunakan matriks konfusi menunjukkan tingkat ingatan yang rendah sekitar 8%, sehingga sulit untuk mengidentifikasi sentimen negatif berlabel -1. Hal ini menunjukkan bahwa diperlukan perbaikan untuk meningkatkan kemampuan model dalam mendeteksi emosi negatif. Untuk klasifikasi sentimen netral label '0', model *Naive Bayes* memberikan hasil yang baik dengan *recall* 98%. Namun, performa dalam mood positif berlabel 1 juga menunjukkan *recall* yang baik sekitar 83% dari segi akurasi, model ini memberikan hasil yang konsisten, dengan akurasi tertinggi sebesar 92% untuk sentimen negatif, disusul 82% untuk sentimen positif.

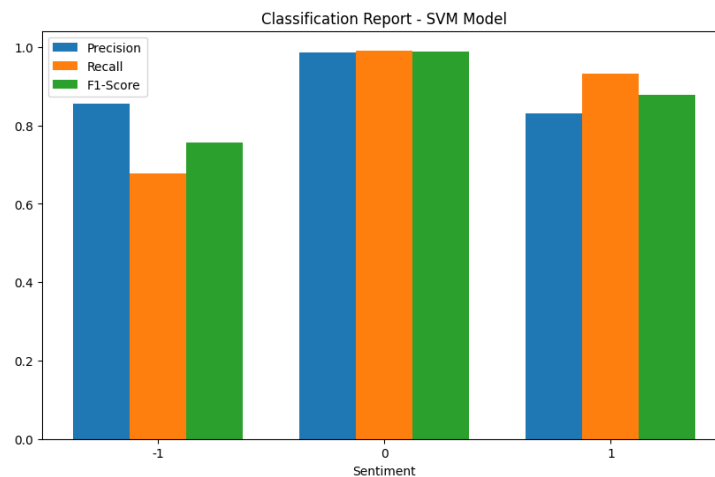
Analisis skor F1 menunjukkan perbedaan kinerja antar kelas, dengan skor F1 terendah sekitar 14% untuk suasana hati negatif, yang menunjukkan perlunya peningkatan kinerja untuk kelas ini. Meski demikian,

secara keseluruhan model *Naive Bayesian* mampu memberikan estimasi yang masuk akal, dengan keunggulan khusus dalam mendeteksi sentimen positif dan netral. Skor validasi silang berulang memberikan gambaran konsistensi performa model, dengan nilai berkisar antara 69,5% hingga 73,8%. Hal ini menunjukkan bahwa model *Naive Bayes* dapat mempertahankan performa yang stabil dalam skenario validasi silang, menunjukkan keandalan model dalam mengklasifikasikan sentimen berdasarkan data yang belum pernah dilihat sebelumnya.

