

Algoritma Support Vector Machine dalam Analisis Sentimen Aplikasi Primer League Pada Google Play Store

Syahrudin¹, Edy Subowo², Haedar Al Fath³, Jufri Nur Haryanto⁴

^{1,2,3,4} Informatika Fakultas Teknik dan Ilmu Komputer Universitas Muhammadiyah Pekalongan
syahrudin@umpp.ac.id, edy.subowo@gmail.com, alfathhaedar8@gmail.com³, jufrihurharyanto@gmail.com⁴

Abstract

This application not only provides the latest information about the English Premier football league but also provides features that enrich the user experience. These reviews provide a clear picture of the app's reliability, quality, and functionality, helping consumers make more informed decisions. From user feedback, we can evaluate the developer's responsiveness to customer input, potential improvements, and progress on the features offered. Therefore, reviews on the Play Store are not only a source of reference, but also a valuable guide for consumers looking for the best English Premier football experience through this app. This research produces an accuracy of 79.67% using the SVM algorithm. These results show that the SVM algorithm can quite accurately analyze the sentiment of reviews of football applications such as the Primer League on the Google Play Store. Analysis of the sentiment of reviews of the Primer League application can provide valuable information for application developers. This information can be used to improve the quality of the application, thereby increasing user satisfaction.

Keyword: Sentiment Analysis, App Reviews, Premier League, Google Play Store, Support Vector Machine

Abstraksi

Aplikasi ini tidak hanya memberikan informasi terkini tentang liga sepakbola Primer Inggris tetapi juga menyediakan fitur-fitur yang memperkaya pengalaman pengguna. Ulasan-ulasan ini memberikan gambaran yang jelas mengenai kehandalan, kualitas, dan fungsionalitas aplikasi, sehingga membantu konsumen membuat keputusan yang lebih terinformasi. Dari feedback pengguna, kita dapat mengevaluasi responsifitas pengembang terhadap masukan pelanggan, potensi perbaikan, dan kemajuan fitur-fitur yang ditawarkan. Oleh karena itu, ulasan di Play Store bukan hanya sebagai sumber referensi, tetapi juga sebagai panduan yang berharga bagi konsumen yang mencari pengalaman sepakbola Primer Inggris yang terbaik melalui aplikasi ini. Penelitian ini menghasilkan akurasi sebesar 79.67% menggunakan algoritma SVM. Hasil ini menunjukkan bahwa algoritma SVM dapat dengan cukup akurat menganalisis sentimen ulasan aplikasi bola seperti Primer League di Google Play Store. Analisis sentimen ulasan aplikasi Primer League dapat memberikan informasi yang berharga bagi pengembang aplikasi. Informasi tersebut dapat digunakan untuk meningkatkan kualitas aplikasi, sehingga dapat meningkatkan kepuasan pengguna.

Kunci : Analisis Sentimen, Ulasan Aplikasi, Primer League, Google Play Store, Support Vector Machine

1. PENDAHULUAN

Dalam era digital yang terus berkembang, aplikasi mobile menjadi sahabat setia bagi pecinta sepakbola, terutama bagi mereka yang memantau setiap detik aksi liga terbaik dunia, seperti Primer League. Salah satu sarana yang memungkinkan pengalaman penggemar sepakbola semakin mendalam adalah aplikasi Primer League yang tersedia di platform Google Play Store.

Aplikasi ini diharapkan tidak hanya menjadi sumber informasi terkini seputar pertandingan, statistik pemain, dan berita terbaru, tetapi juga menyediakan fitur-fitur canggih yang dapat memenuhi kebutuhan pengguna yang semakin tinggi. Dalam ulasan ini, kita akan membahas secara rinci tentang keunggulan, kehandalan, dan keunikan aplikasi Primer League, serta bagaimana aplikasi ini memberikan nilai tambah yang signifikan bagi penggemar sepakbola. Dengan demikian, ulasan ini diharapkan dapat membantu konsumen membuat keputusan cerdas dalam memilih aplikasi yang sesuai dengan keinginan dan kebutuhan mereka dalam menjalani pengalaman sepakbola digital yang tak terlupakan. Dalam menghadapi persaingan ketat di dunia aplikasi mobile, aplikasi Primer League tidak hanya fokus pada informasi dasar tentang pertandingan dan pemain. Melainkan, aplikasi ini juga menghadirkan berbagai fitur inovatif yang menjadikannya unggul di antara aplikasi sejenis. Salah satu fitur unggulan adalah kemampuannya untuk menyajikan konten multimedia berkualitas tinggi, seperti cuplikan gol, wawancara eksklusif, dan highlight menarik.

