

KLASIFIKASI SENTIMEN ULASAN *E-COMMERCE* MENGGUNAKAN *NAÏVE BAYES*, *LIGHTGBM*, DAN *GRU*

Aulia Dwi Yulianti ¹⁾ Tony ²⁾

¹⁾Program Studi Teknik Informatika Fakultas Teknologi Informasi Universitas Tarumanagara

Jl. Letjen S. Parman No. 1, Jakarta Barat

email : aulia.535220178@stu.untar.ac.id

²⁾Program Studi Sistem Informasi Fakultas Teknologi Informasi Universitas Tarumanagara

Jl. Letjen S. Parman No. 1, Jakarta Barat

email : tony@fti.untar.ac.id

ABSTRAK

Ulasan pelanggan pada platform *e-commerce* merupakan data krusial untuk wawasan bisnis, namun volumenya yang besar menuntut adanya analisis sentimen otomatis. Penelitian ini bertujuan untuk merancang dan membangun sebuah prototipe aplikasi web yang fungsional untuk melakukan klasifikasi sentimen pada ulasan *e-commerce* secara otomatis. Sistem ini dirancang untuk mengimplementasikan dan membandingkan kinerja dari tiga model *machine learning* yang berbeda: metode probabilistik *Naïve Bayes*, *ensemble learning LightGBM*, dan *deep learning* sekuensial *Gated Recurrent Unit (GRU)*. Data ulasan (20.100) dikumpulkan dari *Google Play Store* melalui *web scraping* dan melalui tahap pra-pemrosesan teks sebelum diimplementasikan pada ketiga model. Hasil pengujian kinerja menunjukkan adanya *trade-off* antara performa dan efisiensi. *GRU* mencapai *F1-Score* tertinggi (0.8082) namun dengan waktu komputasi terlama (877.72 detik). *Naïve Bayes* menunjukkan Akurasi tertinggi (0.8841) dan waktu pelatihan tercepat (0.027 detik), namun dengan nilai *Recall* terendah (0.7420). Pengujian fungsionalitas (*Black Box Testing*) membuktikan semua fitur berjalan "valid", dan pengujian *System Usability Scale (SUS)* terhadap 30 responden menghasilkan skor rata-rata 93.84 (kategori "Excellent").

Key words

Analisis Sentimen, *GRU*, Klasifikasi, *LightGBM*, *Naïve Bayes*

1. Pendahuluan

Ulasan pelanggan memainkan peran krusial dalam ekosistem pemasaran digital, khususnya pada platform *e-commerce*. Tidak hanya berfungsi sebagai sumber informasi bagi calon pembeli untuk menilai kualitas produk atau layanan secara lebih objektif, ulasan juga

memberikan umpan balik berharga bagi pelaku usaha dalam meningkatkan kualitas layanan, memperbaiki produk, serta memperkuat hubungan dengan pelanggan [1]. Seiring pertumbuhan pesat *e-commerce* di Indonesia, volume ulasan yang dihasilkan setiap hari menjadi sangat besar menghasilkan ribuan ulasan baru setiap harinya, sehingga analisis terhadapnya menjadi sangat krusial untuk mempertahankan daya saing.

Untuk menjawab tantangan tersebut, otomatisasi analisis sentimen menggunakan *machine learning* dengan menerapkan teknik Pemrosesan Bahasa Alami (*Natural Language Processing*) hadir sebagai solusi strategis untuk mengubah data ulasan mentah menjadi wawasan bisnis yang terukur [2]. Keberhasilan implementasi analisis sentimen sangat bergantung pada pemilihan algoritma klasifikasi yang tepat. Hal ini menjadi tantangan tersendiri, terutama saat berhadapan dengan kompleksitas dan nuansa bahasa manusia seperti sarkasme, bahasa sehari-hari, dan konteks kalimat yang ambigu [3]. Penelitian ini bertujuan untuk merancang dan membangun sebuah prototipe aplikasi berbasis web yang fungsional untuk melakukan klasifikasi sentimen ulasan secara otomatis. Secara spesifik, luaran utama dari sistem yang dirancang adalah menyajikan perbandingan kinerja klasifikasi yang dilihat dari akurasi, presisi, recall, *F1-score* serta efisiensi komputasi dari ketiga metode secara objektif.

