

Perbandingan Metode Support Vector Machine

By raka aji

WORD COUNT

3499

TIME SUBMITTED

08-JAN-2026 10:37AM

PAPER ID

119812630

Perbandingan Metode Support Vector Machine dan K-Nearest Neighbor dalam Analisis Sentimen Ulasan Aplikasi MyXL

Raka Aji Nugroho¹⁾; Dwi Remawati²⁾; Teguh Susyanto³⁾; Wawan Laksito Yuly Saptomo⁴⁾

¹⁾Program Studi Sistem Informasi, Tiga Serangkai University Surakarta

²⁾Program Studi Teknologi Informasi, Tiga Serangkai University Surakarta

³⁾Program Studi Sistem Informasi, Tiga Serangkai University Surakarta

⁴⁾Program Studi Informatika, Tiga Serangkai University Surakarta

¹⁾23430009.raka@tsu.ac.id; ²⁾dwirema@tsu.ac.id; ³⁾teguh@tsu.ac.id; ⁴⁾wawanlaksito@tsu.ac.id

ABSTRACT

Sentiment analysis is an essential tool for understanding user perceptions of mobile applications, especially when reviews are unstructured text. This study aims to analyze user reviews of the MyXL application using Support Vector Machine (SVM) and K-Nearest Neighbor (KNN) and to compare their performance in classifying sentiments into positive, negative, and neutral categories. The dataset was obtained from Google Play Store via Kaggle and underwent text preprocessing, including case folding, removal of numbers and punctuation, tokenization, stopword removal, normalization, and stemming. Features were transformed into numerical representations using TF-IDF, and the data was split into training (70%) and testing (30%) sets. Evaluation using accuracy, precision, recall, and F1-score showed that SVM outperformed KNN with an accuracy of 0.743 versus 0.64, particularly in classifying neutral reviews. KNN exhibited higher misclassification in positive and negative classes, while SVM was more stable but tended to be biased toward the neutral class. These results provide insights for application developers to better understand user satisfaction and guide service improvement and feature development.

Keywords : Sentiment Analysis, Support Vector Machine, K-Nearest Neighbor, MyXL

I. PENDAHULUAN

Salah satu fitur Google Play Store yang memungkinkan pengguna memberikan penilaian terhadap suatu aplikasi adalah review, yang menjadi sumber informasi penting bagi pengembang untuk mengevaluasi kualitas layanan dan fitur aplikasi. Namun, ulasan yang tersedia biasanya berbentuk teks bebas yang tidak terstruktur, sehingga sulit untuk digunakan sebagai dasar pengambilan keputusan secara langsung [1].

PT XL Axiata Tbk menawarkan aplikasi self-service MyXL, yang membantu pengguna mengelola layanan telekomunikasi mereka seperti mengaktifkan paket internet, mengecek sisa kuota, dan mendapatkan bantuan melalui FAQ dan Live Chat. Aplikasi MyXL menerima banyak ulasan dengan berbagai karakteristik dari banyak pengguna. Diidentifikasi situasi di mana pengguna meninggalkan ulasan dengan nilai tinggi tetapi konten ulasannya bersifat negatif, atau sebaliknya. Ketidaksesuaian antara penilaian dan konten ulasan menyebabkan informasi tentang perasaan pengguna menjadi kurang jelas dan dapat menyebabkan kesalahan interpretasi [2].

Oleh karena itu, analisis sentimen diperlukan untuk membagi ulasan pengguna ke dalam kategori sentimen positif, negatif, dan netral. Analisis sentimen adalah bagian dari proses pemrosesan teks yang bertujuan untuk mengekstraksi pendapat atau perasaan pengguna dari data teks [1]. Studi sebelumnya menunjukkan bahwa klasifikasi dalam analisis sentimen dapat membantu pengembang aplikasi memahami tingkat kepuasan pengguna secara lebih akurat [1], [2], [3].

Pada tahap awal penelitian ini, data ulasan pengguna aplikasi MyXL dikumpulkan melalui proses scraping dari Google Play Store [2]. Kemudian data dilabelkan dan dimasukkan ke dalam kategori sentimen positif, negatif, dan netral. Tahap preprocessing teks berikutnya dilakukan untuk membersihkan dan memilih data teks. Ini menghasilkan data yang lebih terorganisir dan siap untuk diproses. Selanjutnya, untuk menampilkan teks dalam bentuk numerik, dilakukan pembobotan kata menggunakan metode Term Frequency Inverse Document Frequency (TF-IDF) [4], [5], [6]. Akhir dari penelitian ini adalah proses klasifikasi sentimen menggunakan algoritma Support Vector Machine (SVM) dan K-Nearest

Neighbor (KNN), serta mengevaluasi kinerja kedua metode untuk menentukan metode mana yang memberikan hasil klasifikasi terbaik pada ulasan aplikasi MyXL [2], [7], [8].

29 II. TINJAUAN PUSTAKA

2.1 Analisis Sentimen

Salah satu bagian dari pengolahan teks adalah analisis sentimen, yang bertujuan untuk menemukan dan menempatkan pendapat, emosi, atau sikap pengguna dalam teks ke dalam kategori tertentu, seperti sentimen positif, negatif, atau netral. Analisis sentimen banyak digunakan untuk data ulasan pengguna aplikasi di Google Play Store karena dapat secara otomatis menunjukkan tingkat kepuasan pengguna terhadap layanan atau produk tertentu. Analisis sentimen dapat membantu mengubah data yang tidak terstruktur menjadi informasi yang lebih relevan dan mudah dipahami [4], [5].

