



Perbandingan Performa Algoritma *Support Vector Machine* dan *K-Nearest Neighbors* terhadap Analisis Sentimen Aplikasi Perpustakaan Digital

Ramadhan Bagus Adi Nugroho^{1*}, Mukh Taofik Chulkamdi², Rizki Dwi Romadhona³

¹⁻³ Program Studi Teknik Informatika, Fakultas Teknik dan Informatika, Universitas Islam Balitar

Email: bagusadiblitar@gmail.com^{1*}, chulkamdi@gmail.com², rizkidwiromadhona@unisbablitar.ac.id³

Alamat: Jl. Imam Bonjol No. 16, Jl. Majapahit No.2- 4, Sananwetan, Kec. Sananwetan, Kota Blitar, Jawa Timur 66137

Korespondensi penulis: bagusadiblitar@gmail.com ^{1*}

Abstract: *The increasing use of digital library applications in the era of information technology necessitates a systematic evaluation of user experience. Online user reviews contain valuable information that can be analyzed to measure service effectiveness. This study aims to identify user perceptions through sentiment analysis and compare the performance of the Support Vector Machine (SVM) and K-Nearest Neighbors (KNN) classification algorithms. This applied research collects 6,000 reviews via web scraping from the Google Play Store, focusing on two applications: iPusnas and Gramedia Digital. The data processing begins with text preprocessing to clean the text data, followed by feature extraction using Term Frequency-Inverse Document Frequency (TF-IDF) to identify relevant features before classification with both algorithms. The performance of the algorithms is evaluated using a confusion matrix to measure accuracy based on the number of correct and incorrect predictions. The results show that the SVM algorithm achieves higher accuracy, with 77.78% for iPusnas and 79.34% for Gramedia Digital. In contrast, KNN only achieves 40.06% accuracy for iPusnas and 60.54% for Gramedia Digital. From these results, it can be concluded that SVM outperforms KNN in processing digital library application user reviews. This indicates that SVM is more effective in handling large and complex review data. Based on these findings, it is recommended that digital library application developers utilize SVM in sentiment analysis systems to enhance service quality and user experience. Future research may consider aspect-based approaches or a combination of algorithms to achieve more optimal results in sentiment analysis.*

Keywords: *Digital Library, K-Nearest Neighbors, Sentiment Analysis, Support Vector Machine, TF-IDF*

Abstrak: Meningkatnya penggunaan aplikasi perpustakaan digital di era teknologi informasi mendorong perlunya evaluasi sistematis terhadap pengalaman pengguna. Ulasan pengguna yang tersedia secara daring menyimpan informasi penting yang dapat dianalisis untuk mengukur efektivitas layanan. Penelitian ini bertujuan untuk mengidentifikasi persepsi pengguna melalui analisis sentimen, serta membandingkan performa algoritma klasifikasi Support Vector Machine (SVM) dan K-Nearest Neighbors (KNN). Jenis penelitian ini adalah penelitian terapan yang mengumpulkan 6.000 ulasan melalui web scraping dari Google Play Store, dengan dua aplikasi yang dijadikan objek studi, yaitu iPusnas dan Gramedia Digital. Proses pengolahan data dimulai dengan tahap text preprocessing untuk membersihkan data teks, diikuti dengan pembobotan menggunakan Term Frequency-Inverse Document Frequency (TF-IDF), yang bertujuan untuk mengekstraksi fitur teks yang relevan sebelum dilakukan klasifikasi menggunakan kedua algoritma tersebut. Evaluasi kinerja kedua algoritma dilakukan menggunakan confusion matrix, yang membantu untuk mengukur akurasi klasifikasi berdasarkan jumlah prediksi yang benar dan salah. Hasil penelitian menunjukkan bahwa algoritma SVM menghasilkan akurasi yang lebih tinggi, yaitu 77,78% pada aplikasi iPusnas dan 79,34% pada Gramedia Digital. Sebaliknya, algoritma KNN hanya memperoleh akurasi 40,06% pada iPusnas dan 60,54% pada Gramedia Digital. Dari hasil tersebut, dapat disimpulkan bahwa SVM lebih unggul dalam mengolah data ulasan pengguna aplikasi perpustakaan digital dibandingkan dengan KNN. Hal ini menunjukkan bahwa SVM lebih efektif dalam menangani data teks ulasan yang besar dan kompleks. Berdasarkan temuan ini, disarankan agar pengembang aplikasi perpustakaan digital memanfaatkan SVM dalam sistem analisis sentimen untuk meningkatkan kualitas layanan dan pengalaman pengguna. Penelitian selanjutnya dapat mempertimbangkan pendekatan berbasis aspek atau kombinasi algoritma untuk memperoleh hasil yang lebih optimal dalam analisis sentimen pengguna.

