Walisongo Journal of Information Technology
Vol .., No.. (Year): 00-00. DOI: 10.21580.wjit.2024.6.1.32686
ISSN 2715-0143 (online) | 2714-9048 (printed)

Analisis Sentimen Ulasan Pengguna Aplikasi KBBI pada Google Play Store Menggunakan Metode Naïve Bayes dengan Pembobotan TF-IDF

Via Latifatul Habibah[1]*, Maya Rini Handayani[2]
[12]Universitas Islam Negeri Walisongo Semarang
Corresponding Author : maya@walisongo.ac.id

Abstract: *User reviews on the Google Play Store contain a wide range of feedback regarding app usage. In the case of the Kamus Besar Bahasa Indonesia app, these comments reflect a variety of opinions, making it important to identify sentiment in order to understand user feedback. This study processed 2,319 reviews from the Google Play Store through a series of text preprocessing steps such as data cleaning, tokenization, stop-word removal, stemming, and text reconstruction. Subsequently, the text was converted into a numerical representation using Term Frequency–Inverse Document Frequency (TF-IDF) and classified using the Naïve Bayes algorithm. Testing showed an accuracy of 87.5%. However, the model performed better on positive reviews, with precision, recall, and F1-score values of 0.95, 0.91, and 0.93, respectively, compared with 0.35, 0.46, and 0.40 for negative reviews. This difference indicates that the model was less effective in identifying negative sentiment, potentially due to the dominance of positive reviews in the dataset. These findings suggest that the proposed approach can be used to map user sentiment, although its performance in classifying negative reviews requires further improvement.*

Keyword : *Google Play Store; Kamus Besar Bahasa Indonesia; Naïve Bayes; sentiment analysis; Term Frequency–Inverse Document Frequency*

Copyright ©2024 Walisongo Journal of Information Technology

Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.

1. Pendahuluan

Aplikasi KBBI di Google Play Store memiliki banyak ulasan pengguna yang beragam tentang layanan yang diberikannya. Ulasan pengalaman pengguna dapat tercerminkan terkait akurasi pencarian kata, kemudahan penggunaan, serta stabilitas sistem. Karena banyaknya ulasan, mengidentifikasi opini secara manual menjadi tidak efisien. Akibatnya, analisis sentimen digunakan untuk mengelompokkan opini pengguna secara otomatis (Darmawan et al., 2023; Zhafira et al., 2021).

KBBI memiliki fungsi sebagai kamus digital, tetapi juga menjadi sumber rujukan bahasa yang membantu pengguna memahami kosakata, makna kata, serta penggunaan bahasa Indonesia yang sesuai dengan kaidah. Bukti pemanfaatan KBBI digital melalui dukungan proses pembelajaran bahasa dan meningkatkan pemahaman kosakata pada pengguna (Apriandini et al., 2026). Kegunaan lain selain sebagai pencarian arti kata, aplikasi KBBI juga berperan sebagai

rujukan utama dalam menentukan kata, ejaan, dan makna sehingga banyak dimanfaatkan dalam kegiatan akademik maupun penulisan ilmiah (Syifani et al., 2025). Keberadaan KBBI berbasis digital memberikan kemudahan akses terhadap informasi kebahasaan yang diperlukan pengguna. Pemanfaatan laman KBBI daring dilaporkan mampu membantu pengguna dalam membedakan bentuk baku dan tidak baku serta mendukung menggunakan bahasa Indonesia dengan cara yang sesuai dengan standar (Purba & Aritonang, 2024).

