

Sentiment Analysis of User Reviews on the Line Bank Digital Banking Application Using the Support Vector Machine Method

Rizki Daffa Muttaqin ¹, Ulfa Khaira ^{2*}, Daniel Arsa ³

^{1,2,3} Universitas Jambi, Sistem Informasi, Jl. Jambi - Muara Bulian No.KM. 15, Mendalo Darat, Kec. Jambi Luar Kota, Kabupaten Muaro Jambi, Jambi 36361, Indonesia

Keyword

Line Bank; Sentiment Analysis; Support Vector Machine; User Reviews

*Corresponding Author:

ulfakhaira@unja.ac.id

Abstract

With the increasing popularity of digital banking, user reviews are becoming an important source of data to understand user experiences and sentiment towards the services provided. Line Bank, as a digital bank, also receives various reviews from its users which reflect both satisfaction and dissatisfaction with their services. The problem at hand is how to analyze sentiment from user reviews effectively and accurately to provide better insight into Line Bank's service quality. This research uses the Support Vector Machine machine learning algorithm to analyze the sentiment of Line Bank user reviews. The performance of both methods is evaluated based on the metrics of accuracy, precision, recall, and F1-score. Testing was carried out with several data split scenarios, and the results showed that the Support Vector Machine had good performance. In the 90:10 data sharing scenario, Support Vector Machine achieved 89.61% accuracy. Apart from that, Support Vector Machine also shows good performance in precision, recall and F1-score metrics. Visualization analysis of the results shows the dominance of negative sentiment in Line Bank user reviews, which indicates there is room for service improvement.

1. Pendahuluan

Pada era majunya teknologi khusus pada sektor keuangan sudah banyak perusahaan bank yang mempunyai aplikasi *bank digital*. *Bank digital* merupakan layanan perbankan yang berfungsi untuk memudahkan nasabah dalam melakukan transaksi perbankan kecuali untuk penarikan uang cash [1]. *Bank digital* meningkatkan efisiensi dalam mengelola keuangan pribadi dan pengalaman perbankan yang lebih terjangkau bagi pengguna [2].

Salah satu *bank digital* yang baru dan sedang banyak digunakan adalah *Line bank*. *Line bank* pertama kali hadir di Indonesia pada 10 Juni 2021 dengan menggandeng *LINE Corporation* dan PT Bank KEB Hana [3]. *Line bank* memberikan kemudahan pada proses pembukaan rekening yang dilakukan secara online semua melalui aplikasi smartphone pengguna, kemudahan ini agar pengguna terus bertambah dan memberikan ulasan yang baik kepada aplikasi *Line bank*. Salah satu faktor keberhasilan aplikasi dapat dilihat dari tingkat kepuasan pengguna dalam pelayanan yang diberikan [4]. Ulasan pengguna bisa mempengaruhi bagus atau tidaknya aplikasi tersebut untuk diunduh dan digunakan.

Dari ulasan pengguna tersebut dapat dikumpulkan dan dianalisis sehingga menjadi ilmu pengetahuan yang bisa memberikan keterangan sentimen positif, sentimen negatif maupun netral. Metode analisis sentimen digunakan untuk mengetahui kesimpulan sentimen dari ulasan pengguna aplikasi tersebut. Analisis sentimen adalah sebuah proses menemukan pendapat pengguna tentang beberapa topik atau teks yang disampaikan [5]. Metode yang digunakan dalam penelitian ini yaitu *Support Vector Machine (SVM)*.

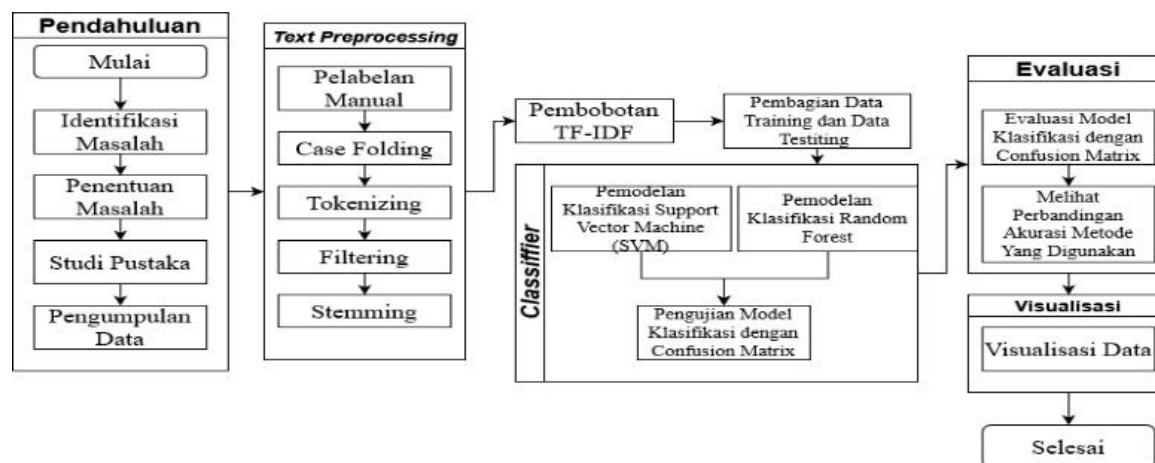
Metode Support Vector Machine telah digunakan pada penelitian dengan judul “Text Mining Untuk Analisis Sentimen Pelanggan Terhadap Layanan Uang Elektronik Menggunakan Algoritma Support Vector Machine”. Dari sebanyak 16 kali percobaan klasifikasi, dengan 4 data testing dan 4 kernel yang berbeda-beda, nilai akurasi terbaik didapatkan oleh kernel linear dengan rasio data 90% (3467 data) untuk data training dan 10% (385 data) untuk data testing dengan nilai akurasi yang didapat yaitu sebesar 0.987 atau 98.7% [6].