Kata kunci: Perpustakaan Digital, K-Nearest Neighbors, Analisis Sentimen, Support Vector Machine, TF-IDF

1. LATAR BELAKANG

Pada era digital yang dimana aplikasi mobile terus berkembang kini menjadi elemen penting dalam kehidupan sehari-hari, dengan jumlah pengguna smartphone yang diperkirakan mencapai lebih dari 3,5 miliar orang di berbagai penjuru dunia pada tahun 2023 (Zhou, 2022). Pertumbuhan pengguna ini mencerminkan betapa besar pengaruh teknologi pada berbagai aspek kehidupan. Berdasarkan data yang diambil dari Google Play, aplikasi-aplikasi ini sering kali memperoleh ribuan hingga jutaan ulasan, yang mencerminkan berbagai pengalaman pengguna yang sangat beragam dan kompleks (Tanggraeni & Sitokdana, 2022). Pengguna sering memberikan umpan balik terkait berbagai aspek aplikasi, mulai dari antarmuka pengguna (UI), fungsionalitas, hingga kestabilan dan kecepatan aplikasi. Oleh sebab itu, untuk mendapatkan wawasan yang lebih dalam mengenai persepsi pengguna, penting untuk menganalisis sentimen yang terkandung dalam ulasan-ulasan ini. Hal ini tidak hanya memberi gambaran mengenai kepuasan pengguna, tetapi juga dapat mengidentifikasi masalah yang sering dihadapi oleh pengguna, sehingga menjadi dasar bagi pengembang untuk melakukan perbaikan.

Dalam konteks aplikasi perpustakaan digital sebagai layanan perpustakaan dalam bentuk digital, analisis sentimen menjadi sangat penting. Perpustakaan digital dirancang dengan tujuan untuk memberikan layanan perpustakaan digital yang gratis serta memudahkan akses ke *e-book* dan e-pustaka. Namun, untuk menilai sejauh mana aplikasi perpustakaan digital ini berhasil memenuhi tujuannya, diperlukan pendekatan yang lebih komprehensif, yakni melalui analisis sentimen yang mendalam terhadap ulasan pengguna. Penelitian-penelitian sebelumnya telah menunjukkan bahwa analisis sentimen bermanfaat untuk membantu dalam mengidentifikasi pola-pola tertentu dalam ulasan pengguna, yang dapat digunakan sebagai alat acuan menilai kelebihan dan kekurangan aplikasi (Zhao, Yang, Bai, & Liu, 2024). Dalam penelitian ini, penulis akan menggunakan algoritma *Support Vector Machine* (SVM) dan *K-Nearest Neighbors* (KNN), yaitu dua algoritma pembelajaran mesin (*machine learning*) yang digunakan untuk tugas klasifikasi dan regresi. Kedua algoritma tersebut dipilih karena mempunyai kelebihan masing-masing dalam menangani data teks.

Menurut statistik yang ada, lebih dari 70% pengguna aplikasi mobile cenderung membaca ulasan pengguna lainnya sebelum memutuskan untuk mengunduh aplikasi (Yu, Gong, & Xia, 2021). Hal ini membuktikan bahwa ulasan pengguna memegang peranan yang sangat penting pada keputusan pengguna, bukan hanya sebagai alat untuk memberikan umpan balik, tetapi juga sebagai indikator kualitas aplikasi itu sendiri. Sehingga analisis sentimen terhadap ulasan yang diberikan oleh pengguna sangat penting untuk memberikan informasi

yang berharga kepada pengembang aplikasi. Sentimen positif bisa menjadi indikator bahwa aplikasi tersebut diterima dengan baik oleh masyarakat, sementara sentimen negatif dapat menunjukkan adanya aspek yang perlu diperbaiki. Oleh karena itu, analisis sentimen bukan hanya memberikan pemahaman tentang kepuasan pengguna, namun juga menyajikan petunjuk yang jelas tentang fitur atau fungsi aplikasi yang perlu ditingkatkan untuk menciptakan pengalaman pengguna yang lebih baik.

Namun, analisis sentimen bukanlah tugas yang mudah. Salah satu tantangan utama yang dihadapi dalam analisis sentimen adalah kerumitan bahasa alami yang digunakan oleh pengguna dalam ulasan mereka. Pengguna sering kali menyampaikan pendapat dengan menggunakan bahasa yang penuh nuansa, yang bisa melibatkan sarkasme, metafora, atau bahkan pernyataan yang ambigu. Terkadang, sebuah kalimat yang terlihat positif bagi sistem analisis bisa jadi mengandung kritik tersembunyi, dan sebaliknya, kalimat yang tampaknya negatif bisa memiliki makna yang lebih kompleks atau bernuansa. Berdasarkan penelitian yang ada, algoritma seperti SVM dan KNN telah terbukti efektif dalam mengklasifikasikan sentimen dalam berbagai konteks, meskipun masih terdapat tantangan untuk meningkatkan akurasi mereka dalam memahami bahasa yang lebih kompleks (Chen, Zhai, Feng, Li, & Wang, 2022). Untuk itu, penelitian ini bertujuan untuk mencari lebih luas penerapan algoritma SVM dan KNN yang lebih canggih untuk memahami keunggulan, kelemahan, serta aplikasi yang tepat dari kedua algoritma ini, sehingga dapat digunakan secara optimal dalam berbagai tugas pembelajaran mesin seperti analisis sentimen.

