



Analisis Sentimen pada Ulasan Aplikasi Access by KAI Berbahasa Indonesia Menggunakan *Word-Embedding* dan *Classical Machine Learning*

R. Damanhuri*, Vito Ahmad Husein

Departemen Informatika, Universitas Diponegoro, Semarang, Indonesia

* Corresponding author: rdamanhuri@students.undip.ac.id

Abstrak

Indonesia memiliki aplikasi perkeretaapian bernama Access by KAI yang dirilis oleh PT Kereta Api Indonesia (KAI). Masyarakat dapat mengunduh dan mengulas aplikasi ini melalui Google Play Store. Rating Access by KAI menurun dari tahun 2022, menandakan bahwa aplikasi belum memenuhi ekspektasi pengguna meskipun telah diperbarui. Ulasan pada platform Google Play Store dapat dianalisis untuk menggali informasi penting, salah satunya adalah sentimen. Penelitian ini melakukan analisis sentimen pada ulasan Access by KAI menggunakan word embedding dengan model Word2Vec untuk ekstraksi fitur dan classical machine learning dengan Naive Bayes dan Logistic Regression untuk algoritma klasifikasi. Metode Logistic Regression lebih baik daripada Naive Bayes dalam hal accuracy dan precision dengan nilai sebesar 68.83% dan 75.49% secara berurutan. Namun, metode Naive Bayes memiliki keunggulan dalam hal recall dengan nilai sebesar 45.07%. Pada penelitian ini, ulasan Access by KAI memiliki sentimen dominan negatif sejumlah 334 data dari total 400 data uji. Kata “mudah” dan “suka” relevan sebagai alasan ulasan bersentimen positif, sedangkan kata “aplikasi”, “bayar”, dan “tiket” relevan sebagai alasan ulasan bersentimen negatif.

Kata kunci : Analisis sentimen, ulasan, access by kai, word embedding, machine learning

Abstract

Indonesia has a railway application called Access by KAI, released by PT Kereta Api Indonesia (KAI). The public can download and review this application through the Google Play Store. The rating of Access by KAI has declined since 2022, indicating that the application has not met user expectations despite being updated. Reviews on the Google Play Store platform can be analyzed to extract important information, one of which is sentiment. This research conducts sentiment analysis on Access by KAI reviews using word embedding with the Word2Vec model for feature extraction and classical machine learning with Naive Bayes and Logistic Regression for classification algorithms. The Logistic Regression method outperforms Naive Bayes in terms of accuracy and precision with values of 68.83% and 75.49% respectively. However, the Naive Bayes method has an advantage in terms of recall with a value of 45.07%. In this study, reviews of Access by KAI have a predominantly negative sentiment, with 334 out of 400 test data reflecting this sentiment. The words "easy" and "like" are relevant as reasons why reviews have positive sentiment, while the words "application", "pay", and "ticket" are relevant as reasons why reviews have negative sentiment.

Keywords : Sentiment analysis, review, access by kai, word embedding, machine learning

1 Pendahuluan

Penggunaan aplikasi *mobile* telah merambah berbagai sektor kehidupan, termasuk transportasi. Pada sektor ini, tepatnya terkait perkeretaapian, Indonesia memiliki aplikasi *mobile* bernama Access by KAI yang dirilis oleh PT Kereta Api Indonesia (KAI). Inovasi ini didasarkan pada kebutuhan masyarakat Indonesia yang tinggi terhadap moda kereta api. Badan Pusat Statistik atau BPS mencatat bahwa jumlah penumpang kereta api di Indonesia sebanyak 277,12 juta orang sepanjang 2022 [1]. Jumlah tersebut meningkat 85,04% dibandingkan dengan tahun sebelumnya yang sebanyak 149,76 juta orang. Dengan Access by KAI, layanan dan informasi perkeretaapian Indonesia dapat diakses oleh calon penumpang dari mana pun melalui ponsel. Inovasi PT KAI ini terus ditingkatkan melalui pembaruan untuk memperoleh kepuasan dan kepercayaan pengguna. Pembaruan terakhir yang dirilis menghadirkan fitur-fitur baru sehingga pengguna dapat menikmati layanan yang makin lengkap melalui Access by KAI. Masyarakat dapat mengunduh aplikasi ini melalui platform distributor aplikasi seperti Google Play Store. Mereka juga dapat memberikan penilaian dan ulasan di sana.