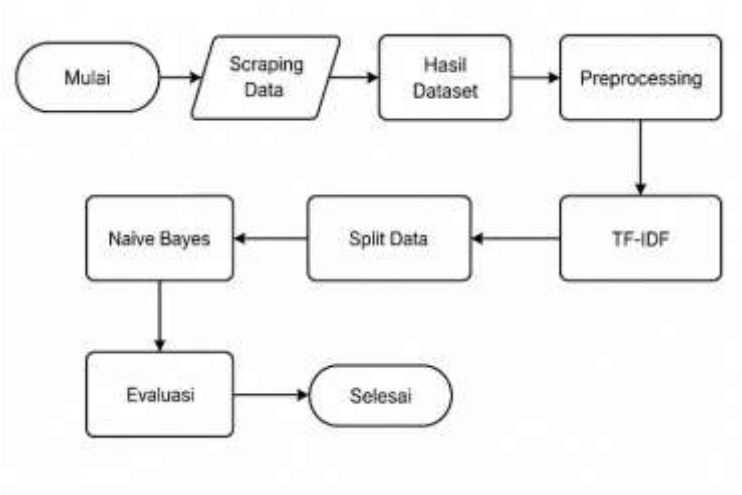
Banyak penelitian analisis sentimen yang diterapkan pada ulasan aplikasi dan media sosial. Penelitian sebelumnya dilakukan pada aplikasi mypertamina (Darmawan et al., 2023), aplikasi Identitas Kependudukan Digital (Nugroho et al., 2025), komentar youtube terkait Kebijakan Kampus Merdeka (Zhafira et al., 2021), serta opini masyarakat mengenai isu kenaikan BIPIH (Wati et al., 2023). Pada penelitian ini algoritma Naïve Bayes menunjukkan mampu melakukan klasifikasi yang baik terhadap teks berbahasa Indonesia. Pada penelitian terdahulu sebagian besar berfokus pada layanan publik, media sosial, dan aplikasi pembayaran digital. Penelitian yang mengkaji sentimen pengguna terhadap aplikasi KBBI masih terbatas. Karakteristik ulasan pada aplikasi KBBI berpotensi menghasilkan pola sentimen yang berbeda karena berkaitan dengan kebutuhan kebahasaan pengguna.

Penggunaan aplikasi KBBI dinilai mampu membantu pengguna menemukan bentuk kata baku dan meningkatkan ketepatan penggunaan bahasa Indonesia. Dengan itu, layanan aplikasi ini harus terus diperiksa berdasarkan pengalaman pengguna melalui ulasan di Google Play Store (Purba & Aritonang, 2024). KBBI adalah aplikasi rujukan resmi untuk bahasa Indonesia yang digunakan oleh siswa, mahasiswa, guru, peneliti, dan orang umum. Kualitas layanan yang ditawarkan oleh aplikasi tersebut memengaruhi seberapa mudah bagi pengguna untuk mendapatkan informasi bahasa yang tepat. Hal itu menjelaskan mengetahui sentimen penggguna merupakan suatu kepentingan bagi pengguna terhadap tanggapan kualitas layanan aplikasi. Selain itu, temuan penelitian dapat digunakan sebagai pedoman untuk studi analisis sentimen pada aplikasi pendidikan dan kebahasaan yang masih relatif terbatas. Penelitian ini memilih Metode Naïve Bayes karena proses klasifikasi yang sederhana dan efektif untuk data teks. Kualitas representasi dokumen ditingkatkan melalui pembobotan TF-IDF untuk menemukan kata yang memiliki kontribusi tinggi terhadap proses klasifikasi. Dipilihnya metode Naïve Bayes dengan pembobotan TF-IDF digunakan untuk menganalisis sentimen ulasan pengguna terhadap aplikasi KBBI di Google Play Store.

2. Metode

Penelitian ini dilaksanakan melalui serangkaian tahapan yang saling terhubung untuk memperoleh model klasifikasi sentimen ulasan aplikasi KBBI. Tahapan penelitian meliputi akuisisi data, pelabelan sentimen, *preprocessing* teks, ekstraksi fitur menggunakan TF-IDF, pembentukan model Naïve Bayes, dan pengukuran kinerja model. Alur tersebut mengacu pada tahapan text mining yang banyak digunakan pada penelitian analisis sentimen berbasis ulasan aplikasi (Darmawan et al., 2023; Hanafi & R, 2024).

