

## Analisis Sentimen Pengguna Aplikasi Livin' by Mandiri Menggunakan Metode Support Vector Machine (SVM) dengan Ekstraksi Fitur TF-IDF dan Word2Vec

Anas Suharman<sup>1</sup>, Mia Kamayani Sulaeman<sup>\*2</sup>

<sup>1,2</sup>Teknologi Industri dan Informatika, Universitas Muhammadiyah Prof. Dr. Hamka, Indonesia  
Email: <sup>1</sup>nassanassuharman@gmail.com, <sup>2</sup>mia.kamayani@uhamka.ac.id

### Abstrak

Pesatnya perkembangan aplikasi mobile banking di Indonesia, termasuk Livin' by Mandiri, menimbulkan kebutuhan untuk memahami respons pengguna secara lebih mendalam. Studi ini bertujuan untuk mengevaluasi sentimen pengguna terhadap aplikasi tersebut berdasarkan ulasan yang diperoleh dari Google Play Store. Penelitian ini menerapkan algoritma Support Vector Machine (SVM) dengan dua teknik ekstraksi fitur, yaitu Term Frequency–Inverse Document Frequency (TF-IDF) dan Word2Vec. Sebanyak 15.000 ulasan dianalisis dan diklasifikasikan ke dalam sentimen positif maupun negatif, setelah melalui tahapan pra-pemrosesan. Hasil analisis menunjukkan bahwa model SVM dengan TF-IDF menghasilkan akurasi 87%, precision 90% untuk sentimen positif, serta recall sebesar 82%. Sebaliknya, pendekatan Word2Vec mencatatkan akurasi 83%, precision positif 92%, dan recall 71%. Temuan ini mengindikasikan bahwa TF-IDF lebih konsisten dalam klasifikasi umum, sedangkan Word2Vec lebih efektif dalam mengenali sentimen negatif.

**Kata kunci:** Analisis Sentimen, Livin' by Mandiri, Mobile Banking, SVM, TF-IDF, Word2Vec.

### *User Sentiment Analysis of the Livin' by Mandiri Application Using the Support Vector Machine (SVM) Method with TF-IDF and Word2Vec Feature Extraction*

### Abstract

The rapid development of mobile banking applications in Indonesia, including Livin' by Mandiri, has created a need to understand user responses in more depth. This study aims to evaluate user sentiment towards the application based on reviews obtained from the Google Play Store. This study applies the Support Vector Machine (SVM) algorithm with two feature extraction techniques, namely Term Frequency–Inverse Document Frequency (TF-IDF) and Word2Vec. A total of 15,000 reviews were analyzed and classified into positive and negative sentiments, after going through a pre-processing stage. The results of the analysis show that the SVM model with TF-IDF produces an accuracy of 87%, a precision of 90% for positive sentiment, and a recall of 82%. In contrast, the Word2Vec approach recorded an accuracy of 83%, a positive precision of 92%, and a recall of 71%. These findings indicate that TF-IDF is more consistent in general classification, while Word2Vec is more effective in recognizing negative sentiments.

**Keywords:** Livin' by Mandiri, Mobile Banking, Sentiment Analysis, SVM, TF-IDF, Word2Vec.

## 1. PENDAHULUAN

Transformasi digital telah membawa perubahan signifikan dalam cara masyarakat mengakses layanan keuangan. Aplikasi mobile banking kini menjadi sarana utama yang digunakan oleh nasabah untuk melakukan transaksi secara cepat dan fleksibel melalui perangkat digital. Salah satu aplikasi yang mengalami pertumbuhan pesat di Indonesia adalah Livin' by Mandiri. Jumlah penggunaannya meningkat dari 15 juta pada tahun 2022 menjadi 19,5 juta pada 2023 [1]. Berdasarkan hasil survei Visa tahun 2023, lebih dari 60% masyarakat sudah tidak lagi mengunjungi kantor cabang, dan 81% di antaranya aktif menggunakan layanan digital banking setiap minggu [2]. Kondisi ini menegaskan bahwa kualitas aplikasi menjadi aspek krusial dalam mempertahankan kenyamanan dan kepuasan pengguna.

Dengan meningkatnya penggunaan aplikasi, semakin banyak pula masukan dari pengguna yang tercermin dalam bentuk ulasan di platform seperti Google Play Store. Informasi ini mencerminkan pengalaman langsung pengguna yang sangat berharga bagi pengembang aplikasi. Untuk menggali informasi tersebut secara efisien, analisis sentimen dapat diterapkan. Teknik ini memungkinkan sistem untuk mendeteksi dan mengklasifikasikan

opini menjadi kategori seperti positif atau negatif secara otomatis dan dalam skala besar [3]. Hasilnya dapat digunakan sebagai dasar untuk memperbaiki layanan dan merespons keluhan pengguna secara lebih tepat sasaran.