Saat ini Access by KAI meraih *rating* rendah di Google Play Store, yakni 2,3 dari 5. Penilaian ini bahkan menurun dari perolehan pada 2022 yang sebesar 3,5 [2]. *Rating* yang menurun di sini menandakan bahwa aplikasi belum mampu memenuhi ekspektasi pengguna secara utuh meskipun telah diperbarui. Hal ini berpotensi memengaruhi citra dan kepercayaan terhadap layanan KAI secara keseluruhan. Penyebab penurunan *rating* pada Google Play Store dapat dipelajari dengan menganalisis ulasan yang diberikan oleh pengguna. Ulasan tersebut dapat menjadi sumber informasi berharga mengenai pengalaman pengguna dan kualitas layanan dan dimanfaatkan oleh penyedia layanan sebagai parameter kepuasan pengguna [3]. Terdapat banyak analisis yang dapat dilakukan untuk menggali informasi penting dari ulasan. Analisis sentimen adalah salah satunya.

Analisis sentimen adalah teknik mengidentifikasi dan menyimpulkan kecenderungan sentimen dan intensitas sentimen dalam teks secara otomatis [4]. Pada ulasan aplikasi, sentimen yang terkandung merupakan persepsi dan pendapat pengguna terhadap aplikasi. Dengan analisis sentimen, ulasan aplikasi akan dikelompokkan menjadi sentimen positif dan negatif, lalu ditelusuri kata dan frasa yang sering muncul untuk setiap sentimen. Pemrosesan kata pada analisis sentimen dapat dilakukan dengan teknik *word embedding* menggunakan Word2Vec. Word2Vec adalah model yang memanfaatkan korpus kata untuk menghasilkan keluaran vektor kata [5]. Kelebihan teknik ini adalah dapat mengatasi masalah sinonim dan homonim yang sering ditemukan dalam tugas NLP [3]. Selain pemrosesan kata, analisis sentimen juga memerlukan algoritma klasifikasi untuk menentukan label sentimen. *Classical machine learning* adalah teknik analisis sentimen yang umum digunakan, terutama untuk data berskala kecil, karena sederhana dan mudah dipahami (eksplanatif) [4]. *Classical machine learning* tersebut di antaranya adalah Naive Bayes yang memiliki keunggulan teoritis dalam hal konsistensi data dan komputasi klasifikasi serta Logistic Regression yang menghitung nilai probabilitas untuk menentukan label [6].

Penelitian ini bertujuan untuk menerapkan analisis sentimen pada ulasan Access by KAI di Google Play Store. Metode yang dipilih adalah *word embedding* menggunakan Word2Vec untuk ekstraksi fitur dan *classical machine learning* menggunakan Naive Bayes dan Logistic Regression untuk algoritma klasifikasi. Kontribusi penelitian ini ialah (1) penerapan gabungan metode ekstraksi fitur *word embedding* dan klasifikasi *classical machine learning* serta (2) penyajian visualisasi hasil analisis sentimen ulasan Access by KAI pada Google Play untuk sampel versi tertentu.

2 Penelitian Terkait

Sentimen analisis pada ulasan aplikasi di Google Play telah banyak dilakukan oleh penelitian-penelitian terdahulu dengan berbagai metode. Penelitian [2] menerapkan *lexicon-based* dengan InSet Lexicon pada analisis sentimen ulasan aplikasi Access by KAI di Google Play Store. Sentimen dibagi menjadi 3 kategori, yakni negatif, netral, dan positif. Penelitian ini tidak menggunakan algoritma *machine learning* sehingga setiap data tidak perlu label awal. Hasil penelitian menunjukkan ulasan didominasi sentimen negatif dengan jumlah 24.700 data dan persentase 60,01%. Visualisasi disajikan dalam bentuk *word cloud*, tetapi secara general. Tidak ada *word cloud* per sentimen sehingga pembahasan kurang mendalam.

Penelitian [5] melakukan analisis sentimen pada ulasan di Google Play untuk 4 aplikasi *marketplace*, yakni Shopee, Tokopedia, Lazada, dan Blibli. Algoritma *machine learning* berupa Naive Bayes dan Logistic Regression digunakan sebagai pengklasifikasi. Pelabelan data dilakukan berdasarkan *rating* ulasan (*rating-based labeling*). Penelitian ini memiliki dua skenario pengelompokan sentimen, yakni 2 kategori dan 3 kategori. Performa terbaik didapatkan dengan algoritma Naive Bayes pada skenario 2 kategori. Nilai *accuracy* mencapai 84.58%, *precision* 84.66%, dan *recall* 84.63%. Tidak ada visualisasi hasil sentimen.

