VOL. 03. NO. 06 NOVEMBER 2026

JOURNAL OF COMPUTER SCIENCE AND ARTIFICIAL INTELLIGENCE (JCSAI)

KLASIFIKASI ANALISIS SENTIMEN ULASAN APLIKASI AL-QUR'AN INDONESIA DI PLAY STORE MENGGUNAKAN METODE NAÏVE BAYES

Muhamad Tri Romandan[1], Rudi Kurniawan[2], Bani Nurhakim[3], Cep Lukman Rohmat[4]

Program Studi Teknik Informatika[12]
Program Studi Manajemen Informatika[3]
Program Studi Rekayasa Perangkat Lunak[4]

STMIK IKMI Cirebon
https://ikmi.ac.id/page/18/?lang=de
dtri38756@gmail.com

(*) Corresponding Author : dtri38756@gmail.com
Published : 20 Juli 2026

Abstract*-This study aims to analyze user sentiment toward the Al-Qur'an Indonesia application on the Google Play Store using the Multinomial Naïve Bayes method. The research data were obtained through a scraping process of user reviews of the Al-Qur'an Indonesia application using Python in Google Colab. A total of 3,000 latest reviews were collected, covering the period from 2023 to 2025. The research stages included data preprocessing consisting of cleaning, case folding, tokenizing, stopword removal, and stemming. After preprocessing, the data were transformed into numerical representations using TF-IDF (Term Frequency–Inverse Document Frequency). The dataset was then divided into training and testing data using the train_test_split method with an 80:20 ratio. The study found an imbalance in the distribution of sentiment classes among positive, neutral, and negative sentiments. Therefore, the SMOTE (Synthetic Minority Oversampling Technique) method was applied to balance the class distribution. After applying SMOTE, each class became balanced with 2,852 positive data, 2,852 neutral data, and 2,852 negative data. The classification process was carried out using the Multinomial Naïve Bayes algorithm. Model evaluation was performed using a confusion matrix, accuracy, precision, recall, and f1-score. The testing results showed that the model achieved an accuracy value of 94.83%. The positive class obtained a precision value of 0.96 and a recall value of 0.99. The results indicate that the combination of preprocessing, TF-IDF, SMOTE, and Multinomial Naïve Bayes was able to produce a fairly good sentiment classification model for analyzing user reviews of the Al-Qur'an Indonesia application on the Google Play Store.*

Keywords: *Sentiment Analysis, Naïve Bayes, TF-IDF, SMOTE, Google Play Store.*

Abstrak—Penelitian ini bertujuan untuk menganalisis sentimen pengguna terhadap aplikasi Al-Qur'an Indonesia di Google Play Store menggunakan metode Multinomial Naïve Bayes. Data penelitian diperoleh melalui proses scraping ulasan pengguna aplikasi Al-Qur'an Indonesia menggunakan Python pada Google Colab. Pengambilan data dilakukan sebanyak 3000 ulasan terbaru dengan rentang tahun 2023–2025.Tahapan penelitian meliputi preprocessing data berupa cleaning, case folding, tokenizing, stopword removal, dan stemming. Setelah preprocessing, data diubah menjadi representasi numerik menggunakan TF-IDF (Term Frequency Inverse Document Frequency). Dataset kemudian dibagi menjadi data latih dan data uji menggunakan metode train_test_split dengan rasio 80:20. Penelitian ini menemukan adanya ketidakseimbangan distribusi data sentimen antara kelas positif, netral, dan negatif. Oleh karena itu diterapkan metode SMOTE (Synthetic Minority Oversampling Technique) untuk menyeimbangkan distribusi kelas. Setelah proses SMOTE dilakukan, jumlah masing-masing kelas menjadi seimbang yaitu 2852 data positif, 2852 data netral, dan 2852 data negatif. Proses klasifikasi dilakukan menggunakan algoritma Multinomial Naïve Bayes. Evaluasi model dilakukan menggunakan confusion matrix, accuracy, precision, recall, dan f1-score. Hasil pengujian menunjukkan bahwa model memperoleh nilai accuracy sebesar 94,83%. Nilai precision kelas positif sebesar 0,96 dengan recall sebesar 0,99. Hasil penelitian menunjukkan bahwa kombinasi preprocessing, TF-IDF, SMOTE, dan Multinomial Naïve Bayes mampu menghasilkan model klasifikasi sentimen yang cukup baik dalam menganalisis ulasan pengguna aplikasi Al-Qur'an Indonesia di Google Play Store.