2.2 Ulasan Aplikasi di Google Playstore

Google Play Store menyediakan fitur ulasan, yang memungkinkan pengguna memberikan ulasan dan komentar terhadap aplikasi yang mereka gunakan. Untuk mengevaluasi kualitas aplikasi, pengembang dapat menggunakan ulasan ini sebagai sumber data penting. Namun, ketidaksesuaian sering terjadi antara penilaian numerik dan isi teks ulasan, yang dapat menyebabkan kesalahan interpretasi apabila hanya melihat penilaian saja tanpa memeriksa teks ulasan secara menyeluruh. Akibatnya, untuk menyelesaikan masalah tersebut, pendekatan analisis sentimen diperlukan [1], [3], [9].

2.3 Text Preprocessing

Preprocessing teks adalah langkah pertama yang sangat penting dalam analisis sentimen. Tujuan dari proses ini adalah untuk membersihkan dan menyiapkan data teks sehingga lebih terorganisir dan siap untuk digunakan dalam proses pembelajaran mesin. Preprocessing yang baik dapat meningkatkan kualitas fitur dan akurasi model klasifikasi sentimen dengan memotong case, menghilangkan tanda baca dan angka, tokenisasi,

menghilangkan stopword, normalisasi kata, dan stemming [4], [5], [6], [10].

44 2.4 Ekstraksi Fitur Menggunakan TF-IDF

Setelah proses preprocessing, data teks harus diubah menjadi representasi numerik sehingga algoritma pembelajaran mesin dapat memrosesnya. Salah satu metode yang paling umum digunakan adalah Term Frequency Inverse Document Frequency (TF-IDF). TF-IDF memberikan bobot pada setiap kata berdasarkan frekuensi kemunculannya dalam sebuah dokumen dan keseluruhan kumpulan dokumen. Ini terbukti efektif dalam representasi teks dan sering digunakan dalam penelitian analisis sentimen berbasis ulasan aplikasi [4], [5].

2.5 Support Virtual Machine

Banyak digunakan dalam klasifikasi teks, seperti analisis sentimen, Support Vector Machine (SVM). Hyperplane terbaik dipilih oleh SVM untuk memisahkan data ke dalam kelas-kelas tertentu secara optimal. SVM memiliki kemampuan untuk menangani data yang sangat besar dan menghasilkan klasifikasi data teks yang baik, yang merupakan salah satu keunggulannya. Akibatnya, SVM sering digunakan sebagai metode pembandingan dalam studi analisis sentiment [2], [6], [11].

49 2.6 K-Nearest Neighbor

Algoritma klasifikasi berbasis instance K-Nearest Neighbor (KNN) menentukan kelas data berdasarkan mayoritas kelas dari k tetangga terdekatnya. Konsep algoritma ini sederhana dan mudah digunakan. KNN sering digunakan sebagai metode pembandingan dalam analisis sentimen karena mampu memberikan hasil klasifikasi yang sangat baik, meskipun memiliki kelemahan dalam hal waktu komputasi pada dataset yang sangat besar [7], [8], [10].

2.7 Penelitian Terdahulu

Beberapa penelitian terdahulu telah membahas analisis sentimen pada ulasan aplikasi di Google Play Store dengan berbagai metode klasifikasi. Penelitian sebelumnya menunjukkan bahwa

Ulasan	Cleaned
Harusnya dikasih bintang 4 bilang terimakasih...	harus kasih bintang bilang terimakasih mau nga...
Tolong dong masalah jaringan hampir setiap har...	dong masalah jaring hampir hari leg parahh jan...
Saya mau komen lagi,mimininn kenapa akhir akhi...	mau komen mimin akhir akhir jaring xl sangat l...
Bonus kouta tiktok 13gb tidak diaktifkan/tidak...	bonus kouta tiktok gb aktif jadi kouta yg...
Sejauh ini bagusA? aja sih apk nya, tpi udh bb...	jauh bagus aja sih apk nya tpi udh bbrp hari s...
...	...
Xi sekarang beda sama dulu, sekarang xi boros...	xi sekarang beda sama dulu sekarang xi boros b...
Sebagai Pelanggan Sella XL Bertahun tahun , Ke...	langgan sella xl tahun tahun kecewa layan xl s...
Suka boros kalo pake sekali beli mahal bgt kat...	suka boros kalo pake sekali beli mahal banget ...

Gambar 2. Hasil Text Preprocessing

Gambar 2 menunjukkan hasil dari proses preprocessing teks, yang menunjukkan bahwa teks ulasan menjadi lebih bersih dan seragam. Teks ulasan sebelum dan sesudah pembersihan berbeda.

3.3 Pelabelan Sentimen

Pelabelan sentimen dilakukan untuk mengelompokkan setiap ulasan ke dalam tiga kelas, yaitu sentimen positif, negatif, dan netral. Ini dilakukan secara manual atau sesuai dengan kriteria tertentu untuk memberikan kelas target yang akan digunakan untuk klasifikasi [3], [7].

```
#menambahkan label sentimen
positif_words =
['mantap','cepat','mudah','ramah','puas',
'bagus','stabil','lancar']
negatif_words =
['jelek','lemot','kesal','buruk','lambat',
'susah','tidak nyaman']
netral_words =
['biasa','standar','lumayan','oke','tida
k buruk','tidak terlalu']

def label_sentimen(text):
    pos_count = sum(1 for word in
    positif_words if word in text)
    neg_count = sum(1 for word in
    negatif_words if word in text)
    net_count = sum(1 for word in
    netral_words if word in text)

    if pos_count > neg_count and
    pos_count > net_count:
        return 'Positif'
    elif neg_count > pos_count and
    neg_count > net_count:
        return 'Negatif'
    elif net_count > pos_count and
    net_count > neg_count:
        return 'Netral'
    else:
```

```
return 'Netral'

df['Keterangan'] =
df['Cleaned'].apply(label_sentimen)
df
```

