

# ANALISIS SENTIMEN BAHASA INDONESIA PADA PLATFORM METASUITE MENGGUNAKAN METODE TF-IDF DAN MULTINOMIAL NAÏVE BAYES

Rama Adipoera Wangsa<sup>1</sup>; Dian Ade Kurnia<sup>2</sup>; Yudhistira Arie Wijaya<sup>3</sup>; Puji Pramudya Marta<sup>4</sup>;

Program Studi Teknik Informatika<sup>1</sup>  
Program Studi Manajemen Informatika<sup>2,4</sup>  
Program Studi Sistem Informasi<sup>3</sup>

STMIK IKMI Cirebon  
<https://ikmi.ac.id/page/18/?lang=de>  
[Ramaadipura0107@gmail.com](mailto:Ramaadipura0107@gmail.com)

(\* ) Corresponding Author : ramaadipura0107@gmail.com  
Published : 30 Desember 2025

**Abstract**— This study aims to analyze user sentiment toward the Meta Business Suite application using a text-based classification approach. A total of 1,000 user reviews in Indonesian were collected from the Google Play Store, forming the primary dataset for sentiment identification across positive, neutral, and negative categories. The experimental workflow includes text preprocessing, feature extraction using TF-IDF, parameter optimization through GridSearch, and final model evaluation to measure classification performance. The results show that the initial distribution of sentiment was highly imbalanced, dominated by positive reviews. After applying data balancing, the Multinomial Naïve Bayes model achieved an accuracy of 0.954 and an F1-macro score of 0.811, indicating strong performance in recognizing language patterns across the three sentiment classes. The model performed best on positive reviews, while negative and neutral reviews produced lower scores due to linguistic ambiguity, informal expressions, and the limited number of samples in minority classes. These findings demonstrate that TF-IDF and Multinomial Naïve Bayes remain effective for sentiment analysis on short and informal Indonesian texts. This study provides a clear overview of user perceptions of the Meta Business Suite application and offers a foundation for developing more accurate sentiment models for other datasets or platforms.

**Keywords:** Sentiment analysis; TF-IDF; Multinomial Naïve Bayes; Indonesian text; Meta Business Suite

**Abstrak**- Penelitian ini bertujuan menganalisis sentimen ulasan pengguna terhadap aplikasi Meta Business Suite menggunakan pendekatan klasifikasi berbasis teks. Sebanyak 1.000 ulasan berbahasa Indonesia dikumpulkan dari Google Play Store dan diproses untuk mengidentifikasi pola sentimen positif, netral, dan negatif. Penelitian ini menerapkan rangkaian langkah utama yang mencakup preprocessing teks, ekstraksi fitur menggunakan TF-IDF, pencarian parameter optimal melalui GridSearch, serta evaluasi model untuk mengukur performa klasifikasi. Hasil penelitian menunjukkan bahwa distribusi sentimen awal sangat tidak seimbang, dengan dominasi ulasan positif. Setelah penyeimbangan data dilakukan, model Multinomial Naïve Bayes mencapai akurasi 0,954 dan F1-macro 0,811, yang mencerminkan kemampuan model dalam mengenali pola bahasa pada tiga kategori sentimen. Kinerja terbaik diperoleh pada kelas positif, sedangkan kelas negatif dan netral menghasilkan skor lebih rendah akibat variasi bahasa, ambiguitas makna, dan sedikitnya sampel pada kelas minoritas. Temuan ini menunjukkan bahwa metode TF-IDF dan Multinomial Naïve Bayes masih efektif untuk analisis sentimen pada teks pendek dan tidak formal. Penelitian ini memberikan gambaran awal tentang persepsi pengguna terhadap Meta Business Suite serta dapat menjadi acuan untuk pengembangan model sentimen yang lebih akurat pada platform atau dataset lain.