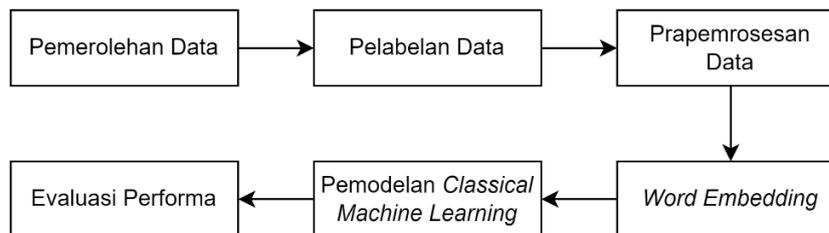
Penelitian [6] melakukan analisis sentimen pada ulasan aplikasi MyPertamina di Google Play Store. Penelitian ini menggunakan algoritma *machine learning* berupa Naive Bayes sebagai pengklasifikasi. Pelabelan data dilakukan berdasarkan *rating* ulasan (*rating-based labeling*). Sentimen dikelompokkan menjadi dua kategori, yaitu positif dan negatif. Hasil analisis menunjukkan bahwa sentimen ulasan aplikasi MyPertamina cenderung bersifat negatif. Dengan menggunakan algoritma Naive Bayes, diperoleh nilai *accuracy* sebesar 87%, *precision* 86%, dan *recall* 90%. Tidak ada visualisasi hasil sentimen.

Penelitian [7] membandingkan kinerja algoritma *classical machine learning* dan *deep learning* pada ulasan hotel situs Traveloka. Algoritma yang diuji meliputi Logistic Regression, Naive Bayes, Support Vector Machine, dan Convolutional Neural Network. Ekstraksi fitur *word embedding* diterapkan dengan model Word2Vec. Pelabelan data dilakukan secara manual dengan sentimen dikelompokkan menjadi dua kategori. Logistic Regression memberikan hasil terbaik untuk *classical machine learning* dengan nilai *accuracy* sebesar 54,4%, *precision* 55%, dan *recall* 54%. Namun, model tersebut belum mampu melampaui performa dari Convolutional Neural Network.

Penelitian ini akan berfokus pada analisis sentimen ulasan Access by KAI di Google Play seperti [2], tetapi dengan menerapkan *classical machine learning*. Algoritma yang digunakan meliputi Naive Bayes dan Logistic Regression. Ekstraksi fitur dengan *word embedding* akan diterapkan seperti pada penelitian [7]. Kategori sentimen yang digunakan adalah positif dan negatif, dengan pelabelan data awal menggunakan gabungan metode *manual labeling* seperti [7] serta *rating-based* seperti [6] dan [5]. Interpretasi akan disajikan melalui visualisasi *word cloud* untuk setiap sentimen.

3 Metode Penelitian

Alur metodologi penelitian diawali dengan pengumpulan, pelabelan dan pemrosesan data. Kemudian dilanjutkan dengan proses word embedding dan pelatihan model serta evaluasi performa seperti yang ditunjukkan pada Gambar 1.



Gambar 1 Alur Metodologi Penelitian

3.1 Pemerolehan Data

Data ulasan aplikasi Access by KAI diperoleh menggunakan metode *scraping* pada Google Play Store. Hasil *scraping* diseleksi hingga terdapat 1000 baris data dan tiga fitur, yakni ulasan, *rating*, dan versi aplikasi. Ulasan terlama bertanggal 08 September 2023, sedangkan ulasan terbaru bertanggal 04 Desember 2023. Data kemudian dibagi menjadi data latih dan uji dengan rasio 0,6: 0,4 dan ketentuan data uji harus ulasan untuk aplikasi versi tertentu yang dipilih, yakni 6.0.40.

3.2 Pelabelan Data

Pelabelan data hanya dilakukan untuk data latih. Data latih yang sebanyak 600 data dibagi menjadi dua bagian sama banyak untuk pelabelan data. Setiap bagian dilabeli dengan metode yang berbeda, yakni *manual labeling* dan *rating-based labeling*. Label yang digunakan berjumlah dua, yakni positif dan negatif. *Manual labeling* dilakukan oleh 2 orang, sedangkan *rating-based labeling* dilakukan dengan ketentuan *rating* 4–5 sebagai label positif dan *rating* 1–3 sebagai label negatif.