Kata Kunci: Analisis Sentimen, Naïve Bayes, TF-IDF, SMOTE, Google Play Store.

INTRODUCTION

Perkembangan pesat dalam informatika dan pemrosesan bahasa alami telah merevolusi cara manusia mengolah, memahami, dan memanfaatkan data pada berbagai sektor kehidupan. Kemajuan metode NLP, mulai dari teknik klasik seperti Naïve Bayes hingga model transformer modern, mencerminkan pergeseran paradigma menuju analisis yang lebih sensitif terhadap konteks dan makna [1]; [2]. Studi empiris menunjukkan bahwa sentimen dalam ulasan daring berpengaruh signifikan terhadap perilaku pengguna dan keputusan adopsi layanan digital [3]. Selain itu, analisis pola panjang–sentimen pada ulasan pengguna juga memberikan wawasan penting untuk desain antarmuka dan pengalaman pengguna [4]. Perkembangan ini menegaskan pentingnya analisis sentimen sebagai pendekatan ilmiah untuk memahami persepsi pengguna dalam ekosistem digital modern.

Transformasi digital telah membawa perubahan signifikan pada praktik keagamaan masyarakat modern, terutama dalam cara individu mengakses sumber pengetahuan dan menjalankan aktivitas religius. Aplikasi keagamaan, termasuk aplikasi Al-Qur'an digital, menjadi semakin populer karena menawarkan akses yang mudah, fleksibel, serta pengalaman interaktif yang relevan dengan budaya penggunaan perangkat mobile masa kini [5]. Analisis bibliometrik mengenai studi Islam digital menunjukkan tren peningkatan riset dalam dekade terakhir, menandakan perhatian akademik yang semakin besar terhadap relasi antara teknologi dan praktik keagamaan [6]. Penelitian tentang mediatization agama juga menyoroti pergeseran otoritas keagamaan menuju ruang digital, yang menuntut literasi digital serta lokalisasi konten agar pesan keagamaan tetap akurat dan sesuai konteks [7]. Sementara itu, kajian mengenai dakwah digital di Indonesia menunjukkan bahwa ekspansi media keagamaan berbasis online turut menghadirkan tantangan terkait privasi, etika, dan verifikasi informasi [8]. Fenomena ini menegaskan pentingnya penelitian yang mengevaluasi persepsi pengguna terhadap aplikasi Al-Qur'an digital, terutama dalam konteks kepercayaan, kualitas konten, dan relevansi praktik keagamaan modern.

Penggunaan aplikasi mobile di Indonesia terus meningkat, namun pemahaman mengenai persepsi pengguna terhadap aplikasi keagamaan masih menghadapi tantangan multidimensional. Kompleksitas linguistik dalam ulasan—seperti penggunaan slang, emotikon, singkatan, serta praktik code-mixing antara bahasa Indonesia dan Inggris—telah terbukti menyulitkan proses tokenisasi dan memengaruhi performa model analisis sentimen apabila preprocessing tidak dilakukan secara optimal [9]Selain itu, kekhawatiran terhadap aspek privasi dan keamanan data pada perangkat Android turut memengaruhi tingkat kepercayaan serta penerimaan pengguna terhadap aplikasi digital, termasuk aplikasi bernuansa keagamaan [10]. Di sisi lain, penelitian pada beberapa aplikasi nasional menunjukkan bahwa analisis sentimen otomatis mampu memberikan wawasan yang lebih terukur, efisien, dan konsisten dibandingkan metode manual, khususnya ketika memproses ribuan ulasan pengguna [11]. Tantangan-tantangan tersebut menegaskan pentingnya pengembangan korpus ulasan aplikasi Islami serta penerapan metode analisis sentimen otomatis yang lebih adaptif terhadap karakteristik bahasa Indonesia dan sensitivitas konteks religius.