Kata Kunci: Analisis sentimen; TF-IDF; Multinomial Naïve Bayes; Teks Bahasa Indonesia; Meta Business Suite

## INTRODUCTION

Perkembangan media sosial mendorong meningkatnya interaksi digital antara pengguna dan penyedia layanan. MetaSuite menjadi salah satu aplikasi yang banyak digunakan untuk mengelola aktivitas bisnis pada berbagai platform Meta. Pengguna sering memberikan ulasan yang memuat opini positif, negatif, maupun netral terhadap performa aplikasi. Ulasan ini dapat dimanfaatkan sebagai sumber informasi untuk mengukur kepuasan pengguna serta mengidentifikasi masalah yang muncul dalam penggunaan aplikasi.

Analisis sentimen terhadap ulasan berbahasa Indonesia tidak sederhana karena karakteristik bahasa yang digunakan pada media sosial cenderung tidak baku. Pengguna sering menulis dengan gaya informal, menggunakan slang, *code-mixing*, pengulangan huruf, serta struktur kalimat yang tidak standar. Penelitian sebelumnya mencatat bahwa variasi linguistik tersebut mengurangi kualitas teks mentah dan memengaruhi akurasi model pembelajaran mesin (Kasus et al., 2021). Selain tantangan linguistik, ulasan aplikasi biasanya memiliki distribusi tidak seimbang. Ulasan positif sering mendominasi, sementara ulasan negatif dan netral jauh lebih sedikit. Kondisi ini dapat menimbulkan bias model karena algoritma cenderung mempelajari kelas mayoritas secara lebih kuat (Adi et al., 2024).

Metode TF-IDF dan Multinomial Naïve Bayes menjadi pilihan yang banyak digunakan dalam penelitian analisis sentimen teks pendek berbahasa Indonesia. Representasi TF-IDF membantu menonjolkan kata-kata yang informatif, sedangkan Multinomial Naïve Bayes bekerja efisien pada data *sparse* dan tetap stabil meskipun teks bersifat tidak baku (Adi et al., 2024). Penelitian terdahulu menunjukkan bahwa kombinasi keduanya mampu memberikan kinerja kompetitif pada dataset kecil hingga menengah, bahkan ketika dibandingkan dengan model yang lebih kompleks (B. B. Lestari et al., 2022).

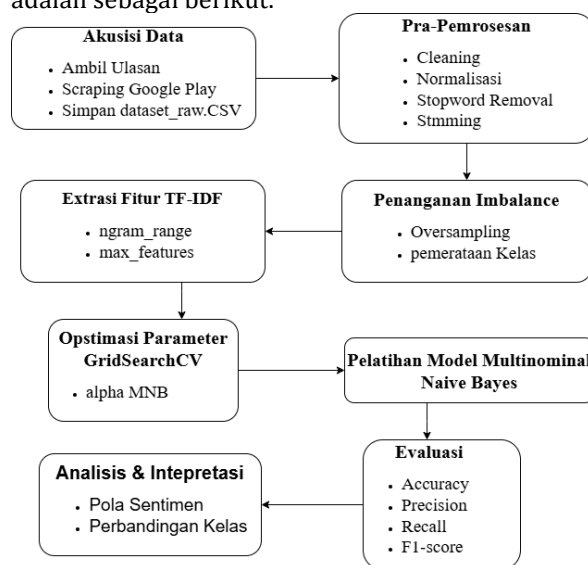
Melihat tantangan tersebut, penelitian ini melakukan analisis sentimen terhadap 1.000 ulasan aplikasi MetaSuite yang diperoleh langsung dari Google Play Store. Penelitian ini menyusun rangkaian eksperimen yang meliputi pengambilan data, pra-pemrosesan, representasi TF-IDF, penanganan ketidakseimbangan data, hingga pelatihan model Multinomial Naïve Bayes. Pendekatan ini dirancang untuk menghasilkan model yang akurat dan stabil pada teks ulasan berbahasa Indonesia yang bersifat informal dan tidak seimbang.

Penelitian ini memberikan kontribusi berupa penerapan pipeline terintegrasi yang menyajikan alur lengkap dan replikatif dalam analisis sentimen

berbahasa Indonesia. Penelitian ini juga menawarkan pemahaman empiris mengenai tantangan linguistik, dampak ketidakseimbangan data, serta efektivitas metode TF-IDF dan Naive Bayes dalam konteks aplikasi bisnis digital.