Sejumlah penelitian terdahulu telah banyak berfokus pada aplikasi-aplikasi komersial seperti *e-commerce* atau aplikasi media sosial, namun penelitian yang secara khusus membahas aplikasi perpustakaan digital seperti iPusnas dan Gramedia Digital masih tergolong sedikit. Padahal, aplikasi-aplikasi ini memiliki peran penting dalam sangat penting dalam mendukung pendidikan, penelitian, serta penyebaran informasi di era digital. Aplikasi perpustakaan digital seringkali memiliki karakteristik yang berbeda dengan aplikasi lainnya yang dirancang khusus untuk mendukung pembelajaran, literasi, dan penelitian dengan menyediakan akses ke koleksi digital yang terorganisasi dan dapat diandalkan. Hal ini bisa dijadikan salah satu alasan aplikasi perpustakaan digital memerlukan analisis sentimen yang lebih mendalam dan relevan (Ardianto & Marhoon, 2020). Penelitian ini mempunyai tujuan untuk mengisi kesenjangan tersebut dengan memberikan analisis yang lebih fokus pada aplikasi perpustakaan digital, khususnya aplikasi iPusnas dan Gramedia Digital, dan memberikan rekomendasi berbasis data untuk pengembangan aplikasi lebih lanjut.

Tujuan penting pada penelitian kali ini adalah melakukan analisis sentimen terhadap ulasan pengguna iPusnas dan Gramedia Digital yang terdapat di *Google Play* dengan menggunakan algoritma SVM dan KNN. Dengan pendekatan ini, diharapkan hasil penelitian tidak hanya memberikan pengetahuan yang lebih baik mengenai penilaian pengguna, tetapi juga memberikan kontribusi baru dalam bidang analisis sentimen. Secara teoritis, penelitian ini akan memperkaya literatur dalam bidang analisis sentimen, khususnya yang berkaitan dengan aplikasi perpustakaan digital. Sedangkan secara praktis, hasil dari analisis sentimen ini dapat dimanfaatkan oleh pengembang aplikasi sebagai sarana meningkatkan kualitas layanan dan pengalaman pengguna, sehingga aplikasi tersebut lebih sesuai dengan harapan masyarakat.

Penelitian ini juga akan membahas bagaimana sentimen pengguna dapat berkaitan dengan fitur-fitur tertentu dari aplikasi perpustakaan nasional. Sebagai contoh, apakah kecepatan aplikasi atau kemudahan penggunaan aplikasi mempengaruhi sentimen positif atau negatif yang muncul dalam ulasan pengguna. Dengan menggunakan teknik analisis sentimen yang tepat, penelitian ini bertujuan untuk mengetahui fitur yang banyak disukai oleh pengguna, serta fitur-fitur yang dirasa kurang optimal dan perlu perbaikan. Dengan hasil ini, pengembang aplikasi dapat memperoleh pemahaman yang lebih baik tentang bagian-bagian mana dari aplikasi yang perlu difokuskan pada pembaruan dan peningkatan di masa mendatang (Malik & Sibaroni, 2022). Selain itu, analisis fitur ini juga dapat memberikan rekomendasi bagi pengembang untuk merancang pengalaman pengguna yang lebih baik, dimana akhirnya dapat meningkatkan kualitas aplikasi secara keseluruhan.

Penelitian ini juga akan mengeksplorasi cara-cara untuk mengoptimalkan algoritma SVM dan KNN dalam analisis sentiment. Algoritma *Support Vector Machine* (SVM) dan *K-Nearest Neighbors* (K-NN) dijalankan secara bersamaan untuk menganalisis sentimen ulasan pengguna aplikasi perpustakaan digital, menggunakan dataset dan tahapan pemrosesan yang sama. Data ulasan terlebih dahulu melalui tahap preprocessing, seperti pembersihan teks, tokenisasi, dan ekstraksi fitur menggunakan teknik seperti TF-IDF. Setelah itu, kedua algoritma dilatih dan diuji dengan data yang sama untuk memastikan evaluasi yang adil sehingga dapat dianalisis algoritma mana yang lebih unggul dalam mengklasifikasikan sentimen pengguna (Muhammad dkk., 2024). Dalam konteks SVM dan KNN, pemilihan parameter yang tepat sangat berpengaruh dalam menciptakan keseimbangan antara bias dan varians pada model, di mana pada SVM, pemilihan nilai C dan jenis kernel yang sesuai menentukan seberapa baik model memisahkan kelas data tanpa terjebak pada overfitting atau underfitting, sementara pada KNN, pemilihan jumlah tetangga (k) yang tepat mempengaruhi kemampuan model untuk menangkap pola lokal data tanpa menjadi terlalu sensitif atau terlalu