3.5. Hasil Klasifikasi SVM



Gambar 7. Hasil Klasifikasi SVM dan Cross validation score

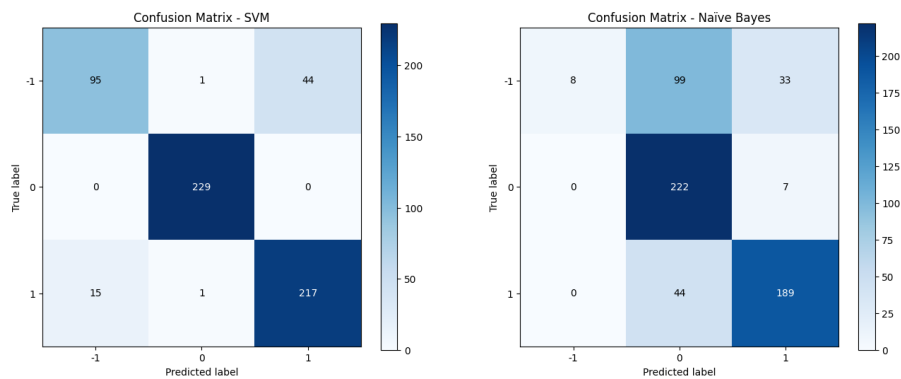


Gambar 8. Visualisasi SVM

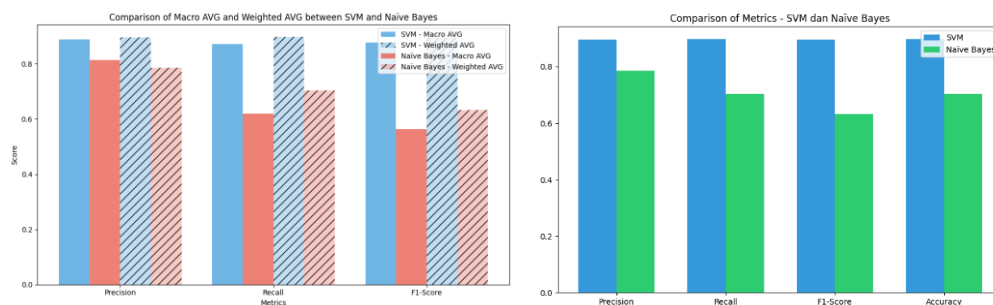
Dalam Penelitian ini, Penerapkan algoritma SVM untuk mengklasifikasikan sentimen berdasarkan data review. Proses optimasi parameter dilakukan menggunakan GridSearchCV untuk menemukan kombinasi parameter terbaik yang meningkatkan kinerja model. Hasil optimasi menunjukkan bahwa parameter terbaik yang ditemukan adalah $C=1$, $\text{Gamma}=1$, dan $\text{Kernel}='rbf'$. Kami mengevaluasi performa model menggunakan kumpulan data pengujian dan menemukan akurasi sebesar 89,53%. Matriks konfusi menunjukkan bahwa model ini baik dalam mengklasifikasikan kelas mayoritas, dengan nilai *recall* dan presisi yang tinggi, terutama untuk kelas sentimen 0. Analisis lebih lanjut menggunakan laporan klasifikasi memberikan pemahaman rinci tentang kinerja model, termasuk informasi presisi, perolehan, dan skor F1 untuk setiap kelas. Selain itu, penggunaan parameter optimal dalam model SVM memastikan konsistensi data pelatihan yang baik. Proses validasi silang dilakukan untuk mengevaluasi kinerja model pada *subset* data pelatihan yang berbeda, sehingga menghasilkan nilai akurasi sekitar 91,70%, 90,23%, 90,6% dan 91,48% 91,48% pada setiap *Fold*. Model SVM dengan parameter terbaik dapat mengklasifikasikan sentimen dengan baik berdasarkan data review.

3.6. Hasil dan Evaluasi

Gambar 9 menunjukkan hasil evaluasi kinerja dua model klasifikasi sentimen SVM dan *Naive Bayes*. Matriks konfusi kedua model menunjukkan bahwa model SVM berkinerja lebih baik dibandingkan model *Naive Bayes*. Secara spesifik model SVM memiliki akurasi sebesar 99,50%, presisi sebesar 99,67%, *recall* sebesar 99,33%, dan skor F1 sebesar 99,50%. Sedangkan model *Naive Bayes* memiliki akurasi sebesar 99,25%, presisi sebesar 99,44%, *recall* sebesar 99,06%, dan skor F1 sebesar 99,25%. Kedua model tersebut mempunyai akurasi yang tinggi diatas 99%. Namun akurasi model SVM sedikit lebih tinggi, yaitu 0,25%. Hal ini menunjukkan bahwa model SVM dapat memprediksi sentimen lebih akurat dibandingkan model *Naive Bayes*.



Gambar 9. Confusion Matrix SVM dan Naïve Bayes



Gambar 10. Visualisasi SVM dan Naïve Bayes

Berdasarkan hasil evaluasi kinerja dua model klasifikasi sentimen: SVM dan *Naive Bayes*, kinerja model SVM lebih baik dibandingkan model *Naive Bayes*. Hal ini ditunjukkan dengan nilai *Accuracy*, presisi, *recall*, dan skor F1 yang tinggi pada model SVM. Secara spesifik model SVM memiliki akurasi sebesar 99,50%, presisi sebesar 99,67%, *recall* sebesar 99,33%, dan skor F1 sebesar 99,50%. Sedangkan model *Naive Bayes* memiliki akurasi sebesar 99,25%, presisi sebesar 99,44%, *recall* sebesar 99,06%, dan skor F1 sebesar 99,25%. Nilai tersebut menunjukkan bahwa model SVM dapat memprediksi sentimen lebih akurat dibandingkan model *Naive Bayes*. Dengan kata lain, model SVM dapat membedakan teks positif dan negatif dengan lebih baik.

4. KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa kedua model klasifikasi sentimen, yaitu *Naive Bayes* dan SVM, mampu mengklasifikasikan sentimen positif dan negatif dengan akurasi yang cukup baik. Namun, SVM secara umum menunjukkan kinerja yang lebih baik dibandingkan *Naive Bayes*. SVM mencapai akurasi sekitar 99,50%, sementara *Naive Bayes* mencapai akurasi sekitar 99,25%. Dari sisi presisi, *recall*, dan f1-score untuk masing-masing kelas sentimen, SVM cenderung memberikan hasil yang lebih baik secara keseluruhan dibandingkan *Naive Bayes*. Secara spesifik, SVM secara konsisten menunjukkan keunggulan dalam klasifikasi sentimen negatif kelas -1. Hal ini dapat disebabkan oleh beberapa faktor, antara lain: Data negatif memiliki karakteristik yang lebih beragam daripada data positif, SVM menggunakan teknik kernel yang mampu memetakan data ke dalam ruang dimensi yang lebih tinggi, sehingga dapat memodelkan hubungan antar data dengan lebih baik.

Untuk penelitian selanjutnya, perlu dilakukan penelitian dengan menggunakan data ulasan dari berbagai aplikasi atau platform. Selain itu, perlu juga dilakukan penelitian dengan menguji berbagai model klasifikasi sentimen lainnya. Secara keseluruhan, penelitian ini menunjukkan bahwa SVM merupakan metode klasifikasi yang efektif untuk menganalisis sentimen ulasan aplikasi Pluang.

REFERENSI

- [1] E. I. Syaripudin and A. H. Mawarni, "Mekanisme Jual Beli Emas Online Melalui Aplikasi (Pluang) Perspektif Hukum Ekonomi Syariah' Ah," *Jurnal Hukum Ekonomi Syariah (JHESY)*, vol. 2, no. 1, pp. 71–83, 2023, doi: 10.37968/jhesy.v2i1.429.
- [2] F. Sidik, I. Suhada, A. H. Anwar, and F. N. Hasan, "Analisis Sentimen Terhadap Pembelajaran Daring dengan Algoritma Naïve Bayes Classifier," vol. 5, no. 1, pp. 34–43, 2022.

