P-ISSN : 2580-4316
https://doi.org/10.37817/ikraith-informatika.v10i2 E-ISSN : 2654-8054

Analisis Sentimen Ulasan Pengguna Aplikasi KitaLulus Berdasarkan Skala Rating Menggunakan Algoritma Support Vector Machine (SVM)

[1]Gheryyan Washesya Syagara, [2]Setyono, [3]Hilmy Aliy Andra Putra
[1,2,3]Program Studi Ilmu Komputer, Fakultas Ilmu Komputer, Universitas Djuanda, Bogor

E-mail: [1]i.2211099@unida.ac.id, [2]setyono@unida.ac.id, [3]hilmy.aliy@unida.ac.id

ABSTRAK

Aplikasi pencarian kerja KitaLulus telah diunduh lebih dari 10 juta kali dan memperoleh rating rata-rata 4,7 di Google Play Store, namun jumlah ulasan yang sangat besar menyulitkan pemahaman persepsi pengguna secara manual. Penelitian ini bertujuan menerapkan algoritma Support Vector Machine (SVM) untuk mengklasifikasikan sentimen ulasan pengguna KitaLulus ke dalam lima kelas berdasarkan skala rating (1-5). Data dikumpulkan melalui teknik web scraping menggunakan library google-play-scraper, menghasilkan 15.000 ulasan yang setelah proses pembersihan data menjadi 10.536 ulasan. Tahapan penelitian mengikuti kerangka kerja Sample, Explore, Modify, Model, Assess (SEMMA), meliputi text preprocessing (cleaning, case folding, tokenization, normalisasi slang, stopword removal, dan stemming), pembobotan fitur dengan TF-IDF, pemodelan SVM dengan 10-fold cross validation, serta optimasi model melalui hyperparameter tuning menggunakan GridSearchCV pada tiga jenis kernel (Linear, Radial Basis Function, dan Polynomial). Hasil pengujian awal dengan LinearSVC memperoleh akurasi rata-rata sebesar 74,17%. Proses hyperparameter tuning memilih kernel RBF dengan C=1 dan gamma=scale sebagai konfigurasi terbaik, yang ketika dievaluasi pada keseluruhan dataset menggunakan cross_val_predict menghasilkan akurasi 70,10%, dengan F1-Score tertinggi pada kelas rating 5 (0,86) dan performa yang masih rendah pada kelas minoritas rating 2 (0,04) akibat ketidakseimbangan distribusi data. Hasil ini menunjukkan bahwa SVM cukup efektif untuk klasifikasi sentimen multikelas pada data ulasan aplikasi, namun penanganan data tidak seimbang perlu menjadi perhatian pada penelitian selanjutnya.

Kata kunci: *analisis sentimen; Support Vector Machine; TF-IDF; ulasan pengguna; KitaLulus*

ABSTRACT

The KitaLulus job-search application has been downloaded more than 10 million times and holds an average rating of 4.7 on the Google Play Store, yet the large volume of user reviews makes manual interpretation of user perception difficult. This study applies the Support Vector Machine (SVM) algorithm to classify the sentiment of KitaLulus user reviews into five classes based on the 1-5 rating scale. Data were collected through web scraping using the google-play-scraper library, yielding 15,000 reviews that were reduced to 10,536 after data cleaning. The research followed the Sample, Explore, Modify, Model, Assess (SEMMA) framework, comprising text preprocessing (cleaning, case folding, tokenization, slang normalization, stopword removal, and stemming), TF-IDF feature weighting, SVM modeling with 10-fold cross validation, and hyperparameter tuning using GridSearchCV across three kernel types (Linear, Radial Basis Function, and Polynomial). The initial LinearSVC evaluation achieved a mean accuracy of 74.17%. Hyperparameter

P-ISSN : 2580-4316
E-ISSN : 2654-8054 https://doi.org/10.37817/ikraith-informatika.v10i2

tuning selected the RBF kernel with C=1 and gamma=scale as the best configuration, which, when evaluated on the full dataset using cross_val_predict, produced an accuracy of 70.10%, with the highest F1-Score on the rating-5 class (0.86) and considerably lower performance on the minority rating-2 class (0.04) due to class imbalance. The study demonstrates that SVM is reasonably effective for multiclass sentiment classification of application reviews, although handling imbalanced data remains an important consideration for future research.

Keyword: *sentiment analysis; Support Vector Machine; TF-IDF; user reviews; KitaLulus*

1. PENDAHULUAN

Transformasi digital telah mengubah proses rekrutmen dari metode tradisional berbasis berkas fisik menjadi platform daring berbasis mobile (Jatmiko et al, 2024). Aplikasi pencarian kerja seperti JobStreet, Glints, Pintarnya, dan KitaLulus kini banyak digunakan sebagai penghubung antara pencari kerja dan penyedia lapangan kerja karena mampu mempercepat proses rekrutmen dan memperluas jangkauan kandidat.