Kode 2. Proses Pelabelan Sentimen

Keterangan
Netral
Netral
Netral
Netral
Positif
...
Netral
Netral
Netral

Gambar 3. Hasil Pelabelan Sentimen

3.4 Ekstraksi Fitur

Ekstraksi fitur digunakan untuk mengubah data teks menjadi representasi numerik yang dapat diproses oleh algoritma pembelajaran mesin. Metode Term Frequency Inverse Document 26:quency (TF-IDF) digunakan. TF-IDF memberikan bobot pada setiap kata berdasarkan frekuensi kemunculannya dalam kedua dokumen dan dataset secara keseluruhan.

```
#menggabungkan fitur untuk metode KNN
dan SVM
vectorizer = TfidfVectorizer()
X_text =
vectorizer.fit_transform(df['Cleaned'])
X_rating = df['Score'].values.reshape(-
1, 1)

#menggabungkan TF-IDF dengan numeric
X_combined = hstack([X_text, X_rating])
y = df['Sentimen_Num']
```

Kode 3. Ekstraksi Fitur

Proses ekstraksi fitur menggunakan metode TF-IDF untuk mengubah data teks hasil preprocessing menjadi representasi numerik ditunjukkan dalam kode 3. Tingkat kepentingan kata dalam dokumen ulasan ditentukan dengan TF-IDF. Metode stacking horizontal menggabungkan vektor TF-IDF dengan fitur teks dan skor penilaian pengguna. Tujuan dari integrasi fitur ini adalah untuk meningkatkan jumlah informasi yang digunakan oleh algoritma

KNN dan SVM untuk melakukan klasifikasi sentiment [4], [5], [6].

3.5 Pembagian Data

Data set telah diekstraksi kemudian dibagi menjadi dua bagian: data latihan (training) dan data uji (testing). Sebagian besar dari data uji digunakan untuk membangun model klasifikasi, sedangkan sebagian besar data latih digunakan untuk mengukur kinerja model [7], [10].

```
#memisahkan data menjadi 70% training
dan 30% testing
X_train, X_test, y_train, y_test =
train_test_split(
    X_combined, y, test_size=0.3,
    random_state=42
)
```

Kode 4. Pembagian Data

Dalam penelitian ini, fungsi `train_test_split` digunakan untuk membagi data menjadi 70% data latihan dan 30% data uji. Nilai `random_state` digunakan untuk memastikan bahwa proses pembagian data konsisten dan replikable. Proses pembagian data yang digunakan dalam penelitian ini digambarkan dalam kode 4.

3.6 Klasifikasi

Dua algoritma digunakan untuk melakukan tahap klasifikasi: K-Nearest Neighbor (KNN) dan Support Vector Machine (SVM). Metode KNN menentukan kelas sentimen berdasarkan tetangga terdekat dengan menggunakan nilai $k = 3$ [7], [10]. Sebaliknya, metode SVM memisahkan data ke dalam kelas sentimen dengan menggunakan kernel linear [2], [3], [5], [6]. Untuk mendapatkan hasil yang dapat dibandingkan, kedua pendekatan digunakan pada data yang sama.

```
Model KNN K3
knn =
KNeighborsClassifier(n_neighbors=3)
knn.fit(X_train, y_train)
y_pred_knn = knn.predict(X_test)

SVM dengan kernel linear
svm = SVC(kernel='linear')
svm.fit(X_train, y_train)
y_pred_svm = svm.predict(X_test)
```

Kode 5. Klasifikasi

Untuk menghasilkan hasil klasifikasi yang dapat dibandingkan secara adil, kedua algoritma digunakan pada dataset yang sama. Kode 5 menunjukkan metode klasifikasi sentimen menggunakan metode KNN dan SVM [8].

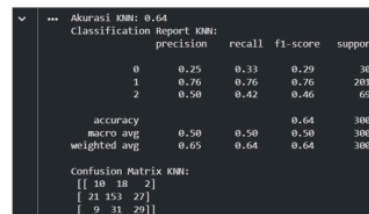
3.7 Evaluasi Model

Evaluasi model dilakukan untuk mengetahui kinerja algoritma KNN dan SVM dalam mengklasifikasikan sentimen ulasan. Metode yang dievaluasi termasuk akurasi, ketepatan, recall, dan skor F1. Hasil evaluasi digunakan untuk membandingkan kinerja kedua algoritma dan menentukan yang memberikan hasil terbaik.

```
#menampilkan akurasi metode KNN
print("Akurasi KNN:",
accuracy_score(y_test, y_pred_knn))
print("Classification Report KNN:\n",
classification_report(y_test,
y_pred_knn))
print("Confusion Matrix KNN:\n",
confusion_matrix(y_test, y_pred_knn))

#menampilkan akurasi metode SVM
print("Akurasi SVM:",
accuracy_score(y_test, y_pred_svm))
print("Classification Report SVM:\n",
classification_report(y_test,
y_pred_svm))
print("Confusion Matrix SVM:\n",
confusion_matrix(y_test, y_pred_svm))
```