Gambar 2.1 Diagram Alur Penelitian



2.1 Pengumpulan Data

Sumber data penelitian berupa ulasan pengguna aplikasi KBBI yang tersedia pada Google Play Store. Proses ekstraksi data dilakukan menggunakan Google Play Scraper yang dijalankan pada lingkungan Google Colaboratory. Data yang berhasil dikumpulkan terdiri atas teks ulasan dan nilai rating yang diberikan pengguna. Pemanfaatan teknik scraping memungkinkan pengambilan data dilakukan secara otomatis dalam jumlah besar sehingga sesuai untuk kebutuhan analisis sentimen (Darmawan et al., 2023; Hanafi & R, 2024).

2.2 Labeling Data

Pelabelan data dilakukan menggunakan atribut rating yang disertakan dengan setiap ulasan pengguna. Karena ulasan menunjukkan tingkat kepuasan pengguna terhadap aplikasi, mereka dikategorikan ke dalam kategori positif dan negatif berdasarkan nilai penilaian mereka. Data hasil pelabelan digunakan sebagai target klasifikasi pada proses analisis sentimen. Pendekatan tersebut banyak diterapkan pada penelitian analisis sentimen ulasan aplikasi karena mampu menghasilkan label yang konsisten dan mudah direproduksi (Apriliyani et al., 2024; Darmawan et al., 2023).

2.3 *Preprocessing* Data

Data ulasan terlebih dahulu dipersiapkan melalui tahap preprocessing sebelum digunakan pada proses klasifikasi. Tahap ini dilakukan untuk menyederhanakan bentuk teks serta mengurangi unsur yang tidak memberikan kontribusi terhadap pembentukan sentimen. Rangkaian proses yang diterapkan meliputi *cleaning text*, *tokenizing*, *stopword removal*, *stemming*, dan *rejoin text* (Pranata et al., 2024; Pratama et al., 2026).

2.3.1 *Cleaning Text*

Tahap *cleaning text* digunakan untuk menyaring elemen yang tidak berkaitan dengan isi ulasan, seperti URL, tanda baca, angka, emoji, karakter khusus, dan simbol lain yang tidak diperlukan. Hasil penyaringan menghasilkan teks yang lebih seragam sehingga proses pengolahan pada tahap berikutnya dapat dilakukan dengan lebih optimal (Pranata et al., 2024).

2.3.2 *Tokenizing*

Pada tahap *tokenizing*, setiap ulasan diuraikan menjadi satuan kata yang berdiri sendiri. Setiap token diperlakukan sebagai unit analisis yang akan digunakan pada proses pengolahan teks berikutnya (Zhafira et al., 2021).

2.3.3 *Stopword Removal*

Tahap *stopword removal* diterapkan untuk mengurangi kemunculan kata-kata umum yang tidak memiliki pengaruh langsung terhadap penentuan sentimen. Proses ini membantu mempertahankan kata yang lebih informatif dengan menghilangkan kata seperti "dan", "yang", "di", dan "ke" sehingga representasi fitur yang terbentuk menjadi lebih relevan untuk proses klasifikasi (Pranata et al., 2024).

2.3.4 *Stemming*

Tahap *stemming* diterapkan untuk mengembalikan kata berimbuhan ke bentuk dasarnya. Penyederhanaan bentuk kata dilakukan agar variasi kata dengan makna yang sama dapat direpresentasikan dalam satu bentuk yang seragam. Algoritma *stemming* bahasa Indonesia digunakan karena mampu meningkatkan kualitas pengolahan dokumen teks berbahasa Indonesia (Pardede & Darmawan, 2025).

2.3.5 *Rejoining Text*

Token hasil *preprocessing* digabungkan kembali menjadi dokumen teks utuh. Hasil *rejoin text* digunakan sebagai masukan pada proses pembobotan TF-IDF.

2.4 Pembagian Data

Setelah proses ekstraksi fitur selesai dilakukan, data dipisahkan menjadi dua kelompok menggunakan skema *train-test split*. Dataset dibagi menjadi dua kelompok data dengan perbandingan 80:20 setelah proses transformasi fitur usai dilakukan untuk kebutuhan pelatihan dan pengujian model. Pengaturan tersebut memungkinkan pengukuran kinerja model dilakukan menggunakan data yang tidak terlibat selama proses pelatihan (Wati et al., 2023).

