

Algoritma Support Vector Machine untuk Analisis Sentimen Ulasan Aplikasi AXISnet di Google Play Store

Edy Subowo^{1*}, Nur Kholidah², Haifan Tribuwono³

^{1,3} Informatika, Fakultas Teknik dan Ilmu Komputer Universitas Muhammadiyah Pekajangan
Pekalongan

² Ekonomi Syariah, Fakultas Teknik dan Ilmu Komputer Universitas Muhammadiyah Pekajangan
Pekalongan

History Artikel

Diterima : 23 November 2023 Disetujui : 07 Januari 2024 Dipublikasikan : 26 April 2024

Abstract

The AXISnet app, available on the Google Play Store, operates as an internet service provider. Despite amassing 10 million downloads, its rating stands at 4.3. Employing the Knowledge Discovery in Database (KDD) method, this study scrutinizes a dataset encompassing 61,895 reviews of the AXISnet app. Among these, 31,742 reviews lean positively, while 30,153 reviews reflect a negative sentiment. Additionally, the evaluation resulted in an 86% accuracy rate, depicting the model's overall precision in classification tasks. The model exhibited a precision of 87%, highlighting its accuracy in identifying true positives among predicted positives. With a recall of 83%, the model demonstrated its ability to accurately classify actual positive instances. The F1-Score, indicating the harmonic mean between precision and recall, reached 85%, signifying a balanced performance between correctly identifying positive classes and minimizing false positives.

Keywords: *AXISnet; Knowledge Discovery in Database; Sentiment Analysis; Scraping; Support Vector Machine*

Abstrak

Aplikasi AXISnet, yang tersedia di Google Play Store, beroperasi sebagai penyedia layanan internet. Meskipun mengumpulkan 10 juta unduhan, peringkatnya berada di 4,3. Dalam ekosistem toko aplikasi, penilaian pengguna dilengkapi dengan ulasan terperinci. Ulasan ini, seringkali tidak terstruktur, merangkum sentimen dan tingkat kepuasan pengguna, sehingga sangat memengaruhi pilihan calon pengguna. Untuk mengukur sentimen pengguna terhadap aplikasi AXISnet, digunakan analisis sentimen menggunakan algoritma Support Vector Machine. Dengan menggunakan metode Knowledge Discovery in Database (KDD), penelitian ini meneliti kumpulan data yang mencakup 61.895 ulasan aplikasi AXISnet. Diantaranya, 31.742 ulasan cenderung positif, sementara 30.153 ulasan mencerminkan sentimen negatif. Selain itu, evaluasi tersebut menghasilkan tingkat akurasi sebesar 86%, yang menggambarkan presisi model secara keseluruhan dalam tugas klasifikasi. Model ini menunjukkan presisi sebesar 87%, yang menyoroti keakuratannya dalam mengidentifikasi hal-hal positif yang sebenarnya di antara hal-hal positif yang diprediksi. Dengan perolehan sebesar 83%, model tersebut menunjukkan kemampuannya dalam mengklasifikasikan kejadian positif aktual secara akurat. Skor F1, yang menunjukkan rata-rata harmonis antara presisi dan perolehan, mencapai 85%, menandakan kinerja seimbang antara mengidentifikasi kelas positif dengan benar dan meminimalkan positif palsu.

Kata Kunci: *AXISnet; Knowledge Discovery in Database; Sentiment Analysis; Scraping; Support Vector Machine*

How to Cite: Subowo, Edy. dkk (2024). Algoritma Support Vector Machine untuk Analisis Sentimen Ulasan Aplikasi AXISnet di Google Play Store. *KOMPUTEK : Jurnal Teknik Universitas Muhammadiyah Ponorogo*, Vol 8 (1): Halaman 36-42

© 2024 Universitas Muhammadiyah Ponorogo. All rights reserved

ISSN 2614-0985 (Print)
ISSN 2614-0977 (Online)

Pendahuluan

Ulasan aplikasi di Google Play Store memiliki sejumlah dampak penting bagi konsumen yang mempertimbangkan untuk menggunakan aplikasi AXISnet. Ada beberapa alasan mengapa ulasan aplikasi ini penting bagi konsumen. Pertama, kredibilitas dan keandalan produk. Ulasan memberikan gambaran tentang seberapa andal dan kredibel produk tersebut. Jika banyak pengguna memberikan ulasan positif, itu menandakan bahwa aplikasi AXISnet dianggap dapat diandalkan dan berkualitas tinggi.