Berdasarkan tujuan tersebut, perancangan ini difokuskan untuk menjawab beberapa rumusan masalah utama. Pertama, bagaimana merancang arsitektur sistem yang mampu mengolah teks ulasan secara otomatis, mulai dari tahap pra-pemrosesan hingga menghasilkan klasifikasi sentimen. Lalu yang kedua, bagaimana mengimplementasikan tiga model klasifikasi (*Naïve Bayes*, *LightGBM*, dan *GRU*) ke dalam sistem yang dirancang agar dapat memprediksi sentimen secara efektif. Terakhir ketiga, bagaimana hasil perbandingan kinerja dan efisiensi komputasi dari ketiga model tersebut ketika diuji pada dataset yang sama

Metode yang digunakan adalah dengan membandingkan tiga model dari paradigma *machine learning* yang berbeda yaitu metode probabilistik *Naïve Bayes*, *ensemble learning* LightGBM, dan *deep learning* sekuensial *Gated Recurrent Unit* (GRU) [4]. Data ulasan publik dikumpulkan dari platform Google Play Store melalui teknik *web scraping*, kemudian melalui tahap pra-pemrosesan sebelum diolah oleh ketiga model. Beberapa penelitian terdahulu telah melakukan analisis sentimen, seperti menggunakan Xgboost [5], IndoBERT [6], dan SVM [7], namun masih terbatas rancangan sistem yang secara komprehensif mengintegrasikan algoritma dari paradigma berbeda pada data ulasan e-commerce berbahasa Indonesia.

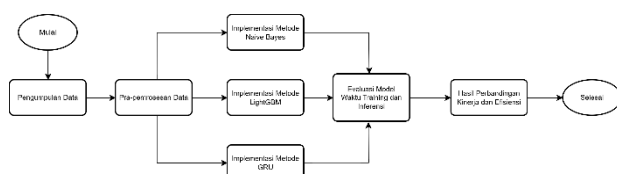
Kegunaan rancangan ini adalah menyediakan platform fungsional untuk menguji dan membandingkan kinerja serta efisiensi dari tiga paradigma model yang berbeda. Hasil perbandingan ini dapat memberikan bukti kuantitatif dan acuan bagi peneliti atau pengembang dalam memilih arsitektur model yang paling optimal untuk tugas klasifikasi ulasan serupa.

2. Metode Penelitian

Sistem klasifikasi sentimen ini dirancang sebagai sebuah prototipe aplikasi berbasis web. Perancangan sistem ini mengadaptasi metodologi *System Development Life Cycle* (SDLC) dengan model *waterfall*, yang mencakup tahap perencanaan, analisis, perancangan, implementasi, dan pengujian.

2.1. Alur dan Perancangan Sistem

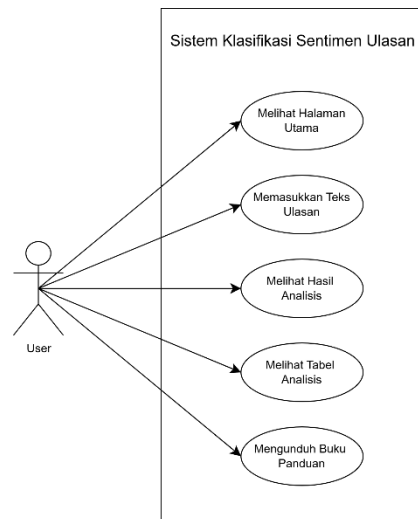
Secara konseptual, alur kerja sistem dimulai dari pengumpulan data, yang kemudian dilanjutkan ke tahap pra-pemrosesan data. Setelah data bersih, data tersebut akan diolah secara bertahap menggunakan tiga implementasi model yang berbeda (*Naïve Bayes*, LightGBM, dan GRU). Kinerja dari ketiga model tersebut kemudian dievaluasi menggunakan metrik yang telah ditentukan, dan hasilnya disajikan sebagai perbandingan akhir. Diagram alur sistem yang dirancang dapat dilihat pada **Gambar 1**.