3.3 Prapemrosesan Data

Prapemrosesan data pada penelitian ini meliputi *case folding*, *cleansing*, tokenisasi, *stopword removal*, dan *stemming*.

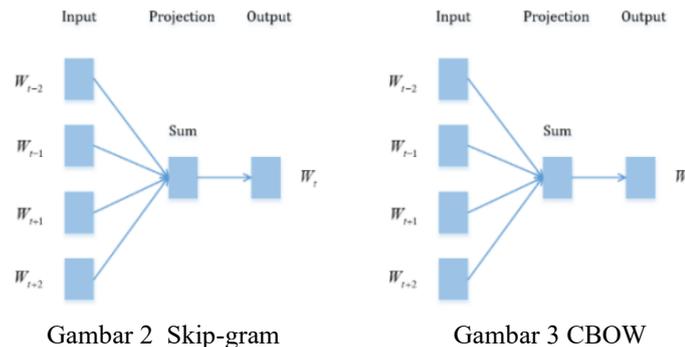
1. *Case folding* adalah proses menyeragamkan seluruh teks ke dalam bentuk huruf kecil agar kata-kata yang sama tetapi memiliki penulisan kapitalisasi berbeda tidak dianggap sebagai kata yang berbeda.
2. Tahap *cleansing* adalah proses membersihkan kalimat dari tanda baca, angka, karakter khusus, dan spasi berlebih.
3. Tokenisasi merupakan proses memecah sebuah kalimat menjadi sekelompok kata yang tersimpan di dalam *array*.
4. *Stopword removal* adalah tahap menghapus kata yang umum muncul, tetapi tidak bermakna atau *stopword*. Pada penelitian ini, daftar *stopword* diambil dari pustaka NLTK [8].
5. *Stemming* merupakan proses mengubah kata berimbuhan menjadi kata dasar. Pada penelitian ini, pustaka Sastrawi [9] digunakan sebagai alat untuk melakukan *stemming*.

3.4 Word Embedding

Word embedding adalah proses memetakan kata-kata ke ruang vektor sehingga menjadi representasi yang mengandung informasi semantik kata [4]. Kata-kata yang mirip secara semantik akan berada lebih dekat satu sama lain di dalam ruang vektor tersebut. Proses *word embedding* dapat dilakukan menggunakan Word2Vec [10]. Ini merupakan algoritma representasi vektor kata

berdasarkan *shallow neural network* yang umum digunakan [4] dan mampu mencapai performa terbaik pada NLP dengan mengelompokkan kata serupa [3].

Word2Vec memiliki dua jenis arsitektur, yaitu model Skip-Gram dan model Continuous Bag of Words (CBOW). Model Skip-gram melakukan prediksi pada rentang sebelum atau sesudah kata saat ini yang masukannya berasal dari kata saat ini, sedangkan model CBOW memprediksi kata saat ini berdasarkan kata konteks [6]. Arsitektur Skip-gram dapat menghasilkan akurasi yang sangat baik untuk kata-kata yang jarang muncul sehingga cocok untuk analisis sentimen dengan representasi kata yang umumnya unik [3]. Jaringan model Skip-gram dan CBOW dapat dilihat pada Gambar 2 dan Gambar 3 secara berturut-turut.



Terdapat dua metode evaluasi pada Word2Vec, yaitu *Hierarchical Softmax* dan *Negative Sampling*. *Hierarchical Softmax* menggunakan pohon biner untuk mewakili seluruh kata di dalam kosakata. Simpul daun akan mewakili kata-kata langka sehingga kata-kata yang jarang muncul mewarisi representasi vektor di dalamnya [3]. Evaluasi ini cocok digunakan untuk analisis sentimen. Penelitian ini akan menerapkan model Word2Vec dengan arsitektur skip-gram dan metode evaluasi *hierarchical softmax*. Adapun dimensi vektor yang dipilih adalah 100 karena data penelitian yang relatif sedikit.

