

# Analisis Sentimen Ulasan Pengguna LinkedIn di Google Playstore Menggunakan Metode Support Vector Machine (SVM)

Parasian Silitonga<sup>1</sup>, Doni El Rezen Purba<sup>2</sup>, Ferri Ojak Immanuel Pardede<sup>3</sup>

<sup>1,2</sup> Universitas Katolik Santo Thomas Medan, Jl. Setia Budi No. 479 F  
Tanjung Sari, Medan, Indonesia

<sup>3</sup> Universitas HKBP Nommensen Pematangsiantar, Jl. Sangnawaluh, No. 4, Pematangsiantar  
e-mail: \* parasianirene@gmail.com

## Abstrak

Penelitian ini dilakukan untuk melihat sentimen ulasan pengguna terhadap aplikasi LinkedIn di google play store. Yang dimana analisis ini dilakukan untuk melihat polaritas tanggapan pengguna terhadap aplikasi LinkedIn berupa tanggapan positif dan negatif. Dimana untuk mengolahnya digunakan metode Support Vector machine, yang klasifikasinya menghasilkan nilai precision untuk kelas “negatif” adalah 92%, dan nilai precision untuk kelas “positif” adalah 86%. Untuk nilai recall kelas “negatif” adalah 90% dan untuk recall kelas “positif” adalah 90%. Untuk nilai f1-score untuk kelas “negatif” adalah 91% dan untuk kelas “positif” adalah 88%. Dengan nilai accuracy adalah 90%.

**Kata kunci :** Support Vector Machine, LinkedIn, Data Mining, Analisis, Sentimen

## Abstract

This research was conducted to see the sentiment of user reviews of the LinkedIn application on the Google Play store. Where this analysis is carried out to see the polarity of user responses to the LinkedIn application in the form of positive and negative responses. Where to process it the Support Vector machine method is used, whose classification results in a precision value for the "negative" class is 92%, and the precision value for the "positive" class is 86%. The recall value for the "negative" class is 90% and the recall value for the "positive" class is 90%. The f1-score value for the "negative" class is 91% and for the "positive" class is 88%. With an accuracy value of 90%..

**Keywords :** Support Vector Machine, LinkedIn, Penambangan Data, Analisis, Sentimen

## 1. PENDAHULUAN

Pada era digital saat ini, masyarakat tidak lagi diharuskan untuk mencari atau menemukan lowongan kerja dari koran, selebaran, mulut ke mulut, dan lain-lain. Pada era sekarang, masyarakat dapat dengan mudah mencari lowongan pekerjaan melalui sosial media maupun melalui aplikasi seperti LinkedIn.

LinkedIn adalah jaringan profesional online terbesar di dunia. Anda bisa menggunakan LinkedIn untuk menemukan lowongan dan kesempatan magang, terhubung dan memperkuat hubungan profesional, dan mempelajari keahlian yang dibutuhkan demi kesuksesan karier. Anda dapat mengakses LinkedIn dari desktop, aplikasi seluler LinkedIn, atau browser web seluler. Profil LinkedIn yang lengkap dapat mendekatkan Anda dengan berbagai peluang, dengan menampilkan cerita profesional yang unik melalui pengalaman, keahlian, dan pendidikan Anda. Juga dapat menggunakan LinkedIn untuk menyelenggarakan acara offline, bergabung dengan grup, menulis artikel, memposting foto dan video, dan banyak lagi [1].

Berdasarkan data dari google play store tanggal 13 November 2023, LinkedIn telah di download lebih dari 1M+ kali dan memiliki rating ulasan pengguna 4,1. Pada google play store tercatat bahwa ada 2.655.267 ulasan pengguna di kolom komentar terhadap pengalaman mereka menggunakan aplikasi LinkedIn. Ulasan pengguna terdiri dari nilai bintang kepuasan dan komentar teks tertulis yang menceritakan pengalaman mereka.

Adapun review pengguna di google play store meliputi opini positif dan negatif, yang berisikan keluhan, kritik, pujian, maupun saran. Berdasarkan hal tersebut, maka akan dilakukan analisis sentimen ulasan pengguna terhadap aplikasi LinkedIn menggunakan metode Support Vector Machine (SVM).

Analisis sentimen adalah sebuah teknik atau cara yang digunakan untuk mengidentifikasi bagaimana sebuah sentimen diekspresikan menggunakan teks dan bagaimana sentimen tersebut bisa

dikategorikan sebagai sentimen positif maupun sentimen negatif. Analisis sentimen merupakan sebuah proses untuk menentukan sentimen atau opini dari seseorang yang diwujudkan dalam bentuk teks dan bisa dikategorikan sebagai sentimen positif atau negatif [2]. Oleh karena itu, penelitian ini bertujuan untuk menganalisis ulasan pengguna LinkedIn di Google Play Store menggunakan metode Support Vector Machine (SVM).