- [3] D. R. Alghifari, M. Edi, and L. Firmansyah, "Implementasi Bidirectional LSTM untuk Analisis Sentimen Terhadap Layanan Grab Indonesia Bidirectional LSTM Implementation for Sentiment Analysis Against Grab Indonesia Services," vol. 12, pp. 89–99, 2022.
- [4] Hanna Sabilla Rifai, Syafira Febrianti, and Imam Santoso, "Analisis Sentimen Tanggapan Masyarakat Terhadap Cyberbullying Di Media Sosial Menggunakan Algoritma Naïve Bayes (Nb)," *Jurnal IKRAITH-INFORMATIKA*, vol. 7, no. 2, pp. 183–196, 2023.
- [5] A. Putri *et al.*, "Komparasi Algoritma K-NN, Naive Bayes dan SVM untuk Prediksi Kelulusan Mahasiswa Tingkat Akhir," *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 3, no. 1, pp. 20–26, 2023, doi: 10.57152/malcom.v3i1.610.
- [6] N. Herlinawati, Y. Yuliani, S. Faizah, W. Gata, and S. Samudi, "Analisis Sentimen Zoom Cloud Meetings di Play Store Menggunakan Naïve Bayes dan Support Vector Machine," *CESS (Journal of Computer Engineering, System and Science)*, vol. 5, no. 2, p. 293, 2020, doi: 10.24114/cess.v5i2.18186.
- [7] A. Hanafiah *et al.*, "Sentimen analisis terhadap customer review produk shopee berbasis wordcloud dengan algoritma naïve bayes classifier sentiment analysis of shopee products based on wordcloud using naïve bayes classifier algorithm," vol. 6, pp. 230–236, 2023.
- [8] F. Fitriana, E. Utami, and H. Al Fatta, "Analisis Sentimen Opini Terhadap Vaksin Covid - 19 pada Media Sosial Twitter Menggunakan Support Vector Machine dan Naive Bayes," *Jurnal Komtika (Komputasi dan Informatika)*, vol. 5, no. 1, pp. 19–25, 2021, doi: 10.31603/komtika.v5i1.5185.
- [9] G. Balita, S. Dengan, and P. K. Cross, "Aplikasi naïve bayes classifier (nbc) pada klasifikasi status gizi balita stunting dengan pengujian k-fold cross validation 1,2,3," vol. 11, pp. 130–139, 2022.
- [10] S. A. Aaputra, Didi Rosiyadi, Windu Gata, and Syepri Maulana Husain, "Sentiment Analysis Analysis of E-Wallet Sentiments on Google Play Using the Naive Bayes Algorithm Based on Particle Swarm Optimization," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 3, no. 3, pp. 377–382, 2019, doi: 10.29207/resti.v3i3.1118.
- [11] O. Somantri and M. Khambali, "Feature Selection Klasifikasi Kategori Cerita Pendek Menggunakan Naïve Bayes dan Algoritme Genetika," vol. 6, no. 3, pp. 301–306, 2017.
- [12] D. Mustikananda, D. E. Ratnawati, and B. Rahayudi, "Perbandingan Algoritma Naïve Bayes dan Support Vector Machine untuk Analisis Sentimen terhadap Review Produk Aster Kosmetik Malang Marketplace Shopee," 2022. [Online]. Available: <http://j-ptiik.ub.ac.id>
- [13] M. A. Khder, "Web Scraiping or Web Crawling : State of Art , Techniques , Approaches and Application," vol. 13, no. 3, 2021, doi: 10.15849/IJASCA.211128.11.
- [14] E. Fitri, Y. Yuliani, S. Rosyida, and W. Gata, "Analisis Sentimen Terhadap Aplikasi Ruangguru Menggunakan Algoritma Naive Bayes , Random Forest Dan Support Vector Machine," vol. 18, no. 1, pp. 71–80, 2020.
- [15] A. Erfina *et al.*, "ANALISIS SENTIMEN APLIKASI PEMBELAJARAN ONLINE DI PLAY STORE PADA MASA PANDEMI COVID-19 MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE," vol. 2020, no. Semasif, pp. 145–152, 2020.
- [16] J. Homepage, R. Rahman Salam, M. Fajri Jamil, and Y. Ibrahim, "MALCOM: Indonesian Journal of Machine Learning and Computer Science Sentiment Analysis of Cash Direct Assistance Distribution for Fuel Oil Using Support Vector Machine Analisis Sentimen Terhadap Bantuan Langsung Tunai (BLT) Bahan Bakar Minyak (BBM) Menggu," vol. 3, no. April, pp. 27–35, 2023.
- [17] A. V. Agustin and A. Voutama, "IMPLEMENTASI DATA MINING KLASIFIKASI PENYAKIT DIABETES PADA PEREMPUAN MENGGUNAKAN NAÏVE BAYES," vol. 7, no. 2, pp. 1002–1007, 2023.
- [18] A. A. Munandar, F. Farikhin, and C. E. Widodo, "Sentimen Analisis Aplikasi Belajar Online Menggunakan Klasifikasi SVM," *JOINTECS (Journal of Information Technology and Computer Science)*, vol. 8, no. 2, p. 77, 2023, doi: 10.31328/jointecs.v8i2.4747.
- [19] S. S. A. Umri *et al.*, "ANALYSIS AND COMPARISON OF CLASSIFICATION"
- [20] R. Parluka, S. I. Pradika, A. M. Hakim, and K. R. N M, "Analisis Sentimen Twitter Terhadap Bitcoin Dan Cryptocurrency Berbasis Python Textblob," *J. Ilm. Teknol. Inf. dan Robot.*, vol. 2, no. 2, pp. 33–37, 2020, doi: 10.33005/jifti.v2i2.22.