Dengan adanya ulasan pada aplikasi memungkinkan calon konsumen melihat pengalaman nyata pengguna lain dengan aplikasi. Mereka bisa mempelajari kelebihan, kekurangan, dan solusi dari sudut pandang pengguna sebelumnya. Pilihan yang lebih baik ulasan membantu calon konsumen membuat keputusan yang lebih baik. Dengan memahami pengalaman orang lain, mereka dapat menilai apakah aplikasi AXISnet sesuai dengan kebutuhan dan preferensi mereka. Keempat, Perbaikan dan Pengembangan Produk Ulasan memberikan umpan balik yang berharga bagi pengembang. Dengan melihat keluhan atau saran dari pengguna, pengembang dapat melakukan perbaikan yang diperlukan untuk meningkatkan kualitas aplikasi.

Ulasan seringkali juga membahas masalah keamanan dan privasi. Ini penting karena pengguna ingin memastikan bahwa data pribadi mereka aman ketika menggunakan aplikasi internet service provider. Pertimbangan harga dan nilai ulasan sering kali juga memberikan gambaran tentang nilai yang diberikan oleh layanan berdasarkan harga yang dibayarkan. Ini membantu konsumen menilai sejauh mana layanan tersebut sepadan dengan biayanya. Perbandingan dengan alternatif juga seringkali dibahas pada ulasan sebuah aplikasi. Hal ini membantu konsumen memilih produk yang paling sesuai dengan kebutuhan mereka.

Secara keseluruhan, ulasan di Google Play Store sangat penting bagi konsumen yang ingin menggunakan aplikasi internet service provider seperti AXISnet karena memberikan wawasan yang jujur dan nyata tentang kualitas, keamanan, nilai, dan pengalaman pengguna yang dapat membantu mereka membuat keputusan yang lebih baik sebelum mengunduh dan menggunakan aplikasi tersebut.

Penilaian pengguna terhadap aplikasi di Google Play Store sering kali disertai dengan ulasan yang berisi opini, menjadi pertimbangan penting bagi calon pengguna (Hasanah dkk., 2024). Ulasan ini seringkali bersifat tidak terstruktur dan sangat banyak, memerlukan teknik analisis untuk mengekstrak informasi yang berguna (Fransiska & Irham Gufroni, 2020). Salah satu teknik yang diterapkan dalam kasus ini adalah analisis sentimen, suatu metode untuk memproses teks guna mendapatkan informasi yang tersembunyi di dalamnya (Aditiya dkk., 2022).

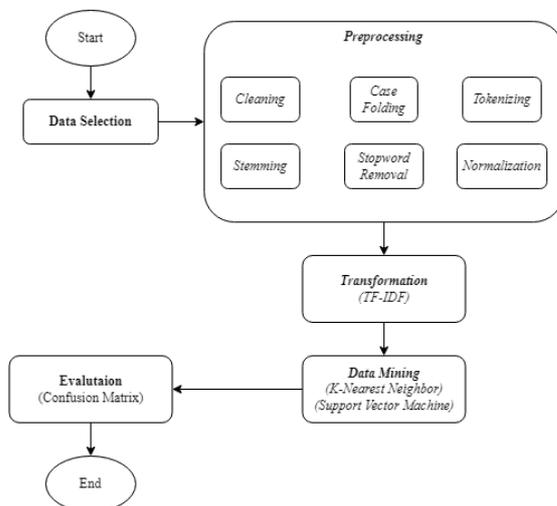
Analisis sentimen menggunakan beberapa algoritma seperti K-Nearest Neighbors (KNN), Naive Bayes, dan Support Vector Machines (SVM). Menurut literatur, SVM seringkali memberikan akurasi tertinggi dibandingkan algoritma lainnya. Misalnya pada penelitian sebelumnya, SVM memiliki akurasi sebesar 81,46%, sedangkan Naive Bayes hanya mampu mencapai 75,41% pada analisis sentimen ulasan bahasa Indonesia di Google Play Store (Ilmawan & Mude, 2020). Dalam penelitian lain yang membandingkan beberapa algoritma, SVM mencapai akurasi rata-rata 96,43% dalam analisis sentimen gadget (Iskandar & Nataliani, 2021). SVM memiliki fungsi untuk mentransformasikan data ke dalam ruang dimensi yang lebih tinggi menggunakan kernel, seperti Linear, Radial Basis Function (RBF), Sigmoid, dan Polynomial (Zuriel & Fahrurrozi, 2021). Pada penelitian lain, kernel RBF menghasilkan akurasi terbaik hingga