## 2. METODE PENELITIAN

Penelitian ini bertujuan untuk menganalisis ulasan pengguna LinkedIn pada Google Play Store menggunakan metode Support Vector Machine, dengan tahapan penelitian seperti yang ada pada gambar 1.



Gambar 1. Kerangka Pemikiran

### 1. Pengumpulan Data

Pengumpulan data (Data mining) adalah metode yang dapat membantu para penggunanya untuk mengakses data yang besar dalam waktu yang relatif singkat. Dengan kata lain, data mining adalah alat dan aplikasi yang menggunakan analisis statistik pada data melalui suatu proses ekstraksi atau penggalian data dan informasi yang sebelumnya tidak diketahui. Data mining adalah proses mengekstraksi data untuk menemukan informasi terbaru dengan mencari pola atau aturan tertentu dari sejumlah data yang sangat besar, sehingga aktivitas penambangan adalah memeriksa database yang berukuran besar untuk menemukan pola atau bentuk yang baru sehingga berguna dalam proses pengambilan keputusan [3]. Data yang diambil untuk penelitian ini berasal dari ulasan pengguna di google play store mengenai aplikasi LinkedIn. Dengan jumlah data ulasan yang dikumpulkan adalah 2000 data.

### 2. Preprocessing

Preprocessing merupakan langkah mengubah data mentah menjadi format yang dapat dipahami adalah bagian dari teknik penambangan data. Masalah seperti data noisy, redundansi, dan kehilangan nilai dapat diatasi dengan pra-pemrosesan data [4]. Cleaning, case folding, tokenizing, stopword, stemming, dan formalisasi adalah beberapa bagian dari tahap ini menurut [5].

1. Cleaning: Proses membersihkan data dari elemen yang tidak penting, seperti kesalahan tata bahasa, konjungsi, tanda baca, dan konten asing.
2. Case folding: metode yang mengubah huruf besar pada kata atau kalimat menjadi huruf kecil.
3. Tokenizing: proses membagi atau memecah frase kata-kata yang awalnya terbentuk menjadi kata-kata individual.
4. Stopword: Seperti kata sambung dan kata ganti orang, istilah-istilah yang tidak memiliki data yang diperlukan akan dihapus selama proses ini.
5. Stemming: proses dengan setiap kata bebas dari imbuhan.

### 3. Term Weighting

Term weighting adalah sebuah metode pembobotan kata (term) untuk memberikan sebuah bobot atau nilai untuk kata (term) yang terkandung dalam sebuah dokumen. Bobot nilai ini menjadi ukuran besarnya jumlah dan tingkat kontribusi sebuah kata (term) untuk penentuan suatu kelas atau kategori dalam suatu dokumen. Terdapat beberapa metode pembobotan kata (term weighting) diantaranya adalah TF, TF-IDF, WIDF, dan TF-RF [6].

#### 3.1 TF-IDF

TF (Term Frequency) adalah frekuensi dari kemunculan sebuah term dalam dokumen yang bersangkutan. IDF (Inverse Document Frequency) merupakan sebuah perhitungan dari bagaimana term didistribusikan secara luas pada koleksi dokumen yang bersangkutan. Semakin besar jumlah

kemunculan suatu term (TF tinggi) dalam dokumen, semakin besar pula bobotnya atau akan memberikan nilai kesesuaian yang semakin besar [7].

#### 4. Support Vectore Machine (SVM)

Support Vector Machine (SVM) merupakan metode yang sering digunakan pada klasifikasi data berupa text dengan tingkat akurasi yang lebih baik. SVM adalah salah satu metode pada supervised learning yang digunakan untuk proses klasifikasi. Pada model klasifikasi, SVM mempunyai konsep yang lebih kuat dan lebih jelas secara matematis. SVM digunakan untuk mencari hyperplane optimal dengan cara memaksimalkan jarak antar kelas. Hyperplane adalah fungsi yang bisa dipakai untuk memisahkan antar kelas. Pada 2-D fungsi yang dipakai untuk klasifikasi antar kelas disebut line whereas, fungsi yang digunakan untuk klasifikasi antar kelas pada 3-D disebut plane similarity, sedangkan fungsi yang dipakai dalam klasifikasi ruang kelas adalah dimensi yang lebih tinggi dinamakan hyperplane [8].