Gambar 1 Alur Kerja Sistem

Pengguna dapat melakukan beberapa fungsionalitas utama, seperti memasukkan teks ulasan yang ingin dianalisis, melihat hasil analisis sentimen secara langsung, serta melihat halaman perbandingan kinerja dan waktu komputasi dari ketiga model yang diimplementasikan. Interaksi pengguna dengan sistem

dibuat menggunakan *Use Case Diagram* seperti yang ditunjukkan pada **Gambar 2**.



Gambar 2 Use Case Diagram

2.2. Data

Data yang digunakan dalam penelitian ini adalah 20.100 ulasan pengguna aplikasi *e-commerce* yang berbahasa Indonesia, yang diambil dari platform Google Play Store melalui teknik *web scraping* dalam rentang periode waktu 2018 hingga 2025. Dari dataset tersebut, dilakukan proses pelabelan sentimen biner berdasarkan fitur *score* sebagai penilaian. Ulasan dengan skor 4 dan 5 diberi label positif, sedangkan ulasan dengan skor 1 dan 2 diberi label negatif. Ulasan dengan skor 3 diabaikan karena dianggap netral, dengan tujuan untuk mempertajam analisis klasifikasi.

Setelah data dilabeli, setiap teks ulasan dari fitur *content* akan melalui tahapan pra-pemrosesan. Tahapan ini meliputi *Case Folding* (mengubah teks menjadi huruf kecil), *Cleansing* (membersihkan teks dari tanda baca, angka, dan karakter non-alfabet), *Stopword Removal* (menghapus kata-kata umum yang tidak memiliki makna penting), *Stemming* (mengubah kata ke bentuk dasarnya menggunakan pustaka Sastrawi), serta Tokenisasi (memecah kalimat menjadi unit kata tunggal). Data yang telah bersih kemudian direpresentasikan ke dalam format numerik yang siap diolah oleh model. Untuk model *Naïve Bayes* dan LightGBM, data direpresentasikan menggunakan metode TF-IDF (*Term Frequency-Inverse Document Frequency*). Sementara untuk model GRU, data diubah menjadi sekuens vektor menggunakan lapisan Embedding dan disamakan panjangnya melalui proses Padding.

2.3. Model Klasifikasi

Penelitian ini membandingkan kinerja tiga model klasifikasi *machine learning* yang dipilih untuk mewakili paradigma yang berbeda. Berikut adalah penjelasan dari ketiga model tersebut:

2.3.1. Naïve Bayes

Naïve Bayes Classifier adalah metode klasifikasi berbasis probabilitas yang digunakan untuk memprediksi kategori seperti sentimen positif atau negatif berdasarkan data yang ada. Model ini bekerja dengan menggunakan Teorema Bayes, yang secara matematis dirumuskan sebagai berikut.

$$P(y|X) = \frac{P(X|y) \cdot P(y)}{P(X)} \quad (1)$$

Keterangan:

1. $P(y|X)$ = Probabilitas posterior
2. $P(X|y)$ = Probabilitas bersyarat kelas atau *Likelihood*
3. $P(y)$ = Probabilitas prior dari hipotesis y
4. $P(X)$ = Probabilitas dari data X

Model ini juga membuat asumsi bahwa setiap kata dianggap independen satu sama lain setelah kelasnya diketahui [8]. Istilah "*naïve*" (lugu) digunakan karena model ini memiliki asumsi sederhana bahwa setiap kata dalam ulasan bersifat independen dan tidak saling mempengaruhi satu sama lain. Meskipun asumsi independensi ini sering kali tidak sepenuhnya benar dalam bahasa manusia, *Naïve Bayes* dikenal sangat cepat, efisien, dan seringkali memberikan hasil yang baik, menjadikannya *baseline* yang populer dalam tugas klasifikasi teks [9].

