

ANALISIS SENTIMEN PADA APLIKASI MOBILE JKN MENGGUNAKAN METODE TEXT MINING DAN NAIVE BAYES

Rijik Emharis¹, Rudi Kurniawan², Bani Nurhakim³, Nisa Dienwatinuris⁴.

Program Studi Teknik Informatika¹²⁴
Program Studi Manajemen Informatika³

STMIK IKMI Cirebon
<https://ikmi.ac.id/page/18/?lang=de>
emharisrizik33@gmail.com

(*) Corresponding Author : emharisrizik33@gmail.com
Published : 30 Mei 2026

Abstract—Digital transformation in the healthcare sector has encouraged the utilization of public service applications, one of which is the Mobile JKN application developed by BPJS Kesehatan. Despite the increasing number of users, many negative reviews on the Google Play Store indicate user dissatisfaction related to login issues, system errors, and suboptimal interface performance. These reviews serve as important data sources for evaluating service quality through computational approaches such as Text Mining and Sentiment Analysis. This study aims to detect and analyze negative sentiment in Mobile JKN user reviews using the Multinomial Naïve Bayes algorithm. The research data consist of approximately 2,000 Indonesian-language reviews obtained through web scraping using the google-play-scraper library. The text processing stages include case folding, cleaning, tokenization, normalization, negation handling, and stemming using Sastrawi. Feature extraction is carried out using TF-IDF, while model evaluation employs accuracy, precision, recall, and F1-score metrics to assess classification performance. The results indicate that the combination of systematic text preprocessing and TF-IDF feature representation is capable of producing an accurate and effective classification model. Multinomial Naïve Bayes also proves efficient in handling high-dimensional text data and provides interpretable results. Further analysis reveals that the majority of user complaints are related to system performance, login processes, and data update issues. These findings are recommended to developers as a basis for improving service quality, particularly in server reliability, interface design, and authentication mechanisms..

Keywords : Sentiment Analysis, Naïve Bayes, Text Mining, Mobile JKN, TF-IDF.

Abstrak—Transformasi digital dalam sektor kesehatan mendorong pemanfaatan aplikasi layanan publik, salah satunya Aplikasi Mobile JKN yang dikembangkan oleh BPJS Kesehatan. Meski penggunaan aplikasi meningkat, banyak ulasan pada Google Play Store yang menunjukkan ketidakpuasan pengguna terkait kendala login, gangguan sistem, dan antarmuka yang kurang optimal. Ulasan tersebut menjadi sumber data penting untuk mengevaluasi kualitas layanan secara komputasional melalui pendekatan Text Mining dan Analisis Sentimen. Penelitian ini bertujuan untuk mendeteksi serta menganalisis sentimen pada ulasan pengguna Mobile JKN menggunakan algoritma Multinomial Naïve Bayes. Data penelitian berupa ±2.000 ulasan berbahasa Indonesia diperoleh melalui web scraping menggunakan pustaka google-play-scraper. Tahapan pengolahan teks meliputi case folding, cleaning, tokenization, normalization, negation handling, dan stemming menggunakan Sastrawi. Ekstraksi fitur dilakukan dengan TF-IDF, sementara evaluasi model menggunakan metrik akurasi, presisi, recall, dan F1-score untuk menilai performa klasifikasi. Hasil penelitian menunjukkan bahwa kombinasi prapengolahan teks yang sistematis dan representasi fitur TF-IDF mampu menghasilkan model klasifikasi yang akurat dan efektif. Multinomial Naïve Bayes terbukti mampu menangani data berdimensi tinggi dan menghasilkan interpretasi yang jelas. Analisis lebih lanjut mengungkap bahwa keluhan dominan pengguna berkaitan dengan performa sistem, proses login, dan pembaruan data. Temuan ini direkomendasikan kepada pengembang sebagai dasar peningkatan kualitas layanan, terutama pada keandalan server, desain antarmuka, serta mekanisme autentikasi.

Kata Kunci : Analisis Sentimen, Naïve Bayes, Text Mining, Mobile JKN, TF-IDF.