#### 5. Evaluasi Model

Evaluasi model adalah proses menganalisis kinerja model machine learning dengan menggunakan berbagai metrik/parameter. Evaluasi model sangat penting dalam menilai efektivitas sebuah model selama fase penelitian awal dan juga berperan dalam pemantauan model kedepannya. Hasil dari algoritma pembelajaran perlu dinilai dan dianalisis dengan benar, sehingga dapat mengevaluasi performa algoritma pembelajaran yang berbeda. Performa klasifikasi diwakili oleh nilai skalar dalam metrik yang berbeda seperti accuracy, sensitivity, dan specifity [9]. Penulis menggunakan metrik evaluasi standar untuk mengukur performa model dengan menggunakan metrik True Positive, True Negative, False Positive, dan False Negative dimana Sebuah true positive adalah hasil di mana model dengan benar memprediksi kelas positif. Demikian pula, sebuah true negative adalah hasil di mana model dengan benar memprediksi kelas negatif. Sebuah false positive adalah hasil di mana model dengan salah memprediksi kelas positif. Dan sebuah false negative adalah hasil di mana model dengan salah memprediksi kelas negatif. Dengan menggunakan metrik - metrik tersebut penulis dapat mengkalkulasikan akurasi, recall, presisi, dan F1-score menggunakan persamaan 1, 2, 3, dan 4 [24].

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$Recall = \frac{TP}{TP + FN} \quad (2)$$

$$Precision = \frac{TP}{TP + FP} \quad (3)$$

$$F1 - Score = \frac{2 * TP}{2 * TP + FP + FN} \quad (4)$$

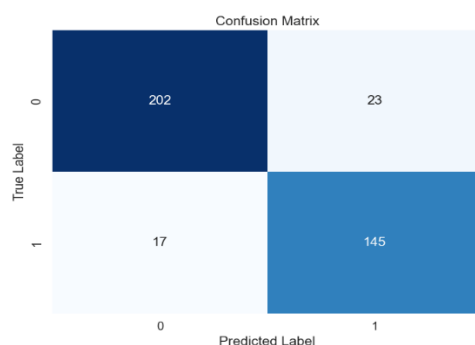
Pada persamaan (1) yaitu accuracy seberapa sering model / mesin membuat prediksi yang benar. Persamaan (2) recall menunjukkan perbandingan antara True Positive (TP) dengan banyaknya data yang sebenarnya positif sedangkan precision yang ditunjukkan pada persamaan (3) mengukur seberapa banyak prediksi mesin positif yang sebenarnya positif. Terakhir, F1 score yang ditunjukkan pada persamaan (4) menggabungkan precision dan recall dalam satu metrik, F1 score sangat penting untuk menentukan performa model machine learning [10].

### 3. HASIL DAN PEMBAHASAN

Pada bagian ini akan membahas mengenai hasil analisis sentimen dengan data yang diperoleh melalui scraping data ulasan pada google play store mengenai aplikasi LinkedIn. Data yang diperoleh ada sebanyak 2000 data ulasan, maka data ulasan yang telah diperoleh akan melalui tahapan preprocessing, yang dimana beberapa tahapan preprocessing yang dilakukan pada penelitian ini yaitu merubah huruf besar menjadi huruf kecil dan menghilangkan seluruh tanda baca pada kalimat (case folding), pemecahan kalimat menjadi kata-kata, frase, atau simbol (tokenization), membersihkan sebuah kalimat atau proses mengubah kata berimbuhan menjadi kata dasar (Stemming). Dari hasil preprocessing maka didapatkan hasil data yang sudah dibersihkan ada sebanyak 1933 data ulasan. Maka dari 1933 data ulasan tersebut akan dibagi menjadi 70% data uji dan 30% data test, yang dimana data

uji ada sebanyak 1546 data, dan data test ada sebanyak 387 data. Dengan pembelahan data tersebut maka akan dilakukan proses TF-IDF maupun klasifikasi SVM, dan menghasilkan keluaran sebagai berikut:

Hasil Confusion Matrix



Gambar 2. Confusion Matrix

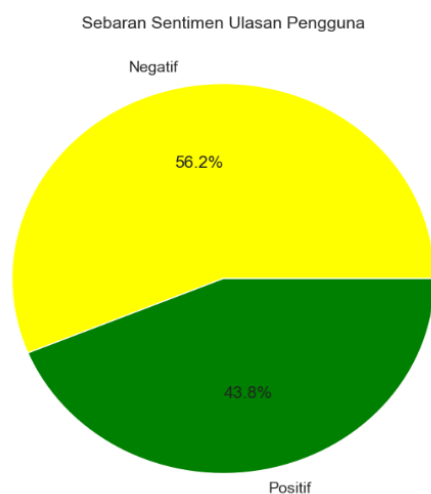
### Evaluasi Model

Pada Evaluasi model akan memperlihatkan hasil dari proses klasifikasi Support Vectore Machine, yang dimana akan memperlihatkan nilai dari accuracy, precision, recall, dan fl-score.