Tantangan linguistik dalam analisis teks berbahasa Indonesia - seperti morfologi aglutinatif, agihan imbuhan, ragam dialek, dan keberagaman idiom/slang - menjadi faktor penting yang harus diperhatikan dalam riset NLP. Indonesia memiliki ratusan bahasa dan dialek, sehingga variasi linguistik memperumit pemrosesan otomatis dan menuntut metode serta korpus yang sesuai [12]. Upaya penyusunan dataset multibahasa seperti NusaX menunjukkan kemajuan dalam representasi bahasa lokal, namun fokusnya lebih ke bahasa daerah dan sentiment umum - sedangkan korpus domain-spesifik (misalnya ulasan aplikasi keagamaan) masih sangat minim [13]. Lebih jauh, pemrosesan teks informal yang kerap mengandung campuran bahasa (Indonesia–Inggris, dialek, slang) dapat menurunkan kinerja model NLP bila tidak ditangani dengan teknik tokenisasi, normalisasi, dan preprocessing khusus [9]. Oleh karena itu, ada kebutuhan mendesak untuk membangun korpus ulasan aplikasi keagamaan berbahasa Indonesia yang mempertimbangkan karakteristik linguistik lokal, serta merancang metodologi yang sesuai untuk tokenisasi, anotasi, dan evaluasi model - untuk meningkatkan akurasi serta relevansi hasil analisis.

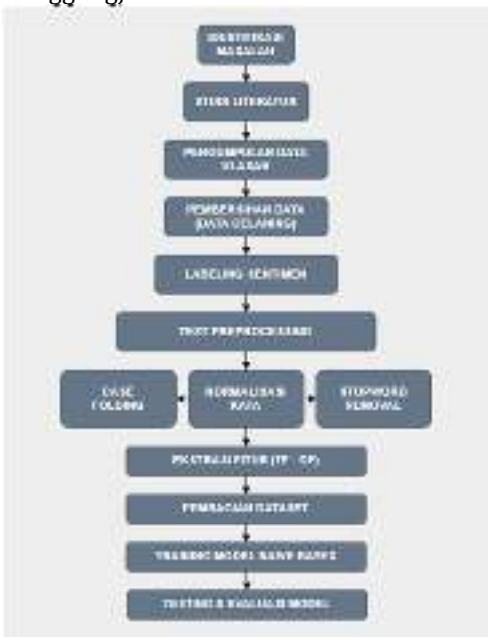
Berbagai penelitian menunjukkan bahwa metode klasik seperti Naïve Bayes, SVM, dan Random Forest tetap menjadi baseline yang efisien dan reprodusibel untuk analisis sentimen. Dalam studi kasus analisis opini di media sosial, ketiga

algoritma tersebut diuji dan menunjukkan performa kompetitif — meskipun SVM sedikit unggul, Naïve Bayes dan Random Forest tetap relevan untuk dataset menengah [14]. Dalam domain ulasan aplikasi dan marketplace, kombinasi preprocessing + TF-IDF + algoritma klasik berhasil mengklasifikasikan sentimen dengan akurasi dan stabilitas yang memadai [15]; [16]. Bahkan ketika dibandingkan dengan pendekatan deep learning lewat semi-supervised learning, metode statistik klasik tetap menunjukkan nilai lebih dari segi efisiensi komputasi, interpretabilitas, dan kebutuhan data label yang relatif kecil — menjadikannya pilihan praktis di banyak skenario [17]. Namun demikian, penelitian secara spesifik terhadap ulasan aplikasi keagamaan — seperti aplikasi Qur'an — masih sangat terbatas, dan belum tersedia korpus domain-spesifik yang mapan. Kesenjangan ini diperparah oleh isu-isu seperti kualitas anotasi, keragaman bahasa informal, dan kebutuhan penyesuaian metodologi yang memperhatikan konteks teologis maupun sosial. Oleh karena itu, studi yang merancang korpus ulasan aplikasi Al-Qur'an dan mengevaluasi performa model baseline seperti Naïve Bayes tetap menjadi kontribusi ilmiah penting dan urgent.