3.5 Pemodelan Classical Machine Learning

3.5.1 Logistic Regression

Logistic Regression adalah salah satu algoritma klasifikasi dalam ranah *machine learning* yang digunakan untuk memprediksi probabilitas variabel dependen yang bersifat kategorikal. Metode ini merupakan bentuk umum dari regresi linier yang diterapkan pada konteks pembelajaran mesin, yang fokusnya adalah memahami hubungan antara beberapa variabel dengan variabel biner atau variabel probabilistik [5]. Metode klasifikasi Logistic Regression menggunakan fungsi logistik untuk memodelkan probabilitas dari berbagai kelas dalam dataset yang diberikan. Algoritma ini menghasilkan prediksi dalam bentuk probabilitas dan setelahnya menggunakan ambang batas tertentu untuk mengklasifikasikan data ke dalam kelas positif, negatif, atau netral.

3.5.2 Naive Bayes

Penelitian ini akan melibatkan penggunaan algoritma Naive Bayes untuk pengolahan data. Naive Bayes adalah suatu metode klasifikasi probabilitas yang sederhana, menggunakan model statistik untuk mengestimasi probabilitas yang sederhana, menggunakan model statistik untuk mengestimasi probabilitas setiap kelas berdasarkan atribut yang teramati. Label kelas diambil dari himpunan terbatas,

dan metode ini memiliki beberapa keunggulan, seperti perhitungan yang mudah, tingkat presisi yang tinggi, dan kecepatan pemrosesan yang baik, terutama saat menangani basis data besar [6]. Rumus untuk Naive Bayes dapat dilihat pada Persamaan (1):

$$P(c|X) = \frac{P(x|c)P(c)}{P(x)} \quad (1)$$

Keterangan:

- c : Hipotesis data X merupakan suatu class spesifik
- X : Data dengan class yang belum diketahui
- $P(c|X)$: Probabilitas hipotesis berdasarkan kondisi X (*posteriori probability*)
- $P(x|c)$: Probabilitas hipotesis x berdasarkan kondisi pada hipotesis c
- $P(x)$: Probabilitas c
- $P(c)$: Probabilitas hipotesis c (*prior probability*)

3.6 Evaluasi Performa

Metrik evaluasi yang digunakan pada penelitian ini meliputi *accuracy*, *precision*, dan *recall*. Rumus *accuracy* dapat dilihat pada Persamaan (2), *precision* pada Persamaan (3), dan *recall* pada Persamaan (4). Pada penelitian ini, ketiga evaluasi dikenakan hanya untuk hasil dengan data latih karena tidak adanya data validasi.

$$accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (2)$$

$$precision = \frac{TP}{TP+FP} \quad (3)$$

$$recall = \frac{TP}{TP+FN} \quad (4)$$

Keterangan:

- TP : Jumlah data positif yang diprediksi dengan benar
- TN : Jumlah data negatif yang diprediksi dengan benar
- FP : Jumlah data negatif yang diprediksi sebagai positif
- FN : Jumlah data positif yang diprediksi sebagai negatif

3.7 Visualisasi

Analisis visual untuk setiap sentimen yang dihasilkan oleh algoritma terbaik dilakukan menggunakan *word cloud*. *Word cloud* merupakan teknik visualisasi dengan menampilkan kata-kata yang paling sering muncul dalam data dan pada kasus ini mewakili suatu sentimen secara visual [8].

4 Hasil dan Pembahasan

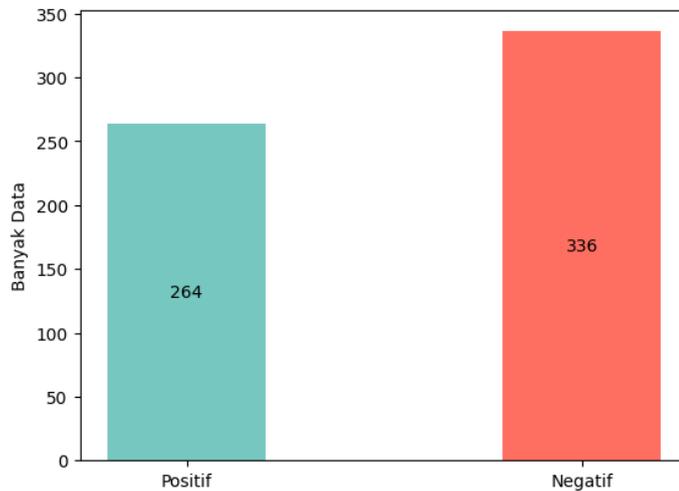
Pada tahap pemerolehan data, hasil akhir yang didapatkan berupa data mentah dengan fitur *content*, *score*, dan *appVersion* yang belum dilabeli. Tabel 1. menampilkan contoh data hasil tahap pemerolehan data.