umum. Selain itu, eksperimen dengan dua metode tersebut pemilihan fitur akan dilakukan untuk mengetahui fitur-fitur mana yang paling mempengaruhi klasifikasi sentimen, serta untuk mengurangi dimensi data yang dapat mempercepat proses klasifikasi (Wankhade, Rao, & Kulkarni, 2022). Dengan pendekatan tersebut, peneliti berharap KNN bisa menghasilkan hasil yang lebih akurat dan efektif, serta dapat diimplementasikan dalam analisis sentimen aplikasi lainnya di masa depan.

Analisis sentimen sangat penting untuk mengidentifikasi aspek-aspek aplikasi yang perlu diperbaiki, sehingga pengembang dapat melakukan penyesuaian yang diperlukan guna memenuhi ekspektasi pengguna. Dengan menerapkan strategi pemasaran berbasis analisis sentimen, aplikasi tidak hanya menjadi lebih menarik, namun juga mampu meningkatkan pengalaman pengguna, yang pada gilirannya memperkuat kepuasan dan loyalitas mereka (Garousi & Cutting, 2021). Penelitian menunjukkan bahwa penerapan analisis sentimen dalam proses pengembangan aplikasi dapat mempercepat tingkat adopsi aplikasi oleh pengguna, karena aplikasi yang responsif terhadap umpan balik cenderung meningkatkan kepuasan pengguna. Dengan menggunakan alat analisis sentimen yang tepat, pengembang dapat dengan sistematis mengukur dan memahami reaksi pengguna, serta segera menanggapi umpan balik mereka, sehingga tercipta siklus umpan balik positif yang menguntungkan antara pengguna dan pengembang (Rawat Rawat, Rawat, Sharma, & Shefali, 2022)

Secara menyeluruh, penelitian ini bertujuan untuk memberikan rekomendasi kebijakan bagi pengembangan aplikasi lebih responsif terhadap kebutuhan pengguna. Dengan memanfaatkan hasil dari analisis sentimen, pengambil kebijakan dapat lebih memahami bagaimana kebijakan dan layanan yang mereka tawarkan diterima oleh masyarakat melalui aplikasi digital. Rekomendasi ini dapat menjadi dasar bagi perbaikan aplikasi di masa depan, agar lebih inklusif, *user-friendly*, dan bermanfaat bagi berbagai lapisan masyarakat (Aftab et al., 2023).

Dengan demikian, hasil dari penelitian dapat bermanfaat dalam bidang analisis sentimen yang signifikan, terutama dalam konteks aplikasi perpustakaan digital di *Google Play*. Dengan pendekatan yang lebih sistematis dan berbasis data, diharapkan hasil penelitian akan memberikan pengetahuan bermanfaat kepada pengembang aplikasi, pembuat kebijakan, serta peneliti yang tertarik pada teknologi informasi dan komunikasi. Dengan demikian, penelitian ini diangkat dengan judul “Perbandingan Algoritma *Support Vector Machine* Dan *K-Nearest Neighbors* Terhadap Analisis Sentimen Aplikasi Perpustakaan Digital”.

2. METODE PENELITIAN

Penelitian ini dilaksanakan mulai Desember 2024 hingga Mei 2025 dengan lokasi pengambilan data di Google Play Store, yang dipilih karena menyediakan ulasan pengguna yang kaya dan beragam terhadap aplikasi iPusnas dan Gramedia Digital. Penelitian ini termasuk jenis terapan dengan tujuan membandingkan efektivitas algoritma Support Vector Machine (SVM) dan K-Nearest Neighbors (KNN) dalam mengklasifikasikan sentimen positif, netral, dan negatif dari ulasan pengguna. Alur penelitian diawali dengan identifikasi masalah, yang berfokus pada kualitas layanan kedua aplikasi, kemudian dilanjutkan dengan pengumpulan data menggunakan teknik web scraping untuk mendapatkan 3.000 ulasan dari masing-masing aplikasi dalam bahasa Indonesia.

Tahap pengumpulan data dilakukan dengan library Python *google-play-scraper*, menghasilkan data berisi *review ID*, *username*, *rating*, teks ulasan, dan tanggal. Data yang diperoleh kemudian diekspor ke format .csv dan dilanjutkan ke tahap *preprocessing* untuk memastikan kualitas dan keseragaman data. Proses *preprocessing* meliputi *cleaning* (menghapus simbol, angka, dan karakter tak relevan), *case folding* (mengubah huruf menjadi huruf kecil), normalisasi kata (mengganti kata tidak baku menjadi baku), *tokenizing* (memecah teks menjadi token), penghapusan *stopword*, dan *stemming* (mengubah kata ke bentuk dasar). Langkah-langkah ini bertujuan mempersiapkan data agar lebih siap untuk analisis sentimen.