Penelitian selanjutnya dengan judul “Perbandingan Algoritma *Support Vector Machine* dan *Random Forest* untuk Analisis Sentimen Terhadap kebijakan Pemerintah Indonesia Terkait Kenaikan Harga BBM Tahun 2022”. Tujuan penelitian ini menentukan algoritma yang memiliki tingkat akurasi lebih baik serta memberikan masukan terhadap pandangan masyarakat terhadap kebijakan tersebut. Hasil penelitian menunjukkan, kedua metode baik algoritma *Support Vector Machine* maupun *Random Forest* memberikan kinerja yang baik, dengan akurasi masing-masing sebesar 77% dan 76%. Meskipun perbedaan akurasi tidak signifikan, *Support Vector Machine* memiliki tingkat akurasi yang sedikit lebih tinggi. [7].

Kemudian penelitian dengan judul “Penerapan Metode Klasifikasi Support Vector Machine (SVM) pada Data Akreditasi Sekolah Dasar (SD) di Kabupaten Magelang”. Klasifikasi Akreditasi Sekolah Dasar (SD) di Kabupaten Magelang menggunakan metode Support Vector Machine (SVM) dengan fungsi kernel Gaussian Radial Basic Function (RBF) menghasilkan akurasi yang sangat baik, yaitu 100% pada data training, di mana 337 SD dari total 337 SD berhasil diklasifikasikan dengan benar sesuai dengan kelas aslinya. Pada data testing, metode ini juga menunjukkan performa yang tinggi dengan tingkat akurasi sebesar 93,9%, di mana 77 SD dari total 82 SD berhasil diklasifikasikan dengan benar sesuai dengan kelas aslinya [8].

Berdasarkan penelitian terdahulu diatas, terdapat nilai akurasi yang dihasilkan dari metode tersebut memiliki nilai akurasi yang berbeda dalam menghasilkan sentimen. Maka, berlandaskan penelitian terdahulu yang penulis telah sampaikan sebelumnya, penulis ingin menggunakan metode tersebut pada penelitian kali ini yaitu menggunakan metode *Support Vector Machine (SVM)*. Penulis melakukan penelitian dengan judul “Analisis Sentimen Ulasan Pengguna Pada Aplikasi Bank Digital Line Bank Menggunakan Metode *Support Vector Machine*”.

2. Metode Penelitian



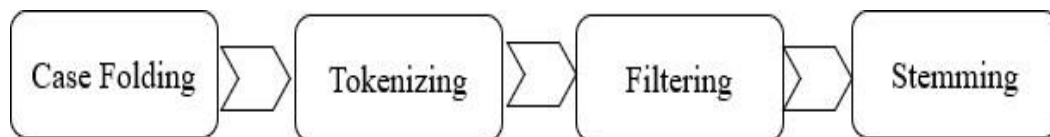
Gambar 1. Alur Penelitian

Dari kerangka kerja penelitian yang dilakukan, peneliti merincikannya sebagai berikut:

Pada tahap pendahuluan, peneliti mengidentifikasi masalah penelitian dan menentukan masalah yang relevan dengan topik yang akan diteliti. Tujuan dari penelitian juga ditetapkan, didukung oleh beberapa penelitian terkait. Literatur studi pustaka digunakan sebagai referensi untuk memahami topik penelitian. Selanjutnya, peneliti mengumpulkan data dengan melakukan *scraping* data ulasan pengguna. Adapun langkah melakukan *scraping* [9], yaitu :

1. *Create Scraping Template*, Pembuat program mempelajari dokumen HTML dari website yang akan diambil informasinya untuk tag HTML yang mengapit informasi yang akan diambil
2. *Explore Site Navigation*, Pembuat program mempelajari teknik navigasi pada website yang akan diambil informasinya untuk ditirukan pada aplikasi web scraper yang akan dibuat.
3. *Automate Navigation and Extraction*, Berdasarkan informasi yang didapat pada langkah 1 dan 2 di atas, aplikasi web scraper dibuat untuk mengotomatisasi pengambilan informasi dari website yang ditentukan
4. *Extracted Data and Package History*, Informasi yang didapat dari langkah 3 disimpan dalam tabel atau tabel-tabel database.