KitaLulus merupakan salah satu platform pencarian kerja populer di Indonesia yang telah diunduh lebih dari 10 juta kali dan memperoleh rating 4,7 di Google Play Store dari 196 ribu ulasan. Banyaknya ulasan yang diberikan pengguna membuka peluang untuk memahami persepsi publik terhadap aplikasi ini melalui analisis sentimen, yaitu teknik text mining yang mengklasifikasikan opini, emosi, atau penilaian pengguna terhadap suatu topik.

Penelitian terdahulu terkait analisis sentimen aplikasi ketenagakerjaan umumnya menggunakan algoritma Naive Bayes. Pada ulasan aplikasi Glints, model Naive Bayes memperoleh akurasi 87% namun masih lemah dalam mendeteksi sentimen negatif (Dinda & Eka, 2025). Untuk mengatasi keterbatasan tersebut, penelitian ini menggunakan algoritma Support Vector Machine (SVM) yang dikenal unggul dalam menangani data teks berdimensi tinggi. Berbeda dari penelitian-penelitian sebelumnya yang umumnya hanya menggunakan satu jenis kernel SVM tanpa cross-validation, penelitian ini mengeksplorasi tiga jenis kernel (Linear, Radial Basis Function, dan Polynomial) disertai hyperparameter tuning dan 10-fold cross validation untuk memperoleh model yang lebih optimal dan evaluasi yang lebih stabil.

Berdasarkan uraian tersebut, penelitian ini bertujuan untuk: (1) menerapkan algoritma SVM dalam analisis sentimen ulasan pengguna aplikasi KitaLulus, (2) mengetahui tingkat akurasi dan performa algoritma SVM berdasarkan metrik accuracy, precision, recall, dan F1-Score, serta (3) mengidentifikasi persepsi pengguna terhadap aplikasi KitaLulus berdasarkan hasil klasifikasi sentimen pada rating 1-5. Objek penelitian dibatasi pada 15.000 ulasan berbahasa Indonesia terbaru dari Google Play Store.

2. LANDASAN TEORI

Analisis Sentimen dan TF-IDF

Analisis sentimen adalah cabang penelitian dalam text mining yang berfokus pada klasifikasi dokumen teks berdasarkan opini, emosi, atau penilaian seseorang terhadap suatu topik (Ali Putra *et al*., 2024). Salah satu pendekatan yang umum digunakan adalah machine learning based approach, yaitu algoritma yang dilatih dari data berlabel dan cenderung memiliki akurasi lebih baik dibandingkan pendekatan berbasis leksikon. Sebelum data teks dapat diproses oleh algoritma machine learning, data perlu

direpresentasikan secara numerik. Term Frequency-Inverse Document Frequency (TF-IDF) adalah metode pembobotan kata yang mengukur seberapa penting suatu kata dalam sebuah dokumen relatif terhadap keseluruhan korpus, dengan memberikan bobot lebih tinggi pada kata yang sering muncul pada satu dokumen tetapi jarang ditemukan pada dokumen lain (Hanani, 2023).

Support Vector Machine (SVM)

Support Vector Machine (SVM) merupakan metode supervised learning yang digunakan untuk analisis data pada tugas klasifikasi dan regresi (Purnamasari *et al*., 2023). Pada kasus klasifikasi, SVM bekerja dengan mencari hyperplane optimal yang memisahkan data ke dalam kelas-kelas berbeda dengan margin maksimum. Titik data yang berada paling dekat dengan hyperplane disebut support vector. Untuk data yang tidak dapat dipisahkan secara linear, SVM memanfaatkan fungsi kernel seperti Linear, Radial Basis Function (RBF), dan Polynomial guna memetakan data ke ruang fitur berdimensi lebih tinggi sehingga proses pemisahan antar kelas menjadi lebih mudah.

Confusion Matrix

Confusion matrix digunakan untuk mengevaluasi hasil klasifikasi dengan membandingkan label aktual dan label hasil prediksi model. Dari matriks ini dapat dihitung metrik accuracy, precision, recall, dan F1-Score. F1-Score, sebagai rata-rata harmonik antara precision dan recall, dinilai lebih representatif pada kondisi distribusi data yang tidak seimbang (*imbalanced* data) dibandingkan metrik akurasi semata.

