

Klasifikasi Sentimen Layanan pada Aplikasi by.U menggunakan Algoritma *Support Vector Machine*

Classification of Service Sentiments on the by.U Application using the Support Vector Machine Algorithm

¹Zulkarnain*, ²Rice Novita, ³Angraini, ⁴Zarnelly

^{1,2,3,4}Program Studi Sistem Informasi, Fakultas Sains dan Teknologi, Universitas Islam Negeri Sultan Syarif Kasim Riau

^{1,2,3,4}Jl. H.R. Soebrantas KM. 15 No. 155 Tuah Madani Kec. Tuah Madani, Pekanbaru, Indonesia

*e-mail: 12150314261@students.uin-suska.ac.id

(received: 28 May 2025, revised: 6 June 2025, accepted: 7 June 2025)

Abstrak

Penelitian ini didasari oleh pentingnya pemahaman opini pengguna terhadap kualitas layanan digital sebagai bahan evaluasi dan peningkatan layanan. Penelitian ini bertujuan untuk melakukan klasifikasi sentimen pengguna terhadap layanan aplikasi by.U dengan memanfaatkan algoritma Support Vector Machine (SVM). Data ulasan diambil dari Google Play Store sebanyak 9.091 data, yang kemudian diproses melalui tahapan pra-pemrosesan seperti cleaning, case folding, tokenisasi, penghapusan kata umum (stopword), dan stemming. Sentimen dibagi menjadi tiga kategori, yaitu positif, negatif, dan netral. Proses pelatihan dan pengujian dilakukan dengan membagi data ke dalam data latih dan data uji dengan proporsi 80:20, serta dievaluasi menggunakan matriks akurasi, presisi, recall, dan F1-score. Hasil evaluasi menunjukkan bahwa algoritma SVM mampu mengklasifikasikan sentimen dengan akurasi sebesar 83%. Model memiliki performa terbaik pada sentimen positif (precision 84%, recall 90%, F1-score 87%) dan negatif (precision 81%, recall 92%, F1-score 86%), sementara sentimen netral masih memiliki kelemahan dengan F1-score hanya sebesar 64%. Hal ini menunjukkan bahwa klasifikasi sentimen netral masih memerlukan penguatan model. Penelitian ini membuktikan bahwa SVM merupakan metode yang efektif untuk menganalisis opini pengguna terhadap layanan digital secara otomatis. Hasil klasifikasi ini dapat dijadikan referensi bagi pengembang dalam melakukan evaluasi dan peningkatan kualitas layanan berdasarkan umpan balik pengguna.

Kata kunci: klasifikasi sentimen, aplikasi by.U, Support Vector Machine, ulasan pengguna, text mining.

Abstract

This study aims to classify user sentiment toward the by.U application service using the Support Vector Machine (SVM) algorithm. The background of this research is based on the importance of understanding user opinions on the quality of digital services as a basis for evaluation and service improvement. Review data was collected from the Google Play Store, totaling 9,091 data points, which were then processed through preprocessing stages such as cleaning, case folding, tokenization, stopword removal, and stemming. Sentiments were categorized into three groups: positive, negative, and neutral. The training and testing process involved dividing the data into training and testing sets with an 80:20 ratio, and evaluation was conducted using metrics such as accuracy, precision, recall, and F1-score. The evaluation results showed that the SVM algorithm achieved an accuracy of 83% in classifying sentiments. The model performed best on positive sentiment (precision 84%, recall 90%, F1-score 87%) and negative sentiment (precision 81%, recall 92%, F1-score 86%), while neutral sentiment still had weaknesses with an F1-score of only 64%. This indicates that neutral sentiment classification still requires model enhancement. This study demonstrates that SVM is an effective method for automatically analyzing user opinions on digital services. These classification results can serve as a reference for developers in evaluating and improving service quality based on user feedback.

Keywords: sentiment classification, by.U application, Support Vector Machine, user reviews, text mining.

<http://sistemasi.ftik.unisi.ac.id>

1 Pendahuluan

Teknologi telekomunikasi terus mengalami perkembangan pesat dalam beberapa dekade terakhir[1]. Salah satu inovasi dalam industri ini adalah kehadiran by.U, sebuah operator seluler digital pertama di Indonesia yang menawarkan layanan berbasis aplikasi, mulai dari pembelian kartu SIM, pemilihan paket data, hingga pengelolaan layanan pelanggan. Dengan konsep serba digital, by.U menargetkan segmen pengguna muda yang menginginkan fleksibilitas dan kemudahan dalam mengatur layanan seluler mereka.