Sejumlah penelitian sebelumnya telah menunjukkan bahwa algoritma Support Vector Machine (SVM) merupakan salah satu metode yang andal dalam analisis sentimen, terutama karena kemampuannya dalam mengelola data teks berdimensi tinggi dan kompleks. Misalnya, penelitian sebelumnya berhasil menerapkan SVM dalam menganalisis ulasan aplikasi BRImo dan BCA Mobile dengan akurasi tinggi, mencapai 94% dan 95% [4]. Selain itu, dua teknik populer dalam ekstraksi fitur teks, yakni TF-IDF dan Word2Vec, juga banyak digunakan di berbagai domain seperti ulasan produk, hotel, dan platform e-commerce [5]. Beberapa penelitian sebelumnya juga telah membahas penggunaan metode ekstraksi fitur dan klasifikasi sentimen membandingkan performa TF-IDF dan Word2Vec dalam klasifikasi teks sentimen masyarakat terhadap produk lokal, dan menemukan bahwa TF-IDF memberikan hasil yang lebih konsisten [6]. Sementara itu, penelitian sebelumnya menggunakan pendekatan Word Embedding dan LSTM untuk menganalisis sentimen ulasan pengguna aplikasi Shopee dan menunjukkan bahwa representasi kata kontekstual dapat meningkatkan akurasi klasifikasi [7]. TF-IDF lebih fokus pada frekuensi kemunculan kata, sedangkan Word2Vec mampu menangkap makna kata dalam konteks yang lebih dalam [8].

Penelitian lain juga membandingkan performa SVM dan Naïve Bayes dalam menganalisis sentimen terhadap isu penghapusan skripsi. SVM unggul dengan akurasi 80% dibandingkan Naïve Bayes (75%), memperkuat posisi SVM sebagai metode klasifikasi yang andal [9]. Kajian lain melakukan analisis sentimen terhadap kebijakan "Merdeka Belajar Kampus Merdeka" menggunakan SVM dan Word2Vec. Studi ini menunjukkan bagaimana representasi kontekstual dari Word2Vec dapat memperkuat klasifikasi terhadap data opini kebijakan [10]. Ini menunjukkan bahwa kombinasi teknik klasifikasi dan representasi teks masih menjadi area aktif untuk dieksplorasi lebih lanjut.

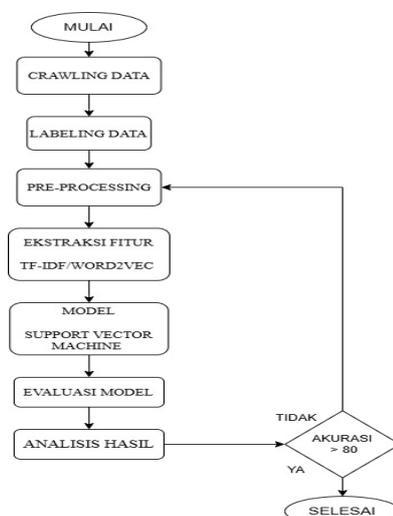
Meski berbagai studi telah membuktikan efektivitas metode SVM dengan TF-IDF maupun Word2Vec, sebagian besar difokuskan pada aplikasi atau domain lain, seperti BRImo, BCA Mobile, serta ulasan produk dan layanan e-commerce. Studi yang secara spesifik menganalisis sentimen pengguna terhadap aplikasi Livin' by Mandiri masih sangat terbatas. Selain itu, belum banyak penelitian yang secara langsung membandingkan dua pendekatan ekstraksi fitur tersebut dalam konteks aplikasi ini. Mengingat karakteristik pengguna dan layanan yang khas dari Livin' by Mandiri, analisis terhadap ulasannya memiliki nilai penting tersendiri.

Penelitian ini bertujuan untuk melakukan klasifikasi sentimen berdasarkan ulasan pengguna aplikasi Livin' by Mandiri yang diperoleh dari Google Play Store. Metode klasifikasi yang digunakan adalah SVM, dengan dua pendekatan ekstraksi fitur, yaitu TF-IDF dan Word2Vec. Penelitian ini bertujuan untuk mengevaluasi dan membandingkan efektivitas kedua teknik tersebut dalam mengenali sentimen positif dan negatif. Diharapkan hasil studi ini dapat menjadi kontribusi untuk pengembangan sistem evaluasi berbasis machine learning guna meningkatkan kualitas layanan digital banking di Indonesia.

## 2. METODE PENELITIAN

Penelitian ini terdiri atas enam tahapan utama yang digambarkan pada gambar 1. Diagram alir tersebut menunjukkan keseluruhan proses penelitian mulai dari pengumpulan data hingga evaluasi model.

### 2.1. Alur Penelitian



Gambar 1. Flowchart Diagram Alur Penelitian

Gambar 1 menunjukkan *flowchart* diagram alur penelitian yang mencakup: pengumpulan data, pelabelan data, pra-pemrosesan, ekstraksi fitur TF-IDF dan Word2Vec, klasifikasi dengan SVM, dan evaluasi model. [11].

### 2.2. Pengumpulan Data

Sumber data pada penelitian ini berasal dari review pengguna terhadap aplikasi Livin' by Mandiri yang diambil dari Google Play Store. Data tersebut diperoleh melalui pemrograman Python untuk melakukan web scraping. Proses pengumpulan data berlangsung antara 5 Februari 2024 hingga 24 Februari 2025, dengan total sebanyak 15.000 ulasan data yang mencakup teks, tanggal, dan rating yang diberikan oleh pengguna [12].