Setelah data terkumpul dari tahap pendahuluan dengan menggunakan teknik *scraping*. Di tahapan ini dilakukan pelabelan data manual dibantu oleh kantor bahasa Jambi terlebih dahulu. Selanjutnya setelah data dilabeli, masuk pada tahap *preprocessing* data, informasi yang telah di *scraping* sebelumnya diubah menjadi format data yang lebih terstruktur. Beberapa langkah yang dilakukan dalam *preprocessing* data yaitu, *case folding*, *tokenizing*, *filtering*, dan *stemming*. Setelah *preprocessing* data selesai dilakukan, hasilnya adalah data yang lebih bersih dan terstruktur, yang akan memudahkan proses klasifikasi yang akan dilakukan selanjutnya.



Gambar 2. Tahapan Preprocessing Data

Untuk mengatasi set data yang tidak seimbang, menggunakan metode *SMOTE* (*Synthetic Minority Over-sampling Technique*). *SMOTE* merupakan teknik yang dimanfaatkan untuk mengatasi masalah *class imbalance problem* (*CIP*) [10]. Pada tahapan klasifikasi peneliti menggunakan 2 metode klasifikasi pada penelitian ini yaitu, *support vector machine* dan *random forest* dengan tujuan untuk mengetahui perbandingan dari kedua metode yang digunakan dan mengetahui metode mana yang terbaik dari keduanya.

Adapun metode yang digunakan pada penelitian ini menggunakan klasifikasi *Support Vector Machine* (*SVM*) adalah sebagai berikut:

1. Dalam implementasi klasifikasi menggunakan model *support vector machine*, setelah melatih model dengan data pelatihan, model tersebut kemudian diuji dengan menggunakan data uji atau data testing.
2. Langkah selanjutnya adalah menghitung kembali akurasi model pada data uji atau data testing untuk mengevaluasi performa model *support vector machine*, hasil dari akurasi terbaik dengan menggunakan metode *support vector machine* ini menghasilkan pembobotan pada tiap kata yang ada.
3. Pembobotan kata yang telah dihitung digunakan dalam proses klasifikasi untuk memprediksi sentimen dari data yang akan diuji.
4. Pembobotan kata ini memiliki nilai yang berbeda, jika angka pembobotan minus (-), maka kata tersebut dikategorikan sebagai kata negatif, jika angka pembobotan lebih besar dari 0, maka kata tersebut dikategorikan sebagai kata positif.

Setelah mendapat hasil pengujian dari tahapan klasifikasi, performa hasil klasifikasi dari metode yang digunakan akan dinilai dan dievaluasi menggunakan *confusion matrix* berdasarkan perhitungan nilai *accuracy*, *precision*, *recall* dan *f1-score*. Bi

Tabel 1. Contoh Penggunaan Tabel Confusin Matrix

		Actual Data	
		Actual positif	Actual negatif
Predicted data	Predicted Positif	True positives (TP)	False positives (FP)
	Predicted Negatif	False negatives (FN)	True negatives (TN)

Setelah melakukan evaluasi terhadap model klasifikasi, masuk ke tahapan penelitian yang terakhir adalah melakukan proses visualisasi data. Setiap jenis data memerlukan teknik visualisasi yang sesuai dengan kebutuhan. Visualisasi data membantu mempermudah interpretasi data dengan meng gambarkannya dalam bentuk grafis atau visual. Visualisasi juga memungkinkan kita untuk melihat pola atau tren yang mungkin tidak terlihat ketika melihat data mentah. Pada Gambar 2. dibawah bisa dilihat contoh visualisasi dengan menggunakan wordcloud.



Gambar 3. Contoh visualisasi dengan menggunakan wordcloud

3. Hasil dan Pembahasan

Tahap awal penelitian ini adalah mengumpulkan data ulasan pengguna tentang aplikasi *Line Bank* dari *Google Play Store* dan didapat 11.915 ulasan dari 2021 hingga 2024.

		content	at	score
0	haloo Keb. Hana, aku suka sekali fitur angka a...	2024-08-17 16:23:17		5
1	Sebaiknya diadakan fitur login dengan biometri...	2024-06-01 03:10:57		3
2	Lebih ditingkatkan lg apk dan menu2 nya. Menu ...	2024-08-06 08:28:36		4
3	Bank gak jelas, kartu udh lama gak dipake pas ...	2024-06-17 17:52:53		1
4	Kartu sampe 3 bulan lebih aja belum sampe, har...	2024-08-09 11:47:42		1
...		...	...	...
11910		❤️	2022-04-11 15:04:06	5
11911	Kereenn.....		2022-10-12 10:43:31	5
11912		👍👍👍👍👍	2022-12-11 02:27:41	1
11913		😊	2021-11-22 00:09:41	2
11914		👍👍👍👍	2023-06-05 17:02:51	2

11915 rows x 3 columns

11915 rows x 3 columns

Gambar 4. Hasil dari scraping data

Untuk data yang akan digunakan yaitu data ulasan pada Januari 2023 hingga Juni 2024 dan didapat hasil 4.237 dataset, yang akan di *Random Sampling* sebanyak 1000 data

	content	score	Year	Month	Day
10808	Aplikasi sampah	1	2023	10	25
129	OTp SDH betul persis dibilang salah Sistem nya...	1	2024	2	12
4791	Kenapa kalo top up ke ovo ada biaya admin 1500...	4	2023	3	1
1391	Line bank, baik sekali, dn deposito nya, tks ya	5	2024	5	30
2104	Aplikasi gak guna cuma buat org biar download ...	1	2023	12	12
...	...	...	...	...	...
9337	Gakatau	3	2023	8	26
7189	aplikasi ini terkesan mengambil data doaaanngg...	1	2023	6	19
2693	Bagus sih, tapi saat mengecek kredibilitas kre...	4	2023	1	21
899	kartu debit tidak pernah sampai, tanya ke cs k...	1	2023	3	3
3381	Selalu gagal untuk registrasi, setelah komplai...	1	2023	8	2