2.5 Evaluasi Model

Pengukuran kinerja model dilakukan menggunakan *confusion matrix*, *accuracy*, *precision*, *recall*, dan *F1-score*. Metrik tersebut digunakan untuk melihat tingkat ketepatan prediksi yang dihasilkan model terhadap data pengujian. Hasil evaluasi memberikan gambaran mengenai kemampuan model dalam membedakan ulasan positif dan negatif pada aplikasi KBBI (Apriliyani et al., 2024; Darmawan et al., 2023).

3. Hasil dan Pembahasan

3.1 Hasil *Preprocesing* Data

Preprocessing dilakukan melalui tahapan *cleaning text, tokenizing, stopword removal, stemming,* dan *rejoin text*. Tahapan tersebut menghasilkan data yang lebih terstruktur sehingga dapat digunakan pada proses pembobotan fitur dan klasifikasi sentimen. Penerapan *preprocessing* bertujuan mengurangi unsur yang tidak relevan pada dokumen serta meningkatkan kualitas representasi teks (Zhafira et al., 2021).

Contoh keluaran *preprocessing* ditunjukkan pada Gambar 3.1 dalam bentuk teks yang telah melalui proses pembersihan dan normalisasi. Teks hasil pengolahan kemudian disimpan pada kolom *final_teks* dan digunakan pada tahap pembobotan TF-IDF.

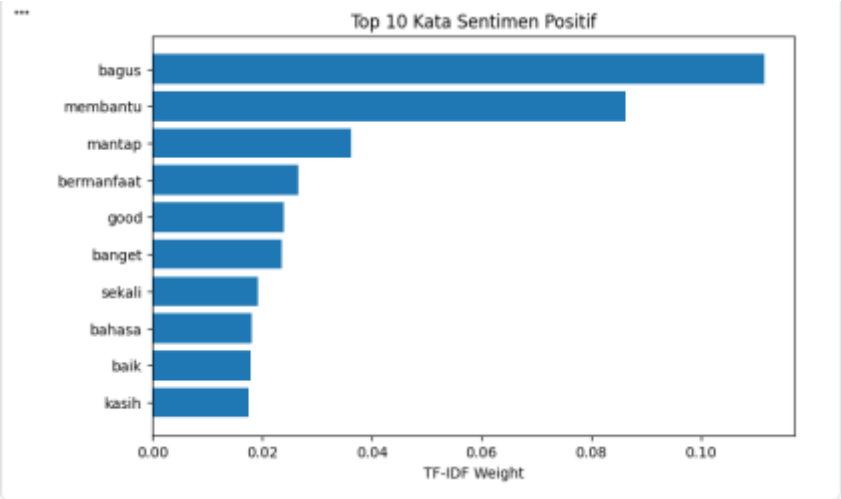
Gambar 3.1 Rekontruksi Teks



3.2 Hasil Pembobotan TF-IDF

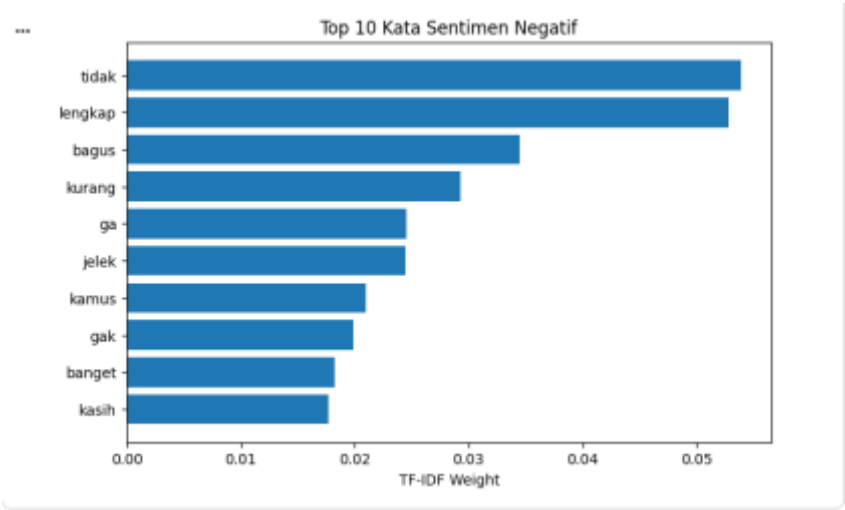
Setelah tahap *preprocessing* selesai, sebanyak 2.319 ulasan diubah menjadi representasi fitur numerik melalui penerapan TF-IDF. Representasi tersebut digunakan sebagai dasar pembentukan fitur pada proses klasifikasi. Sepuluh kata dengan bobot terbesar pada sentimen positif ditampilkan pada Gambar 3.2. Kata bagus, sangat, membantu, dan mantap muncul dengan bobot yang relatif lebih besar. Penelitian terdahulu yang membahas klasifikasi sentimen berdasarkan ulasan pengguna juga menemukan bahwa kata "bagus" muncul dengan bobot yang relatif tinggi (Wibowo, 2022). Kata ini muncul lebih sering dalam kelompok sentimen positif daripada kata lainnya.