INTRODUCTION

Perkembangan pesat di bidang Informatika memberikan dampak besar pada berbagai aspek kehidupan, seperti teknologi, bisnis, pendidikan, kesehatan, dan pemerintahan. Berbagai sektor kini semakin mengandalkan sistem digital untuk mendukung aktivitas sehari-hari. Transformasi digital mengubah cara manusia berinteraksi, mengakses informasi, dan menyelesaikan masalah secara lebih efisien melalui sistem berbasis teknologi informasi. Inovasi seperti Artificial Intelligence, Machine Learning, dan Natural Language Processing menjadi pilar utama dalam menghadirkan solusi yang cerdas dan adaptif terhadap kebutuhan masyarakat modern. Dalam konteks pelayanan publik, penerapan teknologi informasi tidak hanya mempercepat proses administrasi, tetapi juga meningkatkan transparansi dan kepuasan masyarakat terhadap layanan pemerintah. Oleh karena itu, bidang Informatika memegang peran penting dalam mendukung implementasi Smart Governance dan e-Government di Indonesia.

Namun, di tengah kemajuan tersebut, muncul tantangan baru terkait pengelolaan dan evaluasi kualitas layanan digital, khususnya pada aplikasi kesehatan publik seperti Mobile JKN milik BPJS Kesehatan. Aplikasi Mobile JKN dirancang untuk memberikan kemudahan bagi peserta dalam mengakses informasi layanan kesehatan, melakukan pendaftaran, serta mengelola data kepesertaan. Meskipun demikian, berbagai ulasan pengguna pada platform aplikasi menunjukkan keluhan seperti kesulitan login, kegagalan pembaruan data, antarmuka yang kurang intuitif, serta error sistem. Ulasan tersebut mengindikasikan adanya kepuasan dan ketidakpuasan pengguna terhadap performa aplikasi. Di sisi lain, komentar dan ulasan pengguna merupakan sumber informasi yang sangat berharga untuk menilai dan meningkatkan kualitas layanan secara berkelanjutan. Tantangan yang muncul adalah bagaimana mengekstraksi dan menganalisis opini pengguna tersebut secara otomatis dan efisien melalui pendekatan komputasional di bidang Text Mining dan Sentiment Analysis.

Sejumlah penelitian sebelumnya telah mengeksplorasi penerapan metode Naïve Bayes dalam analisis sentimen pada teks berbahasa Indonesia. [1] menemukan bahwa tahapan prapengolahan teks seperti tokenization, stop-word removal, dan normalisasi kata memberikan dampak signifikan terhadap peningkatan akurasi model Naïve Bayes. [2] mengembangkan model RoBERTa-GRU untuk meningkatkan performa analisis sentimen dan menunjukkan bahwa meskipun model berbasis

deep learning memberikan hasil terbaik, Naïve Bayes tetap digunakan sebagai baseline yang efisien dan mudah diinterpretasikan. Penelitian [3] mengenai analisis sentimen pemilihan presiden di Indonesia juga menunjukkan bahwa Naïve Bayes memiliki stabilitas kinerja pada data berbahasa Indonesia. Di sisi lain, [4] melakukan penelitian tentang aspect-based sentiment analysis dan menyimpulkan bahwa metode konvensional seperti Naïve Bayes tetap relevan sebagai acuan metodologis. Secara keseluruhan, temuan-temuan ini menegaskan bahwa Naïve Bayes merupakan algoritma yang kuat dan efisien untuk tugas klasifikasi teks, terutama pada bahasa dengan sumber daya terbatas.

Berdasarkan penelitian-penelitian terdahulu tersebut, penelitian ini bertujuan untuk menganalisis sentimen pada ulasan pengguna aplikasi Mobile JKN dengan menggunakan metode Text Mining dan algoritma Naïve Bayes. Tujuan utama penelitian adalah mengidentifikasi pola keluhan pengguna, menemukan aspek-aspek aplikasi yang sering menjadi sumber masalah, dan memberikan rekomendasi berbasis data untuk peningkatan kualitas layanan. Selain kontribusi praktis bagi BPJS Kesehatan dan pengembang aplikasi Mobile JKN, penelitian ini juga diharapkan memperkaya literatur mengenai analisis sentimen pada aplikasi layanan publik di Indonesia menggunakan metode klasifikasi tradisional yang tetap relevan.