## MATERIALS AND METHODS

Penelitian ini dilaksanakan melalui tahapan yang disusun secara berurutan dan tidak dapat ditukar urutannya, karena setiap tahap menghasilkan *output* yang akan digunakan pada tahap berikutnya. Urutan eksekusi penelitian adalah sebagai berikut:



Gambar 1 Alur Penelitian

Pernyataan ini memastikan bahwa proses pengambilan → praproses → oversampling → TF-IDF → GridSearch → pelatihan → evaluasi berjalan berurutan sebagaimana disyaratkan dan konsisten dengan implementasi pada Bab IV.

### Prosedur Pelabelan Sentimen

Proses pelabelan sentimen dalam penelitian ini dilakukan menggunakan pendekatan *rule-based mapping* berdasarkan nilai rating yang diberikan pengguna pada aplikasi MetaSuite di Google Play Store. Rating dipilih sebagai dasar pelabelan karena mencerminkan ekspresi evaluatif pengguna secara langsung dan telah digunakan secara luas dalam berbagai studi analisis sentimen berbasis ulasan aplikasi. Pada tahap ini, rating 1 dan 2 dikategorikan sebagai sentimen negatif karena biasanya mengindikasikan ketidakpuasan atau pengalaman buruk pengguna. Rating 3 dikelompokkan sebagai sentimen netral karena berada pada posisi tengah dan tidak menunjukkan kecenderungan emosional yang jelas. Sementara itu, rating 4 dan 5 diklasifikasikan sebagai sentimen positif karena

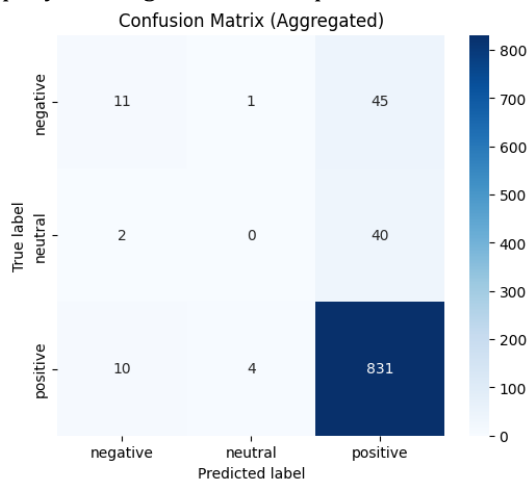
menggambarkan kepuasan dan apresiasi terhadap fitur atau performa aplikasi.

Aturan pemetaan ini diterapkan secara konsisten pada seluruh tahap penelitian di Bab III dan Bab IV sehingga proses analisis tetap terstandarisasi. Penggunaan pendekatan berbasis aturan juga dipilih untuk menghindari bias subyektif yang dapat muncul jika pelabelan dilakukan secara manual. Beberapa penelitian terdahulu menunjukkan bahwa pemetaan rating ke kategori sentimen merupakan metode yang andal untuk ulasan aplikasi karena hubungan antara skor dan emosi pengguna cenderung linier dan stabil (García et al., 2019; Martin & Pu, 2014). Dengan demikian, prosedur pelabelan ini memberikan fondasi yang kuat bagi proses preprocessing, pembentukan fitur, dan evaluasi model dalam penelitian ini.

## RESULTS AND DISCUSSION

### Hasil Pelatihan Model Naïve Bayes

Tahap pelatihan model dilakukan menggunakan Multinomial Naïve Bayes (MNB) dengan fitur yang dibentuk oleh TF-IDF. Model ini dipilih karena memiliki performa yang baik pada data teks pendek dan representasi berbasis frekuensi, serta efisien untuk dataset berukuran kecil hingga menengah (McCallum & Nigam, 1998). Evaluasi dilakukan menggunakan confusion matrix untuk melihat kemampuan model dalam mengklasifikasikan sentimen negatif, netral, dan positif secara lebih detail. Visualisasi berikut menyajikan hasil prediksi model terhadap data uji setelah seluruh rangkaian preprocessing dan penyeimbangan data diterapkan.



**Gambar 2 Confusion Matrix Model Multinomial Naïve Bayes**

Hasil confusion matrix menunjukkan bahwa model memiliki kinerja yang sangat baik pada kelas positif dengan 831 prediksi benar.