Kode 6. Evaluasi Model



Gambar 4. Akurasi Metode KNN

Dengan akurasi 0.64 dan performa terbaik pada kelas 1 (sentimen netral atau positif, tergantung label), KNN memiliki precision dan recall yang lebih rendah pada kelas 0 dan 2, seperti yang ditunjukkan oleh confusion matrix, di mana banyak prediksi kelas 0 dan 2 yang salah diklasifikasikan.

```

Akurasi SVM: 0.7433333333333333
Classification Report SVM:
              precision    recall  f1-score   support

 0             0.40         0.07         0.11         30
 1             0.74         0.95         0.83        201
 2             0.79         0.45         0.57         69

 accuracy             0.74         0.74         0.74        300
 macro avg             0.65         0.49         0.51        300
 weighted avg          0.72         0.74         0.70        300

Confusion Matrix SVM:
[[ 2  28  0]
 [ 3 190  8]
 [ 0 38 31]]
    
```

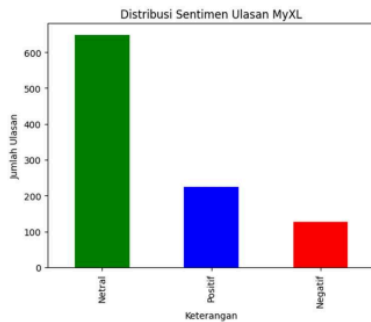
Gambar 5. Akurasi Metode SVM

SVM memiliki akurasi lebih tinggi, yaitu 0.743, dengan precision dan recall terbaik pada kelas 1, serta performa lebih stabil pada kelas 2 dibanding KNN. Confusion matrix menunjukkan bahwa SVM lebih mampu membedakan kelas dengan jumlah data besar (kelas 1), meskipun kelas minor (kelas 0) masih sering salah klasifikasi [7], [8], [10].

IV. HASIL DAN PEMBAHASAN

Untuk membuat interpretasi dan analisis hasil penelitian lebih mudah, visualisasi hasil menampilkan grafik perbandingan akurasi metode KNN dan SVM, confusion matrix, dan distribusi sentimen ulasan pengguna aplikasi MyXL.

4.1 Distribusi Sentimen Ulasan

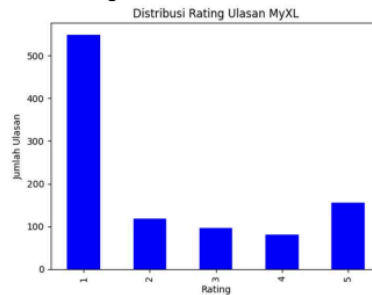


Gambar 6. Visualisasi Berdasarkan Sentimen

Gambar Distribusi Sentimen Ulasan MyXL menunjukkan bahwa sekitar 650 ulasan menunjukkan sentimen netral, dengan 220 ulasan positif dan 130 ulasan negatif, yang menunjukkan bahwa sebagian besar pengguna aplikasi MyXL tidak sepenuhnya puas atau tidak puas. Namun, lebih banyak ulasan positif daripada ulasan negatif, yang menunjukkan persepsi Hasil ini dapat digunakan sebagai dasar evaluasi untuk

meningkatkan pengalaman pengguna dan kualitas layanan.

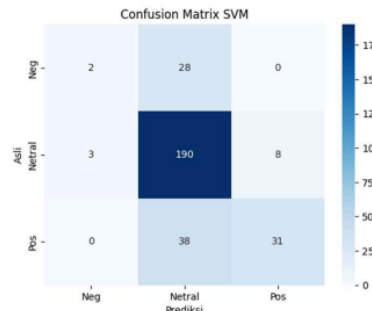
4.2 Distribusi Rating Ulasan



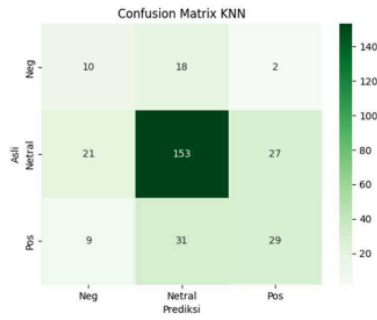
Gambar 7. Rating atau Score

Gambar Distribusi Rating Ulasan MyXL menunjukkan bahwa rating 1 mendominasi dengan sekitar 550 ulasan, diikuti oleh rating 5 dengan 160 ulasan, dan rating 2, 3, dan 4 masing-masing memiliki jumlah yang relatif lebih sedikit dan tersebar secara merata. Dominasi rating rendah menunjukkan bahwa aplikasi MyXL memiliki tingkat ketidakpuasan pengguna yang cukup tinggi, meskipun masih ada beberapa pengguna yang memberikan rating yang sangat baik.

4.3 Confusion Matrixs SVM dan Confusion Matrixs KNN



Gambar 8. Visualisasi Confusion Matrixs SVM



Gambar 9. Visualisasi Confusion Matrixs SVM

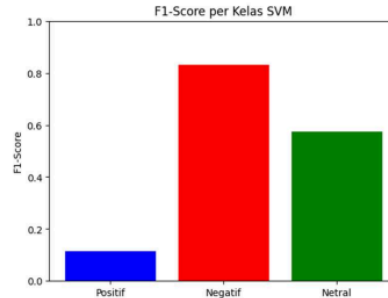
43
Confusion Matrix KNN menunjukkan bahwa model mampu mengklasifikasikan sentimen netral dengan benar sebanyak 153 data, sedangkan klasifikasi benar untuk sentimen negatif dan positif masing-masing sebanyak 10 dan 29 data. Namun, masih terdapat kesalahan klasifikasi yang cukup tinggi, terutama pada sentimen negatif dan positif yang sering diprediksi sebagai netral.

38
Confusion Matrix SVM menunjukkan performa yang lebih baik dalam mengklasifikasikan sentimen netral, dengan 190 data terklasifikasi secara benar. Akan tetapi, SVM juga mengalami kesalahan klasifikasi pada sentimen negatif dan positif, yang sebagian besar diprediksi sebagai netral, menunjukkan kecenderungan bias terhadap kelas netral.