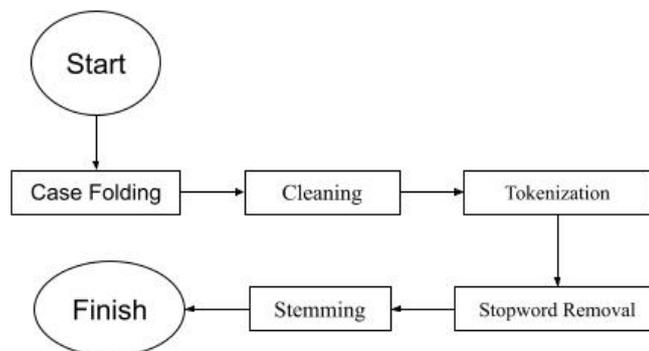
### 2.3. Pelebelan Data

Pada penelitian ini dilanjutkan dengan tahap pelabelan data. Data diberi label berdasarkan sentimen, yang dibagi menjadi dua kategori: sentimen positif untuk rating 4 hingga 5, dan sentimen negatif untuk rating 1 hingga 2. Ulasan dengan label positif diberikan apabila pengguna menyatakan dukungan atau kepuasan terhadap aplikasi, sedangkan ulasan dengan label negatif diberikan apabila terdapat kritik atau keluhan mengenai masalah pada aplikasi, seperti kinerja yang buruk atau kesalahan (error) [13].

### 2.4. Preprocessing Data

Gambar 2 menunjukkan flowchart proses preprocessing data. Pada tahap ini, hasil scraping berupa data mentah diproses melalui beberapa tahapan pra-pemrosesan untuk memastikan kualitas data yang layak dianalisis. Tahapan tersebut meliputi: [14].

1. Case Folding: Mengubah teks menjadi huruf kecil.
2. Pembersihan (Cleaning): Menghapus angka, simbol, tanda baca, dan link.
3. Tokenisasi: Memecah teks menjadi kata-kata.
4. Penghapusan Stopword: Menghapus kata yang tidak bermakna.
5. Stemming: Mengubah kata ke bentuk dasar.



Gambar 2. Flowchart Preprocessing Data

### 2.5. TF-IDF

Term Frequency - Inverse Document Frequency (TF-IDF) merupakan teknik pembobotan kata berdasarkan frekuensi relatif dalam dokumen dan seluruh korpus. Kata yang relevan dengan dokumen tetapi jarang di dokumen lain yang mendapatkan bobot tinggi [15].

$$W_{t,d} = T_{Ft,d} * I_{DFt} \tag{1}$$

Keterangan:

- $W_{t,d}$  = bobot dari t (term) dalam suatu dokumen,  $T_{Ft,d}$  = frekuensi kemunculan t (term) dalam dokumen d, dan
- $I_{DFt}$  = Inverse Document Frequency. Di mana persamaan IDF seperti pada persamaan

$$I_{DFt} = \log \left( \frac{N}{NT} \right) \tag{2}$$

Keterangan:

- $N$  = jumlah dokumen, dibagi dengan  $nt$
- $N$  = jumlah dokumen yang mengandung term

## 2.6. WORD2VEC

Word2Vec digunakan untuk menghasilkan vector numerik berdasarkan konteks kata menggunakan metode skip-gram. Tiap kata direpresentasikan dalam bentuk vector, dan vector untuk satu ulasan diperoleh dari rata-rata vector semua kata dalam ulasan tersebut [16].

Secara matematis, fungsi utama dari Skip-gram adalah memaksimalkan peluang munculnya kata-kata konteks dengan rumus.

$$\frac{1}{T} \sum_{t=1}^T \sum_{-c \leq j \leq c, j \neq 0} \log p(w_{t+j} | w_t) \quad (3)$$

## 2.7. Support Vectore Mechine

Klasifikasi dilakukan menggunakan algoritma Support Vector Machine (SVM) untuk memisahkan data ulasan menjadi dua kelas sentimen: positif dan negatif. SVM dipilih karena terbukti efektif dalam menangani data berdimensi tinggi seperti representasi teks, serta memberikan performa yang lebih baik dalam berbagai penelitian analisis sentimen. Beberapa jenis kernel yang digunakan dalam SVM adalah [17]:

$$\text{Kernel Liner} : k(x, y) = x, y \quad (4)$$

$$\text{Kernel sigmoid} : k(x, y) = \tanh(\sigma(x, y) + c) \quad (5)$$

$$\text{Kernel Polynomial} : k(x, y) = (x, y) \quad (6)$$

$$\text{Kernel RBF} : K(x, y) = \exp\left(\frac{-x-y^2}{2a^2}\right) \quad (7)$$

Parameter SVM yang digunakan:

- Kernel Linear : Dipilih karena teks yang telah dikonversi ke ruang vektor (TF-IDF dan Word2Vec umumnya dapat dipisahkan secara linier.
- C (Regularization) 1.0: Mengontrol trade-off antara kompleksitas model dan kesalahan klasifikasi. Nilai default 1.0 memberikan margin optimal dengan generalisasi yang baik.
- Gamma auto :Untuk kernel linier, gamma tidak terlalu berpengaruh, namun tetap disebutkan untuk menjaga konsistensi pengaruh model

## 2.8. Confusion Matrix

Confusion Matrix yakni tahap pengujian terhadap performa dan tingkat akurasi yang telah dilaksanakan. Confusion Matrix memberikan informasi prediksi dan aktual dari data yang telah dilakukan klasifikasi. Confusion Matrix dapat menentukan hasil dari accuracy, precision dan recall. Confusion Matrix ialah resume hasil dari prediksi masalah klasifikasi. Confusion Matrix menganalisis model klasifikasi yang paling efektif dalam mengidentifikasi kategori data yang berbeda [18].

$$\text{Accuracy} = \frac{TP+TN}{TF+FP+TN+FN} \quad (8)$$

$$\text{Precision} = \frac{TP}{TP+FP} \quad (9)$$

$$\text{Recall} = \frac{TP}{TP+FN} \quad (10)$$

$$F1 - score = 2 \times \frac{precision \times recall}{precision + recall} \quad (11)$$

### 3. HASIL DAN PEMBAHASAN

Penelitian ini membandingkan kinerja algoritma Support Vector Machine (SVM) dengan dua metode ekstraksi fitur, yaitu TF-IDF dan Word2Vec, dalam klasifikasi sentimen pengguna aplikasi *Livin' by Mandiri*. Hasil pengujian disajikan dalam bentuk tabel, grafik, dan analisis berdasarkan metrik evaluasi, yaitu akurasi, precision, recall, F1-score, dan AUC. Diskusi dilengkapi dengan perbandingan hasil penelitian sebelumnya untuk memperkuat interpretasi temuan.