Melihat perkembangan teknologi digital, kompleksitas linguistik Bahasa Indonesia, serta meningkatnya kebutuhan masyarakat terhadap layanan keagamaan berbasis mobile, penelitian mengenai analisis sentimen pada ulasan aplikasi Al-Qur'an menjadi semakin penting. Analisis sentimen tidak hanya memberikan gambaran objektif mengenai tingkat kepuasan pengguna, tetapi juga membantu pengembang dalam mengidentifikasi bug, prioritas fitur, serta peluang peningkatan kualitas aplikasi [18]; [19]. Selain itu, pembangunan korpus domain-spesifik berbahasa Indonesia dapat memperkuat ekosistem NLP lokal dan mendorong kemajuan riset pada bahasa berisiko rendah melalui pendekatan transfer learning dan pengembangan sumber daya linguistik [20]; [21]. Penelitian ini juga berpotensi mendukung inovasi teknologi Islami yang lebih inklusif dan etis, terutama dalam aspek privasi pengguna dan kredibilitas data [22]. Oleh karena itu, analisis sentimen terhadap ulasan aplikasi Al-Qur'an memiliki kontribusi penting baik pada ranah akademik maupun pengembangan aplikasi digital Islami secara praktis.

MATERIALS AND METHODS

Penelitian ini menggunakan pendekatan kuantitatif dengan desain eksperimen komputasi yang berfokus pada pengolahan data teks melalui metode *text mining*. Pendekatan kuantitatif dipilih karena penelitian ini melibatkan proses pengukuran numerik terhadap data ulasan pengguna dan evaluasi kinerja model klasifikasi secara objektif melalui nilai akurasi dan analisis *confusion matrix*. Desain penelitian mengadopsi alur kerja *text mining* yang sistematis untuk melakukan analisis sentimen, dimulai dari pengumpulan data ulasan aplikasi *Al-Qur'an Indonesia*, dilanjutkan dengan tahapan preprocessing teks seperti *cleaning*, *case folding*, *stopword removal*, *tokenizing*, dan *stemming*. Setelah data siap, proses pembobotan dilakukan dengan menggunakan metode TF-IDF untuk mengubah teks menjadi representasi numerik yang dapat diproses oleh algoritma. Tahap berikutnya adalah pembagian data menjadi data pelatihan dan data pengujian dengan proporsi 80:20 untuk menjamin model memperoleh cukup informasi saat dilatih dan diuji secara objektif. Model *Naïve Bayes Classifier* kemudian dilatih menggunakan data pelatihan untuk mempelajari pola-pola sentimen, dan dievaluasi menggunakan data pengujian untuk menilai ketepatannya dalam mengklasifikasikan ulasan. Melalui desain penelitian ini, seluruh proses analisis dilakukan secara komputasional dan terukur, sehingga menghasilkan temuan yang dapat dipertanggungjawabkan secara ilmiah.



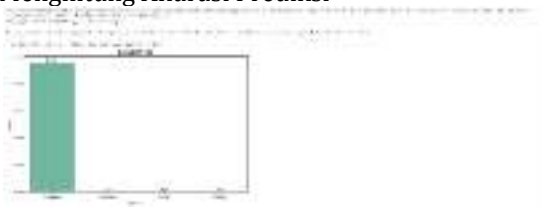
Gambar 1 Jenis Desain Penelitian

Alur penelitian ini disusun secara sistematis untuk memastikan proses analisis sentimen