Penelitian ini menggunakan pendekatan kuantitatif dengan metode supervised learning. Tahapan penelitian dimulai dari pengumpulan ulasan pengguna dari platform aplikasi, kemudian dilanjutkan dengan prapengolahan teks seperti case folding, tokenization, stop-word removal, dan stemming menggunakan pustaka Sastrawi. Selanjutnya, fitur teks diekstraksi menggunakan metode TF-IDF dan dilakukan klasifikasi menggunakan algoritma Multinomial Naïve Bayes. Algoritma ini dipilih karena efisiensinya dalam mengolah data berdimensi tinggi, kemudahan interpretasi, dan kestabilan performanya. Evaluasi dilakukan dengan menggunakan metrik akurasi, presisi, recall, dan F1-score untuk menilai efektivitas model dalam mendeteksi sentimen.

Hasil penelitian ini diharapkan memberikan implikasi teoritis dan praktis yang signifikan. Secara akademik, penelitian ini berkontribusi pada pengembangan kajian terkait penerapan Naïve Bayes dalam analisis sentimen berbahasa Indonesia dan menegaskan relevansinya di era perkembangan model deep learning. Secara praktis, penelitian ini dapat menjadi acuan bagi lembaga publik atau pengembang aplikasi dalam membangun sistem evaluasi otomatis terhadap opini pengguna. Jika

berhasil, temuan penelitian ini berpotensi meningkatkan efektivitas monitoring layanan, mendeteksi permasalahan secara dini, serta mempercepat proses pengambilan keputusan dalam pengembangan aplikasi Mobile JKN. Dengan demikian, penelitian ini diharapkan turut mendukung pembentukan ekosistem digital layanan publik yang lebih responsif dan berorientasi pada kepuasan masyarakat.

MATERIALS AND METHODS

Pendekatan pembelajaran terawasi seperti Naïve Bayes sangat sesuai untuk penelitian ini yang menggunakan data ulasan aplikasi dengan label sentimen yang terstruktur. Gambaran umum desain penelitian disajikan pada Gambar 1 Alur Penelitian.



Gambar 1. Alur Penelitian

Tahap pertama, *Studi Pendahuluan*, berfokus pada proses identifikasi masalah dan penentuan tujuan penelitian. Pada tahap ini, peneliti mengamati fenomena meningkatnya jumlah ulasan pengguna pada platform aplikasi digital seperti Google Playstore yang berisi opini mengenai kinerja dan kepuasan pengguna terhadap suatu aplikasi. Permasalahan utama yang diangkat adalah perlunya metode komputasional untuk menganalisis dan mengklasifikasikan sentimen pengguna secara otomatis. Hasil dari tahap ini berupa rumusan masalah penelitian dan penentuan tujuan untuk mengukur efektivitas algoritma *Naïve Bayes* dalam klasifikasi sentimen berdasarkan data teks ulasan pengguna.

Tahap kedua, *Literatur Review*, bertujuan membangun dasar teoritis penelitian melalui telaah terhadap penelitian terdahulu yang relevan. Kajian mencakup teori dasar *text mining*, konsep analisis sentimen, proses *text preprocessing* untuk Bahasa Indonesia, serta penerapan algoritma *Naïve Bayes* dan model hibrida dalam klasifikasi teks. Selain itu, disusun pula kerangka teori serta metodologi penelitian

berdasarkan referensi ilmiah terkini. Menurut beberapa penelitian terdahulu, kombinasi metode *TF-IDF* dan *Naïve Bayes* memberikan hasil akurasi tinggi dalam klasifikasi teks berbahasa alami karena mempertimbangkan frekuensi kata dan bobot term terhadap keseluruhan dokumen.

Tahap ketiga, *Pendekatan Eksperimen*, merupakan inti dari pelaksanaan penelitian. Data diperoleh melalui proses *scraping* ulasan pengguna dari Google Playstore. Data yang terkumpul kemudian melalui tahapan *text preprocessing* meliputi *case folding*, *cleaning*, *tokenization*, *normalization*, *negation handling*, *stemming*, serta pembentukan representasi menggunakan metode *TF-IDF*. Tahap selanjutnya adalah penerapan algoritma *Naïve Bayes* untuk klasifikasi sentimen menjadi positif, , atau netral. Model dievaluasi menggunakan metrik *accuracy*, *precision*, *recall*, dan *F1-score* untuk menilai performa algoritma. Diagram juga menampilkan urutan alur penelitian secara menyeluruh agar setiap proses dapat dijalankan secara terstruktur.