Ketika kedua model dibandingkan, terlihat bahwa SVM lebih baik dalam klasifikasi daripada KNN, terutama dalam menemukan sentimen netral. Namun, kedua model masih menghadapi masalah dalam membedakan sentimen positif dan negatif, sehingga diperlukan pengembangan tambahan untuk meningkatkan kinerja klasifikasi, seperti penyeimbangan data atau optimasi parameter.

4.4 Report dan F1 Score SVM

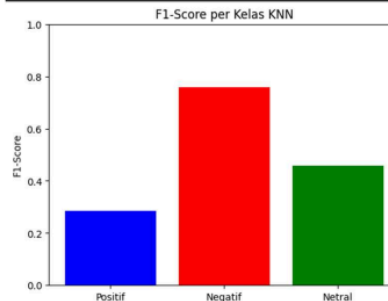
```
Report SVM
precision recall f1-score support
0 0.400000 0.066667 0.114286 30.000000
1 0.742188 0.945274 0.831510 201.000000
2 0.798872 0.449275 0.574874 69.000000
accuracy 0.743333 0.743333 0.743333 0.743333
macro avg 0.645686 0.487072 0.586623 300.000000
weighted avg 0.720086 0.743333 0.780577 300.000000
```



Gambar 10. F1 Score SVM

4.5 Report dan F1 Score KNN

```
Report KNN
precision recall f1-score support
0 0.250000 0.333333 0.285714 30.00
1 0.757426 0.761194 0.759305 201.00
2 0.580000 0.420290 0.455693 69.00
accuracy 0.640000 0.640000 0.640000 0.64
macro avg 0.592475 0.504939 0.500571 300.00
weighted avg 0.647475 0.640000 0.642345 300.00
```



Gambar 11. F1 Score KNN

Tabel 1. Perbandingan F1 Score SVM dan F1 Score KNN

Metode	Positif	Netral	Negatif	Avg
SVM	0.114	0.832	0.574	0.507
KNN	0.286	0.759	0.457	0.501

4.6 Analisis F1-Score per Kelas

a. Kelas Positif

Meskipun KNN memiliki skor F1 yang lebih tinggi pada kelas Positif, nilai keduanya relatif rendah, menunjukkan bahwa kedua model menghadapi masalah dalam mengklasifikasikan data Positif yang termasuk kelas minor. Karena SVM cenderung bias terhadap kelas yang

lebih dominan, mereka tidak berhasil pada kelas ini.

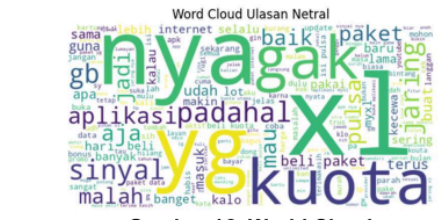
b. Kelas Netral

Dengan dominasi kelas netral dalam dataset, SVM lebih stabil dalam memprediksi kelas netral. SVM juga berhasil pada kelas netral dengan skor F1 yang lebih tinggi daripada KNN.

c. Kelas Negatif

SVM unggul dibandingkan KNN pada kelas negatif, meskipun skor F1 tidak setinggi kelas netral. Ini menunjukkan bahwa kelas minor masih sulit diklasifikasikan karena jumlah datanya lebih sedikit daripada kelas mayoritas.

Hasil evaluasi menunjukkan bahwa SVM dan KNN berbeda dalam hal karakteristik performa. SVM lebih akurat dan stabil, terutama dalam mengklasifikasikan kelas mayoritas (Netral), tetapi masih ada kesalahan pada kelas minor. Model ini lebih tahan terhadap fluktuasi data dan ketidakseimbangan kelas. Sebaliknya, KNN lebih baik untuk menangani kelas minor, seperti Positif, karena klasifikasi berdasarkan mayoritas tetangga terdekat membuatnya sensitif terhadap data minor. KNN mudah digunakan dan mudah digunakan, sehingga cocok untuk percobaan awal atau dataset kecil. Namun, kestabilan dan akurasi keseluruhan lebih rendah ketika dihadapkan pada kelas mayoritas. Oleh karena itu, metode klasifikasi harus disesuaikan dengan tujuan analisis. SVM lebih baik untuk dataset dengan ketidakseimbangan kelas, sedangkan KNN adalah alternatif untuk pendekatan sederhana untuk menunjukkan kelas minor.



Gambar 12. World Cloud

Visualisasi Word Cloud untuk Sentimen Netral dibuat untuk memahami kata-kata dominan yang muncul dalam ulasan Netral aplikasi MyXL. Word Cloud ini menampilkan kata-kata yang paling sering digunakan oleh pengguna, tanpa membawa nilai emosional yang jelas (tidak positif maupun negatif).

34
V. PENUTUP
5.1 Kesimpulan

20 Hasil penelitian menunjukkan bahwa Support Vector Machine (SVM) dan K-Nearest Neighbor (KNN) keduanya memiliki kemampuan untuk melakukan analisis sentimen pada ulasan aplikasi MyXL. SVM lebih baik dengan akurasi 0,743 dibandingkan KNN dengan 0,64, dan SVM lebih stabil dan akurat dalam mengklasifikasikan kelas mayoritas (Netral), tetapi masih ada kesalahan pada kelas minor (Positif dan Negatif). KNN juga lebih baik dalam menangani kelas minor, terutama Menurut distribusi sentimen dan analisis Word Cloud, sebagian besar ulasan bersifat Netral. Oleh karena itu, interpretasi sentimen pengguna harus diperhatikan agar hasil analisis mencerminkan persepsi sebenarnya. Secara keseluruhan, penelitian ini menunjukkan bahwa memilih teknik klasifikasi yang tepat sangat penting, terutama untuk dataset dengan distribusi kelas yang tidak seimbang. Hasil penelitian akan membantu pengembang aplikasi MyXL memahami tingkat kepuasan pengguna dan merencanakan perbaikan layanan dan penambahan fitur di masa depan.