Selain itu, interaktivitas pengguna dengan aplikasi ini ditingkatkan melalui fitur komentar, polling, dan forum diskusi. Hal ini menciptakan komunitas pengguna yang aktif dan berpartisipasi dalam berbagai diskusi seputar dunia sepakbola. Aplikasi ini juga terus memperbarui diri dengan menambahkan fitur-fitur baru berdasarkan umpan balik pengguna, sehingga dapat terus memenuhi ekspektasi yang terus berkembang.

Keamanan dan kenyamanan pengguna menjadi fokus utama, dengan penyediaan opsi personalisasi untuk notifikasi dan tata letak antarmuka. Selain itu, integrasi dengan platform media sosial memungkinkan pengguna untuk berbagi pengalaman mereka secara langsung dengan teman-teman mereka.

2. LANDASAN TEORI

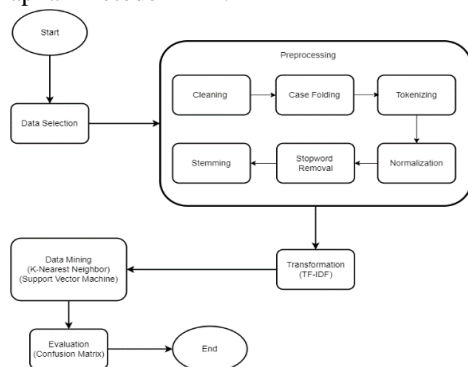
Pemberian rating aplikasi di Google Play Store diikuti dengan ulasan dari para pengguna terhadap aplikasi tersebut. Ulasan tersebut mengandung opini dari para pengguna mengenai aplikasi tersebut dan calon pengguna melihat ulasan dari sebuah aplikasi sebagai pertimbangan sebelum memutuskan untuk menggunakan aplikasi tersebut (Saputra et al., 2019). Ulasan di Google Play Store ini cukup banyak dan tidak terstruktur sehingga diperlukan suatu teknik untuk mengetahui bagaimana ulasan pengguna terhadap aplikasi tersebut (Fransiska & Irham Gufroni, 2020). Oleh karena itu, diperlukan analisis sentimen pada data ulasan. Analisis sentimen adalah teknik untuk mengolah data tekstual untuk memperoleh informasi pada teks (Wahyudi & Kusumawardana, 2021).

Terdapat beberapa algoritma yang dapat diterapkan dalam analisis sentimen seperti K-Nearest Neighbor (KNN), Naïve Bayes, dan Support Vector Machine (SVM). Berdasarkan beberapa referensi, SVM menghasilkan akurasi terbaik dalam melakukan analisis sentimen dibandingkan algoritma klasifikasi yang lain. Pada penelitian sebelumnya SVM memiliki tingkat akurasi yang lebih tinggi sebesar 81,46% dibandingkan Naïve Bayes sebesar 75,41% saat diterapkan untuk analisis sentimen ulasan Bahasa Indonesia di Google Play Store (Ilmawan & Mude, 2020). Pada penelitian lain yang membandingkan SVM, Naïve Bayes, dan KNN untuk analisis sentimen gadget, SVM menghasilkan rata-rata akurasi terbaik sebesar 96,43% (Iskandar & Nataliani, 2021). SVM mempunyai fungsi yang bisa mentransformasikan data ke ruang dimensi yang lebih tinggi yaitu ruang kernel yang disebut dengan fungsi kernel trick sehingga data dapat dipisahkan dengan lebih baik dibandingkan algoritma klasifikasi lainnya (Mukarramah et al., 2021). Beberapa kernel pada SVM yaitu kernel Linear, Radial Basis Function (RBF), Sigmoid, dan Polynomial (Zuriel & Fahrurrozi, 2021). Pada penelitian analisis sentimen ulasan aplikasi Ruangguru dengan membandingkan 3 kernel algoritma SVM diperoleh hasil bahwa kernel linear merupakan kernel terbaik dengan akurasi hingga 89,7% (Irfani et al., 2020). Sedangkan dalam penelitian lainnya akurasi terbaik diperoleh oleh kernel RBF dalam analisis sentimen pada maskapai penerbangan sebesar 84,37% (Husada & Paramita, 2021).