Gambar 3.2 Top 10 Kata Penting TF-IDF Sentimen Positif



Gambar 3.3 menampilkan sepuluh kata dengan bobot TF-IDF tertinggi pada kelompok sentimen negatif. Kata *tidak*, *kurang*, *ga*, *gak*, dan *jelek* memiliki bobot yang relatif lebih tinggi dibandingkan kata lainnya. Bobot yang lebih besar menandakan bahwa kata tersebut lebih sering ditemukan sebagai pembeda pada sentimen negatif.

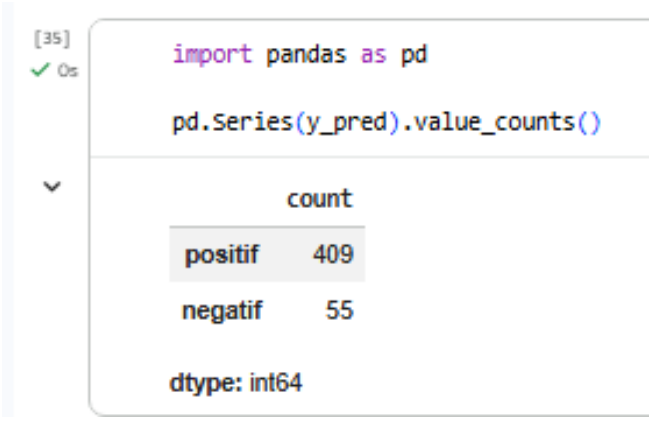
Gambar 3.3 Top 10 Kata Penting TF-IDF Sentimen Negatif



3.3 Hasil Klasifikasi *Naïve Bayes*

Klasifikasi dilakukan pada data TF-IDF menggunakan algoritma Naive Bayes. Pada tahapan berikutnya, model yang dibuat digunakan untuk melakukan prediksi kelas sentimen pada data pengujian. Hasil prediksi menunjukkan bahwa 409 ulasan dianggap sebagai sentimen positif dan 55 ulasan dianggap sebagai sentimen negatif, perbedaan jumlah prediksi antara kedua kelas menunjukkan bahwa sentimen positif lebih mendominasi pada data pengujian. Hasil klasifikasi selanjutnya dievaluasi menggunakan *confusion matrix, accuracy, precision, recall,* dan *F1-score* untuk mengukur kinerja model secara lebih rinci.

Gambar 3.4 Distribusi Hasil Prediksi Naïve Bayes



3.4 Evaluasi Model

Untuk melakukan evaluasi dilakukan menggunakan *accuracy, precision, recall, F1-score,* dan *confusion matrix.* Pengujian menghasilkan akurasi sebesar 87,5%. Sebagian besar ulasan positif berhasil ditempatkan pada kelas yang sesuai.