By.U yang merupakan inovasi dari PT. Telekomunikasi Indonesia[2]. Aplikasi ini dirancang untuk memberikan pengalaman pengguna yang optimal dalam mengakses layanan telekomunikasi dengan memanfaatkan jaringan 4G dan 5G. By.U menawarkan berbagai fitur unggulan, seperti kemudahan registrasi online tanpa perlu mengunjungi gerai fisik, pengelolaan paket data secara fleksibel sesuai kebutuhan pengguna, monitoring penggunaan data secara real-time, serta layanan pelanggan yang responsif melalui aplikasi. Kehadiran aplikasi ini menjadi solusi bagi pengguna yang menginginkan fleksibilitas dan kemudahan dalam mengatur kebutuhan telekomunikasi mereka. Aplikasi digital By.U dikembangkan sebagai solusi atas kendala masyarakat dalam pembelian paket data. Kendala ini umumnya muncul saat pengguna memerlukan paket data secara mendesak namun terhalang karena harus keluar rumah untuk melakukan pembelian[3].

Sebagai layanan berbasis aplikasi, kualitas layanan by.U sangat bergantung pada pengalaman pengguna (user experience) yang mereka rasakan saat menggunakan aplikasi [4]. Pengguna dapat memberikan umpan balik mereka melalui berbagai platform digital, seperti Google Play Store, App Store, media sosial, dan forum diskusi. Google Play Store merupakan platform distribusi konten digital milik Google yang menawarkan berbagai jenis konten seperti aplikasi, gim, dan lainnya. Layanan ini dapat diakses melalui perangkat Android, situs resmi Google, maupun Google TV [5].

Umpan balik yang diberikan oleh pengguna berbentuk teks tidak terstruktur dan jumlahnya sangat besar, sehingga diperlukan teknik khusus untuk mengolahnya agar dapat diekstraksi menjadi informasi yang bermakna. Dalam bidang data science, proses pengolahan teks dalam jumlah besar ini dikenal sebagai text mining[6]. Text mining memungkinkan analisis lebih lanjut terhadap ulasan pengguna dengan cara mengidentifikasi pola, menemukan kata-kata kunci yang sering muncul, serta mengklasifikasikan sentimen pengguna berdasarkan isi ulasan mereka [7]. Dengan demikian, data yang diperoleh dari ulasan pengguna dapat dikategorikan sebagai sentimen positif, netral, atau negatif.

Dalam analisis sentimen, pemilihan algoritma klasifikasi yang tepat sangat berpengaruh terhadap pencapaian tingkat akurasi yang optimal[8]. Analisis sentimen berfokus pada identifikasi dan ekstraksi informasi subjektif dalam teks, dengan tujuan untuk mengetahui sikap atau perasaan penulis pada subjek, baik itu bersifat positif, negatif, maupun netral[9]. Dalam penelitian ini, algoritma Support Vector Machine (SVM) digunakan untuk menganalisis sentimen. SVM adalah salah satu metode pembelajaran yang memiliki kualitas dan akurasi tinggi, sehingga menjadikannya algoritma yang populer di antara algoritma lainnya [10]. Secara umum, SVM adalah linear classifier yang bekerja dengan memisahkan data-data secara linier, tapi seakarang juga dikembangkan lebih lanjut agar dapat menjalankan data non-linier dengan memanfaatkan konsep kernel di ruang yang berdimensi tinggi[11]. SVM adalah algoritma yang kuat dalam pemisahan kelas dengan margin yang optimal, seringkali memberikan kinerja yang baik dalam berbagai tugas klasifikasi, termasuk analisis sentimen[12]. Dibandingkan dengan algoritma lain, SVM terbukti mampu membangun model klasifikasi dengan tingkat akurasi yang lebih unggul [13].

Untuk mendukung kinerja algoritma dalam analisis sentimen, digunakan teknik Term Frequency-Inverse Document Frequency (TF-IDF) sebagai metode transformasi data dari bentuk tekstual menjadi angka[14]. TF-IDF merupakan ukuran statistik digunakan agar menilai seberapa perlu sebuah kata dalam dokumen[15]. Term Frequency (TF) mencerminkan frekuensi munculnya suatu kata dalam dokumen, sedangkan Inverse Document Frequency (IDF) berfungsi menurunkan bobot kata-kata yang umum dan tidak informatif dalam keseluruhan korpus[16]. Kombinasi antara SVM dan TF-IDF terbukti efektif dalam meningkatkan akurasi klasifikasi sentimen.

Penelitian ini bertujuan untuk mengevaluasi kinerja algoritma SVM dalam menganalisis sentimen terhadap layanan yang diberikan oleh aplikasi by.U. Dengan melakukan klasifikasi ini, diharapkan dapat diperoleh pemahaman yang lebih mendalam mengenai efektivitas algoritma tersebut dalam mengelompokkan sentimen pengguna terhadap layanan by.U, baik dalam aspek kualitas

jaringan, fitur aplikasi, maupun layanan pelanggan. Hasil dari penelitian ini diharapkan agar bisa memberikan gambaran yang lebih jelas mengenai persepsi pengguna terhadap layanan by.U serta menjadi dasar bagi pengembangan dan perbaikan layanan di masa mendatang.