#### 3.1. Pengumpulan Data

Tabel 1 menunjukkan hasil pengumpulan data penelitian diperoleh dari ulasan pengguna aplikasi *Livin' by Mandiri* di Google Play Store. Pengumpulan data dilakukan melalui web crawling menggunakan Google Colab dengan kata kunci "*Livin' by Mandiri*". Sebanyak 15.000 ulasan berbahasa Indonesia berhasil dikumpulkan dalam rentang waktu 5 Februari 2024 hingga 24 Februari 2025. Data mencakup opini pengguna terkait pengalaman mereka dalam menggunakan aplikasi mobile banking tersebut.

Tabel 1. Hasil Pengumpulan Data

Ulasan	Rating	Sentimen
Verifikasi wajah susahnya minta ampun sudah puluhan kali gagal terus.. 🤔 Kalo hp nya tidak mendukung tapi aplikasi lain mudah sekali verifikasi langsung ok.. gak perlu berulang ulang	1	Positif
Setelah lewat jam 23.00 semalem ada maintenance pemeliharaan, setelah itu livin mandiri saya ga bisa dibuka lagi, di updates malah jadi downgrade, ini salah satu bank terbesar Indonesia tapi aplikasinya begini.	4	Negatif
Masih ngeri transaksi nominal besar yg kecil aja trtunda dari 07 Febuari sampe skarang gak balik2 udah komfirmasi lewat email katanya tgal 24 tapi blom Balik juga	2	Positif
Aplikasi ndak guna tiap kali ganti hp ndak bisa masuk, untuk masuk ke akun aja, harus ke cabang dulu baru bisa, benar2 aplikasi tidak berguna	5	Negatif

#### 3.2. Preprocessing Data

Pada tahapan selanjutnya dataset yang belum terstruktur dengan baik diproses melalui tahap preprocessing untuk menghasilkan data yang lebih bersih dan terstruktur. Proses ini dilakukan menggunakan Google Collab untuk memudahkan analisis sentimen yang akan dilakukan pada data tersebut.

Tabel 2 menunjukkan hasil *case folding*, yaitu tahap mengubah seluruh text ulasan yang dihasilkan dari pengumpulan data menjadi huruf kecil untuk menyamakan format teks. Proses ini dilakukan dengan mengubah semua huruf karakter kapital menjadi huruf lebih kecil.

Tabel 2. Hasil Proses *Case Folding*

Sebelum	Sesudah
Setelah lewat jam 23.00 semalem ada maintenance pemeliharaan, setelah itu livin mandiri saya ga bisa dibuka lagi, di updates malah jadi downgrade, ini salah satu bank terbesar Indonesia tapi aplikasinya begini.	setelah lewat jam 23.00 semalem ada maintenance pemeliharaan, setelah itu livin mandiri saya ga bisa dibuka lagi, di updates malah jadi downgrade, ini salah satu bank terbesar indonesia tapi aplikasinya begini.

Tabel 3 menunjukkan hasil proses *cleaning*, yaitu tahapan pembersihan data dengan menghapus tanda baca, simbol, angka, serta karakter non-huruf seperti hashtag, mention, retweet, dan tautan. Proses ini bertujuan untuk memastikan hanya teks yang relevan dan layak digunakan dalam analisis sentimen.

Tabel 3. Hasil Proses *Cleaning*

Sebelum	Sesudah
Setelah lewat jam 23.00 semalem ada maintenance pemeliharaan, setelah itu livin mandiri saya ga bisa dibuka lagi, di updates malah jadi downgrade, ini	setelah lewat jam semalem ada maintenance pemeliharaan setelah itu livin mandiri saya ga bisa

salah satu bank terbesar Indonesia tapi aplikasinya begini.	dibuka lagi di updates malah jadi downgrade ini salah satu bank terbesar indonesia tapi aplikasinya begini
--	---

Tabel 4 menunjukkan hasil proses *tokenization*, yaitu tahapan setelah cleaning yang bertujuan memecah kalimat menjadi beberapa kata (token). Dengan proses ini, setiap ulasan dapat diuraikan menjadi unit kata yang lebih kecil, sehingga mempermudah proses analisis lanjutan. Proses tokenisasi ini dilakukan menggunakan operator tokenisasi standar dalam pemrosesan teks.