berjalan terarah dan dapat dipertanggungjawabkan secara ilmiah. Penelitian diawali dengan tahap pengumpulan data, yaitu menghimpun ulasan pengguna aplikasi *Al-Qur'an Indonesia* dari platform Google Play Store sebagai sumber data utama. Setelah data terkumpul, langkah berikutnya adalah melakukan preprocessing teks yang mencakup *cleaning*, *case folding*, *stopword removal*, *tokenizing*, dan *stemming* untuk menghasilkan teks yang lebih bersih dan terstruktur. Data hasil preprocessing kemudian diubah ke bentuk representasi numerik melalui metode pembobotan TF-IDF agar dapat diproses oleh algoritma klasifikasi. Setelah itu, dataset dibagi menjadi dua bagian, yaitu 80% untuk data pelatihan dan 20% untuk data pengujian menggunakan metode *train-test split*. Tahap berikutnya adalah melatih model *Naïve Bayes Classifier* dengan memanfaatkan data pelatihan yang telah dibobot, sehingga model dapat mempelajari pola-pola sentimen yang ada pada ulasan. Setelah model dilatih, dilakukan tahap evaluasi menggunakan *confusion matrix* dan nilai akurasi untuk menilai performa model dalam memprediksi sentimen pada data pengujian. Alur penelitian kemudian diakhiri dengan penarikan hasil prediksi dan penyimpulan performa model berdasarkan evaluasi yang telah dilakukan. Melalui alur penelitian yang terstruktur ini, analisis sentimen terhadap ulasan pengguna dapat dilakukan secara efektif dan menghasilkan temuan yang akurat.

RESULTS AND DISCUSSION

Menghitung Akurasi Prediksi



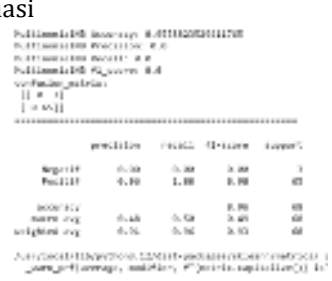
Gambar 2 Menghitung Akurasi Prediksi

Hasil evaluasi model yang ditampilkan pada grafik menunjukkan adanya ketidakseimbangan performa antar metrik yang sangat signifikan. Nilai akurasi terlihat sangat tinggi, yakni mencapai sekitar 0.958, yang sekilas memberikan gambaran bahwa model bekerja dengan sangat baik dalam melakukan prediksi. Namun, ketika diperhatikan lebih dalam, metrik lainnya seperti precision, recall, dan F1-score masing-masing bernilai 0.0, yang menunjukkan adanya masalah serius dalam performa model, terutama terkait dengan kemampuan mendeteksi kelas tertentu. Situasi ini biasanya muncul ketika model tidak pernah memprediksi kelas positif atau minoritas sama sekali. Dengan kata lain, model hanya memberikan prediksi pada satu kelas (biasanya kelas mayoritas), sehingga meskipun sebagian besar data berhasil diklasifikasikan dengan benar—dan memberikan kesan akurasi tinggi—model gagal total dalam mengidentifikasi atau mengenali kelas lain yang mungkin lebih penting secara analitis. Kondisi ini dikenal sebagai accuracy paradox, di mana tingginya akurasi tidak mencerminkan performa model yang sebenarnya dalam tugas klasifikasi.

Nilai precision = 0 menunjukkan bahwa tidak ada satu pun prediksi kelas positif yang benar, sementara recall = 0 mengindikasikan bahwa dari seluruh data yang seharusnya diklasifikasikan sebagai positif, tidak satu pun yang berhasil dideteksi oleh model. Hal ini berdampak langsung pada nilai F1-score, yang juga menjadi 0 karena tidak adanya keseimbangan antara precision dan recall. Ketidakmampuan model dalam memprediksi kelas minoritas mengindikasikan dua kemungkinan utama: pertama, distorsi distribusi kelas (class imbalance) yang sangat tinggi, sehingga model lebih mudah 'bermain aman' dengan selalu memprediksi kelas mayoritas; kedua, parameter, fitur, atau algoritma yang digunakan belum optimal, sehingga model belum mampu menangkap pola-pola yang mewakili kelas minoritas secara efektif.

Secara keseluruhan, visualisasi ini memberikan sinyal bahwa meskipun akurasi terlihat impresif, performa model sebenarnya sangat kurang memadai dari perspektif kemampuan klasifikasi yang menyeluruh. Oleh karena itu, dibutuhkan evaluasi ulang terhadap dataset dan pendekatan pemodelan, seperti penerapan teknik oversampling/undersampling, penggunaan class weight, eksplorasi algoritma lain, atau tuning hyperparameter agar model dapat lebih seimbang dan mampu mendeteksi seluruh kelas dengan lebih akurat. Dengan demikian, analisis ini menegaskan bahwa akurasi tinggi tidak boleh menjadi satu-satunya rujukan dalam menilai keefektifan model, khususnya pada kasus klasifikasi dengan distribusi kelas yang tidak seimbang.