2 Tinjauan Literatur

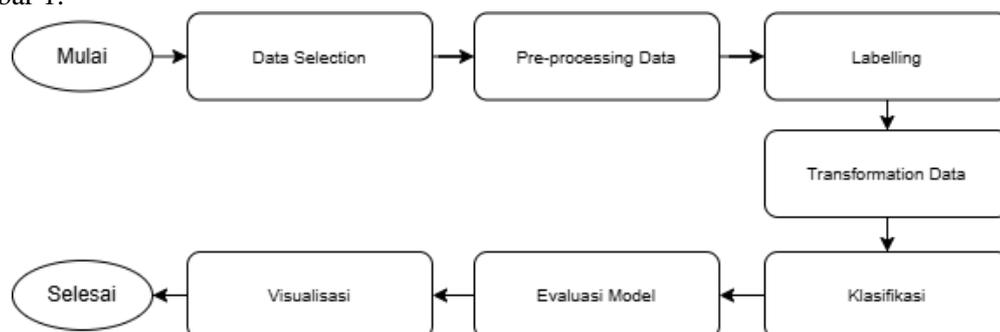
Berbagai penelitian telah mengkaji penerapan analisis sentimen dalam berbagai konteks menggunakan algoritma Support Vector Machine (SVM). Salah satunya penelitian oleh [17], yang menganalisis sentimen pengguna Twitter terhadap layanan pinjaman online. Penelitian ini menggunakan data ulasan dari media sosial yang diklasifikasikan menjadi sentimen positif dan negatif, dengan hasil akurasi sebesar 62% dan proporsi sentimen negatif sebesar 59%. Hasil ini menunjukkan bahwa SVM mampu mengidentifikasi persepsi masyarakat terhadap layanan keuangan digital secara cukup efektif.

Selanjutnya, penelitian [18] menganalisis sentimen pengguna terhadap layanan Gojek berdasarkan 1.500 tweet yang telah melalui proses preprocessing dan pelabelan, baik secara manual maupun menggunakan metode sentiment scoring. Hasil klasifikasi menggunakan kernel Radial Basis Function (RBF) pada algoritma SVM menunjukkan akurasi sebesar 79,19% dan akurasi kappa tertinggi mencapai 21% untuk data berlabel sentiment scoring. Temuan ini memperkuat peran SVM dalam menangani data opini pengguna dari media sosial secara efektif. Penelitian lain [19] juga menggunakan algoritma SVM untuk mengklasifikasikan sentimen masyarakat terhadap isu korupsi benih lobster tahun 2020 berdasarkan data Twitter. Dengan jumlah data sebanyak 156 tweet, model menghasilkan akurasi sebesar 84,21%, sensitivitas 73,38%, dan spesifisitas 82,10%. Penelitian ini menyatakan bahwa SVM tidak hanya efektif dalam konteks layanan digital, tetapi juga dalam menganalisis opini publik terhadap isu-isu sosial dan politik. Dan juga pada penelitian [20] menerapkan algoritma Support Vector Machine (SVM) untuk analisis sentimen terkait Covid-19 pada data Twitter. Data sebanyak 8577 tweet diproses menggunakan TF-IDF dan diuji dengan tiga kernel SVM: Linear, RBF, dan Polynomial. Hasilnya, kernel RBF memberikan akurasi tertinggi sebesar 89,66%. Penelitian ini menunjukkan bahwa kombinasi SVM dan TF-IDF efektif untuk klasifikasi sentimen teks berbahasa Indonesia.

Berbeda dari penelitian-penelitian sebelumnya yang banyak menggunakan data dari media sosial seperti Twitter dan berfokus pada isu keuangan, transportasi, sosial, maupun kesehatan, penelitian ini berkontribusi dalam konteks layanan telekomunikasi digital, khususnya pada aplikasi by.U. Dengan menggunakan dataset yang lebih besar, yakni 9.091 data ulasan dari Google Play Store, serta mencakup tiga kategori sentimen (positif, negatif, dan netral), penelitian ini memberikan gambaran yang lebih luas dan kompleks terhadap persepsi pengguna. Selain itu, penelitian ini juga mengkaji performa klasifikasi terhadap sentimen netral yang selama ini dianggap lebih menantang, serta menyajikan evaluasi menggunakan metrik akurasi, precision, recall, dan f1-score untuk masing-masing kelas, sehingga memberikan kontribusi yang lebih komprehensif dibanding studi-studi sebelumnya.

3 Metode Penelitian

Berikut ini adalah metodologi penelitian yang akan diterapkan untuk penelitian ini, ditunjukkan pada Gambar 1.