1000 rows x 6 columns

Gambar 5. Hasil dari random sampling data yang akan digunakan

Dapat dilihat hasil dari pelabelan manual oleh tim ahli di kantor bahasa provinsi Jambi, dengan hasil : Sentimen positif 331, sentimen negatif 607 dan sentimen netral 62. Selanjutnya setelah data dilabelkan masuk ketahap *preprocessing data*.

	content	neg.		Day	Sentiment	Keterangan:	
1	Line bank, baik sekali, dn deposito nya, tks ya	1	2023	10	25	2	Merah: Sentimen negatif = 607
2	Semua fitur bagus ..is the best bank...	1	2024	2	12	2	Biru : Sentimen positif = 331
3	masah Allah aplikasi sangat membantu dan bermanfaat	4	2023	3	1	2	Kuning: Netral = 62
4	Saya menggunakan bank ini karena transaksi free transaksi ke semua bank	5	2024	5	30	2	
5	perlu di coba dulu kalau bagus akan sy tamba bintang bintang lagi	1	2023	12	12	2	
6	Okelah sbku terimakasih atas informasinya sangat bermanfaat sekali bagi dunia ini juga bisa melakukan hal yang baik untuk meningkatkan hasil belajar siswa dalam belajar yang baik	5	2024	4	21	2	
7	sahabatku syg oke 😊😊😊	1	2023	9	2	2	
8	Aplikasi nya sangat membantu dan mudah mendaftar nya dan juga banyak promo nya saya membeli token listrik tanpa biaya admin sama sekali dan juga untuk mendapatkan kartu	1	2024	5	31	2	
9	debitnya sangat cepat, terima kasih Line bank semoga makin banyak promosi lainnya.	5	2023	6	29	2	
10	Mantap	5	2023	7	2	2	
11	Line Bank,Bank digital yg mempermudah segala transaksi,lancar selama pakai GK pernah bermasalh ada KTA juga the best semoga sukses dan tambah inovasi2 terbaru	5	2023	7	2	2	
12	Good bank	1	2024	2	22	2	
13	Ok	2	2023	4	26	2	
14	Good	5	2023	1	23	2	
15	Simple... semua dalam satu genggam	1	2023	5	7	2	
16	Terbaik, bebas tarik via ATM manapun. Dapat cashback pula 25k. Bebas TF ke bank mana pun.	1	2023	12	21	2	
17	Andalan ku bagt	1	2023	3	20	2	
18	Menurut saya ini paling bagus, udh gratis biaya admin, transfer kemana mana bisa keren	3	2023	6	26	2	
19	makasih banyak line bank	5	2024	2	24	2	
20	Mantap						
21	Memudhkn transaksi						
22	Bintang 5 gak ada kendala selama make, senang nya tarik tunai bisa d atm apapun bebas admin						
23	hehehehe	5	2024	5	4	2	

Gambar 6. Hasil pelabelan data oleh kantor bahasa provinsi Jambi

Tabel 2. Tahapan Cleaning(Case Folding)

Contoh Data	
Dataset Awal	Tahapan Cleaning
Saya baru mau coba pakai LINE Bank, Senggol dong🙏🏻 Semoga amanah,aman dan terpercaya🙏🏻🙏🏻🙏🏻🙏🏻	saya baru mau coba pakai line bank senggol dong semoga amanah aman dan terpercaya

Tabel 3. Tahapan Tokenizing

Contoh Data	
Tahapan Cleaning	Tahapan Tokenizing
saya baru mau coba pakai line bank senggol dong semoga amanah aman dan terpercaya	['saya', 'baru', 'mau', 'coba', 'pakai', 'line', 'bank', 'senggol', 'dong', 'semoga', 'amanah', 'aman', 'dan', 'terpercaya']

Tabel 4. Tahapan Filtering

Contoh Data	
Tahapan Tokenizing	Tahapan Filtering
['saya', 'baru', 'mau', 'coba', 'pakai', 'line', 'bank', 'senggol', 'dong', 'semoga', 'amanah', 'aman', 'dan', 'terpercaya']	['coba', 'pakai', 'line', 'bank', 'senggol', 'semoga', 'amanah', 'aman', 'terpercaya']

Tabel 5. Tahapan Stemming

Contoh Data	
Tahapan Filtering	Tahapan Stemming
['coba', 'pakai', 'line', 'bank', 'senggol', 'semoga', 'amanah', 'aman', 'terpercaya']	['coba', 'pakai', 'line', 'bank', 'senggol', 'moga', 'amanah', 'aman', 'percaya']