5.2 Saran

Agar model menjadi lebih akurat, terutama untuk penelitian mendatang, disarankan untuk menggunakan dataset

yang lebih besar dan seimbang. Ini terutama berlaku untuk kelas minor (positif dan negatif). Untuk meningkatkan kinerja, metode klasifikasi tambahan seperti Random Forest, Naive Bayes, atau deep learning dapat dicoba. Selain itu, metadata ulasan, seperti waktu pembuatan, lokasi, dan jenis perangkat, dapat diperkaya dengan menambah informasi yang digunakan model. Untuk membuat model lebih responsif terhadap berbagai bahasa ulasan pengguna, disarankan agar proses preprocessing teks dioptimalkan serta penyesuaian hyperparameter SVM dan KNN. Terakhir, pengembang aplikasi dapat menggunakan sistem pemantauan sentimen secara real-time untuk membantu mereka membuat keputusan cepat.

DAFTAR PUSTAKA

- [1] I. Aida Sapitri, M. Fikry, F. Sains dan Teknologi, and U. Islam Negeri Sultan Syarif Kasim Riau, "PENGLASIFIKASIAN SENTIMEN ULASAN APLIKASI WHATSAPP PADA GOOGLE PLAY STORE MENGGUNAKAN SUPPORT VECTOR MACHINE," *Jurnal TEKINKOM*, vol. 6, no. 1, 2023, doi: 10.37600/tekinkom.v6i1.1773.
- [2] D. Diandra Audiansyah, D. Eka Ratnawati, and B. Trias Hanggara, "Analisis Sentimen Aplikasi MyXL menggunakan Metode Support Vector Machine berdasarkan Ulasan Pengguna di Google Play Store," 2022. [Online]. Available: <http://j-ptiik.ub.ac.id>
- [3] I. Garuda, P. Alva, E. Zuliarso, and U. Stikubank, "ANALISIS SENTIMEN BERDASARKAN ULASAN GOOGLE PLAY STORE PADA APLIKASI DANA MENGGUNAKAN METODE SUPPORT VECTOR MACHINE (SVM) SENTIMENT ANALYSIS BASED ON GOOGLE PLAY STORE REVIEWS OF THE DANA APPLICATION USING THE SUPPORT VECTOR MACHINE (SVM) METHOD," *Journal of Information Technology and Computer Science (INTECOMS)*, vol. 8, no. 5, 2025.
- [4] S. F. Kadir and A. Fairuzabadi, "Analisis Sentimen Ulasan Shopee di Google Play dengan TF-IDF dan Logistic Regression," *RIGGS: Journal of Artificial Intelligence and Digital Business*, vol. 4, no. 2, pp. 7940–57945, Jul. 2025, doi: 10.31004/riggs.v4i2.2850.
- [5] N. Putri Husain, A. Febriana Syam, and R. Mustikosari, "Analisis Sentimen Ulasan Pengguna Tiktok pada Google Play Store Berbasis TF-IDF dan Support Vector Machine," 2024. [Online]. Available: <https://images.app.goo.gl/hC6494uW637VmYVW9>
- [6] A. Suharman and M. Kamayani Sulaeman, "Analisis Sentimen Pengguna Aplikasi Livin' by Mandiri Menggunakan Metode Support Vector Machine (SVM) dengan Ekstraksi Fitur TF-IDF dan Word2Vec," *Jurnal Pendidikan dan Teknologi Indonesia*, vol. 5, no. 8, pp. 2201–2212, Aug. 2025, doi: 10.52436/1.jpti.941.
- [7] "Analisis Sentimen Pada Ulasan Aplikasi Home Credit dengan Metode SVM dan KNN", [Online]. Available: <https://ejournal.mediaantartika.id/index.php/jka>
- [8] Ramadhan Bagus Adi Nugroho, Mukh Taofik Chulkamdi, and Rizki Dwi Romadhona, "Perbandingan Performa Algoritma Support Vector Machine dan K-Nearest Neighbors terhadap Analisis Sentimen Aplikasi Perpustakaan Digital," *Jurnal Ilmiah Teknik Informatika dan Komunikasi*, vol. 5, no. 3, pp. 115–126, Aug. 2025, doi: 10.55606/juitik.v5i3.1522.
- [9] D. Fajar Nawulansih, N. Ceisa Santi, and I. Aristia Sa'ida, "Analisis Sentimen

Ulasan Aplikasi DANA di Google Play Store: Penerapan Support Vector Machine dan Synthetic Minority Over-sampling Technique," *Jurnal Pendidikan dan Teknologi Indonesia*, vol. 5, no. 9, pp. 2660–2671, Sep. 2025, doi: 10.52436/1.jpti.1053.