Gambar 1. Metode penelitian

3.1 Data Selection

Pada tahap ini, data diambil pada ulasan user aplikasi by.U di Google Play Store. Proses ini melibatkan teknik web scraping untuk mengumpulkan review (ulasan) menggunakan tools seperti Google Collaboratory dan bahasa pemrograman Python. Ulasan yang dikumpulkan mencakup berbagai tanggapan pengguna mengenai aplikasi, mulai dari kualitas jaringan, fitur aplikasi, hingga layanan pelanggan, dengan rentang waktu pengambilan data yang telah ditentukan.

Setelah data ulasan berhasil dikumpulkan dan diseleksi, dilakukan proses pelabelan atau labeling. Proses ini melibatkan pemberian kategori pada setiap ulasan (misalnya: positif, negatif, atau netral) untuk memudahkan pengelompokan, klasifikasi, dan analisis lebih lanjut. Labeling merupakan langkah krusial dalam aplikasi data mining, khususnya dalam supervised learning, di mana model pembelajaran mesin (machine learning) dilatih menggunakan data berlabel agar dapat memprediksi label pada data baru yang belum dikenal sebelumnya.

3.2 Pre-processing Data

Pada tahap ini dilakukan pengurangan kata untuk menghilangkan noise. ini melibatkan lima proses utama, yaitu:

- a. Cleaning, Menghapus karakter yang bukan diinginkan, yaitu tanda baca, angka, dan simbol khusus yang tidak relevan dengan analisis.
- b. Case Folding, yakni proses mengganti seluruh huruf besar pada teks menjadi huruf kecil, guna menghindari perbedaan makna antara kata yang ditulis menggunakan huruf besar dan huruf kecil.
- c. Tokenizing, yaitu memisahkan teks menjadi satuan-satuan kata atau token agar dapat dianalisis secara terstruktur.
- d. Stopword Removal, merupakan tahap penghilangan kata-kata umum yang memiliki nilai informasi rendah dan terlalu sering muncul dalam teks.
- e. Stemming, merupakan metode mengubah kata pada bentuk dasarnya dengan cara menghapus imbuhan, baik berupa awalan, sisipan, akhiran, maupun konfiks, agar analisis kata lebih konsisten.

3.3 Labelling

Pada tahap ini dilakukan proses pemberian label atau kategori pada data untuk memudahkan pengelompokan, klasifikasi, dan analisis lebih lanjut. Proses labelling dilakukan secara manual oleh pakar, yaitu Anjasmara, S.Pd., guna memastikan ketepatan klasifikasi sentimen sesuai dengan konteks ulasan. Labelling sangat penting dalam banyak aplikasi data mining, terutama dalam supervised learning, di mana model pembelajaran mesin dilatih menggunakan data berlabel untuk memprediksi label pada data baru yang tidak dikenal.

3.4 Transformation Data

Setelah melalui tahap preprocessing data, data teks yang telah dibersihkan kemudian ditransformasikan ke dalam bentuk fitur numerik menggunakan metode Term Frequency-Inverse Document Frequency (TF-IDF). TF-IDF digunakan untuk mengukur pentingnya suatu kata didalam sebuah dokumen relatif terhadap seluruh korpus. Transformation data ini bertujuan untuk merubah data dalam bentuk yang sesuai dengan proses oleh data mining [21]. Metode ini menghasilkan representasi vektor dari setiap dokumen, di mana nilai setiap kata mencerminkan frekuensi kemunculannya yang disesuaikan dengan seberapa jarang kata tersebut muncul di dokumen lain. Hasil dari transformasi ini berupa matriks fitur yang digunakan sebagai input untuk proses klasifikasi sentimen menggunakan algoritma SVM.

3.5 Klasifikasi

Pada tahap ini dilakukan proses klasifikasi sentimen terhadap ulasan pengguna aplikasi by.U dengan memanfaatkan algoritma SVM. Dataset yang digunakan dalam penelitian ini dibagi menjadi dua kelompok, yaitu data latih (train-set) dan data uji (test-set), dengan proporsi masing-masing sebesar 80% dan 20% dari total data ulasan yang tersedia. Algoritma SVM diimplementasikan untuk mengklasifikasikan sentimen pengguna ke dalam tiga kategori, yaitu positif, negatif, dan netral. Data latih digunakan agar membangun model pembelajaran yang nantinya dijadikan dasar dalam proses pengujian terhadap data uji. SVM bekerja dengan membentuk sebuah hyperplane yang berfungsi memisahkan kelas-kelas sentimen secara optimal berdasarkan fitur yang diperoleh dari teks ulasan.

3.6 Evaluasi Model

Tahapan ini dilakukan untuk mengevaluasi kinerja algoritma klasifikasi yang dipakai pada penelitian, yaitu SVM. Evaluasi dilakukan menggunakan Confusion Matrix guna memperoleh nilai akurasi, presisi, recall, serta f1-score berdasarkan data yang telah melalui proses pelatihan.