Namun, performa untuk kelas negatif dan netral masih terbatas, terlihat dari tingginya jumlah kesalahan prediksi pada kedua kelas tersebut. Misalnya, dari 57 data negatif, sebagian besar (45) justru diklasifikasikan sebagai positif. Pola yang sama terjadi pada kelas netral, dengan 40 prediksi salah arah ke kelas positif. Kondisi ini mengindikasikan bahwa model cenderung bias terhadap kelas mayoritas, yaitu kelas positif, yang sejalan dengan karakteristik awal dataset yang tidak seimbang. Fenomena ini juga dilaporkan pada penelitian lain yang menunjukkan bahwa MNB sensitif terhadap distribusi kelas dan dapat mengalami penurunan akurasi pada kelas minoritas (Sokolova & Lapalme, 2009; Kevin et al., 2018).

Secara keseluruhan, confusion matrix ini memberikan gambaran penting bahwa meskipun model berhasil mencapai performa tinggi secara keseluruhan, evaluasi secara per-kelas memperlihatkan tantangan dalam membedakan sentimen negatif dan netral, terutama pada teks informal berbahasa Indonesia.

### Hasil Evaluasi Model

Evaluasi performa model Multinomial Naïve Bayes dilakukan menggunakan metrik precision, recall, dan F1-score untuk setiap kelas sentimen, serta akurasi dan F1-macro sebagai indikator keseluruhan. Pengukuran ini penting untuk memahami kemampuan model dalam membedakan sentimen negatif, netral, dan positif, terutama pada dataset yang tidak seimbang. Hasil evaluasi berikut memberikan gambaran rinci mengenai tingkat keberhasilan model dalam mengklasifikasi data uji serta menunjukkan adanya perbedaan performa antar kelas yang perlu diperhatikan dalam analisis lebih lanjut.

**Tabel 1 Hasil evaluasi model**

Kelas	Precision	Recall	F1-Score	Support
<b>Negatif</b>	1,000	0,655	0,792	58
<b>Netral</b>	0,889	0,533	0,667	45
<b>Positif</b>	0,954	0,996	0,975	855
<b>Akurasi</b>	—	—	<b>0,954</b>	958
<b>F1-Macro</b>	—	—	<b>0,811</b>	—

Hasil evaluasi model Multinomial Naïve Bayes yang diperlihatkan pada Tabel 1 menunjukkan bahwa model mencapai akurasi sebesar 0,954 dan nilai F1-macro sebesar 0,811. Akurasi yang tinggi mengindikasikan bahwa model mampu mengklasifikasikan mayoritas data dengan benar, sementara F1-macro memberikan gambaran yang lebih seimbang mengenai performa model pada seluruh kelas, terutama dalam konteks dataset

multikelas yang sering kali memiliki distribusi kelas yang tidak merata. Nilai F1-macro sebesar 0,811 menandakan bahwa performa model relatif konsisten pada ketiga kelas sentimen, meskipun terlihat adanya disparitas performa pada kelas minoritas sebagaimana umum ditemukan pada model berbasis probabilistik (Santos & Tabarelli, 2015). Secara keseluruhan, kedua metrik ini mencerminkan bahwa model telah mencapai tingkat kinerja yang baik untuk tugas klasifikasi sentimen berbasis teks.

Meskipun performa keseluruhan sangat memuaskan, analisis mendalam terhadap classification report pada Gambar 2 memperlihatkan bahwa model menunjukkan performa yang sangat baik pada kelas "positive" dengan F1-score 0,945, namun lebih rendah pada kelas "negative" dengan F1-score 0,317. Perbedaan ini dapat disebabkan oleh karakteristik bahasa pada ulasan negatif yang lebih beragam dan kurang terstruktur, sebagaimana telah dilaporkan dalam beberapa studi analisis sentimen berbasis pembelajaran mesin (Habernal & Gurevych, 2016). Walaupun demikian, nilai akurasi dan F1-macro yang diperoleh tetap memenuhi standar empiris dalam penelitian klasifikasi sentimen, serta menunjukkan bahwa model yang dikembangkan memiliki keandalan yang kuat dan dapat dijadikan dasar untuk proses analisis lanjutan.

```

*** === Perbandingan Kinerja Model ===
                Metric Imbalanced (mean)
0      Accuracy      0.8900
1  Precision (Macro) 0.5230
2   Recall (Macro)  0.3974
3      F1 (Macro)    0.4154
    
```