Setelah *preprocessing*, dilakukan pelabelan data menggunakan kamus leksikon positif dan negatif untuk menentukan skor sentimen setiap ulasan. Skor positif menunjukkan sentimen positif, skor negatif untuk sentimen negatif, dan skor nol untuk sentimen netral. Dataset dibagi menjadi dua bagian, yaitu data latih (80%) dan data uji (20%), untuk melatih serta menguji model. Proses pelatihan model dilakukan dengan teknik ekstraksi fitur TF-IDF, seleksi fitur menggunakan *SelectKBest*, dan reduksi dimensi dengan PCA. Algoritma SVM dilatih untuk mencari *hyperplane* pemisah, sedangkan KNN menggunakan metode *k nearest neighbors* untuk menentukan label berdasarkan mayoritas tetangga terdekat (Damayanti & Nuzuli, 2023).

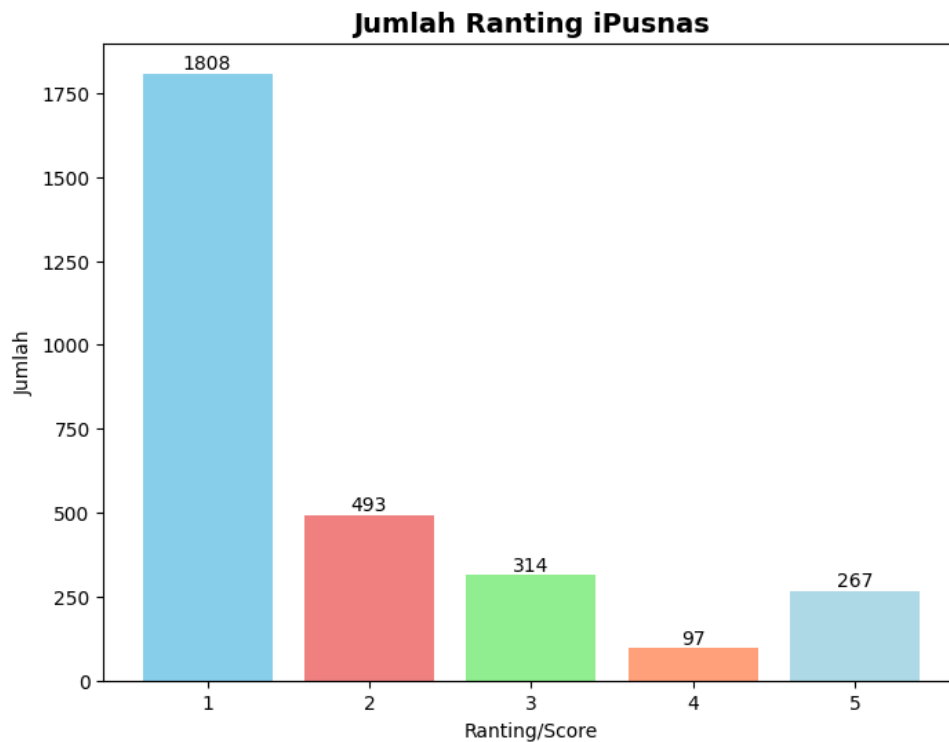
Evaluasi model dilakukan menggunakan data uji dengan metrik *confusion matrix*, akurasi, presisi, recall, dan F1-score untuk menilai kinerja kedua algoritma. Hasil evaluasi dikategorikan berdasarkan tabel kriteria, mulai dari *Excellent Classification* hingga *Failure*. Flowchart sistem juga dibuat untuk menggambarkan alur analisis sentimen, mulai dari pengumpulan data hingga klasifikasi akhir, sehingga dapat memberikan gambaran komprehensif tentang proses yang dilakukan. Penelitian ini diharapkan membantu pengembang aplikasi dalam meningkatkan kualitas layanan serta memberikan perbandingan yang jelas mengenai efektivitas SVM dan KNN dalam analisis sentimen.