3.7 Visualisasi

Visualisasi data adalah proses penyajian data dalam bentuk grafis atau visual, seperti grafik, diagram, dan chart, dengan tujuan untuk memudahkan pemahaman terhadap informasi yang disajikan. visualisasi yang umum digunakan dalam penelitian sentimen meliputi diagram batang untuk jumlah sentimen positif, negatif, dan netral, dan word cloud agar menunjukkan kata-kata yang sering muncul. Dengan visualisasi yang tepat, hasil analisis menjadi lebih komunikatif dan mudah dipahami oleh berbagai pihak, termasuk yang tidak memiliki latar belakang teknis.

4 Hasil dan Pembahasan

4.1 Data Selection

Penelitian ini memakai data yang dikumpulkan pada ulasan pengguna aplikasi by.U di Google Play Store. Data yang telah crawling berupa tanggal ulasan dan isi ulasan (review) yang diposting oleh pengguna aplikasi. Total data yang berhasil dikumpulkan mencapai 9091 ulasan. Berikut ini adalah data sampel yang digunakan dalam penelitian ini dan disajikan pada Tabel 1.

Tabel 1. Pengumpulan data

	Tanggal	Content
1	24/02/2025 11:47:49	jujur saya sangat suka dengan provider satu ini.. tapi entah kenapa sekarang sering banyak kendala, mulai dari aplikasi yg terkadang suka crash, atau error.. terakhir saat saya melakukan pembelian kuota via e-wallet malah tidak masuk.. tolong untuk kualitasnya lebih diperbaiki lagi.. terimakasih...
2	24/02/2025 11:41:17	Aneh bgt, ngelag terus, kuota ga bisa dipake. Kacau bgt rugi, udah ngikutin cara2 nya tapi aplikasi suka error
...
9090	22/01/2024 0:01:21	bagus, sediakan data perjam (paket warnet) dong yang murah
9091	22/01/2024 23:39:48	Sangat Bagus dan mudah digunakan

4.2 Preprocessing Data

Data yang telah di kumpul selanjutnya akan dilakukan preprocessing data yang dimulai dari cleaning, case folding, tokenizing, stopword removal dan stemming. Hasil dari preprocessing data bisa di lihat di Tabel 2.

Tabel 2. Pengumpulan data

Tahapan	Hasil
Dataset	jujur saya sangat suka dengan provider satu ini.. tapi entah kenapa sekarang sering banyak kendala, mulai dari aplikasi yg terkadang suka crash, atau error.. terakhir saat saya melakukan pembelian kuota via e-wallet malah tidak masuk.. tolong untuk kualitasnya lebih diperbaiki lagi.. terimakasih...
Cleaning	jujur saya sangat suka dengan provider satu ini tapi entah kenapa sekarang sering banyak kendala mulai dari aplikasi yg terkadang suka crash atau error terakhir saat saya melakukan pembelian kuota via ewallet malah tidak masuk tolong untuk kualitasnya lebih diperbaiki lagi terimakasih
Case Folding	jujur saya sangat suka dengan provider satu ini tapi entah kenapa sekarang sering banyak kendala mulai dari aplikasi yg terkadang suka crash atau error terakhir saat saya melakukan pembelian kuota via ewallet malah tidak masuk tolong untuk kualitasnya lebih diperbaiki lagi terimakasih
Tokenizing	['jujur', 'saya', 'sangat', 'suka', 'dengan', 'provider', 'satu', 'ini', 'tapi', 'entah',

	'kenapa', 'sekarang', 'sering', 'banyak', 'kendala', 'mulai', 'dari', 'aplikasi', 'yg', 'terkadang', 'suka', 'crash', 'atau', 'error', 'terakhir', 'saat', 'saya', 'melakukan', 'pembelian', 'kuota', 'via', 'ewallet', 'malah', 'tidak', 'masuk', 'tolong', 'untuk', 'kualitasnya', 'lebih', 'diperbaiki', 'lagi', 'terimakasih']
Stopword Removal	['jujur', 'suka', 'provider', 'kendala', 'aplikasi', 'terkadang', 'suka', 'crash', 'error', 'pembelian', 'kuota', 'via', 'ewallet', 'masuk', 'tolong', 'kualitasnya', 'diperbaiki', 'terimakasih']
Stemming	['jujur', 'suka', 'provider', 'kendala', 'aplikasi', 'terkadang', 'suka', 'crash', 'error', 'beli', 'kuota', 'via', 'ewallet', 'masuk', 'tolong', 'kualitas', 'baik', 'terimakasih']

4.3 Labelling

Tahapan labelling dilakukan untuk mengklasifikasikan data ulasan ke dalam tiga kategori sentimen, yaitu positif, negatif, dan netral. Berikut data yang telah diberikan label bisa di lihat di Tabel 3.