Berdasarkan penjelasan di atas, karena belum ada hasil pasti mengenai akurasi kernel terbaik dalam melakukan analisis sentimen, maka pada penelitian ini akan dilakukan perbandingan akurasi antara kernel Linear, kernel RBF, kernel Sigmoid, dan kernel Polynomial pada algoritma SVM dalam melakukan analisis sentimen ulasan aplikasi PRIMER LEAGUE untuk mengetahui bagaimana sentimen para pengguna terhadap aplikasi PRIMER LEAGUE di Google Play Store dan akurasi kernel terbaik.

3. METODOLOGI PENELITIAN

Penelitian ini menggunakan metode KDD (Knowledge Discovery in Database) karena memiliki keunggulan proses pengidentifikasian pola yang terorganisir dari sekumpulan data yang kompleks, sehingga datanya mudah untuk dipahami (Ramos et al., 2021; Rizki et al., 2021). Gambar 1 menunjukkan alur penelitian yang menerapkan metode KDD.



Gambar 1 Alur Penelitian

Alur penelitian pada gambar 1 dijelaskan sebagai berikut.

1. Data Selection

Data selection merupakan tahapan pertama dalam KDD, pada tahap ini akan dilakukan pengumpulan, seleksi, dan pelabelan data. Data dikumpulkan dari situs Google Play Store pada tanggal 1 Januari 2020 sampai 30 Desember 2023

2. Preprocessing

Pada tahap preprocessing ini, data mentah yang telah terkumpul akan diolah menjadi data yang dapat digunakan pada tahap selanjutnya. Tahap preprocessing terdiri dari 6 proses yaitu:

- Cleaning*, merupakan proses untuk menghilangkan atribut yang tidak berpengaruh terhadap klasifikasi yaitu tanda baca, karakter kosong, dan emoji.
- Case Folding*, merupakan proses untuk mengubah semua huruf yang ada pada dokumen menjadi huruf kecil.
- Tokenizing*, merupakan proses untuk memisahkan kalimat menjadi beberapa bagian kata.
- Normalization*, merupakan proses untuk memperbaiki kesalahan yang ada pada kata seperti ejaan yang salah agar kata yang memiliki makna sama menjadi setara.
- Stopword Removal*, merupakan proses untuk menghilangkan katakata yang dianggap tidak berpengaruh terhadap kalimat.
- Stemming*, merupakan proses untuk mengubah katakata yang ada menjadi bentuk kata dasar.

3. Transformation

Tahap transformation adalah tahap mengubah data menjadi bentuk yang dapat diolah pada tahapan data mining. Pada tahap ini akan dilakukan pembagian data ke dalam 3 skenario terlebih dahulu yaitu skenario 1 (90% data training dan 10% data testing), skenario 2 (80% data training dan 10% data testing), skenario 3 (90% data training dan 10% data testing). Setelah itu akan dilakukan pembobotan kata dengan TF-IDF yang berguna untuk mengubah data berupa teks menjadi vektor bobot. Perhitungan metode TFIDF dinyatakan dalam persamaan sebagai berikut.

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

4. Data Mining

Pada tahap data mining akan dilakukan proses klasifikasi sentimen pada data ulasan dengan menggunakan empat kernel algoritma Support Vector Machine yaitu Kernel Linear, Kernel RBF, Kernel Sigmoid, dan Kernel Polynomial.

5. Evaluation

Tahap evaluasi akan menggunakan confusion matrix untuk mengetahui performansi dari setiap kernel algoritma SVM dengan hasil evaluasi berupa nilai accuracy, precision, recall, dan f1-score.