Tabel 4. Hasil Proses *Tokenization*

Sebelum	Sesudah
Setelah lewat jam 23.00 semalem ada maintenance pemeliharaan, setelah itu livin mandiri saya ga bisa dibuka lagi, di updates malah jadi downgrade, ini salah satu bank terbesar Indonesia tapi aplikasinya begini.	['setelah', 'lewat', 'jam', 'semalem', 'ada', 'maintenance', 'pemeliharaan', 'setelah', 'itu', 'livin', 'mandiri', 'saya', 'ga', 'bisa', 'dibuka', 'lagi', 'di', 'updates', 'malah', 'jadi', 'downgrade', 'ini', 'salah', 'satu', 'bank', 'terbesar', 'indonesia', 'tapi', 'aplikasinya', 'begini']

Tabel 5 menunjukkan hasil proses *stopword removal*, yaitu rangkaian preprocessing. Pada tahap ini, kata-kata umum yang tidak memiliki makna penting untuk analisis, seperti “dan”, “yang”, atau “di”, dihapus dari teks. Proses dilakukan dengan menggunakan operator *filter by length* untuk membatasi jumlah kata yang dipertahankan. Hasilnya ditampilkan sebagai teks yang hanya mengandung kata-kata bermakna.

Tabel 5. Hasil Proses *Stopword Removal*

Sebelum	Sesudah
Setelah lewat jam 23.00 semalem ada maintenance pemeliharaan, setelah itu livin mandiri saya ga bisa dibuka lagi, di updates malah jadi downgrade, ini salah satu bank terbesar Indonesia tapi aplikasinya begini.	jam semalem maintenance pemeliharaan livin mandiri ga dibuka updates downgrade salah bank terbesar indonesia aplikasinya

Tabel 6 menunjukkan hasil proses *stemming*, yaitu tahap mengubah kata-kata dalam ulasan menjadi bentuk dasar. Proses ini menghilangkan imbuhan pada kata sehingga diperoleh bentuk dasar yang konsisten dan bermakna, yang memudahkan analisis sentimen.

Tabel 6. Hasil Proses *Stemming*

Sebelum	Sesudah
Setelah lewat jam 23.00 semalem ada maintenance pemeliharaan, setelah itu livin mandiri saya ga bisa dibuka lagi, di updates malah jadi downgrade, ini salah satu bank terbesar Indonesia tapi aplikasinya begini.	jam semalem maintenance pemeliharaan livin mandiri ga buka updates downgrade salah bank besar indonesia aplikasi

### 3.4. Pembagian Data

Tabel 7 menunjukkan hasil proses pembagian data menjadi dua subset, yaitu 80% data latih dan 20% data uji. Pembagian ini bertujuan untuk melatih model menggunakan data latih dan menguji performanya dengan data uji. Parameter `random_state=42` digunakan untuk memastikan konsistensi hasil pembagian pada setiap eksekusi.

Tabel 7. Hasil Proses Pembagian Data

Data Latih	Data Test	Total
12.000	3.000	15.000

### 3.5. Ekstraksi Fitur TF-IDF

Tabel 8 menunjukkan hasil proses ekstraksi fitur menggunakan TF-IDF. Pada tahap ini, setiap kata dalam ulasan dikonversi menjadi bobot numerik yang mencerminkan tingkat kepentingannya dalam dokumen. Setiap baris mewakili satu ulasan, sedangkan setiap kolom menunjukkan kata unik. Kata yang sering muncul dalam satu

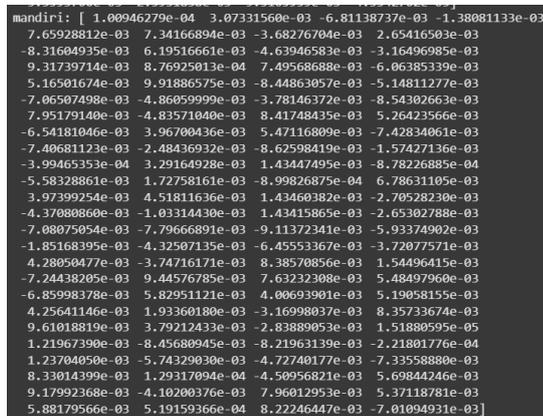
ulasan tetapi jarang di ulasan lain akan memiliki bobot lebih tinggi. Nilai TF-IDF nol menunjukkan kata tersebut tidak terdapat dalam ulasan tersebut.

Tabel 8. Hasil Proses Ekstraksi Fitur TF-IDF

No	Aplikasi	Sering	Error	Selalu	Dibuka	Transfer	Verifikasi
0	0.256729	0.345942	0.000000	0.000000	0.000000	0.000000	0.000000
1	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
2	0.282424	0.380567	0.000000	0.000000	0.000000	0.000000	0.000000
3	0.189460	0.000000	0.455292	0.000000	0.000000	0.307555	0.000000
4	0.000000	0.000000	0.000000	0.230134	0.392858	0.000000	0.000000
5	0.182614	0.000000	0.000000	0.000000	0.000000	0.000000	0.296441

### 3.6. Ekstraksi Fitur WOR2VEC

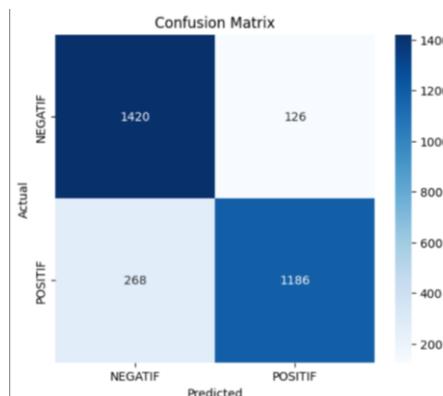
Gambar 3 menunjukkan hasil proses ekstraksi fitur menggunakan Word2Vec. Pada tahap ini, setiap kata dalam ulasan diubah menjadi vektor numerik berdimensi tinggi yang merepresentasikan makna berdasarkan konteksnya. Model Word2Vec dilatih dengan metode Continuous Bag-of-Words dan Skip-gram untuk menghasilkan representasi vektor kata yang digunakan sebagai fitur dalam klasifikasi.