3. METODOLOGI

Penelitian ini menggunakan pendekatan Sample, Explore, Modify, Model, Assess (SEMMA), yaitu kerangka kerja data mining yang memfasilitasi proses eksplorasi, pembersihan, dan pemodelan data secara sistematis. Objek penelitian adalah ulasan pengguna aplikasi KitaLulus yang diperoleh dari Google Play Store, dengan lokasi penelitian bersifat virtual karena seluruh data dikumpulkan secara daring.

Pengumpulan Data dan Pelabelan

Data primer berupa ulasan pengguna dikumpulkan melalui teknik web scraping menggunakan library google-play-scraper pada Python, menghasilkan 15.000 ulasan terbaru dalam rentang Mei 2023 hingga Juni 2026. Setelah proses penghapusan data duplikat dan baris kosong, jumlah data yang digunakan pada tahap analisis menjadi 10.536 ulasan. Data selanjutnya diberi label berdasarkan skor rating, menghasilkan lima kelas kategori (kelas 1 sampai 5).

Pre-processing dan Pembobotan TF-IDF

Tahap pre-processing meliputi text cleaning, case folding, tokenization, normalisasi slang menggunakan kamus colloquial-indonesian-lexicon, stopword removal, dan stemming menggunakan library Sastrawi. Data teks bersih selanjutnya diubah menjadi representasi numerik menggunakan TF-IDF melalui TfidfVectorizer dari library scikit-learn dengan pembatasan 4.000 fitur (max_features), menghasilkan matriks berdimensi 10.536 x 4.000 sebagai data input (X), dengan skor rating 1-5 sebagai data target (y).

Pemodelan, Tuning, dan Evaluasi

Algoritma SVM digunakan dengan pendekatan Support Vector Classification (SVC) karena penelitian ini merupakan klasifikasi multikelas. Model diuji menggunakan pendekatan LinearSVC dengan 10-fold cross validation sebagai evaluasi awal, kemudian dioptimasi melalui hyperparameter tuning menggunakan GridSearchCV yang menguji kombinasi kernel Linear, Radial Basis Function (RBF), dan Polynomial dengan parameter C, gamma, dan degree pada sampel 3.000 data. Model terbaik hasil tuning dievaluasi ulang pada seluruh 10.536 data menggunakan metode cross_val_predict dengan 10-fold cross

P-ISSN : 2580-4316
E-ISSN : 2654-8054 https://doi.org/10.37817/ikraith-informatika.v10i2

validation, sehingga setiap data memperoleh prediksi dari model yang tidak pernah melihat data tersebut saat pelatihan. Evaluasi dilakukan menggunakan confusion matrix serta metrik accuracy, precision, recall, dan F1-Score.

4. HASIL DAN PEMBAHASAN

Eksplorasi Data

Hasil scraping menghasilkan dataset dengan empat kolom utama (username, content, score, at). Setelah pembersihan duplikat, jumlah data menjadi 10.536 ulasan dengan skor rata-rata (mean) 4,33 dan standar deviasi 1,28. Ringkasan statistik deskriptif ditampilkan pada Tabel 1, sedangkan distribusi kelas rating ditampilkan pada Tabel 2.

Tabel 1. Ringkasan statistik data

Statistik	Nilai
Jumlah data	10.536
Rata-rata (Mean)	4,33
Std. Deviasi	1,28
Skor Min	1
Median	5
Skor Max	5

Tabel 2. Distribusi data berdasarkan skor rating

Skor	Jumlah	Persentase
1	995	9,4%
2	337	3,2%
3	553	5,2%
4	947	9,0%
5	7.704	73,1%
Total	10.536	100%

Distribusi data pada Tabel 2 menunjukkan dominasi kuat pada kelas rating 5 (73,1%), sedangkan kelas rating 2 hanya memiliki 3,2% dari total data. Kondisi ini mengindikasikan adanya ketidakseimbangan data (imbalanced data) yang berpotensi memengaruhi performa klasifikasi, khususnya pada kelas minoritas.

Pre-processing dan Pembobotan TF-IDF

Tahap pre-processing berhasil mengubah teks ulasan mentah menjadi bentuk yang lebih terstruktur. Sebagai contoh, kata slang seperti “bgt” dinormalisasi menjadi “banget”, dan kata berimbuhan seperti “membantu” diubah menjadi bentuk dasarnya “bantu” melalui proses stemming. Hasil transformasi TF-IDF menghasilkan matriks berdimensi 10.536 x 4.000, dengan kata-kata seperti “sangat” (bobot rata-rata 0,0804), “kerja” (0,0748), dan “aplikasi” (0,0734) memiliki bobot tertinggi karena kemunculannya yang spesifik dan relevan dengan konteks ulasan aplikasi pencarian kerja.