Setelah data melewati tahap *preprocessing* selanjutnya menyeimbangkan data yang masih belum seimbang dengan menggunakan teknik *SMOTE*. *SMOTE* (Synthetic Minority Over-sampling Technique) adalah metode oversampling yang digunakan dalam machine learning untuk menangani ketidakseimbangan kelas dalam dataset, khususnya pada masalah klasifikasi di mana kelas minoritas memiliki frekuensi yang rendah dibandingkan dengan kelas mayoritas. Dengan memodifikasi dataset yang tidak seimbang, *SMOTE* membantu mengurangi bias yang disebabkan oleh perbedaan jumlah sampel antara kelas mayoritas dan minoritas, sehingga meningkatkan kemampuan model klasifikasi untuk mengenali kelas minoritas dengan lebih baik.

```
[49] # Menampilkan jumlah kelas sebelum SMOTE
for i in range(3):
    print(f"Jumlah kelas {i} sebelum SMOTE:", sum(y == i))
# Menampilkan jumlah kelas setelah SMOTE
for i in range(3):
    print(f"Jumlah kelas {i} setelah SMOTE:", sum(y_smote == i))
```


 Jumlah kelas 0 sebelum SMOTE: 607
 Jumlah kelas 1 sebelum SMOTE: 62
 Jumlah kelas 2 sebelum SMOTE: 331
 Jumlah kelas 0 setelah SMOTE: 607
 Jumlah kelas 1 setelah SMOTE: 607
 Jumlah kelas 2 setelah SMOTE: 607

Gambar 7. Jumlah dataset sebelum dan sesudah menggunakan SMOTE

Setelah data diseimbangkan dengan teknik *SMOTE*, langkah berikutnya adalah membagi data menjadi dua bagian: data latih (training set) dan data uji (testing set). Akan ada tiga skenario pembagian, yaitu 90:10, 80:20, dan 70:30.

```

✓ [39] # Train the Support Vector Machine model
0s svm = SVC(kernel='rbf', gamma=0.5)
    svm.fit(X_train, y_train)

    # Test the Support Vector Machine model
    y_pred_svm = svm.predict(X_test)

    accuracy_svm = accuracy_score(y_test, y_pred_svm)
    print("Support Vector Machine Accuracy:", accuracy_svm)
    print(classification_report(y_test, y_pred_svm))

```

Gambar 8. Code python model klasifikasi data menggunakan support vector machine

Pada Gambar 8. dapat dilihat kernel yang digunakan *RBF* dan gamma 0,5. Kernel sendiri adalah fungsi yang digunakan dalam *SVM* untuk memetakan data ke ruang dimensi yang lebih tinggi. Tujuannya adalah untuk membuat data yang awalnya tidak dapat dipisahkan secara linear menjadi dapat dipisahkan secara linear di ruang dimensi yang lebih tinggi. Sedangkan gamma adalah parameter yang mengontrol seberapa jauh pengaruh satu data poin terhadap data poin lainnya. Nilai gamma yang tinggi akan membuat model sangat sensitif terhadap setiap data poin, sehingga dapat menyebabkan *overfitting*. Pada tabel dibawah akan dilakukan pengujian *SVM* dengan berbagai macam kernel dan gamma.

Tabel 6. Perbandingan hasil kernel

No	Kernel yang digunakan	Hasil
1	<i>Linear</i>	81.96
2	<i>Polynomial (gamma 0.1)</i>	38.79
3	<i>Polynomial (gamma 0.2)</i>	38.79
4	<i>Polynomial (gamma 0.3)</i>	43.16
5	<i>Polynomial (gamma 0.4)</i>	46.99
6	<i>Polynomial (gamma 0.5)</i>	49.18
7	<i>RBF (gamma 0.1)</i>	84.69
8	<i>RBF (gamma 0.2)</i>	84.69
9	<i>RBF (gamma 0.3)</i>	85.79
10	<i>RBF (gamma 0.4)</i>	89.07
11	<i>RBF (gamma 0.5)</i>	89.61
12	<i>Sigmoid (gamma 0.1)</i>	68.85
13	<i>Sigmoid (gamma 0.2)</i>	75.40
14	<i>Sigmoid (gamma 0.3)</i>	76.50
15	<i>Sigmoid (gamma 0.4)</i>	78.14
16	<i>Sigmoid (gamma 0.5)</i>	80.32

Tabel 6. memberikan gambaran perbandingan kinerja berbagai jenis kernel (*Linear*, *Polynomial*, *RBF*, dan *Sigmoid*) dengan variasi nilai gamma pada model *Support Vector Machine (SVM)*. Berdasarkan Tabel 6. nilai gamma memiliki pengaruh signifikan terhadap kinerja model. Semakin besar nilai gamma, akurasi model cenderung meningkat hingga titik tertentu, setelah itu akurasi dapat menurun karena model menjadi *overfitting* [11]. Sehingga hasil akurasi yang didapatkan adalah sebagai berikut:

	data_train	data_tes	support_vector_machine
0	90	10	89.61
1	80	20	87.12
2	70	30	86.10

Gambar 9. Hasil akurasi dari tiap rasio

Tabel 7. Hasil confusion matrix Support Vector Machine

		Predict class		
Actual Class	Positif	Positif	Netral	Negatif
		56	1	7
	Netral	1	57	9
	Negatif	1	0	51

Tabel 7 menampilkan *confusion matrix* berukuran 3x3 yang menunjukkan hasil klasifikasi untuk tiga kelas: positif, netral, dan negatif. Dari matriks ini, kita bisa melihat bahwa model berhasil mengklasifikasikan 56 data sebagai positif, 57 data sebagai netral, dan 51 data sebagai negatif dengan benar. Namun, model juga melakukan kesalahan klasifikasi, yaitu 7 data diklasifikasikan sebagai negatif padahal seharusnya positif (False Positive), dan 1 data diklasifikasikan sebagai positif padahal seharusnya negatif (False Negative).