Gambar 3. Hasil Proses Ekstraksi Fitur Word2Vec

### 3.7. Hasil Kinerja Model TF-IDF

Gambar 4 menunjukkan hasil proses confusion matrix SVM dan TF-IDF. Analisis sentimen dilakukan dengan bahasa pemrograman Python dan dievaluasi menggunakan metrik accuracy, precision, recall, F1-score, dan AUC. Confusion matrix ini menggambarkan perbandingan hasil prediksi model dengan label asli.



Gambar 4. Hasil Proses Confusion Matrix SVM dan TF-IDF

Confusion matrix menunjukkan bahwa model berhasil mengklasifikasikan, 1.420 ulasan negatif dan 1.186 ulasan positif dengan benar. Selain itu model melakukan kesalahan prediksi 126 data negatif yang diklasifikasikan sebagai positif (false positive). Dan 268 data positif yang diklasifikasikan sebagai negatif (false negative).

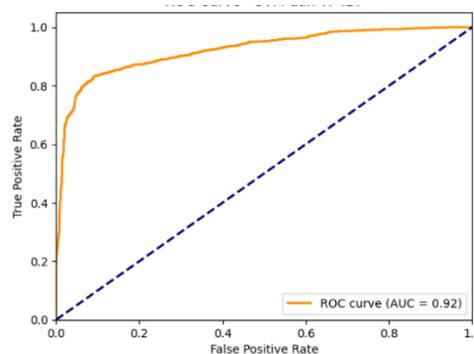
Tabel 9 menunjukkan hasil evaluasi kinerja model SVM dengan fitur TF-IDF. Model ini menunjukkan performa yang baik dalam mengklasifikasikan sentimen pengguna. Berdasarkan pengujian terhadap 3.000 data uji, diperoleh hasil evaluasi sebagaimana ditampilkan pada tabel berikut.

**Tabel 9. Hasil Evaluasi Kinerja SVM dan TF-IDF**

	<b>Precision</b>	<b>Recall</b>	<b>F1-Score</b>	<b>Support</b>
Negatif	0.84	0.92	0.88	1546
Positif	0.90	0.82	0.86	1454
Accuray			0.87	3000
Macro Avg	0.87	0.87	0.87	3000
Wighted Avg	0.87	0.87	0.87	3000
Score AUC			0.92	

Model SVM dengan TF-IDF menunjukkan kinerja yang seimbang dalam mengklasifikasikan kedua sentimen. Precision yang tinggi pada sentimen positif (90%) menunjukkan efektivitas model dalam mendeteksi ulasan positif. Recall yang tinggi pada sentimen negatif (92%) menunjukkan model yang cukup sensitif dalam menangkap ulasan negatif. Secara keseluruhan, model menghasilkan akurasi sebesar 87% dan nilai AUC 0,92% yang menunjukkan performa yang sangat baik dalam membedakan kedua kategori sentiment.

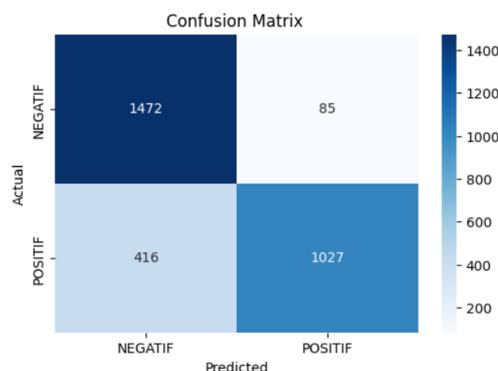
Gambar 5 menunjukkan hasil kurva AUC dari model SVM dengan fitur TF-IDF, dengan nilai AUC sebesar 0,92. Nilai in mencerminkan kemampuan model yang sangat baik dalam membedakan sentimen positif dan negatif.



Gambar 5. Hasil Kurva AUC

### 3.8. Analisis Kinerja Model SVM dan WORD2VEC

Gambar 6 menunjukkan confusion matrix SVM dan Word2VEC. Analisis dilakukan dengan bahasa pemrograman Python, dan evaluasi kinerja menggunakan metrik accuracy, precision, recall, F1-score, serta AUC. Confusion matrix digunakan untuk membandingkan hasil prediksi model dengan label sebenarnya.



Gambar 6. Hasil Confusion Matrix SVM dan Word2VEC

Confusion matrix menunjukkan bahwa model berhasil mengklasifikasikan, 1.027 ulasan positif dan 1.472 ulasan positif dengan benar. Selain itu model melakukan kesalahan prediksi 85 data negatif yang diklasifikasikan sebagai positif (false positive). Dan data positif yang diklasifikasikan sebagai negatif (false negative).