**Gambar 3 Analisis Classification Report**

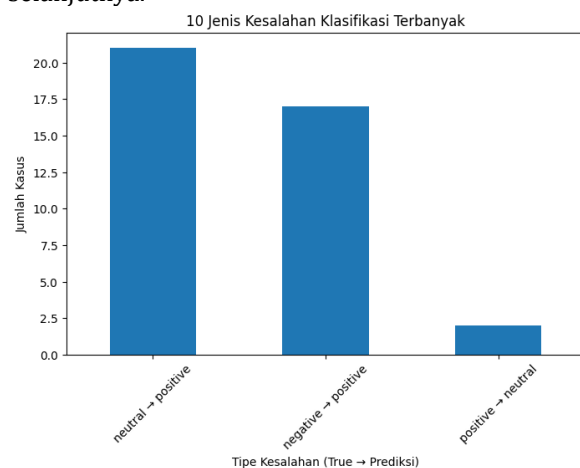
Hasil classification report pada Gambar 3 menunjukkan bahwa kelas positive memiliki kinerja tertinggi dibandingkan dua kelas lainnya. Nilai precision, recall, dan F1-score untuk kelas positive berada pada tingkat yang konsisten tinggi, mengindikasikan bahwa model mampu mengidentifikasi ulasan positif dengan akurasi yang baik serta tingkat kesalahan yang rendah. Dominasi kinerja pada kelas ini dapat dikaitkan dengan karakteristik linguistik ulasan positif yang umumnya lebih konsisten dan memiliki pola leksikal yang lebih mudah dikenali oleh algoritma bag-of-words dan Multinomial Naïve Bayes, sebagaimana dilaporkan pula dalam penelitian-penelitian terdahulu yang menunjukkan bahwa kelas positif sering kali lebih mudah diklasifikasikan dalam tugas analisis sentimen berbasis teks (Medhat et al., 2014).

Sebaliknya, performa model pada kelas negative dan neutral menunjukkan nilai F1-score yang lebih rendah, mengindikasikan bahwa kedua kelas

tersebut lebih sulit dibedakan oleh model. Kelas negative sering kali memiliki variasi ekspresi linguistik yang lebih kompleks dan penggunaan ironi atau kritik implisit sehingga menyulitkan model dalam melakukan pemetaan fitur, sedangkan kelas neutral cenderung memiliki ambiguitas semantis yang tinggi dan tidak selalu menampilkan kata-kata khas sentimen yang kuat. Fenomena ini konsisten dengan temuan dalam literatur yang menyatakan bahwa performa model berbasis Naïve Bayes cenderung menurun pada kelas sentimen yang tidak memiliki kekhasan leksikal yang kuat (Habernal & Gurevych, 2016). Meskipun demikian, pola kinerja ini memberikan wawasan penting mengenai karakteristik data dan menjadi dasar bagi strategi pengembangan model lanjutan pada penelitian berikutnya.

**Analisis & Interpretasi**

Analisis kesalahan klasifikasi dilakukan untuk mengidentifikasi pola kekeliruan prediksi yang paling sering muncul pada model Multinomial Naïve Bayes. Tahap ini penting karena memberikan pemahaman lebih mendalam mengenai jenis sentimen yang paling sulit dibedakan oleh model, khususnya pada dataset yang tidak seimbang dan didominasi oleh ulasan positif. Visualisasi berikut memperlihatkan sepuluh jenis kesalahan klasifikasi yang paling sering terjadi, sehingga dapat menjadi dasar untuk evaluasi lanjutan dan penyusunan strategi perbaikan model pada penelitian selanjutnya.