Tabel 3. Labelling data

Dataset	Labelling
1 jujur saya sangat suka dengan provider satu ini.. tapi entah kenapa sekarang sering banyak kendala, mulai dari aplikasi yg terkadang suka crash, atau error.. terakhir saat saya melakukan pembelian kuota via e-wallet malah tidak masuk.. tolong untuk kualitasnya lebih diperbaiki lagi.. terimakasih...	Negatif
2 Aneh bgt, ngelag terus, kuota ga bisa dipake. Kacau bgt rugi, udah ngikutin cara2 nya tapi aplikasi suka error	Negatif
...	
9090 bagus, sediakan data perjam (paket warnet) dong yang murah	Positif
9091 Sangat Bagus dan mudah digunakan	Positif

4.4 Transformation Data

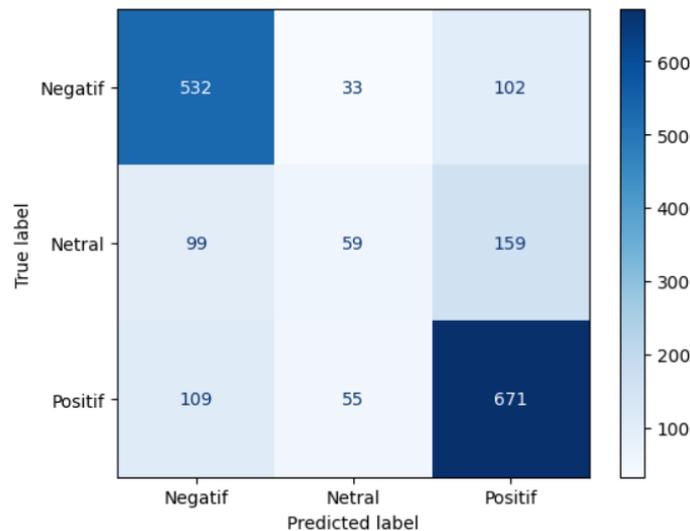
Setelah melalui proses preprocessing, data diberi bobot menggunakan pendekatan TF- IDF. Nilai bobot suatu kata akan semakin tinggi apabila kata tersebut sering muncul dalam suatu dokumen, namun tetap memperhitungkan tingkat kelangkaannya di seluruh kumpulan dokumen. Hasil transformasi TF-IDF ditunjukkan pada Tabel 4.

Tabel 4. Tranformation data

	aja	aplikasi	sinyal	beli	tolong
1	0	0	0	0	0
2	0	0.440696	0.168624	0	0.172188
3	0.162541	0.312519	0	0	0
4	0	0.282782	0.162302	0.153746	0
5	0	0.368912	0	0.200574	0

4.5 Evaluation

Penelitian ini menggunakan 9.091 data ulasan pengguna aplikasi by.U yang digunakan berjumlah 9091 data, yang telah diklasifikasikan ke dalam tiga kategori sentimen, yaitu positif sebanyak 4044 data, negatif sebanyak 3359 data, dan netral sebanyak 1688 data. Data ini kemudian diproses dan dianalisis menggunakan algoritma SVM untuk melakukan klasifikasi sentimen. Hasil dari proses klasifikasi ini akan dievaluasi menggunakan metrik seperti akurasi dan confusion matrix, guna mengetahui sejauh mana model mampu mengidentifikasi sentimen dengan tepat berdasarkan ulasan yang diberikan oleh pengguna. Confusion matrix dari hasil klasifikasi bisa di lihat pada Gambar 2 berikut:



Gambar 2. Confusion maxtrik

Confusion matrix pada Gambar menunjukkan bahwa model klasifikasi sentimen memiliki kinerja yang baik pada kelas Positif dan Negatif, namun lemah pada kelas Netral. Model berhasil mengklasifikasikan 532 dari 667 data Negatif dan 671 dari 835 data Positif dengan benar, tetapi hanya 59 dari 317 data Netral yang diklasifikasikan dengan tepat. Hal ini mengindikasikan bahwa model cenderung mengalami kebingungan dalam membedakan sentimen Netral, yang dapat menurunkan performa keseluruhan, khususnya dalam hal presisi dan recall pada kelas tersebut. Selanjutnya data dari confusion matrix tersebut dapat dimanfaatkan untuk menghitung metrik evaluasi seperti precision, recall, akurasi, spesifisitas, dan F1 Score. Hasil evaluasi dapat dilihat pada Tabel 5 berikut:

Tabel 5. Evaluasi Performa Klasifikasi

	Precision	Recall	F1-Score
Negatif	81%	92%	86%
Netral	88%	50%	64%
Positif	84%	90%	87%

Accuracy = 83%

Hasil klasifikasi sentimen ulasan aplikasi by.U menggunakan algoritma Support Vector Machine (SVM) menunjukkan akurasi sebesar 83%. Model mampu mengenali sentimen negatif dan positif dengan cukup baik. Untuk sentimen negatif, precision sebesar 81%, recall 92%, dan F1-score 86%, menunjukkan model sangat baik dalam mendeteksi ulasan negatif. Pada sentimen positif, precision 84%, recall 90%, dan F1-score 87%, menunjukkan performa yang seimbang dan akurat.