3. HASIL DAN PEMBAHASAN

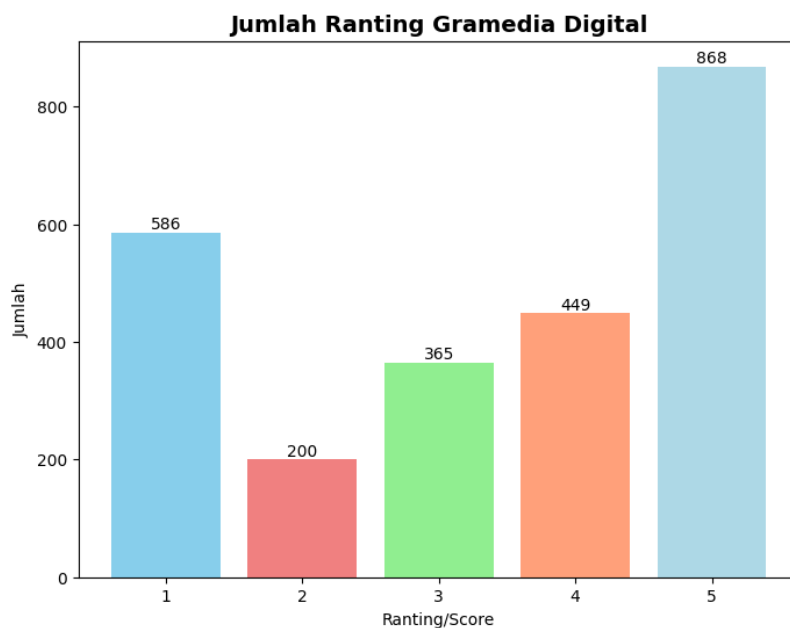
Hasil

Penelitian ini melakukan analisis perbandingan rating pengguna terhadap dua aplikasi perpustakaan digital, yaitu iPusnas dan Gramedia Digital. Data diperoleh dari jumlah rating yang diberikan pengguna pada masing-masing aplikasi, yang dikategorikan berdasarkan skala 1 hingga 5.

Visualisasi diagram batang menunjukkan distribusi jumlah rating sebagai berikut:



Gambar 1. Rating Aplikasi iPusnas



Gambar 2. Rating Aplikasi Gramedia Digital

Berdasarkan data tersebut, iPusnas didominasi oleh rating bintang 1, dengan total sebanyak 1808 pengguna memberikan penilaian terendah. Hal ini menunjukkan bahwa sebagian besar pengguna merasa sangat tidak puas terhadap kinerja aplikasi tersebut. Jumlah rating bintang 2 dan 3 pun cukup tinggi, yaitu masing-masing 493 dan 314, sedangkan rating tertinggi (bintang 5) hanya diberikan oleh 267 pengguna.

Sebaliknya, aplikasi Gramedia Digital menunjukkan tren yang berbanding terbalik. Aplikasi ini memperoleh rating bintang 5 tertinggi, yakni sebanyak 1641 ulasan. Rating bintang 4 juga cukup tinggi 236, sedangkan rating bintang 1 hanya sebesar 327. Ini menandakan bahwa pengguna umumnya merasa puas dengan performa dan layanan yang ditawarkan Gramedia Digital.

Dari hasil ini dapat disimpulkan bahwa Gramedia Digital memiliki persepsi pengguna yang lebih positif dibandingkan iPusnas, sehingga mampu membangun loyalitas dan tingkat kepercayaan yang lebih tinggi. Oleh karena itu, iPusnas perlu melakukan perbaikan menyeluruh berdasarkan umpan balik pengguna, baik dalam aspek teknis, desain, maupun layanan konten digital.

Pembahasan

Berdasarkan penelitian yang telah dilakukan dapat diketahui bahwa data dikumpulkan dari ulasan pengguna aplikasi iPusnas dan Gramedia Digital yang tersedia di *Google Play Store*. Proses pengambilan data dilakukan dengan teknik *web scraping* menggunakan pustaka *google-play-scraper* yang diimplementasikan dalam bahasa *Python*. Sebanyak 3000 ulasan dikumpulkan dari masing-masing aplikasi iPusnas dan Gramedia Digital. Tahap selanjutnya yang dilakukan antara lain *text preprocessing* dan klasifikasi dengan metode SVM dan KNN serta pengujian dengan confusion matrix.

Berdasarkan hasil evaluasi yang dilakukan terhadap algoritma *Support Vector Machine* dan *K-Nearest Neighbors* dalam klasifikasi sentimen ulasan pengguna aplikasi iPusnas dan Gramedia Digital, dapat disimpulkan bahwa algoritma SVM memiliki kinerja yang lebih unggul dibandingkan KNN. Hal ini ditunjukkan oleh nilai akurasi, *precision*, *recall*, dan *f1-score* yang secara konsisten lebih tinggi pada kedua jenis data ulasan. Pada aplikasi iPusnas, SVM berhasil mencapai akurasi sebesar 77,78%, jauh lebih tinggi dibandingkan KNN yang hanya mencapai 40,06%. Begitu juga pada aplikasi Gramedia Digital, SVM memperoleh akurasi 79,34%, mengungguli KNN dengan 60,54%.