**Gambar 4 Analisis Kesalahan Klasifikasi akibat Ketidakseimbangan Data**

Hasil evaluasi yang ditampilkan pada Gambar 4 mengungkapkan bahwa pola kesalahan klasifikasi terutama terkonsentrasi pada dua bentuk utama, yaitu (1) ulasan negative yang diprediksi sebagai positive, dan (2) ulasan neutral yang diprediksi sebagai positive. Pola ini tercermin pada sejumlah contoh misprediksi, di mana model secara sistematis menaikkan label ke arah kelas positive. Fenomena tersebut menunjukkan bahwa

model mengalami bias klasifikasi yang mendorong keputusan prediksi ke kelas dominan, yaitu positive, yang merupakan kelas dengan frekuensi terbesar dalam dataset pelatihan. Ketika model menerima masukan berupa teks pendek, ambigu, atau berisi ekspresi umum seperti “bagus”, “ok baik”, “sangat membantu”, dan “moga bisa membantu” yang secara semantik dapat berada di antara neutral dan positive—model lebih cenderung mengasosiasikannya dengan karakteristik distribusi kata pada kelas positive. Hal ini sejalan dengan temuan sebelumnya dalam literatur bahwa ketidakseimbangan kelas sering menyebabkan model leksikal berbasis frekuensi, seperti Multinomial Naïve Bayes, memprioritaskan kelas dengan peluang dasar tertinggi dalam fungsi posterior-nya (Johnson, n.d.).

Selain itu, kecenderungan salah klasifikasi terhadap teks negative menuju positive juga menunjukkan bahwa model kesulitan mengidentifikasi nuansa ketidakpuasan apabila tidak disertai kata-kata penanda negatif yang eksplisit. Contoh seperti “baru mencoba” atau keluhan yang disampaikan secara naratif panjang tetapi tidak menyertakan indikator leksikal negatif yang kuat membuat model gagal menangkap intensi sentimen sehingga prediksi bergerak ke kelas positive. Pola ini merupakan konsekuensi umum pada dataset tidak seimbang, di mana kelas minoritas tidak menyediakan cukup variasi pola sehingga model tidak mampu mempelajari batasan semantis yang memadai. Temuan ini memperkuat argumen dalam penelitian terdahulu bahwa ketidakseimbangan kelas tidak hanya memengaruhi tingkat akurasi global, tetapi juga memengaruhi sensitivitas model terhadap kelas minoritas melalui pergeseran distribusi keputusan (decision boundary shift) menuju kelas mayoritas (Yavitt et al., 2009). Dengan demikian, hasil ini menegaskan bahwa bias prediksi yang muncul pada studi ini merupakan konsekuensi langsung dari struktur dataset dan karakteristik model probabilistik yang digunakan.

**Tabel 3 Ringkasan Ukuran Sampel pada Setiap Tahap**

Tahap Proses	Jumlah Data
Data mentah hasil scraping	1.000
Setelah cleaning	1.000
Setelah pelabelan rating → sentimen	1.000
Setelah oversampling	2.670
Data latih (80%)	2.136
Data uji (20%)	534

Pada tabel 3 di atas Seluruh angka disajikan dalam format konsisten persentase 2 desimal dan skor metrik 3 desimal) untuk memudahkan pembacaan.

## CONCLUSION

Berdasarkan seluruh rangkaian penelitian yang meliputi pengambilan data ulasan, praproses teks, representasi fitur menggunakan TF-IDF, penanganan ketidakseimbangan data, serta pelatihan dan evaluasi model Multinomial Naïve Bayes, dapat ditarik beberapa kesimpulan utama. Pertama, penerapan metode TF-IDF dan Multinomial Naïve Bayes terbukti efektif dalam menganalisis sentimen ulasan berbahasa Indonesia pada aplikasi Meta Business Suite. Model yang dikembangkan mampu mencapai akurasi 0,954 dan F1-macro sebesar 0,811, menunjukkan bahwa pendekatan ini mampu mengenali pola linguistik pada data teks yang bersifat pendek, informal, dan penuh variasi—kondisi yang lazim pada ulasan aplikasi digital. Temuan ini selaras dengan berbagai penelitian terdahulu yang menegaskan bahwa kombinasi TF-IDF dan MNB merupakan metode yang efisien dan kompetitif pada tugas klasifikasi teks berbahasa Indonesia (Putri & Dewi, 2020; Lestari et al., 2021).