Hasil Evaluasi



Gambar 3 Hasil Evaluasi

Hasil evaluasi model Multinomial Naive Bayes menunjukkan adanya kecenderungan prediksi yang sangat kuat ke kelas *positif*. Hal ini tercermin dari confusion matrix, di mana seluruh 65 data yang benar-benar termasuk kelas positif berhasil diprediksi dengan benar (true positive = 65). Namun, tiga data yang seharusnya masuk kelas negatif sepenuhnya salah diklasifikasikan sebagai positif (false positive = 3). Kondisi ini menghasilkan nilai akurasi yang tinggi, yaitu sekitar 0.958, tetapi performa pada kelas negatif sangat buruk, sehingga menyebabkan nilai precision, recall, dan F1-score untuk kelas negatif menjadi 0.00.

Pada kelas positif, model tampak bekerja sangat baik, dengan precision sebesar 0.96, recall 1.00, dan F1-score 0.98, menunjukkan bahwa model hampir sempurna dalam mengenali dan mengklasifikasikan data positif. Akan tetapi, performa yang tampak cemerlang pada kelas positif justru menutupi kegagalan mendasar model dalam mendeteksi kelas negatif. Ketidakseimbangan ini tercermin pula dalam nilai macro average, di mana precision dan recall hanya berada di kisaran 0.48 dan 0.50, menandakan performa keseluruhan model masih sangat lemah ketika kedua kelas dihitung secara setara.

Ketimpangan performa ini kemungkinan muncul karena ketidakseimbangan data (class imbalance), di mana kelas positif jauh lebih dominan dibandingkan kelas negatif. Dengan total 68 data, hanya 3 data yang masuk kelas negatif, sehingga model lebih mudah "bermain aman" dengan memprediksi semua data sebagai positif. Akibatnya, meskipun akurasi terlihat tinggi, model tidak memiliki kemampuan identifikasi terhadap kelas minoritas. Ini mengindikasikan bahwa akurasi saja tidak dapat dijadikan indikator kehandalan model dalam konteks klasifikasi berimbang.

Dalam konteks penelitian dan aplikasi nyata, performa seperti ini sangat berisiko, terutama pada kasus yang menuntut sensitivitas tinggi terhadap kelas negatif. Oleh karena itu, perbaikan perlu difokuskan pada peningkatan kemampuan model dalam menangani kelas minoritas, seperti melalui teknik resampling (SMOTE atau undersampling), pemberian bobot kelas (*class weighting*), atau penggunaan algoritma lain yang lebih sensitif terhadap data tidak seimbang. Langkah-langkah tersebut akan membantu menghasilkan model yang lebih stabil, adil, dan memiliki performa yang merata pada kedua kelas.

CONCLUSION

Penelitian berjudul "Klasifikasi Analisis Sentimen Ulasan Aplikasi Al-Qur'an Indonesia di Play Store Menggunakan Metode Naïve Bayes" telah dilakukan melalui serangkaian proses ilmiah mulai dari pengumpulan data, pembersihan data (*preprocessing*), ekstraksi fitur menggunakan TF–IDF, hingga pelatihan model Multinomial Naïve Bayes. Berdasarkan hasil implementasi dan evaluasi, diperoleh beberapa kesimpulan penting sebagai berikut:

Dataset hasil *scraping* berjumlah 338 ulasan, seluruhnya dapat digunakan setelah tahap preprocessing karena tidak ditemukan duplikasi, data kosong, maupun ulasan yang berisi emoji saja. Hal ini menunjukkan bahwa sumber data memiliki kualitas baik dan sesuai untuk dilakukan analisis lebih lanjut. Dominasi 326 ulasan positif (96,4%) menunjukkan bahwa aplikasi *Al-Qur'an Indonesia* sangat disukai oleh pengguna Play Store, terutama karena kelengkapan fitur, kemudahan navigasi, dan kualitas murottal. Sementara itu, hanya 12 ulasan negatif (3,6%), yang umumnya terkait masalah teknis, bug aplikasi, atau ketidakakuratan terjemahan.