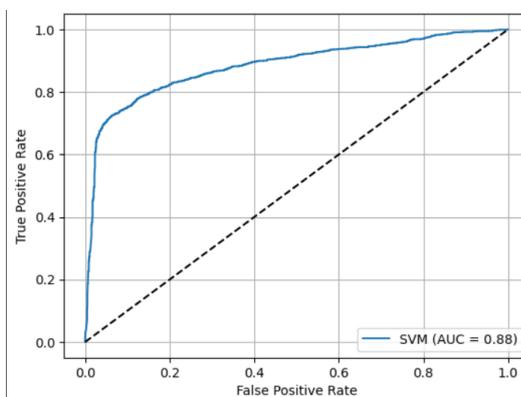
Tabel 10 menunjukkan hasil evaluasi kinerja model SVM dengan fitur Word2Vec. Model ini menunjukkan kecenderungan yang kuat dalam mendeteksi sentimen negatif, namun memiliki kelemahan dalam mengidentifikasi sentimen positif secara konsisten.

Tabel 10. Evaluasi Kinerja SVM dan Word2Vec

	Precision	Recall	F1-Score	Support
Negatif	0.78	0.95	0.85	1557
Positif	0.92	0.71	0.80	1443
Accuray			0.83	3000
Macro Avg	0.85	0.83	0.83	3000
Wighted Avg	0.85	0.83	0.83	3000
Score AUC			0.88	

Model SVM dengan Word2Vec menunjukkan kemampuan yang sangat baik dalam mendeteksi ulasan negatif, ditandai dengan recall sebesar 95% pada kategori negatif. Precision pada kategori positif juga cukup tinggi, yaitu 92% menunjukkan akurasi prediksi yang baik untuk ulasan positif. Namun, recall pada kategori positif relatif lebih rendah di (71%), yang berarti beberapa ulasan positif tidak terdeteksi dengan tepat. Secara keseluruhan, model menghasilkan akurasi di angka 83% dan AUC 0.88%, menunjukkan performa yang baik namun sedikit di bawah pendekatan TF-IDF.

Gambar 7 menunjukkan hasil kurva AUC pada model SVM dengan fitur Word2Vec menghasilkan nilai 0,88. Meskipun sedikit lebih rendah dari TF-IDF, model menunjukkan performa klasifikasi yang baik.



Gambar 7. Hasil Kurva AUC WORD2VEC

Tabel 11 menunjukkan hasil perbandingan kinerja model. TF-IDF unggul dalam akurasi (87%) dan keseimbangan klasifikasi. Sementara itu, Word2Vec lebih baik dalam mendeteksi sentimen negatif dengan recall sebesar (95%).

Tabel 11. Perbandingan Kinerja Model

Matrix	SVM + TF-IDF	SVM + Word2Vec	Terbaik
Akurasi	0.87 %	0.83%	TF-IDF
Precision Positif	0.90 %	0.92%	Word2Vec
Recall Positif	0.82 %	0.71%	TF-IDF
Precision Negatif	0,84 %	0.78%	TF-IDF
Recall Negatif	0.92 %	0.95%	Word2Vec
Score AUC	0.922	0.884	TF-IDF



#### 4. KESIMPULAN

Sesuai dengan hasil dalam ini penelitian yang telah dilakukan pada Analisis Sentimen *SVM* dengan dua pendekatan aspek berbeda, yaitu *TF-IDF* dan *Word2Vec*, untuk analisis sentimen ulasan pengguna aplikasi *Livin' by Mandiri* dari *Google Play Store*, berikut adalah tujuan utama dari eksperimen yang dilakukan:

Penelitian ini menunjukkan bahwa *SVM* dengan fitur *TF-IDF* merupakan kombinasi terbaik untuk analisis sentimen ulasan aplikasi *Livin' by Mandiri*. Model ini mencapai akurasi 87%, precision 90%, recall 82%, dan *AUC* .922, sehingga menunjukkan performa stabil dalam klasifikasi sentimen.

Sementara itu model dengan fitur *Word2Vec* memberikan precision positif 92% dan recall negatif 95% yang menunjukkan keunggulan dalam mendeteksi ulasan bernada negatif. Kedua metode memiliki kekuatan masing-masing yang dapat disesuaikan dengan tujuan analisis.

Hasil penelitian ini bermanfaat dalam pengembangan layanan aplikasi mobile banking, terutama untuk memantau opini pengguna secara otomatis, mempercepat respons terhadap keluhan, serta sebagai dasar perbaikan fitur dan strategi pengembangan aplikasi.