2.3.2. LightGBM

Light Gradient Boosting Machine (LightGBM) adalah kerangka kerja *machine learning* berbasis algoritma *Gradient Boosting Decision Tree* (GBDT). Algoritma ini dirancang agar efisien dan skalabel, khususnya untuk dataset besar dengan banyak fitur. GBDT membangun pohon keputusan secara berurutan, di mana setiap pohon baru memperbaiki kesalahan (residual) dari pohon sebelumnya. Untuk klasifikasi biner, residual ini atau juga disebut gradien dihitung sebagai selisih antara label asli dan probabilitas prediksi sebelumnya, dengan rumus:

$$r_i = y_i - p_i \quad (2)$$

Keterangan:

1. r_i = Nilai gradien (residual) untuk data ke- i
2. y_i = Label asli untuk data ke- i (0=negatif, 1=positif)
3. p_i = Probabilitas hasil prediksi sebelumnya

Setiap pohon baru kemudian dilatih untuk memprediksi residual ini. Nilai output untuk setiap daun (*leaf*) pada pohon baru tersebut dihitung menggunakan rumus:

$$\text{Nilai Output} = \frac{\sum \text{Residual}}{\sum (p_{\text{sebelumnya}} \times (1 - p_{\text{sebelumnya}}))} \quad (3)$$

Keterangan:

1. $\sum \text{Residual}$ = Jumlah gradien (residual) dalam satu *leaf*
2. $p_{\text{sebelumnya}}$ = Probabilitas prediksi dari iterasi sebelumnya
3. $\sum (p_{\text{sebelumnya}} \times (1 - p_{\text{sebelumnya}}))$ = varians Bernoulli, ekuivalen dengan Hessian

Hasil akhirnya adalah gabungan pohon-pohon tersebut sehingga prediksi menjadi lebih akurat [10]. LightGBM juga dikenal mampu menangani data tidak seimbang dan mengurangi risiko *overfitting*. Keunggulan ini membuat LightGBM menjadi salah satu algoritma populer dalam berbagai tugas klasifikasi modern, termasuk analisis sentimen [11].

2.3.3. GRU

Gated Recurrent Unit (GRU) adalah salah satu varian dari *Recurrent Neural Network* (RNN) yang dirancang khusus untuk memproses data sekuensial seperti teks. Model ini hanya menggunakan dua gerbang, yaitu *update gate* dan *reset gate*, untuk mengatur aliran informasi [12].

$$z_t = \sigma(W_z x_t + U_z h_{t-1} + b_z) \quad (4)$$

$$r_t = \sigma(W_r x_t + U_r h_{t-1} + b_r) \quad (5)$$

Keterangan:

1. z_t = Nilai *update gate*
2. r_t = Nilai *reset gate*
3. σ = Fungsi sigmoid
4. W_z = Bobot input pada *update gate*
5. x_t = Input pada waktu t
6. U_z = Bobot *hidden state* lama pada *update gate*
7. h_{t-1} = *Hidden state* dari waktu sebelumnya
8. b_z = Bias untuk *update gate*

Kedua gerbang ini digunakan untuk menghitung *hidden state* akhir (h_t) yang menggabungkan informasi lama dan baru. Hal ini membuat GRU lebih efisien secara komputasi dibandingkan algoritma *deep learning* lainnya karena memiliki jumlah parameter yang lebih sedikit, namun tetap efektif dalam menangkap ketergantungan jangka panjang dalam sebuah kalimat [13]. Kemampuan ini sangat bermanfaat pada klasifikasi ulasan, karena GRU dapat memahami bagaimana kata-kata dalam ulasan saling mempengaruhi secara kontekstual. Sebagai contoh, model dapat memahami bahwa kata "tidak" yang muncul di awal kalimat akan mengubah arti kata "bagus" di akhir kalimat, sehingga menghasilkan klasifikasi sentimen yang lebih akurat [14].

3. Hasil dan Pembahasan

3.1. Hasil Pengujian

Pengujian sistem dilakukan melalui beberapa pendekatan, yaitu pengujian kinerja model klasifikasi, *Black Box Testing* untuk fungsionalitas, dan *System Usability Scale* (SUS) untuk mengukur kemudahan penggunaan aplikasi.

3.1.1. Hasil Uji Metode Klasifikasi

Pengujian kinerja bertujuan untuk mengevaluasi kemampuan prediktif dan efisiensi komputasi dari ketiga model yaitu *Naïve Bayes*, *LightGBM*, dan *GRU*. Pengujian ini dilakukan pada *test set* yang terdiri dari 3.669 ulasan (20% dari total dataset). Hasil evaluasi kinerja model diukur menggunakan metrik Akurasi, Presisi, Recall, dan F1-Score, dan juga dengan efisiensi komputasi (waktu pelatihan dan inferensi) yang disajikan pada **Tabel 1** dan **Tabel 2**. Model diurutkan berdasarkan F1-Score tertinggi sebagai indikator kinerja yang paling seimbang.