- [10] Y. Y. Zandroto, A. V. Vitianingsih, A. L. Maukar, N. K. Hikmawati, and R. Hamidan, "Sentiment Analysis of BCA Mobile App Reviews Using K-Nearest Neighbour and Support Vector Machine Algorithm," *Indonesian Journal of Artificial Intelligence and Data Mining*, vol. 8, no. 2, p. 448, Aug. 2025, doi: 10.24014/ijaidm.v8i2.37773.
- [11] R. B. Afandi, T. Fajar Nurdiansyah, A. N. Ramadhani, and A. P. Sari, "JIP (Jurnal Informatika Polinema) IMPLEMENTASI SUPPORT VECTOR MACHINE UNTUK ANALISIS SENTIMEN APLIKASI 'MPStore-Super App UMKM'".
- [12] A. Syaifudin Agil Rafsanjani *et al.*, "Sistemasi: Jurnal Sistem Informasi Analisis Sentimen Ulasan Pengguna Aplikasi KitaLulus pada Google Play Store dengan menggunakan algoritma Support Vector Machine (SVM) Sentiment Analysis of User Reviews of the KitaLulus Application on Google Play Store using the Support Vector Machine (SVM) Algorithm." [Online]. Available: <http://sistemasi.ftik.unisi.ac.id>

Perbandingan Metode Support Vector Machine

ORIGINALITY REPORT

24%

SIMILARITY INDEX

PRIMARY SOURCES

1	jurnal.umuslim.ac.id Internet	56 words — 2%
2	ipm2kpe.or.id Internet	42 words — 1%
3	www.geeksforgeeks.org Internet	35 words — 1%
4	Andriani Nurian, Muhammad Samsul Ma'arif, Indira Nur Amalia, Chaerur Rozikin. "ANALISIS SENTIMEN PENGGUNA APLIKASI SHOPEE PADA SITUS GOOGLE PLAY MENGGUNAKAN NAIVE BAYES CLASSIFIER", Jurnal Informatika dan Teknik Elektro Terapan, 2024 Crossref	29 words — 1%
5	Mawadatul Maulidah, Angga Ardiansyah, Suleman Suleman, Lina Putri Gemilang, Novi Fitria Indriarti. "Analisis Sentimen Pada Ulasan Aplikasi Superbank Dengan Metode Support Vector Machine Dan Naive Bayes", Indonesian Journal on Software Engineering (IJSE), 2024 Crossref	28 words — 1%
6	Septian Cahyo Widiyanto, Bebas Widada, Kumaratih Sandradewi, Dwi Remawati. "IMPLEMENTASI METODE CERTAINTY FAKTOR DALAM SISTEM PAKAR UNTUK DIAGNOSA KERUSAKAN LAPTOP", Jurnal Teknologi Informasi dan Komunikasi (TIKomSiN), 2025 Crossref	27 words — 1%
7	jurnal.polibatam.ac.id Internet	

25 words — 1%

8 www.upgrad.com
Internet

24 words — 1%

9 www.jetir.org
Internet

22 words — 1%

10 gitlab.sliit.lk
Internet

21 words — 1%

11 j-ptiik.ub.ac.id
Internet

21 words — 1%

12 journal.unj.ac.id
Internet

21 words — 1%

13 codepointtech.com
Internet

20 words — 1%

14 journals.gesociety.org
Internet

18 words — 1%

15 Munawar Amin, Muhammad Hasbi, Bebas Widada, Bramasto Wiryawan Yudanto. "IMPLEMENTASI MONITORING JARINGAN INTERNET PADA MIKROTIK MENGGUNAKAN NOTIFIKASI BOT TELEGRAM DI SMK", Jurnal Teknologi Informasi dan Komunikasi (TIKomSiN), 2025
Crossref

17 words — 1%

16 Saiful Nur Budiman, Sri Lesanti, Erwan. "Analisis Sentimen Berdasarkan Hasil Review Lokasi Google Map Menggunakan Natural Language Toolkit TextBlob dan Naïve Bayes", JAMI: Jurnal Ahli Muda Indonesia, 2024
Crossref

17 words — 1%

17 Muhamad Ali Zaenal Abidin. "Evaluasi Sentimen Ulasan Pengguna CGV Cinemas Indonesia Menggunakan Metode Naïve Bayes dan Support Vector

16 words — < 1%

-
- 18 repository.its.ac.id
Internet 16 words — < 1%
-
- 19 T. Mariprasath, Kumar Reddy Cheepati, Marco Rivera. "Practical Guide to Machine Learning, NLP, and Generative AI: Libraries, Algorithms, and Applications", River Publishers, 2024
Publications 13 words — < 1%
-
- 20 ejournal.uksw.edu
Internet 12 words — < 1%
-
- 21 sistemasi.ftik.unisi.ac.id
Internet 12 words — < 1%
-
- 22 www.iotpe.com
Internet 12 words — < 1%
-
- 23 Muhammad Rizky Julianto, Yuma Akbar, Tri Wahyudi. "Analisis Sentimen Respon Publik Terhadap Program Internet Gratis di Platform X Melalui Pendekatan Algoritma Naïve Bayes", Jurnal Indonesia : Manajemen Informatika dan Komunikasi, 2024
Crossref 11 words — < 1%
-
- 24 journal.eng.unila.ac.id
Internet 11 words — < 1%
-
- 25 repository.uin-suska.ac.id
Internet 11 words — < 1%
-
- 26 technologic.polytechnic.astra.ac.id
Internet 11 words — < 1%
-
- 27 Rahayu Noveandini, Maria Sri Wulandari, Farhan Rasyad. "Penerapan Model LSTM pada Analisis 10 words — < 1%

Sentimen Ulasan Pengguna Aplikasi Shopee Google Play Store",
JURNAL FASILKOM, 2025

Crossref

28 journal.aira.or.id 10 words — < 1%
Internet

29 repository.unsulbar.ac.id 10 words — < 1%
Internet

30 Alfi Syahri, Angraini Angraini, Fitriani Muttakin.
"Sentimen Analisis Pada Ulasan Aplikasi Ajaib Di
Google Play Store Dengan Algoritma Support Vector Machine",
INOVTEK Polbeng - Seri Informatika, 2024
Crossref