4. HASIL DAN PEMBAHASAN

1. Data Selection

Pengumpulan data ulasan aplikasi Primer League yang dilakukan dengan teknik scraping pada tanggal 1 Januari 2020 sampai 30 Desember 2023 menggunakan library google play scraper menghasilkan sebanyak 1330 data. Data yang telah terkumpul memiliki 5 atribut yaitu reviewId, userName, at, content, dan score. Selanjutnya dilakukan penyeleksian atribut yang digunakan yaitu hanya content. Kemudian dilakukan pelabelan data ulasan secara manual dan divalidasi oleh ahli Bahasa dan Sastra Indonesia. Pelabelan data tersebut menghasilkan 776 ulasan positif dan 554 ulasan negatif. Hasil pada tahap data selection dapat dilihat pada gambar 2.

	content	score	sentime
0	Kenapa gak bisa dibuka aplikasi punya saya	1	
1	👎👎👎👎 tidak ada live streaming 🤔🤔🤔	1	
2	Apk eror sudah buat team malah suruh buat lagi...	1	
3	Update terlalu sering.. makan memori, bikin la...	1	
4	Jangan di download, aplikasi ga mutu	1	

Gambar 2 Hasil data Selection

2. Preprocessing

Pada tahap ini dilakukan untuk mengubah data agar sesuai dan dapat diolah pada tahapan selanjutnya. Contoh hasil dari 6 proses pada tahap preprocessing yang telah dilakukan dapat dilihat pada tabel 1.

Tabel 1. Contoh Hasil Preprocessing

Proses	Hasil
Data Ulasan	Setelah update apk malah gak bisa login
Cleaning	Setelah update apk malah gak bisa login
Case Folding	setelah update apk malah gak bisa login
Tokenizing	['setelah', 'update', 'apk', 'malah', 'gak', 'bisa', 'login']
Normalization	['setelah', 'update', 'apk', 'malah', 'gak', 'bisa', 'login']
Stopword Removal	['update', 'apk', 'malah', 'gak', 'login']
Stemming	['updat', 'apk', 'malah', 'gak', 'login']

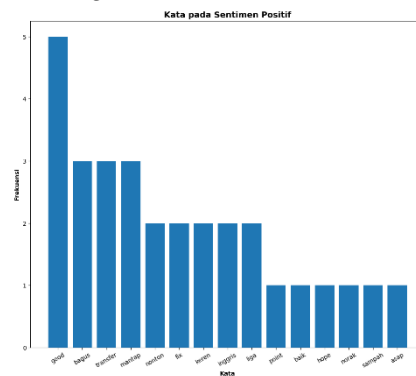
3. Transformation

Tahap transformation ini diawali dengan melakukan pembagian data menjadi data training dan data testing dalam 3 skenario. Pembagian data ini menggunakan fungsi *train test split* dengan parameter *random_state=42*. Tabel 2 menunjukkan hasil pembagian data dalam 3 skenario.

Tabel 2. Hasil Pembagian Data

Skenario Rasio Perbandingan	Data Training	Data Testing
80 :20	1064	266

Setelah data terbagi menjadi data training dan data testing, selanjutnya dilakukan pembobotan kata dengan metode TF-IDF. Gambar 3 merupakan contoh hasil pembobotan kata dengan TF-IDF pada data *training* skenario 1 (90:10)



Gambar 3 Hasil TF-IDF

4. Data Mining

Pada tahap ini dilakukan klasifikasi sentimen dengan tiga skenario pembagian data yang telah dilakukan menggunakan empat kernel pada algoritma *Support Vector Machine*. Hasil klasifikasi dipaparkan pada tabel 3.

Tabel 3. Hasil Klasifikasi SVM

Skenario	Kernel	Accuracy
90 : 20	Linear	79.67%

5. Evaluation

Tahap ini dilakukan evaluasi dengan confusion matrix. Tabel 4 memaparkan hasil perbandingan dari seluruh evaluasi menggunakan *confusion matrix* terhadap setiap skenario dari tahap data mining dengan 4 kernel algoritma *Support Vector Machine*

Tabel 4. Evaluasi Confusion Matrix

Skenario	Kernel	Accuracy	Precision	Recall	F1-score
90 :20	Linear	79.67%	83.33%	71.43%	76.92%