Preprocessing mencakup *case folding, cleaning text, tokenizing, stopword removal,* dan *stemming Bahasa Indonesia*. Tahapan ini sangat penting untuk menormalkan data ulasan yang awalnya memiliki karakteristik informal, mengandung simbol, angka, dan variasi penulisan. Setelah preprocessing, data menjadi lebih bersih dan siap untuk diekstraksi menjadi fitur numerik. Hal ini berpengaruh langsung pada kualitas hasil klasifikasi.

Metode TF–IDF memberikan bobot terhadap kata-kata penting dalam ulasan. Kata yang muncul sering di seluruh dokumen memperoleh bobot rendah, sedangkan kata yang memiliki makna spesifik bagi satu ulasan diberikan bobot lebih tinggi. Representasi ini sangat cocok untuk model klasifikasi berbasis Multinomial Naïve Bayes karena model ini bekerja dengan pola distribusi frekuensi kata.

Model menghasilkan akurasi 95,8%, namun seluruh data uji diprediksi sebagai positif. Nilai *precision, recall, dan F1-score kelas negatif adalah 0.00*, yang berarti model tidak mampu mengenali satu pun ulasan negatif. Masalah ini muncul karena ketidakseimbangan kelas yang sangat besar (positif 326 vs negatif 12). Kondisi ini membuat model cenderung mempelajari pola dari kelas positif saja sehingga terjadi fenomena model collapse.

Naïve Bayes secara teori bekerja optimal ketika distribusi data relatif seimbang. Pada kasus penelitian ini, model hanya mengenali konteks yang cenderung positif, sehingga ulasan negatif yang jumlahnya sangat sedikit tidak dapat dipelajari

secara efektif. Dengan demikian, meskipun akurasi tinggi, namun model tidak layak digunakan untuk sistem analisis sentimen yang membutuhkan deteksi kedua kelas secara akurat.

Akurasi tinggi tidak menjamin kegunaan model di dunia nyata. Untuk aplikasi dengan tujuan memonitor kualitas layanan, ulasan negatif jauh lebih penting karena memberi wawasan terhadap keluhan dan masalah aplikasi. Karena model tidak mendeteksi ulasan negatif, maka hasil yang diberikan tidak mencerminkan kondisi sesungguhnya. Dengan demikian, model Naïve Bayes dalam penelitian ini belum siap diterapkan sebagai sistem analisis sentimen otomatis untuk kebutuhan evaluasi aplikasi Al-Qur'an Indonesia.

REFERENCE

[1] Y. Mao, Q. Liu, and Y. Zhang, "Journal of King Saud University - Computer and Sentiment analysis methods , applications , and challenges : A systematic literature review," *Journal of King Saud University - Computer and Information Sciences*, vol. 36, no. 4, p. 102048, 2024, doi: 10.1016/j.jksuci.2024.102048.

[2] P. Narayanan, V. Mukund, and S. Sanjana, "The Sentiment Problem : A Critical Survey towards Deconstructing Sentiment Analysis," pp. 13743–13763, 2023.

[3] Q. Wang, W. Zhang, J. Li, F. Mai, and Z. Ma, "Effect of online review sentiment on product sales: The moderating role of review credibility perception," *Computers in Human Behavior*, vol. 133, p. 107272, 2022, doi: https://doi.org/10.1016/j.chb.2022.107272.

[4] T. Guzsvinecz, "Computers in Human Behavior Length and sentiment analysis of reviews about top-level video game genres on the steam platform ˝ cs," vol. 149, no. October 2022, 2023, doi: 10.1016/j.chb.2023.107955.

[5] D. Amalia, I. Desyanti, K. N. Salsabila, N. I. Azyan, and R. Rizkya, "Peran Aplikasi Al-Qur ' an Digital dalam Meningkatkan Minat Membaca di Kalangan Mahasiswa," vol. 3, no. 2, 2025.

[6] P. Islam, A. C. Kamilla, S. Anwar, and R. N. Fatimah, "Analisis Bibliometrik Terhadap Tren dan Perkembangan Penelitian Pendidikan Islam di Perguruan Tinggi di Indonesia (2019-2024)," vol. 10, 2025.