Table 1. Klasifikasi SVM

	precision	recall	f1-score
Negatif	0.92	0.90	0.91
Positif	0.86	0.90	0.88
Accuracy	0.90		
macro avg	0.89	0.90	0.89
weighted avg	0.90	0.90	0.90

Untuk gambaran hasil sebaran ulasan pengguna terhadap aplikasi LinkedIn, akan digambarkan dalam bentuk plot diagram pie yang akan memperlihatkan polaritas nilai positif dan negatif. Dan juga untuk menganalisa frekuensi sebaran kata untuk sentimen positif dan negatif akan menggunakan wordcloud. Dari wordcloud akan menampilkan kata-kata apa saja yang sering muncul dengan index jika ukuran font-size katanya besar itu diartikan bahwa kata tersebut memiliki frekuensi kemunculan yang lebih sering dibandingkan dengan kata yang lain.



Gambar 2. Sebaran Sentimen Ulasan Pengguna



Gambar 3. Wordcloud Negatif



Gambar 3. Wordcloud Positif

#### 4. KESIMPULAN

Berdasarkan hasil pembahasab yang telah didapat, maka dapat diambil kesimpulan bahwa, hasil analisis sentimen ulasan pengguna aplikasi LinkedIn yang dilakukan menggunakan model SVM, nilai precision untuk kelas “negatif” adalah 92%, dan nilai precision untuk kelas “positif” adalah 86%. Untuk nilai recall kelas “negatif” adalah 90% dan untuk recall kelas “positif” adalah 90%. Untuk nilai f1-score untuk kelas “negatif” adalah 91% dan untuk kelas “positif” adalah 88%. Dengan nilai accuracy adalah 90%.

#### DAFTAR PUSTAKA

- [1] M. Fauzia, “Apa Itu LinkedIn dan Apa Fungsinya dalam Mengembangkan Karir?,” Kompas.com, 28 08 2021. [Online]. Available: <https://money.kompas.com/read/2021/08/28/174052326/apa-itu-linkedin-dan-apa-fungsinya-dalam-mengembangkan-karir?page=all>. [Diakses 17 11 2023].
- [2] M. N. Akbar, N. H. Rusydi, M. H. H., N. S. Ramadhanti dan Erfiana, “Sentiment Analysis Terhadap Review Aplikasi Maxim di Google Play Store Menggunakan Support Vector Machine(SVM),” *AGENTS*, vol. II, pp. 1-8, 2022.
- [3] M. L. Siahaan dan Z. Sitomorang, “Algoritma Naïve Bayes Dalam Penentuan Bantuan Renovasi Rumah Di Desa Sialang Buah,” *KAKIFIKOM*, vol. 05, pp. 71-81, 2023.
- [4] B. Gunawan, H. S. Pratiwi dan E. E. Pratama, “Sistem Analisis Sentimen pada Ulasan Produk Menggunakan Metode Naive Bayes,” *JEPIN*, vol. 4, pp. 113-118, 2018.
- [5] R. Nanda, E. Haerani, S. K. Gusti dan S. Ramadhani, “Klasifikasi Berita Menggunakan Metode Support Vector Machine,” *Jurnal Nasional Komputasi dan Teknologi Informasi*, vol. 5, pp. 269-278, 2022.
- [6] A. Deolika, Kusriani dan E. T. Luthfi, “ANALISIS PEMBOBOTAN KATA PADA KLASIFIKASI TEXT MINING,” *Jurnal Teknologi Informasi*, vol. 3, pp. 179-184, 2019.
- [7] A. Z. Rasyida, I. D. Wijaya dan Y. Yunhasnawa, “ANALISIS SENTIMEN KUALITAS LAYANAN ONLINE MARKETPLACE DI INDONESIA MENGGUNAKAN METODE

SUPPORT VECTOR MACHINE,” *SEMINAR INFORMATIKA APLIKATIF POLINEMA (SIAP)*, pp. 70-75, 2020.

- [8] R. W. Pratiwi, S. Febbi, Dairoh, D. I. Af'idah, Q. R. A dan A. G. F, “Analisis Sentimen Pada Review Skincare Female Daily Menggunakan Metode Support Vector Machine (SVM),” *Journal of Informatics, Information System, Software Engineering and Applications*, vol. 1, pp. 40-46, 2021.
- [9] M. I. Amal, E. S. Rahmasita, E. Suryaputra dan N. A. Rakhmawati, “Analisis Klasifikasi Sentimen Terhadap Isu Kebocoran Data Kartu Identitas Ponsel di Twitter,” *Jurnal Teknik Informatika dan Sistem Informasi*, vol. 8, pp. 645-660, 2022.
- [10] S. G. Alexander, A. T. Ananto, I. P. A. P. M. Purnama, B. L. L. Habibullah dan N. A. Rakhmawati, “Analisis Sentimen Opini Masyarakat Indonesia Terhadap Konten Deepfake Tokoh Publik,” *KAKIFIKOM*, vol. 5, pp. 95-102, 2023.