Evaluasi Awal Model SVM

Pengujian awal menggunakan LinearSVC dengan 10-fold cross validation pada seluruh 10.536 data menghasilkan akurasi rata-rata sebesar 74,17% (standar deviasi 0,0129), dengan rentang akurasi antar fold antara 72,68% hingga 76,92%. Nilai standar deviasi yang kecil menunjukkan performa model yang stabil pada setiap pembagian data, mengindikasikan kemampuan generalisasi yang cukup baik.

Hyperparameter Tuning

Proses hyperparameter tuning menggunakan GridSearchCV pada sampel 3.000 data menguji tiga kernel dengan variasi parameter C, gamma, dan degree. Ringkasan skor terbaik tiap kernel ditampilkan pada Tabel 3.

Tabel 3. Perbandingan skor terbaik antar kernel SVM

Kernel	Parameter	Skor
Linear	C=1	63,1%
RBF	C=1, gamma=scale	68,1%*
Polynomial	C=1, deg=2	67,2%

*Kernel RBF (C=1, gamma=scale) terpilih sebagai konfigurasi terbaik (best_params_) oleh GridSearchCV. Skor ini lebih rendah dibandingkan hasil evaluasi awal LinearSVC pada seluruh data (74,17%) karena tiga faktor: (1) tuning hanya menggunakan sampel 3.000 data dari 10.536 data; (2) perbedaan implementasi, yaitu LinearSVC yang dioptimasi khusus untuk data berdimensi tinggi dibandingkan SVC yang lebih umum; serta (3) karakteristik data TF-IDF yang sparse dan berdimensi tinggi, sehingga kernel linear sudah cukup

ekspresif untuk memisahkan kelas tanpa transformasi non-linear tambahan.

Evaluasi Model Akhir

Model terbaik dari hasil tuning (kernel RBF, C=1, gamma=scale) dievaluasi ulang pada keseluruhan 10.536 data menggunakan cross_val_predict dengan 10-fold cross validation, menghasilkan akurasi keseluruhan sebesar 70,10%. Rincian performa tiap kelas ditampilkan pada Tabel 4.

Tabel 4. Laporan klasifikasi per kelas (10-fold CV)

Kelas	Prec.	Rec.	F1	N
1	0,51	0,69	0,59	995
2	0,07	0,03	0,04	337
3	0,17	0,20	0,18	553
4	0,17	0,23	0,20	947
5	0,89	0,83	0,86	7704
Macro	0,36	0,40	0,37	10536
W.avg	0,73	0,70	0,71	10536

Model menunjukkan performa terbaik pada kelas rating 5 dengan F1-Score 0,86, didukung oleh jumlah data terbesar (7.704 sampel). Sebaliknya, kelas rating 2 memiliki F1-Score terendah (0,04) akibat jumlah sampel yang sangat terbatas (337 data). Perbedaan besar antara macro average F1-Score (0,37) dan weighted average F1-Score (0,71) secara langsung mencerminkan tingkat ketidakseimbangan distribusi data dalam penelitian ini.

Analisis confusion matrix menunjukkan bahwa kesalahan klasifikasi terbesar terjadi antara kelas rating 4 dan 5, yaitu 489 kasus rating 4 diprediksi sebagai rating 5, dan 827 kasus rating 5 diprediksi sebagai rating 4. Hal ini mengindikasikan tumpang tindih semantik (semantic overlap) pada ulasan dengan tingkat kepuasan yang berdekatan, karena pengguna cenderung menggunakan kosakata positif yang serupa pada kedua rating tersebut. Sementara itu, dominasi prediksi benar pada kelas rating 5 (6.359 dari 7.704 data, atau 83%) menunjukkan model cukup mampu mengenali sentimen positif yang kuat.

Evaluasi overfitting dan underfitting melalui perbandingan skor data latih dan data uji pada setiap fold menunjukkan rata-rata akurasi data latih sebesar 88,15% dan data uji sebesar 73,80%, dengan selisih 14,35%. Selisih ini mengindikasikan adanya overfitting ringan yang wajar terjadi pada dataset tidak seimbang, namun belum tergolong overfitting yang parah karena akurasi data uji tetap konsisten dan berada pada tingkat yang cukup baik di setiap fold (standar deviasi hanya 0,0115).