Tahap keempat, *Kesimpulan dan Rekomendasi*, merupakan bagian akhir dari penelitian. Pada tahap ini, peneliti melakukan analisis hasil eksperimen untuk menarik kesimpulan mengenai efektivitas model *Naïve Bayes* dalam mengklasifikasikan sentimen pengguna terhadap aplikasi. Hasil penelitian ini diharapkan mampu memberikan kontribusi dalam pengembangan sistem analisis sentimen otomatis berbasis Bahasa Indonesia. Rekomendasi diberikan untuk penelitian selanjutnya, seperti penggunaan algoritma pembandingan (*Support Vector Machine*, *Random Forest*), penerapan representasi fitur berbasis *word embedding* (*Word2Vec*, *BERT*), serta peningkatan jumlah dan keragaman data agar hasil klasifikasi lebih akurat dan adaptif terhadap bahasa informal pengguna.

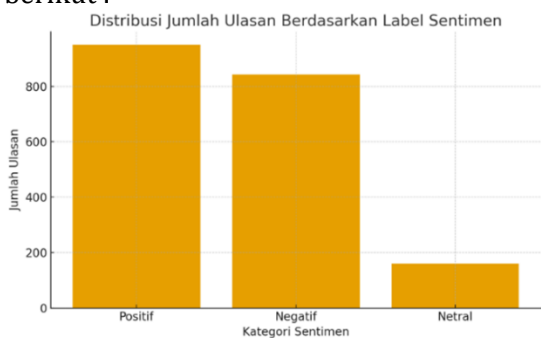
RESULTS AND DISCUSSION

Pada tahap pelabelan data, pelabelan data dalam penelitian ini dilakukan secara manual oleh peneliti berdasarkan isi teks ulasan pengguna. Proses pelabelan dilakukan dengan membaca setiap ulasan dan mengklasifikasikannya ke dalam kategori sentimen berdasarkan konteks kalimat yang mengandung keluhan, kritik, atau ketidakpuasan terhadap aplikasi.

Penelitian ini tidak melibatkan ahli (expert judgment) dalam proses pelabelan, sehingga pelabelan bersifat subjektif berdasarkan interpretasi peneliti. Untuk meminimalkan bias, peneliti menggunakan pedoman

pelabelan yang konsisten berdasarkan indikator sentimen negatif seperti keluhan sistem, error aplikasi, dan kesulitan penggunaan. Pendekatan ini umum digunakan dalam penelitian awal (baseline study) pada analisis sentimen.

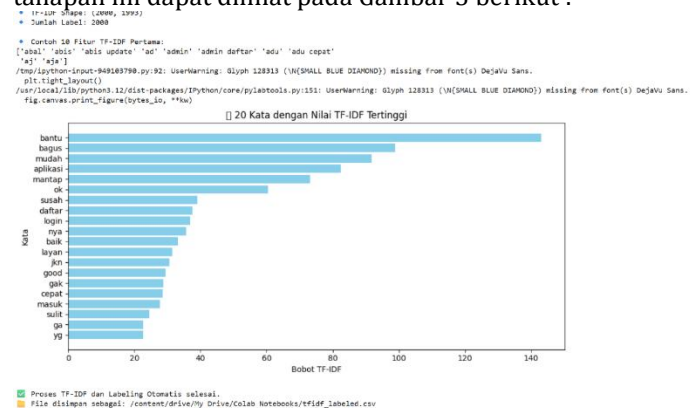
Dari total 2.000 ulasan yang dianalisis, terdapat 47 ulasan yang dihapus karena dianggap tidak relevan, ambigu, atau tidak dapat ditentukan sentimennya secara linguistik. Setelah proses kurasi akhir, diperoleh 1.953 ulasan yang siap digunakan untuk pelatihan dan evaluasi model. Distribusi label menunjukkan bahwa kategori positif mendominasi dengan 950 ulasan, disusul 843 ulasan, dan 160 ulasan netral. Ketimpangan ini menggambarkan bahwa meskipun banyak keluhan muncul terhadap aplikasi Mobile JKN, pengguna yang memberikan apresiasi positif masih lebih tinggi. Visualisasi diagram batang memperkuat gambaran tersebut dengan menunjukkan perbedaan yang jelas antara tiga kategori sentimen. Distribusi ini juga menjadi pertimbangan penting pada tahap pemodelan karena ketidakseimbangan kelas dapat memengaruhi performa algoritma Naïve Bayes. Implementasi tahapan ini dapat dilihat pada Gambar 2. berikut :



Gambar 2. Pelabelan.