#### DAFTAR PUSTAKA

- [1] F. Riza, B. Kurniawan, P. Studi, T. Informatika, I. Teknologi, and B. Utomo, "KLASIFIKASI MACHINE LEARNING UNTUK ANALISIS SENTIMEN PENGGUNA MOBILE BANKING LIVIN' BY MANDIRI PADA GOOGLE PLAY STORE," (Jurnal ilmiah), vol. 18, no. 2, pp. 116–122, 2024.
- [2] visa, "Visa Consumer Payment Attitudes Study 2023," Visa Inc. (Laporan online), [Online]. Available: <https://www.visa.co.id/about-visa/newsroom/press-releases.413578.html>
- [3] D. M. W. Powers, "Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation," pp. 37–63, (Preprint arXiv), 2020, [Online]. Available: <http://arxiv.org/abs/2010.16061>
- [4] M. Rizky Pratama, Y. R. Ramadhan, and M. A. Komara, "Analisis sentimen BRImo dan BCA Mobile menggunakan Support Vector Machine dan Lexicon Based," *Jutisi J. Ilm. Tek. Inform. dan Sist. Inf.*, (Jurnal ilmiah), vol. 12, no. 3, pp. 1439–1450, 2023.
- [5] V. W. D. Thomas and F. Rumaisa, "Analisis Sentimen Ulasan Hotel Bahasa Indonesia Menggunakan Support Vector Machine dan TF-IDF," *J. Media Inform. Budidarma*, (Jurnal ilmiah), vol. 6, no. 3, p. 1767, 2022, doi: 10.30865/mib.v6i3.4218.
- [6] I. Hendrawan Rifky, E. Utami, and A. Hartanto Dwi, "Analisis Perbandingan Metode Tf-Idf dan Word2vec pada Klasifikasi Teks Sentimen Masyarakat Terhadap Produk Lokal di Indonesia," *Smart Comp Jurnalnya Orang Pint. Komput.*, (Jurnal ilmiah), vol. 11, no. 3, pp. 497–503, 2022, doi: 10.30591/smartcomp.v11i3.3902.
- [7] I. Azizah, I. Cholissodin, and N. Yudistira, "Analisis Sentimen Ulasan Pengguna Aplikasi Shopee di Google Play menggunakan Metode Word Embedding dan Long Short Term Memory (LSTM)," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, (Jurnal ilmiah), vol. 7, no. 5, pp. 2453–2459, 2023.
- [8] Nadia Ristya Dewi, E. Yulia Puspaningrum, and H. Maulana, "Analisis Sentimen Tweet Vaksinasi Covid-19 Menggunakan RNN Dengan Metode TF-IDF Dan Word2Vec," *J. Inform. dan Sist. Inf.*, (Jurnal ilmiah), vol. 3, no. 1, pp. 56–65, 2022, doi: 10.33005/jifosi.v3i1.449.
- [9] R. Yunita and M. Kamayani, "Perbandingan Algoritma SVM Dan Naïve Bayes Pada Analisis Sentimen Penghapusan Kewajiban Skripsi," *Indones. J. Comput. Sci.*, (Jurnal ilmiah), vol. 12, no. 5, pp. 2879–2890, 2023, doi: 10.33022/ijcs.v12i5.3415.
- [10] N. Rezki, "Analisis Sentimen Kebijakan Merdeka Belajar Kampus Merdeka Menggunakan Mesin Vektor Pendukung Dengan Ekstraksi Fitur ...," (Skripsi, Universitas Hasanuddin), 2022, [Online]. Available: [http://repository.unhas.ac.id/id/eprint/23689/%0Ahttp://repository.unhas.ac.id/id/eprint/23689/4/H0511810\\_26\\_skripsi\\_30-08-2022.pdf](http://repository.unhas.ac.id/id/eprint/23689/%0Ahttp://repository.unhas.ac.id/id/eprint/23689/4/H0511810_26_skripsi_30-08-2022.pdf)
- [11] J. E. Br Sinulingga and H. C. K. Sitorus, "Analisis Sentimen Opini Masyarakat terhadap Film Horor Indonesia Menggunakan Metode SVM dan TF-IDF," *J. Manaj. Inform.*, (Jurnal ilmiah), vol. 14, no. 1, pp. 42–53, 2024, doi: 10.34010/jamika.v14i1.11946.
- [12] M. I. Petiwi, A. Triayudi, and I. D. Sholihati, *Analisis Sentimen Gofood Berdasarkan Twitter Menggunakan Metode Naïve Bayes dan Support Vector Machine*, (Jurnal ilmiah), vol. 6, no. 1. 2022. doi: 10.30865/mib.v6i1.3530.
- [13] Y. S. Triyantono, S. Al Faraby, and M. Dwifabri, "Analisis Sentimen Terhadap Ulasan Film Menggunakan Word2Vec dan SVM," *eProceedings Engineering.*, (Prosiding konferensi), vol. 8, no. 4, p. 4136, 2021.

- [14] E. Hokijuliandy, H. Napitupulu, and F. Firdaniza, “Analisis Sentimen Menggunakan Metode Klasifikasi Support Vector Machine (SVM) dan Seleksi Fitur Chi-Square,” *SisInfo J. Sist. Inf. dan Inform.*, (Jurnal ilmiah), vol. 5, no. 2, pp. 40–49, 2023, doi: 10.37278/sisinfo.v5i2.670.
- [15] D. E. Cahyani and I. Patasik, “Performance comparison of TF-IDF and Word2Vec models for emotion text classification,” (Jurnal ilmiah), vol. 10, no. 5, pp. 2780–2788, 2021, doi: 10.11591/eei.v10i5.3157.
- [16] H. F. Naufal and E. B. Setiawan, “Ekspansi Fitur Pada Analisis Sentimen Twitter Dengan Pendekatan Metode Word2Vec,” *e-Proceeding Engineering.*, (Prosiding konferensi), vol. 8, no. 5, pp. 10339–10349, 2021, [Online]. Available: <https://dev.twitter.com>
- [17] W. Andriyani, Y. Astuti, and B. A. Wisesa, “Analisis Sentimen pada Ulasan Produk dengan SVM dan Word2Vec Sentiment Analysis on Product Reviews with SVM and Word2Vec,” *JIKO: Jurnal Informatika dan Komputer* (Jurnal ilmiah), no. 1, pp. 173–185, 2024, doi: 10.26798/jiko.v8i1.1498.
- [18] A. B. Naibaho, D. Anggraini, and B. A. Wardijono, “Penerapan SVM dan Word2Vec untuk Analisis Sentimen Ulasan Pengguna Aplikasi DANA,” *J. Ilm. KOMPUTASI*, (Jurnal ilmiah), vol. 21, no. 4, pp. 319–330, 2021.