31 Mohamad Syafri Tuloli, Titin Seh Kinanti, Lanto
Ningrayati Amali. "Perbandingan Algoritma C4.5,
Naive Bayes, dan K- Nearest Neighbors untuk Prediksi Penyakit
Jantung", Jambura Journal of Informatics, 2025
Crossref

32 Reza Zulfiqri, Betha Nurina Sari, Tesa Nur Padilah.
"ANALISIS SENTIMEN ULASAN PENGGUNA
APLIKASI MEDIA SOSIAL INSTAGRAM PADA SITUS GOOGLE
PLAY STORE MENGGUNAKAN NAÏVE BAYES CLASSIFIER", Jurnal
Informatika dan Teknik Elektro Terapan, 2024
Crossref

33 Rihadatul 'Aisy, Tiara Devina Putri, Meilani Audi
Kustanti, Muhajirul Fajri, Ceysha Diva Ratu
Pramudya. "Classification of Fish Species Using the K-Nearest
Neighbor (KNN) Method: A Case Study at Bintan Center Market,
Tanjungpinang", Jurnal Sustainable: Jurnal Hasil Penelitian dan
Industri Terapan, 2024
Crossref

34 docplayer.info 9 words — < 1%
Internet

-
- 35 download.bibis.ir Internet 9 words — < 1%
-
- 36 e-journal.hamzanwadi.ac.id Internet 9 words — < 1%
-
- 37 ejurnal.methodist.ac.id Internet 9 words — < 1%
-
- 38 jurnal.kdi.or.id Internet 9 words — < 1%
-
- 39 www.researchgate.net Internet 9 words — < 1%
-
- 40 Adika Surya Perdana, Istian Kriya Almanfaluti, Alshaf Pebrianggara. "Klasifikasi Ulasan Pelanggan pada Aplikasi MitraShopee Menggunakan Metode Support Vector Machine dan Naïve Bayes", Teknologi, 2025 Crossref 8 words — < 1%
-
- 41 Andriani Nurian. "ANALISIS SENTIMEN ULASAN PENGGUNA APLIKASI GOOGLE PLAY MENGGUNAKAN NAÏVE BAYES", Jurnal Informatika dan Teknik Elektro Terapan, 2023 Crossref 8 words — < 1%
-
- 42 Dany Pratmanto, Fandhilah Fandhilah, Rousyati Rousyati, Sopian Aji. "Evaluasi Kinerja Naive Bayes dan K-Nearest Neighbors pada Analisis Sentimen Ulasan Aplikasi SnackVideo", Indonesian Journal on Software Engineering (IJSE), 2025 Crossref 8 words — < 1%
-
- 43 Ivan Aditya Nugraha, Irving Vitra Papatungan. "Analisis Sentimen Video Review Starlink Indonesia Menggunakan Pendekatan Lexicon Dictionary Bahasa Indonesia Dan Algoritma Random Forest", Innovative: Journal Of Social Science Research, 2025 Crossref 8 words — < 1%

44	Readel Lantang Beryl Mangkey, Gladly Caren Rorimpandey, SONDY Campvid Kumajas. "Analisis Prostitusi Online Pada Aplikasi Michat Menggunakan Algoritma Naïve Bayes Dan Framework NIST", Jurnal Minfo Polgan, 2025 Crossref	8 words — < 1%
45	ejournal.itn.ac.id Internet	8 words — < 1%
46	etheses.uin-malang.ac.id Internet	8 words — < 1%
47	garuda.ristekbrin.go.id Internet	8 words — < 1%
48	jurnal.stkipppgritulungagung.ac.id Internet	8 words — < 1%
49	jutif.if.unsoed.ac.id Internet	8 words — < 1%
50	kc.umn.ac.id Internet	8 words — < 1%
51	ojs.ninetyjournal.com Internet	8 words — < 1%
52	repository.upnvj.ac.id Internet	8 words — < 1%
53	teknologiliterasi.blogspot.com Internet	8 words — < 1%
54	www.coursehero.com Internet	8 words — < 1%
55	Muhammad Rizki Syafapri, Elin Haerani, Iwan Iskandar, Liza Afriyanti. "Klasifikasi sentimen terhadap larangan pernikahan beda agama menggunakan	7 words — < 1%

metode Naive Bayes Classifier", Jurnal CoSciTech (Computer Science and Information Technology), 2024

Crossref

56 repository.ub.ac.id 7 words — < 1%
Internet

57 Hendry Cipta Husada, Adi Suryaputra Paramita. 6 words — < 1%
"Analisis Sentimen Pada Maskapai Penerbangan di Platform Twitter Menggunakan Algoritma Support Vector Machine (SVM)", Teknika, 2021

Crossref

58 Mohammad Bayu Anggara. "Comparison of Naive Bayes and SVM Methods in Sentiment Analysis of User Reviews on the RSUD AL IHSAN Mobile Application", Competitive, 2025 6 words — < 1%

Crossref

59 Nuri Cahyono, Anggista Oktavia Praneswara. 6 words — < 1%
"Analisis Sentimen Ulasan Aplikasi TikTok Shop Seller Center di Google Playstore Menggunakan Algoritma Naive Bayes", Indonesian Journal of Computer Science, 2023

Crossref

EXCLUDE QUOTES ON

EXCLUDE SOURCES OFF

EXCLUDE BIBLIOGRAPHY ON

EXCLUDE MATCHES OFF