Pernyataan hasil tersebut sejalan dengan penelitian dari (Dea Safryda et al., 2023) Model analisis sentimen terbaik menggunakan algoritma SVM dengan pelabelan manual dan proporsi data latih:uji 90:10, menghasilkan akurasi, presisi, dan *recall* sebesar 93%, serta *f1-score* 92%. SVM terbukti lebih optimal dibandingkan KNN dalam 7 dari 10 kasus, sementara pelabelan manual menghasilkan performa lebih baik namun kurang efisien dibandingkan pelabelan leksikon.

Penelitian dari (Busrayan & Andrianingsih, 2025) menunjukkan tingkat akurasi algoritma dalam analisis sentimen, Naive Bayes menunjukkan performansi yang cukup baik dengan *precision*, *recall*, dan *accuracy* sekitar 0.59 hingga 0.63, namun memiliki *f1-score* yang lebih rendah yaitu 0.53. SVM unggul di semua metrik dengan nilai *precision*, *recall*, *f1-score*, dan *accuracy* yang semuanya mencapai 0.80. KNN memiliki performa yang lebih rendah dibandingkan SVM dan *Naive Bayes*, dengan nilai *precision*, *recall*, *f1-score*, dan *accuracy* sekitar 0.33 hingga 0.60. Secara keseluruhan, SVM adalah model yang paling kuat berdasarkan metrik yang digunakan, sedangkan KNN menunjukkan performansi yang paling lemah dalam hal ini.

4. KESIMPULAN DAN SARAN

Kesimpulan

Berdasarkan hasil dan pembahasan di atas dapat disimpulkan bahwa penerapan algoritma Support Vector Machine (SVM) dalam analisis sentimen dilakukan melalui serangkaian tahapan mulai dari pengumpulan data ulasan pengguna aplikasi iPusnas dan Gramedia Digital. Data ulasan ini kemudian diproses dengan tahapan *text preprocessing* seperti *cleaning*, *case folding*, *stopword*, normalisasi kata, *tokenizing*, *stemming*. Setelah tahap *preprocessing* selesai, data teks dikonversi ke bentuk numerik menggunakan teknik TF-IDF agar dapat diolah oleh algoritma kemudian dilatih dengan data pelatihan dan diuji menggunakan data uji untuk mengidentifikasi pola sentimen. Hasil dari pelatihan menunjukkan bahwa SVM memberikan kinerja klasifikasi yang cukup baik, dengan akurasi sebesar 77,78% pada aplikasi iPusnas dan 79,34% pada Gramedia Digital. Nilai *precision*, *recall*, dan *f1-score* yang juga cukup baik dan konsisten di angka 0,78–0,79.

Pada pendekatan *K-Nearest Neighbors* proses analisis sentimen dilakukan terhadap data ulasan yang telah melalui tahap pembersihan dan transformasi TF-IDF. KNN menentukan sentimen berdasarkan jarak terdekat dari data baru ke sejumlah tetangga *k* pada data pelatihan. Kelas sentimen dari data baru ditentukan berdasarkan mayoritas label dari tetangga tersebut.

Pada aplikasi iPusnas, KNN hanya mencatatkan akurasi 40,06%, dengan precision 0,68, recall 0,40, dan f1-score 0,38. Sementara di Gramedia Digital, kinerjanya sedikit lebih baik dengan akurasi 60,54%, precision 0,64, recall 0,61, dan f1-score 0,60.

Berdasarkan hasil pengujian, algoritma *Support Vector Machine* secara konsisten menunjukkan kinerja yang lebih unggul dibandingkan *K-Nearest Neighbors* dalam analisis sentimen. Penelitian ini menunjukkan bahwa SVM memiliki tingkat akurasi yang lebih baik dibandingkan dengan KNN. SVM mendapatkan akurasi sebesar 77,78% pada aplikasi iPusnas dan 79,34% pada Gramedia Digital, sedangkan KNN hanya mencatatkan akurasi 40,06%, pada iPusnas dan Gramedia Digital mendapat nilai akurasi 60,54%. Oleh karena itu, SVM dapat dianggap sebagai pilihan algoritma yang lebih tepat untuk analisis sentimen ulasan aplikasi perpustakaan digital.

Saran

Penulis menyarankan pengembangan penelitian lebih lanjut agar menghasilkan hasil performa yang lebih baik berdasarkan hasil yang sudah didapatkan seperti sebagai berikut:

- a. Pada penelitian selanjutnya dapat membandingkan algoritma *machine learning* yang lain seperti *Naïve Bayes*, *Random Forest*, dan lain-lain.
- b. Bahasa yang digunakan dalam penelitian ini terbatas pada Bahasa Indonesia. Oleh karena itu, disarankan agar penelitian selanjutnya dapat memperluas cakupan dengan menambahkan ulasan dalam Bahasa Inggris atau bahasa asing lainnya guna meningkatkan generalisasi dan keberagaman data dalam analisis sentiment.