Tahap ekstraksi fitur menggunakan metode Term Frequency-Inverse Document Frequency (TF-IDF) menghasilkan representasi numerik yang memperlihatkan tingkat kepentingan setiap kata dalam keseluruhan korpus ulasan Mobile JKN. Hasil visualisasi menunjukkan bahwa kata “bantu” memiliki bobot TF-IDF tertinggi, diikuti oleh kata “bagus”, “mudah”, dan “aplikasi”, yang mengindikasikan bahwa ulasan pengguna banyak menyoroti aspek kemudahan penggunaan serta bantuan yang diberikan oleh aplikasi. Kata lain seperti “susah”, “login”, “gak”,

“sulit”, dan “masuk” juga muncul dengan bobot cukup tinggi, mencerminkan adanya keluhan terkait proses login dan hambatan teknis yang sering dialami pengguna. Pola ini memperlihatkan bahwa TF-IDF mampu menonjolkan kata-kata yang paling relevan terhadap topik keluhan dan kepuasan pengguna, sehingga sangat membantu algoritma Naïve Bayes dalam membedakan sentimen positif dan secara lebih akurat. Selain itu, penyebaran kata dengan bobot tinggi menunjukkan kombinasi antara sentimen positif (misal: bagus, mantap, mudah) dan sentimen (misal: susah, sulit, gak masuk), yang memperkuat relevansi fitur TF-IDF dalam memetakan beragam pengalaman pengguna terhadap aplikasi Mobile JKN. Implementasi tahapan ini dapat dilihat pada Gambar 3 berikut :



Gambar 3 Hasil Implementasi Tahap Ekstraksi Fitur TF IDF

Evaluasi performa model Naïve Bayes menunjukkan bahwa model mampu melakukan klasifikasi sentimen ulasan Mobile JKN dengan tingkat akurasi keseluruhan sebesar 79,67%. Nilai ini mengindikasikan bahwa sebagian besar prediksi model sudah sesuai dengan label aktual pada dataset. Pada kelas *positif*, model memperoleh precision sebesar 0.88 dan recall 0.75, yang berarti model sangat baik dalam mengenali ulasan positif dan sebagian besar prediksi positif sesuai dengan kenyataan. Kinerja terbaik ditunjukkan pada kelas *netral*, dengan recall 0.92 dan f1-score 0.83, menggambarkan bahwa model sangat efektif dalam mendeteksi ulasan yang memang cenderung bernada netral. Sementara itu, kelas menunjukkan recall yang lebih rendah yaitu 0.45, meskipun precision-nya cukup tinggi (0.85), yang menandakan bahwa model cenderung berhati-hati dalam memprediksi sentimen namun cukup akurat ketika melakukannya.

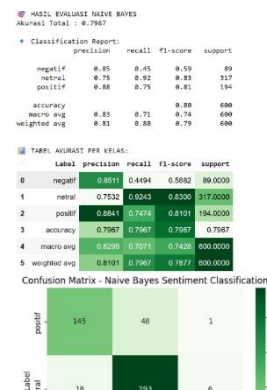
Confusion matrix memberikan gambaran lebih jelas mengenai pola kesalahan prediksi model. Pada kelas positif, model memprediksi benar sebanyak 145 ulasan, namun masih terdapat 48 ulasan positif yang salah diklasifikasikan sebagai netral. Pada

kelas netral, model menampilkan performa yang sangat stabil dengan 293 prediksi benar, dan hanya sejumlah kecil ulasan keliru dipetakan ke kategori positif atau . Sebaliknya, kesalahan prediksi paling banyak terjadi pada kelas , di mana jumlah ulasan yang diprediksi dengan benar hanya 40 ulasan, sedangkan 49 ulasan justru diklasifikasikan sebagai netral. Pola ini menunjukkan bahwa model masih mengalami kesulitan membedakan sentimen dengan sentimen netral, yang kemungkinan disebabkan oleh kemiripan kata-kata yang digunakan atau ketidakseimbangan jumlah data antar kelas.

Evaluasi terhadap algoritma Multinomial Naïve Bayes menunjukkan bahwa model berhasil mencapai akurasi 79,67%, yang menandakan performa yang cukup baik untuk analisis sentimen berbasis teks ulasan pengguna Mobile JKN. Berdasarkan *classification report*, model tampil optimal pada kelas *netral* dengan recall 0.92 dan f1-score 0.83, menunjukkan kemampuan tinggi dalam mengenali ulasan bernada netral. Kelas *positif* juga menunjukkan kinerja stabil dengan precision 0.88 dan recall 0.75, sehingga model dapat mengidentifikasi sentimen positif secara cukup konsisten.