Tabel 3.1 Evaluasi Model

```
from sklearn.metrics import accuracy_score, classification_report

y_pred = model.predict(X_test)

print("Akurasi:", accuracy_score(y_test, y_pred))
print(classification_report(y_test, y_pred))
```

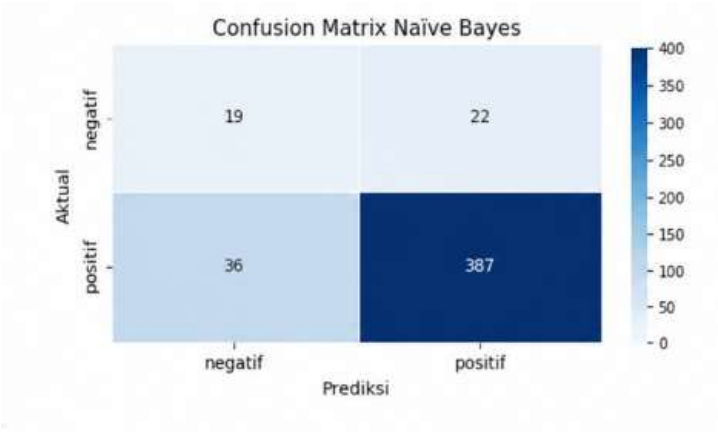
```
Akurasi: 0.875
              precision    recall  f1-score   support

     negatif       0.35      0.46      0.40        41
     positif       0.95      0.91      0.93       423

    accuracy                           0.88       464
   macro avg       0.65      0.69      0.66       464
weighted avg       0.89      0.88      0.88       464
```

Tabel 3.1 menunjukkan hasil evaluasi pada masing-masing kelas sentimen. *Precision, recall,* dan *F1-score* pada sentimen positif masing-masing bernilai 0,95; 0,91; dan 0,93. Kinerja model pada sentimen positif lebih tinggi dibandingkan sentimen negatif. Nilai *precision*, *recall*, dan *F1-score* di kelas negatif masing-masing sebesar 0,35; 0,46; dan 0,40. Ketimpangan jumlah data pada masing-masing kelas memengaruhi hasil evaluasi yang diperoleh, karena jumlah ulasan positif lebih banyak dibandingkan ulasan negatif. Perbedaan nilai *precision*, *recall*, dan *F1-score* dapat dipengaruhi oleh distribusi data dan karakteristik masing-masing kelas sentimen yang digunakan dalam proses klasifikasi (Fattah & Purnawansyah, 2022).

Gambar 3.4 Confusion Matrix



Berdasarkan *confusion matrix* pada Gambar 3.4, Terdapat 387 ulasan positif yang berhasil diprediksi sebagai positif dan 19 ulasan negatif yang berhasil diprediksi sebagai negatif, sedangkan 22 ulasan negatif diprediksi sebagai positif dan 36 ulasan positif diprediksi sebagai negatif. Jumlah prediksi yang tepat lebih banyak ditemukan pada kelas positif dibandingkan kelas negatif. *Confusion matrix* digunakan untuk menghitung jumlah prediksi yang sesuai dengan label sebenarnya dan yang tidak sesuai. Ini menunjukkan bagaimana model klasifikasi bekerja. Kelas negatif pada penelitian ini mengalami kesalahan prediksi yang lebih besar daripada kelas positif. Perbedaan tersebut dapat dikaitkan dengan komposisi dataset yang didominasi oleh ulasan positif sehingga model memperoleh lebih banyak contoh data positif selama proses pelatihan (Hendrawan et al., 2022).

3.5 Pembahasan

Akurasi sebesar 87,5% memperlihatkan bahwa model masih dapat membedakan sentimen positif dan negatif pada ulasan pengguna aplikasi KBBI. Nilai tersebut menunjukkan bahwa fitur yang dihasilkan dari pembobotan TF-IDF mampu membantu proses pengenalan pola pada data teks. Algoritma Naïve Bayes telah banyak dimanfaatkan untuk mengolah opini pengguna dalam bentuk ulasan dan komentar pada berbagai platform digital (Komarudin & Hilda, 2024). Penggunaan TF-IDF sebagai metode pembobotan kata juga banyak diterapkan karena mampu menghasilkan representasi fitur yang lebih informatif pada proses klasifikasi teks (Sevsa & Wahyudi, 2019). Kombinasi TF-IDF dan Naïve Bayes telah digunakan pada berbagai penelitian analisis sentimen berbasis ulasan pengguna dan memberikan hasil klasifikasi yang baik (Darmawan et al., 2023). Selain itu, tahapan *preprocessing* membantu mengurangi variasi penulisan sehingga kata-kata yang memiliki makna serupa dapat direpresentasikan dalam bentuk yang lebih konsisten (Zhafira et al., 2021). Kondisi tersebut membantu model memanfaatkan fitur yang terbentuk secara lebih optimal selama proses klasifikasi.