84,37% dalam analisis sentimen pada maskapai penerbangan (Husada & Paramita, 2021)

Oleh karena itu, penelitian ini bertujuan untuk mengukur tingkat akurasi dengan menggunakan kernel linear dan melakukan evaluasi dengan metode k-fold cross-validation pada algoritma SVM dalam menganalisis sentimen ulasan aplikasi AXISnet di Google Play Store. Tujuan utamanya adalah untuk memperoleh pemahaman yang lebih baik terkait sentimen pengguna terhadap aplikasi AXISnet serta mengevaluasi kinerja kernel linear dalam konteks analisis sentimen pada platform tersebut

Metode

Penelitian ini memanfaatkan pendekatan *Knowledge Discovery in Database* (KDD) karena unggul dalam mengidentifikasi pola yang terstruktur dari data yang kompleks, memudahkan pemahaman terhadap informasi yang terkandung dalam data (Balahadia dkk., 2021). Gambar 1 mendeskripsikan langkah-langkah penelitian yang mengadopsi pendekatan KDD.



Gambar 1 Alur Penelitian

Alur penelitian pada gambar 1 dijelaskan sebagai berikut.

2.1. Data Selection

Seleksi data merupakan tahap pertama KDD. Fase ini melibatkan pengumpulan data, seleksi, dan pelabelan.

Data dikumpulkan dari website Google Play Store pada tanggal 4 Desember 2021 hingga 4 Desember 2023.

2.2. Preprocessing

Pada tahap preprocessing ini, data mentah yang dikumpulkan diolah menjadi data yang dapat digunakan pada tahap berikutnya. Tahap preprocessing terdiri dari enam proses yaitu a. Pembersihan menghapus atribut yang tidak memberikan kontribusi signifikan terhadap klasifikasi, seperti: Contoh: tanda baca, spasi, emoji. SM Melipat huruf besar dan kecil mengubah semua karakter dalam dokumen menjadi huruf kecil. c. Tokenisasi adalah langkah membagi kalimat menjadi beberapa bagian kata. d. Normalisasi merupakan upaya untuk memperbaiki kesalahan kata, misalnya salah ejaan, sehingga kata yang memiliki makna serupa menjadi setara. e. Penghapusan kata berhenti melibatkan penghapusan kata-kata yang dianggap tidak memiliki dampak signifikan dalam sebuah kalimat. f. Stemming adalah proses mengubah suatu kata menjadi bentuk dasarnya, sehingga memberikan kontribusi terhadap konsistensi dan generalisasinya.

2.3. Transformasi

Tahap transformasi adalah tahap dimana data diubah ke dalam format yang dapat diproses oleh tahap data mining. Pada tahap ini data dibagi menjadi 80 data untuk pelatihan dan 20 data untuk pengujian. Pembobotan kata kemudian dilakukan dengan menggunakan Term Frekuensi-Inverse Document Frekuensi (TF-IDF). Ini berguna untuk mengubah data teks menjadi vektor bobot. Perhitungan metode TF-IDF dinyatakan d.

$$wt = TF_{t,d} \times IDF_t = TF_{t,d} \times \log \frac{N}{DF_t}$$

2.4. Data Mining

Pada tahap data mining akan dilakukan proses klasifikasi sentimen pada data ulasan

dengan menggunakan algoritma SVM dengan kernel linear dipilih sebagai model untuk analisis sentimen. Model ini melibatkan pembelajaran dari dataset yang telah diproses untuk memprediksi sentimen ulasan pengguna. Berikut merupakan persamaan dari kernel linear:

$$K(x_i, x) = xix$$

2.5. Evaluation

Tahap evaluasi akan menggunakan metode *k-fold cross-validation* untuk menguji performa model dengan hasil evaluasi berupa nilai *accuracy*, *precision*, *recall*, dan *f1-score*. Evaluasi dilakukan menggunakan metrik akurasi dan *confusion matrix* untuk memahami sejauh mana model dapat memprediksi sentimen ulasan dengan tepat.