Tabel 4 menunjukkan bahwa hasil evaluasi terbaik diperoleh pada skenario 1 dengan menggunakan kernel *Linear* yang menghasilkan akurasi sebesar 79.67%, *precision* 83.33%, *recall* 71.43%, dan *f1-score* 76.92%. Kata-kata yang sering muncul pada data ulasan divisualisasikan dalam bentuk word cloud. Gambar 4 menunjukkan hasil visualisasi ulasan positif, sedangkan Gambar 5 menunjukkan



Gambar 4 Word Cloud Ulasan Positif



Gambar 5 Word Cloud Ulasan Negatif

Berdasarkan gambar 4 dan 5, katakata yang sering muncul pada ulasan positif yaitu “bagus”, “Good”, dan “mantap”. Sedangkan kata-kata yang sering muncul pada ulasan negatif yaitu “error”, “gabisa”, “lama”.

5. KESIMPULAN DAN SARAN

Berdasarkan hasil penelitian yang kami lakukan, dapat disimpulkan bahwa sentimen pengguna terhadap aplikasi PRIMER LEAGUE terbagi menjadi 776 ulasan positif dan 554 ulasan negatif. Visualisasi kata-kata pada ulasan positif mengungkapkan bahwa kata-kata seperti "bagus", "Good", "mantap", dan "bagus" sering muncul. Sebaliknya, pada ulasan negatif, kata-kata seperti "lama", "gabisa", "error", dan "iklan" menjadi kata-kata yang sering muncul.

Selanjutnya, evaluasi kinerja empat kernel algoritma Support Vector Machine dalam menganalisis sentimen ulasan aplikasi PRIMER LEAGUE dilakukan dengan tiga skenario split data. Hasil terbaik diperoleh pada skenario 1 dengan rasio perbandingan data training sebesar 80% dan data testing sebesar 20%. Dalam skenario ini, akurasi mencapai 79.67%, presisi sebesar 83.33%, recall sebesar 71.43%, dan f1-score sebesar 76.92%.

Temuan ini menunjukkan bahwa algoritma Support Vector Machine efektif dalam menganalisis sentimen ulasan aplikasi PRIMER LEAGUE, terutama pada skenario dengan pembagian data training dan testing sebesar 80% dan 20%. Hasil ini memberikan gambaran yang kuat terkait dengan performa algoritma dalam mengklasifikasikan sentimen pengguna dengan akurasi dan presisi yang tinggi. Adapun saran untuk penelitian selanjutnya yaitu:

1. Meluaskan cakupan penelitian dengan memeriksa sentimen pengguna pada platform lain atau versi aplikasi PRIMER LEAGUE yang berbeda
2. Memperkaya penelitian dengan penggunaan dataset yang lebih luas dan representatif

3. Menganalisis bagaimana sentimen pengguna berubah sebagai respons terhadap perubahan atau pembaruan pada aplikasi PRIMER LEAGUE.