Bobot tertinggi pada sentimen positif ditemukan pada kata *bagus*, *sangat*, *membantu*, dan *mantap*. Sebaliknya, kelompok sentimen negatif memuat kosakata yang berkaitan dengan kendala penggunaan aplikasi. Perbedaan kosakata tersebut menjadi salah satu dasar yang digunakan model untuk membedakan kedua kelas sentimen. Perbedaan nilai evaluasi terlihat antara kelas positif dan kelas negatif. Kondisi ini berkaitan dengan jumlah ulasan positif yang jauh lebih banyak dalam dataset. Kondisi tersebut membuat model menerima lebih banyak contoh data positif selama pelatihan (Darmawan et al., 2023).

Hasil penelitian ini memperlihatkan bahwa kombinasi TF-IDF dan Naïve Bayes masih relevan digunakan untuk mengolah data ulasan pengguna pada platform digital. Selain mampu mengidentifikasi kecenderungan sentimen pengguna, hasil klasifikasi juga dapat menjadi bahan masukan bagi pengembang aplikasi dalam melakukan evaluasi dan peningkatan layanan. Temuan serupa juga ditemukan pada penelitian analisis sentimen aplikasi Identitas Kependudukan Digital dan aplikasi Buzzbreak yang memanfaatkan ulasan pengguna sebagai sumber evaluasi kualitas aplikasi (Nugroho et al., 2025; Wibowo, 2022).

4. Kesimpulan

Metode yang digunakan untuk penelitian ini yakni TF-IDF dan Naïve Bayes untuk mengklasifikasikan sentimen ulasan pengguna aplikasi KBBI. Akurasi yang diperoleh mencapai 87,5% pada data pengujian. Nilai evaluasi antar kelas tidak sama karena komposisi dataset didominasi oleh ulasan positif. Namun demikian, proses klasifikasi masih memiliki kemampuan untuk membedakan ulasan positif dan negatif berdasarkan karakteristik kata yang terbentuk selama proses pembelajaran model. Hasil tersebut memperlihatkan bahwa TF-IDF dan Naïve Bayes dapat digunakan untuk melakukan analisis sentimen ulasan pada aplikasi KBBI.

5. Referensi

Apriandini, N. A., Ramly, & Nurhusna. (2026). Pemanfaatan KBBI Digital dalam Pembelajaran Bahasa Indonesia di SMA Sekolah Indonesia Kota Kinabalu Malaysia. *IRFANI*, *22*(1), 335–342.

Apriliyani, M., Musyaffaq, M. I., Nur'Aini, S., Handayani, M. R., & Umam, K. (2024). Implementasi Analisis Sentimen pada Ulasan Aplikasi Duolingo di Google Playstore Menggunakan Algoritma Naïve Bayes. *AITI: Jurnal Teknologi Informasi*, *21*(2), 298–311.

Darmawan, G., Alam, S., & Sulistyo, M. I. (2023). Analisis Sentimen Berdasarkan Ulasan Pengguna Aplikasi MyPertamina Pada Google Playstore Menggunakan Metode Naïve Bayes. *STORAGE: Jurnal Ilmiah Teknik Dan Ilmu Komputer*, *2*(3), 100–108. https://doi.org/10.55123

Fattah, St. F., & Purnawansyah. (2022). Analisis Sentimen Terhadap Body Shaming pada Twitter Menggunakan Metode Naïve Bayes Classifier. *Indonesian Journal of Data and Science (IJODAS)*, *3*(2), 61–71.