Namun, pada sentimen netral, meskipun precision tinggi (88%), recall-nya rendah 50% sehingga F1-score hanya 64%. Ini menunjukkan bahwa model kesulitan dalam mengenali sebagian besar data netral. Hal ini kemungkinan disebabkan oleh jumlah data netral yang lebih sedikit atau kemiripan kata dengan sentimen lain. Secara umum, SVM bekerja efektif untuk sentimen positif dan negatif pada data by.U.

4.6 Visualisasi Data

Visualisasi data sangat diperlukan untuk memperoleh serta menyampaikan tanggapan dari pengguna terhadap layanan aplikasi by.U. Tahap ini bertujuan untuk memperlihatkan persepsi dan opini pengguna yang diperoleh dari hasil klasifikasi sentimen, sekaligus mendukung pengembang dalam memahami serta meningkatkan fitur-fitur yang telah disediakan pada aplikasi. Visualisasi dilakukan dengan pendekatan word cloud, dapat menyajikan kata-kata dengan frekuensi kemunculan tertinggi dalam ulasan pengguna pada masing-masing jenis sentimen, yaitu positif, netral, dan negatif. Hasil visualisasi sentimen dapat dilihat pada Gambar 3 di bawah ini.

Referensi

- [1] J. Kristiyono, "Budaya internet: Perkembangan Teknologi Informasi dan Komunikasi dalam mendukung Penggunaan Media di Masyarakat," *Scriptura*, Vol. 5, No. 1, hlm. 23–30, 2015.
- [2] A. W. Lestari, A. Setiadarma, dan W. Prihatiningsih, "Pengaruh *Social Media Content* Instagram@ BYU. ID terhadap Brand Image Provider Digital BY. U," *Ikon–Jurnal Ilmiah Ilmu Komunikasi*, Vol. 27, No. 3, hlm. 223–234, 2022. <https://doi.org/10.37817/ikon.v27i3.2610>
- [3] A. L. Hananto, M. Guntur, T. Tukino, dan B. Priyatna, "Perencanaan Arsitektur *Enterprise* pada Layanan Telekomunikasi Digital Telkomsel By. U menggunakan Zachman Framework," *Jurnal Accounting Information System (AIMS)*, Vol. 7, No. 1, hlm. 9–16, 2024. <https://doi.org/10.32627/aims.v7i1.916>
- [4] H. R. Ginaris, A. Pratama, dan E. M. Safitri, "Analisis Kepuasan Pengguna Aplikasi Mobile By. U menggunakan Metode EUCS (*End User Computing Satisfaction*) pada Pengguna Provider By. U di Surabaya," *Global Leadership Organizational Research in Management*, Vol. 2, No. 1, hlm. 106–119, 2024.
- [5] I. A. Sapitri, Y. Yusra, dan M. Fikry, "Pengklasifikasian Sentimen Ulasan Aplikasi *Whatsapp* pada Google Play Store menggunakan *Support Vector Machine*," *Jurnal Tekinkom (Teknik Informasi dan Komputer)*, Vol. 6, No. 1, hlm. 1–7, 2023. <https://doi.org/10.37600/tekinkom.v6i1.773>
- [6] A. F. Firdaus dan W. I. Firdaus, "Text Mining dan Pola Algoritma dalam Penyelesaian Masalah Informasi:(Sebuah Ulasan)," *JUPITER (Jurnal Penelitian Ilmu dan Teknologi Komputer)*, Vol. 13, No. 1, hlm. 66–78, 2021.
- [7] M. Eiband, S. T. Völkel, D. Buschek, S. Cook, dan H. Hussmann, "A method and analysis to elicit user-reported problems in intelligent everyday applications," *ACM Transactions on Interactive Intelligent Systems (TIIS)*, Vol. 10, No. 4, hlm. 1–27, 2020. <https://doi.org/10.1145/3370927>
- [8] R. Vincent, I. Maulana, dan O. Komarudin, "Perbandingan Klasifikasi *Naive Bayes* Dan *Support Vector Machine* dalam Analisis Sentimen dengan *Multiclass* di *Twitter*," *JATI (Jurnal Mahasiswa Teknik Informatika)*, Vol. 7, No. 4, hlm. 2496–2505, 2023. <https://doi.org/10.30591/jpit.v6i3.2870>
- [9] S. Styawati, N. Hendrastuty, dan A. R. Isnain, "Analisis Sentimen Masyarakat terhadap Program Kartu Prakerja pada *Twitter* dengan Metode *Support Vector Machine*," *Jurnal Informatika: Jurnal Pengembangan IT*, Vol. 6, No. 3, hlm. 150–155, 2021.
- [10] H. Tuhuteru dan A. Iriani, "Analisis Sentimen Perusahaan Listrik Negara Cabang Ambon menggunakan Metode *Support Vector Machine* dan *Naive Bayes Classifier*," *Jurnal Informatika: Jurnal Pengembangan IT*, Vol. 3, No. 3, hlm. 394–401, Mei 2018, doi: 10.30591/jpit.v3i3.977.
- [11] P. A. Octaviani, Y. Wilandari, dan D. Ispriyanti, "Penerapan Metode Klasifikasi *Support Vector Machine* (SVM) pada Data Akreditasi Sekolah Dasar (SD) di Kabupaten Magelang," Vol. 3, No. 4, hlm. 811–820, 2014, [Daring]. Tersedia pada: <http://ejournal-s1.undip.ac.id/index.php/gaussian>
- [12] R. Astuti, R. A. Husen, A. Triono, M. K. Anam, P. S. T. I. S. A. Riau, dan J. K. M. P. Indah, "Peningkatan Metode *Support Vector Machines* (SVM) pada Data *Child-free* menggunakan *Oversampling*," 2023.
- [13] A. M. Puspitasari, D. E. Ratnawati, dan A. W. Widodo, "Klasifikasi Penyakit Gigi dan Mulut menggunakan Metode *Support Vector Machine*," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, Vol. 2, No. 2, hlm. 802–810, 2018.
- [14] F. A. K. Muhammad dan R. E. Putra, "Perbandingan Analisis Sentimen untuk Prediksi Kepuasan Pelanggan Kedai Kopi di Kofind menggunakan Algoritma SVM dan *Naive Bayes*," *Journal of Informatics and Computer Science (JINACS)*, hlm. 1039–1048, 2025.
- [15] D. Atika, A. A. Aldino, J. P. A. No, L. Ratu, dan K. Kedaton, "Term Frequency-Inverse Document Frequency *Support Vector Machine* untuk Analisis Sentimen Opini Masyarakat terhadap Tekanan Mental pada Media Sosial *Twitter*," *J. Teknol. dan Sist. Inf*, Vol. 3, No. 4, hlm. 86–97, 2022.