Hasil evaluasi model menunjukkan bahwa nilai presisi untuk kelas positif sebesar 98%, kelas netral sebesar 85%, dan kelas negatif sebesar 88%. Ini menunjukkan bahwa proporsi prediksi yang benar dari total prediksi sangat tinggi untuk kelas positif. Sementara itu, tingkat keberhasilan model dalam menemukan kembali informasi yang benar (recall) adalah sebesar 76% untuk kelas positif, 98% untuk kelas netral, dan 97% untuk kelas negatif. Rata-rata nilai precision yang dihasilkan adalah 90%, nilai recall sebesar 90%, dan nilai F1-Score sebesar 89%.

Dari hasil confusion matrix metode *Support Vector Machine*, maka bisa disimpulkan sebagai berikut:

Tabel 8. Hasil metode klasifikasi

Metode	Support Vector Machine
Accuracy	89,61%
Precision	90%
Recall	90%
F1-Score	89%

Setelah didapat model *SVM*, langkah berikutnya adalah menganalisis hasil sentimen pengguna berdasarkan model *Support Vector Machine* yang telah dilatih sebelumnya. Proses ini melibatkan penerapan model *Support Vector Machine* terbaik untuk memprediksi kategori sentimen (positif, netral, atau negatif) pada data ulasan pengguna *Line Bank*.

```
[73] from collections import Counter

# Hitung jumlah sentimen
sentiment_counts = Counter(sentiments)

# Tampilkan hasil
print(sentiment_counts)
```

➞ Counter({'Negatif': 634, 'Positif': 257, 'Netral': 109})

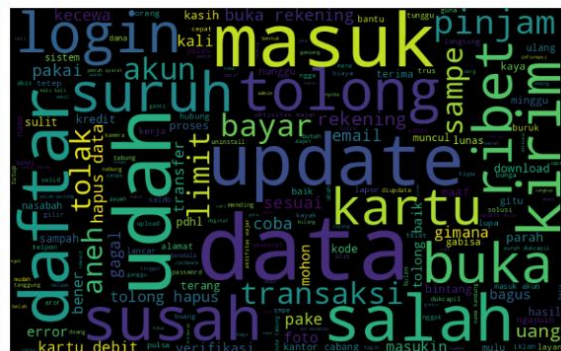
Gambar 10. Hasil sentiment

Pada Gambar 10. merupakan proses dari pengujian model SVM yang sudah dilatih sebelumnya untuk menampilkan hasil dari sentimen ulasan pengguna *Line Bank*. Hasil sentimen dari ulasan pengguna *Line Bank*, dengan hasil Negatif 634, Positif 257 dan Netral 109.

Langkah terakhir dalam penelitian ini dilakukan yaitu, melakukan visualisasi *wordcloud* menggunakan *Google Colab* untuk menampilkan ulasan sentimen positif dan sentimen negatif.



Gambar 11. Sentimen Positif



Gambar 12. Sentimen Negatif

Hasil penelitian menunjukkan bahwa ulasan pengguna aplikasi *Line Bank* didominasi oleh sentimen negatif. Temuan ini mengindikasikan bahwa sebagian besar pengguna lebih terdorong untuk menyampaikan pengalaman yang kurang memuaskan dibandingkan pengalaman positif ketika memberikan ulasan pada aplikasi digital banking. Fenomena ini juga ditemukan pada berbagai penelitian terkait analisis sentimen pada ulasan pengguna aplikasi atau layanan digital, di mana komentar negatif sering kali muncul lebih dominan karena pengguna cenderung menyampaikan keluhan terkait masalah teknis, kualitas layanan, maupun pengalaman penggunaan aplikasi. Penelitian terbaru menunjukkan bahwa analisis sentimen terhadap ulasan pengguna dapat memberikan gambaran penting mengenai persepsi publik terhadap kualitas layanan digital serta membantu organisasi dalam mengidentifikasi aspek layanan yang perlu ditingkatkan [12], [13].

Untuk mengatasi permasalahan ketidakseimbangan data pada proses klasifikasi sentimen, penelitian ini menerapkan metode Synthetic Minority Oversampling Technique (SMOTE). Metode ini digunakan untuk menghasilkan data sintesis pada kelas minoritas sehingga distribusi data menjadi lebih seimbang. Penelitian sebelumnya menunjukkan bahwa penerapan SMOTE mampu meningkatkan kinerja model klasifikasi dengan memperbaiki representasi kelas minoritas pada data pelatihan. Dengan menyeimbangkan distribusi kelas, model dapat mempelajari pola pada setiap kelas secara lebih optimal sehingga meningkatkan nilai precision, recall, dan F1-score pada proses klasifikasi [14], [15].