Hanafi, M. R., & R, R. K. (2024). Analisis Sentimen pada Ulasan Aplikasi Sirekap di Google Play Menggunakan Algoritma Naive Bayes. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, *4*(4), 1578–1586. https://doi.org/10.57152/malcom.v4i4.1693

Hendrawan, I. R., Utami, E., & Hartanto, A. D. (2022). Analisis Perbandingan Metode TF-IDF dan Word2Vec pada Klasifikasi Teks Sentimen Masyarakat Terhadap Produk Lokal di Indonesia. *Smart Comp*, *11*(3), 497–503.

Komarudin, A., & Hilda, A. M. (2024). Analisis Sentimen Ulasan Aplikasi Identitas Kependudukan Digital Pada Play Store Menggunakan Metode Naïve Bayes. *Computer Science (CO-SCIENCE)*, *4*(1), 28–36.

Nugroho, A. K., Hayati, L. N., & Jabir, S. R. (2025). Analisis Perbandingan Metode Naive Bayes dan Random Forest pada Klasifikasi Sentimen Publik terhadap Aplikasi Identitas Kependudukan Digital (IKD). *Jurnal Algoritma*, *22*(2), 619–631. https://doi.org/10.33364/algoritma/v.22-2.2729

Pardede, J., & Darmawan, D. (2025). PERBANDINGAN ALGORITMA STEMMING PORTER, SASTRAWI, IDRIS, DAN ARIFIN & SETIONO PADA DOKUMEN TEKS BAHASA INDONESIA. *Jurnal Teknologi Informasi Dan Ilmu Komputer (JTIIK)*, *12*(1), 69–76. https://doi.org/10.25126/jtiik.2025128860

Pranata, A., Rudiman, & Verdikha, N. A. (2024). Klasifikasi Teks Quick Count Pemilihan Presiden 2024 pada Twitter Menggunakan Metode TF-IDF dan Naive Bayes. *Jurnal Informatika Terpadu*, *10*(2), 93–100.

Pratama, M. A., Nur'aini, S., Handayani, M. R., & Umam, K. (2026). Automated Abstractive Summarization and Entity Extraction of Online News Information Using mT5 and BERT. *Information Technology Education Journal*, 260–276. https://doi.org/10.59562/intec.v5i2.276

Purba, H. S. R., & Aritonang, D. R. (2024). Peningkatan Kemampuan Bahasa Baku Guru dengan Penggunaan Media Laman KBBI Daring di SD Muhammadiyah Se-Kota Padangsidimpuan. *Jurnal Lingue: Bahasa, Budaya Dan Sastra*, *6*(2), 220–231.

Sevsa, B. A., & Wahyudi, M. D. R. (2019). Analisis Sentimen pada Indeks Kinerja Dosen Fakultas SAINTEK UIN Sunan Kalijaga Menggunakan Naive Bayes Classifier. *Jurnal Buana Informatika*, *10*(2), 112–123.

Syifani, A., Rialni, D. A. P., Maysarah, Hijrah, N., & Arini, R. (2025). Eksplorasi Peran KBBI Daring Sebagai Sumber Rujukan Leksikal Bagi Mahasiswa Pendidikan Bahasa dan Sastra Indonesia. *PENTAS: Jurnal Ilmiah Pendidikan Bahasa Dan Sastra Indonesia*, *11*(2), 79–86.

Wati, R., Ernawati, S., & Rachmi, H. (2023). Pembobotan TF-IDF Menggunakan Naïve Bayes Pada Sentimen Masyarakat Mengenai Isu Kenaikan BIPIH. *Jurnal Manajemen Informatika (JAMIKA)*, *13*(1), 84–93. https://doi.org/10.34010/jamika.v13i1.9424

Wibowo, D. P. S. dan W. (2022). Analisis Sentimen Ulasan Aplikasi Buzzbreak Menggunakan Naïve Bayes Classifier Pada Situs Google Play Store. *Jurnal Sains Dan Seni ITS*, *11*(2), D190–D196.

Zhafira, D. F., Rahayudi, B., & Indriati. (2021). Analisis Sentimen Kebijakan Kampus Merdeka Menggunakan Naive Bayes dan Pembobotan TF-IDF Berdasarkan Komentar Pada Youtube. *Jurnal Sistem Informasi, Teknologi Informasi, Dan Edukasi Sistem Informasi (JUST-SI)*, *2*(1), 55–63.