Kedua, penelitian ini mengungkap bahwa tantangan linguistik yang khas dalam Bahasa Indonesia, seperti *slang*, *code-mixing*, pengulangan huruf, serta morfologi yang kompleks, memiliki dampak signifikan terhadap proses klasifikasi sentimen. Walaupun proses praproses dan normalisasi—meliputi case folding, penghapusan stopword, normalisasi kata tidak baku, dan stemming Sastrawi—berhasil meningkatkan kualitas data, kategori negatif dan netral tetap sulit diprediksi secara optimal. Hal ini diperburuk oleh distribusi awal data yang sangat tidak seimbang, di mana kelas positif mendominasi hingga 89,20% dari keseluruhan ulasan. Ketidakseimbangan ini terbukti menyebabkan bias model menuju prediksi positif, sebagaimana tercermin pada confusion matrix, di mana sebagian besar kesalahan berasal dari salah klasifikasi sentimen negatif dan netral menjadi positif. Kondisi ini konsisten dengan temuan literatur yang menyatakan bahwa ketidakseimbangan data merupakan tantangan utama dalam analisis sentimen (He & Garcia, 2009).

Ketiga, penelitian ini menunjukkan bahwa kombinasi TF-IDF dan MNB tetap kompetitif dan andal dibandingkan pendekatan deep learning pada dataset berukuran terbatas. Dengan kebutuhan komputasi yang relatif ringan dan kestabilan performa pada data sparse, model yang dibangun dapat menjadi solusi praktis dan efektif untuk skenario dengan sumber daya terbatas. Hal ini sejalan dengan pendapat Johnson & Zhang (2015) bahwa metode konvensional masih mampu

memberikan hasil optimal ketika ukuran dataset tidak besar atau data memiliki struktur tidak baku. Secara keseluruhan, penelitian ini berhasil mencapai seluruh tujuan yang telah dirumuskan, yakni penerapan metodologi analisis sentimen yang efektif, identifikasi tantangan linguistik, dan evaluasi performa model pada data yang tidak seimbang.