Hasil pengujian menunjukkan bahwa Support Vector Machine (SVM) dengan kernel Radial Basis Function (RBF) memberikan performa terbaik dengan tingkat akurasi sebesar 89,61%, lebih tinggi dibandingkan kernel linear, polynomial, maupun sigmoid. Hal ini menunjukkan bahwa kernel RBF lebih mampu menangkap hubungan non-linear dalam data teks sehingga menghasilkan pemisahan kelas yang lebih optimal. Temuan ini sejalan dengan penelitian terbaru yang menunjukkan bahwa SVM masih menjadi salah satu algoritma yang efektif dalam analisis sentimen berbasis teks karena kemampuannya dalam menangani data berdimensi tinggi dan menghasilkan model klasifikasi yang stabil [16].

Berdasarkan evaluasi menggunakan confusion matrix, model menunjukkan performa yang cukup stabil dengan rata-rata precision sebesar 90%, recall sebesar 90%, dan F1-score sebesar 89%. Model mampu mengenali kelas netral (98%) dan negatif (97%) dengan sangat baik, namun performa pada kelas positif masih lebih rendah (76%). Kondisi ini sering ditemukan dalam penelitian analisis sentimen karena ekspresi sentimen positif cenderung lebih beragam dan sering kali memiliki makna yang ambigu, sehingga lebih sulit dikenali oleh model klasifikasi. Penelitian lain juga menunjukkan bahwa variasi bahasa dalam ulasan pengguna dapat mempengaruhi akurasi klasifikasi, terutama pada kelas sentimen positif yang sering kali memiliki bentuk ekspresi yang lebih kompleks [17], [18].

Secara keseluruhan, tahapan penelitian yang meliputi proses scraping data, preprocessing teks, penerapan SMOTE, serta tuning parameter pada model SVM berhasil menghasilkan model klasifikasi sentimen tiga kelas yang cukup andal. Temuan utama berupa dominasi sentimen negatif memberikan informasi penting bagi pihak Line Bank untuk melakukan evaluasi dan peningkatan kualitas layanan aplikasi. Dari sisi metodologis, penelitian ini juga menegaskan bahwa penggunaan SVM dengan kernel RBF dan pengaturan parameter yang optimal merupakan pendekatan yang efektif untuk analisis sentimen berbasis teks pada ulasan pengguna aplikasi digital. Penelitian lain juga menunjukkan bahwa kombinasi teknik preprocessing, penyeimbangan data, serta pemilihan kernel yang tepat dapat meningkatkan performa klasifikasi sentimen secara signifikan pada berbagai domain ulasan pengguna [19], [20].

4. Penutup

Berdasarkan pada penelitian yang dilakukan, maka dapat diperoleh kesimpulan dari penelitian “Analisis Sentimen Ulasan Pengguna Pada Aplikasi Bank Digital Line Bank Dengan Menggunakan Metode *Support Vector Machine*” yaitu, Hasil penelitian menunjukkan bahwa pada skenario split data 90:10, algoritma *Support Vector Machine (SVM)* memiliki kinerja yang baik dalam mengklasifikasikan sentimen, dengan nilai akurasi, presisi, recall, dan F1-score yang tinggi, dengan nilai akurasi 89,61%, presisi 90%, recall 90%, F1-score 89%. Selain itu, hasil analisis juga mengindikasikan dominasi sentimen negatif pada dataset yang diteliti dengan nilai negatif 634, positif 2257 dan netral 109. Penelitian selanjutnya disarankan untuk mengembangkan metode yang lebih robust terhadap perubahan data pelatihan dan mampu mengatasi ambiguitas kata-kata polisemi untuk meningkatkan akurasi analisis sentimen.

5. Referensi

- [1] D. P. Wulandari, “Faktor-Faktor Yang Mempengaruhi Minat Nasabah Untuk Menggunakan Layanan Mobile Banking Pada PT Bank Syariah Mandiri Kantor Cabang Lubuk Pakam,” *Skripsi*, p. 1, 2018.
- [2] K. I. Ruslim, P. P. Adikara, and Indriati, “Analisis Sentimen Pada Ulasan Aplikasi Mobile Banking Menggunakan Metode Support Vector Machine dan Lexicon Based Features,” *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 3, no. 7, pp. 6694–6702, 2019.
- [3] Husni Shabri, “Transformasi Digital Industri Perbankan Syariah Indonesia,” *El-Kahfi / Journal of Islamic Economics*, vol. 3, no. 02, pp. 1–7, 2022, doi: 10.58958/elkahfi.v3i02.88.
- [4] H. Anisyah and T. Sutabri, “Analisa kepuasan pengguna terhadap Aplikasi ThatQuiz dengan Metode WebQual 4.0,” *Smatika Jurnal*, vol. 13, no. 01, pp. 131–139, 2023, doi: 10.32664/smatika.v13i01.763.