- [16] J. A. Septian, T. M. Fachrudin, dan A. Nugroho, “Analisis Sentimen Pengguna Twitter Terhadap Polemik Persepakbolaan Indonesia menggunakan Pembobotan TF-IDF dan *K-Nearest Neighbor*,” *INSYST: Journal of Intelligent System and Computation*, Vol. 1, No. 1, hlm. 43–49, 2019. <https://doi.org/10.52985/insyst.v1i1.36>
- [17] D. S. Utami dan A. Erfina, “Analisis Sentimen Pinjaman Online di *Twitter* menggunakan *Algoritma Support Vector Machine (SVM)*,” dalam *Prosiding Seminar Nasional Sistem Informasi dan Manajemen Informatika Universitas Nusa Putra*, 2021, hlm. 299–305.
- [18] N. Fitriyah, B. Warsito, dan I. M. Di Asih, “Analisis Sentimen Gojek pada Media Sosial *Twitter* dengan Klasifikasi *Support Vector Machine (SVM)*,” *Jurnal Gaussian*, Vol. 9, No. 3, hlm. 376–390, 2020.
- [19] B. Pamungkas, M. E. Purbaya, dan D. J. AK, “Analisis Sentimen *Twitter* menggunakan Metode *Support Vector Machine (SVM)* pada Kasus Benih Lobster 2020,” *Journal of Informatics Information System Software Engineering and Applications (INISTA)*, Vol. 3, No. 2, hlm. 10–20, 2021.
- [20] A. W. Nugroho, “Analisis Sentimen menggunakan *Algoritma Support Vector Machine* pada *Covid_19 Sentiment Analysis using the Support Vector Machine Algorithm on Covid_19*,” *SISTMASI*, 2024. [Daring]. Tersedia Pada: [Http://Sistemasi.Ftik.Unisi.Ac.Id](http://Sistemasi.Ftik.Unisi.Ac.Id)
- [21] I. W. Saputro dan B. W. Sari, “Uji Performa *Algoritma Naïve Bayes* untuk Prediksi Masa Studi Mahasiswa *Naïve Bayes Algorithm Performance Test for Student Study Prediction*,” *Citec Journal*, Vol. 6, No. 1, 2019.