Visualisasi Word Cloud

Visualisasi word cloud pada setiap kategori rating memperlihatkan pola kosakata yang berbeda secara konsisten dengan hasil klasifikasi. Pada rating 1, kata-kata yang dominan seperti “enggak jelas” dan “tipu” mencerminkan ketidakpuasan pengguna. Rating 2 dan 3 didominasi kata-kata netral hingga negatif ringan seperti “belum” dan “enggak”. Sebaliknya, rating 4 dan 5 didominasi kata-kata positif seperti “bagus”, “sangat”, “bantu”, dan “mudah”, yang menunjukkan tingkat kepuasan pengguna yang tinggi terhadap fitur pencarian kerja pada aplikasi KitaLulus. Pola ini memperkuat kebenaran hasil klasifikasi SVM, di mana kosakata yang digunakan pengguna secara konsisten mencerminkan skor rating yang diberikan.

5. KESIMPULAN

Penelitian ini berhasil menerapkan algoritma Support Vector Machine (SVM) untuk mengklasifikasikan sentimen 10.536 ulasan pengguna aplikasi KitaLulus ke dalam lima kelas berdasarkan skala rating, mengikuti kerangka kerja SEMMA secara sistematis mulai dari pengumpulan data, pre-processing, pembobotan TF-IDF, pemodelan, hyperparameter tuning, hingga evaluasi akhir. Model LinearSVC pada evaluasi awal memperoleh akurasi rata-rata 74,17%, sedangkan model terbaik hasil tuning (kernel RBF, C=1, gamma=scale) memperoleh akurasi 70,10% saat dievaluasi pada keseluruhan data. Model menunjukkan performa sangat baik pada kelas mayoritas rating 5 (F1-Score 0,86), namun performa masih

rendah pada kelas minoritas rating 2 dan 3, terutama akibat ketidakseimbangan distribusi data yang signifikan (rating 5 mendominasi 73,1% data) serta tumpang tindih semantik antar kelas rating yang berdekatan. Temuan ini menegaskan bahwa SVM cukup efektif untuk klasifikasi sentimen multikelas pada data ulasan aplikasi, dengan catatan bahwa penanganan data tidak seimbang, misalnya melalui teknik oversampling atau pembobotan kelas, perlu menjadi perhatian utama untuk meningkatkan performa pada kelas minoritas.

DAFTAR PUSTAKA

A Ilemobayo, J., Durodola, O., Alade, O., J Awotunde, O., T Olanrewaju, A., Falana, O., Ogungbire, A., Osinuga, A., Ogunbiyi, D., Ifeanyi, A., E Odezuligbo, I., & E Edu, O. (2024). Hyperparameter tuning in machine learning: A comprehensive review. Journal of Engineering Research and Reports, 26(6), 388-395.

Agung, A., Daniswara, A., Kadek, I., & Nuryana, D. (2023). Data preprocessing pola pada penilaian mahasiswa program profesi guru. Journal of Informatics and Computer Science, 5.

Anggraini, D., et al. (2025). Analisis sentimen pengguna terhadap aplikasi Glints di Google Play Store menggunakan metode Naive Bayes. JMA, 3(5).

Azis, H., Purnawansyah, P., Fattah, F., & Putri, I. P. (2020). Performa klasifikasi KNN dan cross validation pada data pasien pengidap penyakit jantung. ILKOM Jurnal Ilmiah, 12(2), 81-86.

Idris, I. S. K., Mustofa, Y. A., & Salihi, I. A. (2023). Analisis sentimen terhadap penggunaan aplikasi Shopee menggunakan algoritma Support Vector Machine (SVM). Journal of Information Science, 5(6), 33.

Karal, O. (2020). Performance comparison of different kernel functions in SVM for different k value in k-fold cross-validation. Proceedings - 2020 Innovations in Intelligent Systems and Applications Conference, ASYU 2020.

Medhat, W., Hassan, A., & Korashy, H. (2014). Sentiment analysis algorithms and applications: A survey. Ain Shams Engineering Journal, 5(4), 1093-1113.

Mitchell, R. (2018). Web scraping with Python: Collecting more data from the modern web. O'Reilly Media.

Mujtahid, F. A., Zaman, B., & Aji, G. G. (2026). Analisis klasifikasi sentimen neobank: Perbandingan konfigurasi N-Gram pada TF-IDF menggunakan Naive Bayes dan SVM. Jurnal TIN, 6(12), 2375-2383.

Nur'aini, S. (2023). Transformasi era digital: Peluang menggali pekerjaan dan tantangan terhadap meningkatnya pengangguran.

Purnamasari, D., Bayu, A., Desy, A., Fanka, W. A. P., Reza, A., Safrila, M., Yanda, O. N., & Hidayati, U. (2023). Pengantar metode analisis sentimen.

Wainer, J., & Cawley, G. (2017). Empirical evaluation of resampling procedures for optimising SVM hyperparameters. Journal of Machine Learning Research, 18.