Sebaliknya, performa pada kelas masih menjadi tantangan. Nilai recall hanya 0.45, yang berarti banyak ulasan yang salah diklasifikasikan, terutama sebagai netral. Meskipun demikian, precision yang tinggi yaitu 0.85 menunjukkan bahwa prediksi yang dikeluarkan model sebagian besar benar. Confusion matrix memperkuat temuan ini, di mana kelas memiliki jumlah prediksi benar paling rendah dan mengalami misclassifications yang cukup signifikan ke kelas netral.

Secara keseluruhan, model Naïve Bayes efektif dalam mendeteksi sentimen positif dan netral, namun masih memerlukan peningkatan dalam menangani sentimen . Tantangan ini dapat diatasi melalui penyeimbangan data, penyesuaian fitur TF-IDF, atau penggunaan model yang lebih kompleks pada tahap pengembangan lanjutan. Model ini tetap layak sebagai baseline yang cepat, ringan, dan cukup akurat untuk analisis sentimen awal pada ulasan Aplikasi Mobile JKN.



Gambar 4 Evaluasi Model Naive Bayes

CONCLUSION

Kompleksitas bahasa informal pada ulasan Mobile JKN seperti penggunaan slang, singkatan, serta struktur kalimat tidak baku berhasil ditangani melalui tahapan prapengolahan yang komprehensif, yaitu case folding, cleaning, tokenization, normalization, stopwords removal, dan stemming dengan Sastrawi. Proses ini memastikan teks menjadi bersih, konsisten, dan siap diekstraksi menggunakan TF-IDF sehingga menghasilkan fitur yang representatif untuk kebutuhan klasifikasi sentimen.

Pemilihan Multinomial Naïve Bayes sebagai metode klasifikasi yang terbukti efektif dalam mendeteksi sentimen . Algoritma ini mampu mengolah data teks berdimensi tinggi dan sparse secara efisien. Model menunjukkan kemampuan yang baik dalam mengenali pola sentimen ketika fitur diekstraksi dengan TF-IDF. Meskipun beberapa ulasan masih salah diprediksi sebagai netral karena ambiguitas bahasa, model tetap menjadi baseline yang kuat sebagaimana didukung oleh penelitian terdahulu ((Palomino & Aider, 2022); (Tan et al., 2023)).

Evaluasi performa model menggunakan metrik terukur, yaitu accuracy, precision, recall, dan F1-score. Hasil evaluasi menunjukkan bahwa model memiliki kinerja yang cukup andal dalam mendeteksi sentimen , meskipun recall kelas masih rendah akibat ketidakseimbangan kelas dan variasi negasi yang belum tertangani optimal. Temuan pada confusion matrix mempertegas adanya mis-klasifikasi pada ulasan tertentu. Secara keseluruhan, performa model dinilai layak digunakan sebagai alat bantu analisis otomatis untuk memetakan keluhan pengguna Mobile JKN.

REFERENCE

- [1] M. A. Palomino And F. Aider, "Evaluating The Effectiveness Of Text Pre-Processing In Sentiment Analysis," *Applied Sciences*, Vol.

- 12, No. 17, P. 8765, 2022, Doi: 10.3390/App12178765.
- [2] K. L. Tan, C. P. Lee, And K. M. Lim, "Roberta-Gru: A Hybrid Deep Learning Model For Enhanced Sentiment Analysis," *Applied Sciences*, Vol. 13, No. 6, P. 3915, 2023, Doi: 10.3390/App13063915.
- [3] A. N. Ma'aly, D. Pramesti, A. D. Fathurahman, And H. Fakhurroja, "Exploring Sentiment Analysis For The Indonesian Presidential Election Through Online Reviews Using Multi-Label Classification With A Deep Learning Algorithm," *Information*, Vol. 15, No. 11, P. 705, 2024, Doi: 10.3390/Info15110705.
- [4] A. Jazuli, W. Widowati, And R. Kusumaningrum, "Optimizing Aspect-Based Sentiment Analysis Using Bert For Comprehensive Analysis Of Indonesian Student Feedback," *Applied Sciences*, Vol. 15, No. 1, P. 172, 2024, Doi: 10.3390/App15010172.