- [5] F. Bei and S. Sudin, "Analisis Sentimen Aplikasi Tiket Online Di Play Store Menggunakan Metode Support Vector Machine (Svm)," *Sismatik*, vol. 01, no. 01, pp. 91–97, 2021.
- [6] F. Romadoni, Y. Umaidah, and B. N. Sari, "Text Mining Untuk Analisis Sentimen Pelanggan Terhadap Layanan Uang Elektronik Menggunakan Algoritma Support Vector Machine," *Jurnal Sisfokom (Sistem Informasi dan Komputer)*, vol. 9, no. 2, pp. 247–253, 2020, doi: 10.32736/sisfokom.v9i2.903.
- [7] M. SAMANTRI, "Perbandingan Algoritma Support Vector Machine Dan Random Forest Untuk Analisis Sentimen Terhadap Kebijakan Pemerintah ...," vol. 8, no. 1, pp. 1–9, 2023.
- [8] P. A. Octaviani, Yuciana Wilandari, and D. Ispriyanti, "Penerapan Metode Klasifikasi Support Vector Machine (SVM) pada Data Akreditasi Sekolah Dasar (SD) di Kabupaten Magelang," *Jurnal Gaussian*, vol. 3, no. 8, pp. 811–820, 2014.
- [9] A. Josi, L. A. Abdillah, and Suryayusra, "Penerapan teknik web scraping pada mesin pencari artikel ilmiah," 2014.
- [10] C. Cahyaningtyas, Y. Nataliani, and I. R. Widiyari, "Analisis Sentimen Pada Rating Aplikasi Shopee Menggunakan Metode Decision Tree Berbasis SMOTE," *Aiti*, vol. 18, no. 2, pp. 173–184, 2021, doi: 10.24246/aiti.v18i2.173-184.
- [11] P. Fremmuzar and A. Baita, "Uji Kernel SVM dalam Analisis Sentimen Terhadap Layanan Telkomsel di Media Sosial Twitter," *Komputika : Jurnal Sistem Komputer*, vol. 12, no. 2, pp. 57–66, 2023, doi: 10.34010/komputika.v12i2.9460.
- [12] A. Fatma Putri Indah Sari, A. Eviyanti, and I. Ratna Indra Astutik, "Analisis Sentimen Komentar YouTube MV K-Pop Menggunakan Naïve Bayes dan SVM: Studi Kasus Jung Jaehyun 'Horizon,'" *SMATIKA : STIKI Informatika Jurnal*, vol. 15, no. 2, p. 15, Dec. 2025, doi: 10.32664/smatika.v15i02.1691.
- [13] I. Fahrur Rozi, I. Maulidia, M. Hani, R. Arianto, D. Rizky Yuniarto, and A. Yuli Ananta, "Comparison of feature extraction in support vector machine (SVM) based sentiment analysis system," *Jurnal Ilmiah Kursor*, vol. 13, no. 1, 2025.
- [14] D. Dwi Saputra, R. Fahlapi, A. Yadi Kuntoro, and T. Asra, "Analisis Sentimen Terhadap Twitter Direktorat Jenderal Bea dan Cukai Menggunakan komparasi Algoritma Naïve Bayes dan Support Vector Machine," *J-INTECH (Journal of Information and Technology)*, vol. 12, no. 2, Dec. 2024.
- [15] D. Andriyani, A. Faqih, and S. E. Permana, "Journal of Artificial Intelligence and Engineering Applications The Effect of SMOTE Application on Support Vector Machine Performance in Sentiment Classification on Imbalanced Datasets," *Journal of Artificial Intelligence and Engineering Application*, vol. 4, no. 2, Feb. 2025, [Online]. Available: <https://ioinformatic.org/>
- [16] A. Nafi', A. T. J. Harjanta, B. A. Herlambang, and S. Fahmi, "Analisis Sentimen Review Pelanggan Lazada dengan Sastrawi Stemmer dan SVM-PSO untuk Memahami Respon Pengguna," *J-INTECH (Journal of Information and Technology)*, vol. 12, no. 2, Dec. 2024.
- [17] J. R. Jim, M. A. R. Talukder, P. Malakar, M. M. Kabir, K. Nur, and M. F. Mridha, "Recent advancements and challenges of NLP-based sentiment analysis: A state-of-the-art review," Mar. 01, 2024, *Elsevier Ltd*. doi: 10.1016/j.nlp.2024.100059.
- [18] H. Suhendar, C. Slamet, and U. Syaripudin, "Analisis Sentimen Hasil Transkripsi Audio Berbahasa Indonesia Menggunakan T5 (Text-to-Text Transfer Transformer)," *SMATIKA JURNAL*, vol. 15, no. 01, pp. 115–125, Jun. 2025, doi: 10.32664/smatika.v15i01.1521.

- [19] S. T. Andini, A. Eviyanti, and H. Setiawan, "Analisis Sentimen Pengguna Aplikasi Tantan: Perbandingan Kinerja Metode Naive Bayes dan SVM," *SMATIKA : STIKI Informatika Jurnal*, vol. 15, no. 2, pp. 396–407, 2025, doi: 10.32664/smatika.v15i02.1692.
- [20] P. P. Putra, M. K. Anam, A. S. Chan, A. Hadi, N. Hendri, and A. Masnur, "Optimizing Sentiment Analysis on Imbalanced Hotel Review Data Using SMOTE and Ensemble Machine Learning Techniques," *Journal of Applied Data Sciences*, vol. 6, no. 2, pp. 936–951, May 2025, doi: 10.47738/jads.v6i2.618.