Tabel 1 Contoh data hasil tahap pemerolehan data

<i>content</i>	<i>score</i>	<i>appVersion</i>
sering error. Jadwal kosong terus	1	6.0.31

Pada tahap pelabelan data, fitur *sentiment* ditambahkan pada data dengan isian berupa label sentimen yang ditentukan dengan dua metode, yakni *manual labeling* dan *rating-based labeling*.

Kedua metode digunakan pada dua bagian yang sama banyak sehingga dari 600 data latih, *manual labeling* dilakukan pada 300 data dan *rating-based labeling* dilakukan pada 300 data. Distribusi akhir label pada data latih ditunjukkan pada Gambar 4, sedangkan contoh data hasil tahap pelabelan data dapat dilihat pada Tabel 2.



Gambar 4 Distribusi Label Data Latih

Tabel 2 Contoh data hasil tahap pelabelan data

<i>content</i>	<i>score</i>	<i>appVersion</i>	<i>sentiment</i>
sering error. Jadwal kosong terus	1	6.0.31	<i>negative</i>

Pada tahap prapemrosesan data, hasil akhir berupa data yang telah dikenai metode-metode prapemrosesan, yakni *case folding*, *cleansing*, tokenisasi, *stop-word removal*, dan *stemming*. Perubahan data dalam tahap ini dapat dilihat pada Tabel 3.

Tabel 3 Perubahan data dalam prapemrosesan data

<i>content</i>	<i>casefolding</i>	<i>cleansing</i>	<i>tokenisasi</i>	<i>stop-word removal</i>	<i>stemming</i>
sering error. Jadwal kosong terus	sering error. jadwal kosong terus	sering error jadwal kosong terus	[sering, error, jadwal, kosong, terus]	[error, jadwal, kosong]	[error, jadwal, kosong]

Pada tahap *word embedding*, hasil akhir berupa nilai jumlahan representasi numerik dari data hasil *stemming* yang didapatkan dengan model Word2Vec. Tabel 4. menampilkan contoh data hasil tahap *word embedding*.

Tabel 4 Contoh data hasil tahap *word embedding*

<i>stemming</i>	<i>word embedding</i>
[error, jadwal, kosong]	0.220693

4.1 Skenario 1: Logistic Regression

Hasil evaluasi untuk skenario 1 dengan model Logistic Regression ditampilkan pada Tabel 5. Jumlah data pada TP, TN, FP, FN digunakan untuk menghitung *accuracy*, *precision*, dan *recall*. Nilai *accuracy* menunjukkan bahwa sekitar 68.83% dari keseluruhan data berhasil diklasifikasikan dengan tepat oleh model. Nilai *precision* menunjukkan bahwa 75.49% dari prediksi positif oleh model benar-

benar relevan. Sementara itu, nilai *recall* menunjukkan bahwa sekitar 43.18% dari keseluruhan kasus positif berhasil dideteksi oleh model.

Tabel 5 Hasil evaluasi model Logistic Regression

TP	TN	FP	FN	<i>accuracy</i>	<i>precision</i>	<i>recall</i>
114	299	37	150	0.688333	0.754967	0.431818

4.2 Skenario 2: Naive Bayes

Hasil evaluasi untuk skenario 2 dengan model Naive Bayes ditampilkan pada Tabel 6. Jumlah data pada TP, TN, FP, FN digunakan untuk menghitung *accuracy*, *precision*, dan *recall*. Nilai *accuracy* menunjukkan bahwa sekitar 68.33% dari keseluruhan data berhasil diklasifikasikan dengan tepat oleh model. Nilai *precision* menunjukkan bahwa sekitar 72.56% dari prediksi positif oleh model benar-benar relevan. Sementara itu, nilai *recall* menunjukkan bahwa sekitar 45.07% dari keseluruhan kasus positif berhasil dideteksi oleh model.