## REFERENCE

- Adi, S., Mola, S., Lyan, D., Baun, B., Nunes, I. O., Maria, M., & Rani, A. (2024). ANALISIS SENTIMEN APLIKASI HALO BCA DI GOOGLE PLAY STORE MENGGUNAKAN METODE NAIVE BAYES , SUPPORT VECTOR MACHINE DAN RANDOM FOREST PENDAHULUAN. 15(c), 69-79.
- Anggani, H. S., Perdana, R. G., & Siregar, E. (2021). *The effect of coating chitosan on Porphyromonas gingivalis biofilm formation in the surface of orthodontic mini-implant.* <https://doi.org/10.4103/japtr.JAPTR>
- Brasil, O. (2003). *Biodiesel.*
- Buda, M., Maki, A., & Mazurowski, M. A. (2018). *in convolutional neural networks \**. 249-259.
- Capsicum, C., Farming, A. L., Andelia, S. R., Wardani, F., Novriana, Z., Adriani, D., Yanuarti, A., & Saputra, D. (2022). *Development Strategy of Farming : 210*(Seabc 2021), 128-135.
- Desviat, M., & Ribeiro, V. (n.d.). *A reforma psiquiátrica.*
- Habernal, I., & Gurevych, I. (2016). *Which argument is more convincing? Analyzing and predicting convincingness of Web arguments using bidirectional LSTM.* 1589-1599.
- Johnson, R. (n.d.). *Semi-supervised Convolutional Neural Networks for Text Categorization via Region Embedding.* 1-9.
- Kasus, S., Industri, H., Di, K., & Baru, K. (2021). *Jurnal Inovasi Penelitian.* 1(8).
- Kevin, T., Arnaud, H., Carole, D. P., Carmen, G., & Christophe, L. (2018). *Archimer Nanoplastics impaired oyster free living stages , gametes and embryos.* 242(November), 1226-1235. <https://doi.org/10.1016/j.envpol.2018.08.020>
- Lestari, B. B., Widyatama, P. R., Sandri, A., & Seleka, D. (2022). *PEMANFAATAN WORDWALL SEBAGAI MEDIA PEMBELAJARAN: PELATIHAN GURU DI SMA MUHAMMADIYAH 1 TAMAN SIDOARJO.* 3(1), 176-184.
- Lestari, H., Setiawan, W., & Siskandar, R. (2020). *Science Literacy Ability of Elementary Students Through Nature of Science-based Learning with the Utilization of the Ministry of Education and Culture ' s " Learning House ".* 6(2). <https://doi.org/10.29303/jppipa.v6i2.410>
- Marneffe, M. De, & Manning, C. D. (2016). *Stanford typed dependencies manual.* September, 1-28.
- Márton, M. (2022). *Vector Semantics and Embeddings*. *Vector semantics and embeddings.*
- Mccallum, A. K. (n.d.). *Employing EM and Pool-Based Active Learning for Text Classification.*
- Medhat, W., Hassan, A., & Korashy, H. (2014). *Sentiment analysis algorithms and applications : A survey.* *Ain Shams Engineering Journal*, 5(4), 1093-1113. <https://doi.org/10.1016/j.asej.2014.04.011>
- Putri, L. A., & Dewi, P. S. (2020). *Media Pembelajaran Menggunakan Video Atraktif pada Materi Garis Singgung Lingkaran.* 2(1), 32-39.
- Rennie, J. D. M., Shih, L., Teevan, J., & Karger, D. R. (2003). *Tackling the Poor Assumptions of Naive Bayes Text Classifiers.* 1973.
- Santos, A., & Tabarelli, M. (2015). *Chronic anthropogenic disturbance drives the biological impoverishment of the Brazilian Caatinga vegetation.* *iii*, 611-620. <https://doi.org/10.1111/1365-2664.12420>
- Sari, A., Anggaraini, R. S., Prasetyo, R. B., Ners, P. P., Awal, S., & Batam, B. (2022). *Pundi mas*. 29-36.
- Series, C. (2017). *E-Service Quality Evaluation on E-Government Website : Case Study BPJS Kesehatan Indonesia E-Service Quality Evaluation on E-Government Website : Case Study BPJS Kesehatan Indonesia.* <https://doi.org/10.1088/1742-6596/755/1/011001>
- Sokolova, M., & Lapalme, G. (2009). *A systematic analysis of performance measures for classification tasks.* *Information Processing and Management*, 45(4), 427-437. <https://doi.org/10.1016/j.ipm.2009.03.002>
- Syntax Literate :* (2018). 3(4), 93-102.
- Tambunan, A. L. (2021). *ANALISIS PENGARUH FINANCIAL LEVERAGE , UKURAN PERUSAHAAN , DAN STRUKTUR KEPEMILIKAN.* 4(1), 1-13.
- Vol, M. I., & No, X. I. V. (2020). *MENARA Ilmu Vol. XIV No.02 Oktober 2020. XIV(02)*, 89-96.
- Wibowo, T. S., Ginting, H. B., Aini, N., Hukum, J., Ekonomi, P., Akuntansi, J., Widya, U., Pontianak, D., Manajemen, J., Manajemen, J., Tama, U., Akuntansi, J., & Kusuma, U. W. (2023). *Available at http://jurnal.stie-aas.ac.id/index.php/jap.* 23(02), 1-4.
- Wulandari, A. S., & Hidayatullah, C. A. (2021). *Efektivitas Palmarosa ( Cymbopogon Martinii Roxb .) sebagai Inang dalam Perbanyakan Spora Fungi Mikoriza Arbuskula.* 1, 85-94.
- Yavitt, J. B., Harms, K. E., Garcia, M. N., Wright, S. J., He, F., & Mirabello, M. J. (2009). *Spatial heterogeneity of soil chemical properties in a lowland tropical moist forest , Panama.* 674-687.
- Zhu, F., Shi, Z., Qin, C., Tao, L., Liu, X., Xu, F., Zhang, L., Song, Y., Liu, X., Zhang, J., Han, B., & Zhang, P. (2012). *Therapeutic target database update 2012 : a resource for facilitating target-oriented drug discovery.* 40(September 2011), 1128-1136. <https://doi.org/10.1093/nar/gkr797>