Hasil dan Diskusi

Data ulasan AXISnet dikumpulkan melalui teknik *scraping* dari tanggal 4 Desember 2021 hingga 4 Desember 2023 menggunakan pustaka *google play scraper*, menghasilkan sebanyak 337.870 entri.

Dataset yang terhimpun ini terdiri dari lima atribut utama: *reviewId*, *userName*, *at*, *content*, dan *score*. Untuk proses selanjutnya, hanya atribut *content* yang dipilih untuk digunakan. Selanjutnya, dilakukan proses penandaan atau pelabelan terhadap ulasan berdasarkan skor yang diberikan. Ulasan dengan skor ≤ 3 bintang dikategorikan sebagai ulasan negatif, sedangkan ulasan dengan skor ≥ 4 bintang dikategorikan sebagai ulasan positif. Proses pelabelan ini menghasilkan distribusi data sebanyak 202.522 ulasan positif dan 135.348 ulasan negatif. Detail mengenai tahapan pemilihan data ini dapat dilihat pada Gambar 2.

3.1 Preprocessing

Pada fase ini, dilakukan penyesuaian data agar memenuhi syarat dan dapat diproses pada langkah-langkah berikutnya. Hasil dari enam proses dalam tahapan *preprocessing* dapat diamati pada Tabel 1 sebagai contoh dari hasil yang telah diolah.

	Ulasan	Sentimen
0	jospokoke	Positif
1	saya sangat suka dengan anda apakah anda mau j...	Positif
2	mantapnya	Negatif
3	suka	Negatif
4	Aplikasi apan ini, aplikasi Axis saya tidak bi...	Negatif
...
337865	gajelas makin kesini apk nya	Negatif
337866	Ini kenapa ya saya beli kuota 1 GB 1 hari di a...	Negatif
337867	lag bgt jaringan mu bro, kadang ilang" sendiri...	Negatif
337868	Ko masukin AIGO error ?? Ada apakah Axis net?!...	Negatif
337869	Kenapa kouta nya gak bisa di masukin lol	Negatif

337870 rows × 2 columns

Gambar 2 Hasil Data Selection

Tabel 1 Contoh Hasil Preprocessing

Proses	Hasil
Data	Semuanya jadi mudah pake AXISnet
Ulasan	bagus bangett terimakasih AXISnet
Cleaning	Semuanya jadi mudah pake AXISnet bagus bangett terimakasih AXISnet
Case Folding	semuanya jadi mudah pake axisnet bagus bangett terimakasih axisnet
Tokenizing	['semuanya', 'jadi', 'mudah', 'pake', 'axisnet', 'bagus', 'bangett', 'terimakasih', 'axisnet']
Normaliza tion	['semuanya', 'jadi', 'mudah', 'pakai', 'axisnet', 'bagus', 'banget', 'terima', 'kasih', 'axisnet']
Stopword Removal	['mudah', 'pakai', 'bagus', 'banget', 'terima', 'kasih,']
Stemming	['mudah', 'pakai', 'bagus', 'banget', 'terima', 'kasih,']

3.2 Transformation

Tahap transformasi dimulai dengan memisahkan dataset menjadi dua bagian, yakni *data training* dan *data testing*. Proses pemisahan *dataset* ini dilakukan menggunakan fungsi *train_test_split* dengan pengaturan parameter *random_state=42*. Perintah ini penting dalam tahap transformasi data karena memungkinkan pengelompokan data yang lebih terstruktur untuk keperluan latihan dan pengujian model secara efisien. Tabel 2 menunjukkan hasil pembagian data.