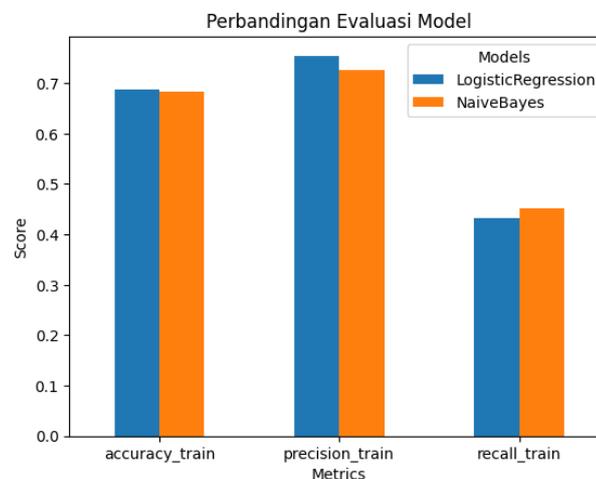
Tabel 6 Hasil evaluasi model Naive Bayes

TP	TN	FP	FN	<i>accuracy</i>	<i>precision</i>	<i>recall</i>
119	291	45	145	0.683333	0.725610	0.450758

4.3 Skenario 3: Perbandingan Hasil Skenario 1 dan 2

Hasil evaluasi untuk skenario 3 berupa perbandingan model Logistic Regression dan Naive Bayes ditampilkan pada Gambar 5. *Accuracy* tertinggi didapatkan oleh model Logistic Regression dengan nilai 0.688333. Ini berarti kemampuan klasifikasi data Logistic Regression secara keseluruhan lebih baik dibandingkan Naive Bayes. *Precision* Logistic Regression juga menunjukkan performa unggul dengan nilai 0.754967, mengindikasikan bahwa prediksi positif dari model ini lebih akurat dalam mencerminkan sentimen positif yang sebenarnya.

Sebaliknya, Logistic Regression tertinggal dari Naive Bayes pada metrik *recall* yang mendapatkan nilai sebesar 0.450758. Nilai ini menunjukkan bahwa Naive Bayes lebih sensitif dalam menangkap keseluruhan sentimen positif, meskipun prediksinya tidak seakurat Logistic Regression. Berdasarkan hasil tersebut, Logistic Regression dipilih sebagai model terbaik untuk memprediksi data uji karena unggul dalam *accuracy* dan *precision*. Konsistensi dan keakuratan prediksi Logistic Regression cocok dan penting untuk diterapkan pada penelitian ini.



Gambar 5 Perbandingan Evaluasi Model

4.4 Analisis Visual

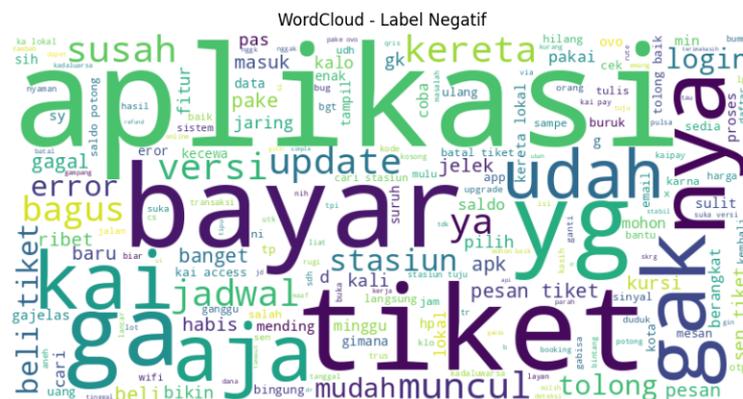
Data uji diprediksi dengan model Logistic Regression. Hasil menunjukkan bahwa sentimen negatif berjumlah 334 data dan positif berjumlah 66 data. Mayoritas ulasan aplikasi Access by KAI untuk versi aplikasi 6.0.40.

Analisis visual kemudian dilakukan untuk setiap label menggunakan WordCloud. Untuk label positif, kata yang sering muncul ditunjukkan oleh Gambar 6. Kata-kata tersebut antara lain “log”, “masuk”, “mudah”, “suka”, “kai”, dan “ga”. Kata “mudah” dan “suka” relevan sebagai alasan ulasan bersentimen positif. Adapun “log”, “masuk”, “kai”, dan “ga” tidak relevan.



Gambar 6 WordCloud Label Positif

Untuk label negatif, kata yang sering muncul ditunjukkan oleh Gambar 7. Kata-kata tersebut antara lain “aplikasi”, “tiket”, “bayar”, “ga”, “gak”, “yg”, “nya”, dan “kai”. Kata “aplikasi”, “bayar”, dan “tiket” relevan sebagai alasan ulasan bersentimen negatif. Adapun “ga”, “gak”, “yg”, “nya”, dan “kai” tidak relevan.