5. UCAPAN TERIMA KASIH

Terima kasih kepada institusi dan perpustakaan digital yang telah menyediakan data, para dosen pembimbing yang memberikan arahan dan masukan konstruktif, serta rekan-rekan yang turut membantu dalam proses pengolahan data dan pengujian algoritma. Terima kasih juga kepada keluarga dan sahabat atas dukungan moral yang tiada henti sehingga penelitian ini dapat terselesaikan dengan baik.

DAFTAR REFERENSI

- Aftab, F., Bazai, S. U., Marjan, S., Baloch, L., Aslam, S., Amphawan, A., & Neo, T. K. (2023). A comprehensive survey on sentiment analysis techniques. *International Journal of Technology*, 14(6), 1288-1298. <https://doi.org/10.14716/ijtech.v14i6.6632>
- Ardianto, R., & Marhoon, H. M. (2020). Understanding user sentiment: Analysis of SATUSEHAT application reviews on Google Play Store. *Journal of Advanced Health Informatics Research*, 1(2), 83-94. <https://doi.org/10.59247/jahir.v1i2.44>
- Busrayan, I., & Andrianingsih. (2025). Analisis sentimen pelanggan terhadap aplikasi Wondr by BNI menggunakan Naive Bayes, Support Vector Machine (SVM), dan K-Nearest Neighbor (KNN), 2(2), 263-274.
- Chen, H., Zhai, Z., Feng, F., Li, R., & Wang, X. (2022). Enhanced multi-channel graph convolutional network for aspect sentiment triplet extraction. *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, 1, 2974-2985. <https://doi.org/10.18653/v1/2022.acl-long.212>
- Damayanti, D., & Nuzuli, A. K. (2023). Studi kasus implementasi teknologi chatbot sebagai asisten virtual dalam menjawab pertanyaan mahasiswa di lingkungan kampus. *Reslaj: Religion Education Social Laa Roiba Journal*, 6(3), 1178-1192. <https://doi.org/10.47467/reslaj.v6i3.4858>
- Dea Safryda, P., Nina, S., & Apriade, V. (2023). Analisis sentimen dan pemodelan ulasan aplikasi AdaKami menggunakan algoritma SVM dan KNN. *Journal Sensi*, 09(02).
- Garousi, V., & Cutting, D. (2021). What do users think of the UK's three COVID-19 contact-tracing apps? A comparative analysis. *BMJ Health and Care Informatics*, 28(1), 1-7. <https://doi.org/10.1136/bmjhci-2021-100320>
- Malik, R. A. A., & Sibaroni, Y. (2022). Multi-aspect sentiment analysis of TikTok application usage using FastText feature expansion and CNN method. *Journal of Computer System and Informatics (JoSYC)*, 3(4), 277-285. <https://doi.org/10.47065/josyc.v3i4.2033>
- Muhammad Adin, P., Rochana Prih, H., & Rian Adam, R. (2024). Analisis sentimen berbasis aspek pada ulasan pengguna aplikasi Starbucks menggunakan algoritma Support Vector Machine. *Journal of Internet and Software Engineering (JISE)*, 13(1), 659. <https://doi.org/10.35889/jutisi.v13i1.1826>
- Rawat, D., Rawat, K., Sharma, A., & Shefali, S. (2022). A study of educational technology startups in India. *Journal of Technology Management for Growing Economies*, 14(1), 16-6. <https://doi.org/10.15415/jtmge/2023.141001>
- Tanggraeni, A. I., & Sitokdana, M. N. N. (2022). Analisis sentimen aplikasi e-government pada Google Play menggunakan algoritma Naïve Bayes. *JATISI (Jurnal Teknik Informatika Dan Sistem Informasi)*, 9(2), 785-795. <https://doi.org/10.35957/jatisi.v9i2.1835>
- Wankhade, M., Rao, A. C. S., & Kulkarni, C. (2022). A survey on sentiment analysis methods, applications, and challenges. In *Artificial Intelligence Review (Vol. 55, Issue 7)*. Springer Netherlands. <https://doi.org/10.1007/s10462-022-10144-1>

- Yu, J., Gong, C., & Xia, R. (2021). Cross-domain review generation for aspect-based sentiment analysis. *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*, 4767-4777. <https://doi.org/10.18653/v1/2021.findings-acl.421>
- Zhao, H., Yang, M., Bai, X., & Liu, H. (2024). A survey on multimodal aspect-based sentiment analysis. *IEEE Access*, 12(November 2023), 12039-12052. <https://doi.org/10.1109/ACCESS.2024.3354844>
- Zhou, L. (2022). Research on quantitative model of brand recognition based on sentiment analysis of big data. *Frontiers in Psychology*, 13(May), 1-11. <https://doi.org/10.3389/fpsyg.2022.915443>