[7] A. Setyowati, "DIGITAL RELIGION DAN RELIGIUSITAS MILENIAL: STUDI PERGESERAN OTORITAS KEAGAMAAN DI DUNIA MEDIA BARU (NEW MEDIA WORDLS)," 2023.

[8] H. Bin Khalifa, "Journal of Islamic Ethics," 2025.

[9] N. H. Setiono and Y. Sari, "Exploring the Impact of Back-Translation on BERT ' s Performance in Sentiment Analysis of Code-Mixed Language Data," vol. 19, no. 2, pp. 223–234, 2025, doi: 10.22146/ijccs.104757.

[10] G. Prambudi, R. Kusuma, A. Hardiyanto, and M. Sholahudin, "Pengukuran Kesadaran Keamanan Informasi Dan Privasi Pada Pengguna Smartphone Android Vivo Di Indonesia," pp. 456–460, 2024.

[11] F. Y. A, "Indonesian Sentiment Analysis towards MyPertamina Application Reviews by Utilizing Machine Learning Algorithms," vol. 8106, pp. 80–91, 2022, doi: 10.20895/INISTA.V5I1.838.

[12] A. F. Aji *et al.*, "One Country , 700 + Languages : NLP Challenges for Underrepresented Languages and Dialects in Indonesia".

[13] A. N. A. S, T. R. Eugene, and E. Nurhayati, "Tantangan Linguistik dalam Pengimplementasian Big Data Berbahasa Indonesia pada Robot Humanoid : Tinjauan dan Rekomendasi," vol. 1, no. 1, pp. 1–9, 2025.

[14] P. Cahyani and L. Abdillah, "Perbandingan Performa Algoritma Naïve Bayes , SVM dan Random Forest : Studi Kasus Analisis Sentimen Pengguna Sosial Media X," vol. 11, no. 02, pp. 12–21, 2024.

[15] F. I. Komputer and U. A. Purwokerto, "Analisis Sentimen Ulasan Pengguna Aplikasi Sinaga Mobile pada Google Play Store Menggunakan Algoritma Naive Bayes Sentiment Analysis of Sinaga Mobile App User Reviews on Google Play Store Using Naive Bayes Algorithm," vol. 5, no. 6, pp. 1635–1645, 2025.

[16] E. L. Ruskan, D. R. Indah, U. Sriwijaya, S. Informasi, K. Ilir, and S. Selatan, "COMPARISON OF SVM AND NAIVE BAYES ALGORITHMS IN SENTIMENT ANALYSIS OF USER REVIEWS ON PERBANDINGAN ALGORITMA SVM DAN NAIVE BAYES," vol. 10, no. 3, pp. 1623–1633, 2025.

[17] R. Husaini, N. H. Cahyana, I. Wiendijarti, and A. S. Aribowo, "Comparison of Semi-Supervised Learning Performance in

Indonesian Sentiment Analysis : An Empirical Study between Statistical Machine Learning and Deep Learning Approaches," vol. 4, no. 1, 2025.

[18] S. Agung and A. Santara, "Implementasi Text Mining untuk Analisis Review pada Aplikasi Crowdfunding LX dan ST Menggunakan Metode Sentiment Analysis," vol. 2, no. 1, pp. 124–130, 2024.

[19] C. Schinckus, M. Gasparin, and W. Green, "Opening the black boxes: financial algorithms and multi-paradigmatic research in information technology," *Journal of Systems and Information Technology*, vol. 24, no. 3, pp. 284–303, Jul. 2022, doi: 10.1108/JSIT-01-2020-0006.

[20] S. Yang, L. Cui, R. Ning, D. Wu, and Y. Zhang, "Challenges to Open-Domain Constituency Parsing," pp. 112–127, 2022.

[21] Y. Zhang and R. Jin, "Understanding Bag-of-Words Model : A Statistical Framework".

[22] H. Liu, P. Patras, and D. J. L. Id, "On the data privacy practices of Android OEMs," pp. 1–15, 2023, doi: 10.1371/journal.pone.0279942.