Gambar 7 WordCloud Label Negatif

5 Kesimpulan

Metode Logistic Regression lebih baik dibandingkan dengan Naive Bayes dalam hal *accuracy* dan *precision* dengan nilai sebesar 68.83% dan 75.49% secara berurutan pada ulasan Access by KAI. Namun, metode Naive Bayes memiliki keunggulan dalam hal *recall* dengan nilai sebesar 45.07%. Ulasan Access by KAI versi aplikasi 6.0.40 sebagai data uji pada penelitian ini memiliki sentimen dominan negatif sejumlah 334 data dari total 400 data. Kata “mudah” dan “suka” relevan sebagai

alasan ulasan bersentimen positif, sedangkan kata “aplikasi”, “bayar”, dan “tiket” relevan sebagai alasan ulasan bersentimen negatif.

Daftar Pustaka

- [1] Badan Pusat Statistik (BPS), *Jumlah Penumpang Kereta Api (Ribu Orang), 2022*, Badan Pusat Statistik (BPS), 2022.
- [2] R. D. Wahyuni and A. N. Utomo, "Using The Lexicon Method for Analysis Sentiments on KAI Access Application Reviews on Google Play Store," *Jurnal Rekayasa Informasi*, vol. 11, no. 2, pp. 134-145, 2 10 2022.
- [3] R. P. Nawangsari, R. Kusumaningrum and A. Wibowo, "Word2Vec for Indonesian Sentiment Analysis towards Hotel Reviews: An Evaluation Study," *Procedia Computer Science*, vol. 157, pp. 360-366, 2019. [doi: 10.1016/j.procs.2019.08.178](https://doi.org/10.1016/j.procs.2019.08.178)
- [4] Y. Jin, K. Cheng, X. Wang and L. Cai, "A Review of Text Sentiment Analysis Methods and Applications," *Frontiers in Business, Economics and Management*, vol. 10, no. 1, pp. 58-64, 2023. [doi: 10.54097/fbem.v10i1.10171](https://doi.org/10.54097/fbem.v10i1.10171)
- [5] R. Kusumaningrum, I. Z. Nisa, R. P. Nawangsari and A. Wibowo, "Sentiment analysis of Indonesian hotel reviews: from classical machine learning to deep learning," *International Journal of Advances in Intelligent Informatics*, vol. 7, no. 3, pp. 292-303, 30 11 2021. [doi: 10.26555/ijain.v7i3.737](https://doi.org/10.26555/ijain.v7i3.737)
- [6] S. A. H. Bahtiar, C. K. Dewa and A. Luthfi, "Comparison of Naïve Bayes and Logistic Regression in Sentiment Analysis on Marketplace Reviews Using Rating-Based Labeling," *Journal of Information Systems and Informatics*, vol. 5, no. 3, pp. 915-927, 2023. [doi: 10.51519/journalisi.v5i3.539](https://doi.org/10.51519/journalisi.v5i3.539)
- [7] R. Maulana, A. Voutama and T. Ridwan, "Analisis Sentimen Ulasan Aplikasi MyPertamina pada Google Play Store Menggunakan Algoritma NBC," *Jurnal Teknologi Terpadu*, vol. 9, no. 1, pp. 42-48, 2023. [doi: 10.54914/jtt.v9i1.609](https://doi.org/10.54914/jtt.v9i1.609)
- [8] Bird, Steven, E. Loper and E. Klein, *Natural Language Processing with Python*, O'Reilly Media Inc, 2019.
- [9] A. Librarian, "High Quality Stemmer Library for Indonesian Language (Bahasa)," 2017. [Online]. Available: <https://github.com/sastrawi>
- [10] T. Mikolov, G. Corrado, K. Chen and J. Dean, "Efficient Estimation of Word Representations in Vector Space," Arizona, 2013. [doi: 10.48550/arXiv.1301.3781](https://doi.org/10.48550/arXiv.1301.3781)
- [11] L. E. Pradana and Y. Ruldeviyani, "Sentiment Analysis of Nanovest Investment Application Using Naive Bayes Algorithm," *Jurnal Nasional Pendidikan Teknik Informatika*, vol. 12, no. 2, pp. 283-293, 2023. [doi: 10.23887/janapati.v12i2.62302](https://doi.org/10.23887/janapati.v12i2.62302)