Gambar 5 Word Cloud Ulasan Negatif

Dari visualisasi pada Gambar 4 dan 5, terlihat bahwa kata-kata yang sering kali muncul dalam ulasan positif mencakup "bagus", "mantap", "bonus", dan "gratis". Sementara pada ulasan negatif, kata-kata yang sering muncul meliputi "kuota", "sinyal", "paket" dan "jaring".

Kesimpulan

Penelitian ini menegaskan bahwa melalui penerapan teknik preprocessing yang cermat, termasuk eliminasi stopword, koreksi penulisan, dan penerapan stemming, bersamaan dengan penggunaan model SVM berbasis kernel linear, berhasil menghasilkan model yang mampu memprediksi sentimen ulasan pengguna pada aplikasi AXISnet dengan tingkat akurasi sebesar 86%. Evaluasi lebih lanjut menunjukkan nilai precision sebesar 87%, recall sebesar 83%, dan F1-Score sebesar 85%. Hasil evaluasi ini memberikan konfirmasi bahwa model memberikan hasil yang dapat diandalkan dalam menganalisis sentimen pengguna, memberikan kontribusi berharga untuk pemahaman yang lebih mendalam terhadap preferensi pengguna. Implementasi teknik ini dapat membantu dalam merinci sentimen pengguna, memberikan wawasan yang berharga untuk pengembangan dan perbaikan aplikasi di masa depan.

Referensi

[1] Aditiya, P., Enri, U., & Maulana, I. (2022). Analisis Sentimen Ulasan Pengguna Aplikasi Myim3 Pada Situs Google Play Menggunakan Support Vector Machine. *JURIKOM (Jurnal Riset Komputer)*, 9(4), 1020. <https://doi.org/10.30865/jurikom.v9i4.4673>

[2] Balahadia, F., Vinluan, A., Gonzales, D., &

Ballera, M. (2021). Application of Spatiotemporal Analysis and Knowledge Discovery for Databases in the Bureau of Fire Protection as Incident Report System: Tool for Improving Fire Services. *International Journal of Computing Sciences Research*, 5(1), 519–533. <https://doi.org/10.25147/ijcsr.2017.001.1.56>

[3] Fransiska, S., & Irham Gufroni, A. (2020). Sentiment Analysis Provider by.U on Google Play Store Reviews with TF-IDF and Support Vector Machine (SVM) Method. *Scientific Journal of Informatics*, 7(2), 2407–7658.

[4] Hasanah, A. N., & Sari, B. N. (2024). ANALISIS SENTIMEN ULASAN PENGGUNA APLIKASI JASA OJEK ONLINE MAXIM PADA GOOGLE PLAY DENGAN METODE NAÏVE BAYES CLASSIFIER. *Jurnal Informatika Dan Teknik Elektro Terapan*, 12(1), 90–96. <https://doi.org/10.23960/jitet.v12i1.3628>

[5] Husada, H. C., & Paramita, A. S. (2021). Analisis Sentimen Pada Maskapai Penerbangan di Platform Twitter Menggunakan Algoritma Support Vector Machine (SVM). *Teknika*, 10(1), 18–26. <https://doi.org/10.34148/teknika.v10i1.311>

[6] Ilmawan, L. B., & Mude, M. A. (2020). Perbandingan Metode Klasifikasi Support Vector Machine dan Naïve Bayes untuk Analisis Sentimen pada Ulasan Tekstual di Google Play Store. *ILKOM Jurnal Ilmiah*, 12(2), 154–161. <https://doi.org/10.33096/ilkom.v12i2.597.154-161>

[7] Iskandar, J. W., & Nataliani, Y. (2021). Perbandingan Naïve Bayes, SVM, dan k-NN untuk Analisis Sentimen Gadget Berbasis Aspek. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 5(6), 1120–1126. <https://doi.org/10.29207/resti.v5i6.3588>

[8] Zuriel, H. P. P., & Fahrurrozi, A. (2021). Implementasi Algoritma Klasifikasi Support Vector Machine Untuk Analisa Sentimen Pengguna Twitter Terhadap Kebijakan Psbb. *Jurnal Ilmiah Informatika Komputer*, 26(2), 149–162. <https://doi.org/10.35760/ik.2021.v26i2>