Daftar Pustaka

- Artanto, F. A. (2024). Support Vector Machine Berbasis Particle Swarm Optimization Pada Analisis Sentimen Anggota KPPS. *Jurnal FASILKOM (Teknologi InFormASi Dan ILmu KOMputer)*, 14(1), 75–79. <https://doi.org/https://doi.org/10.37859/jf.v14i1.6795>
- Fathudin, A., Artanto, F. A., Safla, N. A., & Wibowo, D. (2024). Decision Tree Berbasis SMOTE dalam Analisis Sentimen Penggunaan Artificial Intelligence untuk Skripsi. *Remik: Riset Dan E-Jurnal Manajemen Informatika Komputer*, 8(April), 494–505.
- Fathudin, A., Febrianto, M. Y., Artanto, F. A., Hadinata, M. W. N., & Fahlevi, R. (2022). Algoritma Decision Tree C.45 dalam analisa kelulusan mahasiswa Program Studi Manajemen Informatika UMPP. *Jurnal Ilmiah Ilmu Komputer Fakultas Ilmu Komputer Universitas Al Asyariah Mandar*, 8(2), 83–86.
- Fathudin, A., Khambali, A., Artanto, F. A., & Zade, N. A. P. (2023). Implementasi Algoritma Clustering K-Means Dalam Pengelompokan Mahasiswa Studi Kasus (Prodi Manajemen Informatika). *Jurnal Minfo Polgan*, 12(2), 777–783. <https://doi.org/10.33395/jmp.v12i2.12494>
- Fibrianda, M. F., & Bhawiyuga, A. (2018). Analisis Perbandingan Akurasi Deteksi Serangan Pada Jaringan Komputer Dengan Metode Naïve Bayes Dan Support Vector Machine (SVM). *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 2(9), 3112-3123.
- Handayanto, A., Latifa, K., Saputro, N. D., & Waliansyah, R. R. (2019). Analisis dan Penerapan Algoritma Support Vector Machine (SVM) dalam Data Mining untuk Menunjang Strategi Promosi. *JUITA: Jurnal Informatika*, 7(2), 71-79.
- Isnain, A. R., Sakti, A. I., Alita, D., & Marga, N. S. (2021). Sentimen Analisis Publik Terhadap Kebijakan Lockdown Pemerintah Jakarta Menggunakan Algoritma Svm. *Jurnal Data Mining Dan Sistem Informasi*, 2(1), 31-37.
- 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.v12.i2.597.154-161>
- Irfani, F. F., Triyanto, M., Hartanto, A. D., & Kusnawi. (2020). Analisis Sentimen Review Aplikasi Ruangguru Menggunakan Algoritma Support Vector Machine. *JBMI (Jurnal Bisnis, Manajemen, Dan Informatika)*, 16(3), 258–266. <https://doi.org/10.26487/jbmi.v16i3.8607>
- Mukarramah, R., Atmajaya, D., & Ilmawan, L. B. (2021). Performance comparison of support vector machine (SVM) with linear kernel and polynomial kernel for

- multiclass sentiment analysis on twitter. *ILKOM Jurnal Ilmiah*, 13(2), 168–174. <https://doi.org/10.33096/ilkom.v13.i2.851.168-174>
- Octaviani, P. A., Wilandari, Y., & Ispriyanti, D. (2014). Penerapan Metode Klasifikasi Support Vector Machine (SVM) pada Data Akreditasi Sekolah Dasar (SD) di Kabupaten Magelang. *Jurnal Gaussian*, 3(4), 811–820.
- Perwitasari, A. S. (2021). Tumbuh 100%, Mola kini punya lebih dari 1 juta pelanggan berbayar di Indonesia. *Kontan.Co.Id*. <https://industri.kontan.co.id/news/tumbuh-100-mola-kini-punya-lebih-dari-1-juta-pelangganberbayar-di-indonesia>
- Ramos, S., Soares, J., Cembranel, S. S., Tavares, I., Foroozandeh, Z., Vale, Z., & Fernandes, R. (2021). Data mining techniques for electricity customer characterization. *Procedia Computer Science*, 186, 475–488. <https://doi.org/10.1016/j.procs.2021.04.168>
- Rosyadi, I., Artanto, F. A., Rahmawati, S. E., Tri, H., & Joyo, B. (2022). Decision Tree Dalam Analisis Keputusan Pembelian Program Pada Perkumpulan Penggiat Programmer Indonesia. *Jurnal Fasilkom*, XII(III), 141–144.
- Rosyadi, I., Kusumawardani, H. H., Artanto, F. A., Hardani, A. A. A., & Nafilaturrosyidah, F. (2023). Clustering K-Means Dalam Pengelompokan Penjualan Produk Pada RTO Group. *Teknomatika*, 13(02), 55–60. <http://ojs.palcomtech.ac.id/index.php/teknomatika/article/view/618/439>
- Sarbaini, S., Cynthia, E. P., & Arifandy, M. I. (2021). Pengelompokan Diabetic Macular Edema Berbasis Citra Retina Mata Menggunakan Fuzzy Learning Vector Quantization (FLVQ). *SITEKIN: Jurnal Sains, Teknologi dan Industri*, 19(1), 75–80.
- Saputra, S. A., Rosiyadi, D., Gata, W., & Husain, S. M. (2019). Sentiment Analysis Analysis of E-Wallet Sentiments on Google Play Using the Naive Bayes Algorithm Based on Particle Swarm Optimization. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 3(3), 377–382. <https://doi.org/10.29207/resti.v3i3.1118>