Tabel 1 Hasil Evaluasi Kinerja

Metode	Akurasi	Presisi	Recall	F1-Score
GRU	0.8852	0.8252	0.7984	0.8116
<i>Naïve Bayes</i>	0.8841	0.8646	0.7420	0.7986
<i>LightGBM</i>	0.8555	0.7831	0.7376	0.7597

Tabel 2 Hasil Evaluasi Kinerja

Metode	Training (s)	Inference (s)
GRU	877.72	17.19
<i>Naïve Bayes</i>	0.027	0.0031
<i>LightGBM</i>	47.68	0.1058

3.1.2. Hasil Uji *Black Box Testing*

Black Box Testing adalah metode pengujian perangkat lunak yang fokus pada pengujian fungsionalitas aplikasi tanpa memperhatikan struktur internal atau cara kerja kode di dalamnya [15]. Pengujian meliputi fungsionalitas halaman beranda, analisis, tabel analisis, dan tentang. Setiap skenario uji dilakukan secara manual dengan memberikan input ulasan teks, menekan tombol analisis, serta memverifikasi hasil keluaran sistem apakah sesuai dengan rancangan. Hasil pengujian ditunjukkan pada **Tabel 3**.

Tabel 3 Hasil *Black Box Testing*

No.	Fitur yang Diuji	Skenario Uji Coba	Hasil yang Diharapkan	Status
1.	Beranda	Membuka aplikasi	Halaman beranda	valid

No.	Fitur yang Diuji	Skenario Uji Coba	Hasil yang Diharapkan	Status
		dan menampilkan halaman utama	muncul dengan deskripsi aplikasi	
2.	Analisis	Memasukkan ulasan "Tokopedia sangat membantu dan mudah digunakan"	Sistem menampilkan hasil Positif	valid
3.	Analisis	Memasukkan ulasan "Aplikasi sering error dan lambat"	Sistem menampilkan hasil Negatif	valid
4.	Analisis	Menekan tombol Analisis tanpa teks (input kosong)	Sistem menampilkan pesan "Masukkan teks terlebih dahulu."	valid
5.	Tabel Analisis	Membuka menu Tabel Analisis	Grafik dan tabel perbandingan dan tabel efisiensi komputasi muncul	valid
6.	Tabel Analisis	Menekan tombol Unduh Hasil Analisis	File CSV berhasil diunduh	valid
7.	Tentang	Mengunduh file manual book	File PDF berhasil diunduh	valid

3.1.3. Hasil Uji *System Usability Scale* (SUS)

System Usability Scale (SUS) adalah kuesioner yang digunakan untuk mengukur *usability* atau kegunaan suatu sistem dari sudut pandang subjektif pengguna. SUS terdiri dari 10 pernyataan yang dinilai menggunakan skala likert 1 (Sangat Tidak Setuju) sampai 5 (Sangat Setuju). Skor ini memberikan gambaran tingkat kegunaan sistem, di mana nilai lebih tinggi menunjukkan *usability* yang lebih baik [16]. Target dari responden SUS ini yaitu 30 responden yang akan mencoba menggunakan seluruh fitur utama aplikasi.

3.2. Pembahasan

Tahap ini menyajikan analisis dan interpretasi terhadap hasil-hasil pengujian yang telah dilakukan. Pembahasan difokuskan pada analisis perbandingan kinerja dan efisiensi model klasifikasi, diikuti dengan analisis hasil pengujian fungsionalitas dan *usability* aplikasi.

3.2.1. Pembahasan Metode Klasifikasi

Hasil pengujian kinerja model pada **Tabel 1** dan **Tabel 2** menunjukkan hasil yang signifikan. Ketiga metode mampu melakukan klasifikasi sentimen dengan baik, di mana semua model mencapai akurasi di atas 85%. Namun, terdapat perbedaan jelas dalam keseimbangan presisi dan recall, serta efisiensi komputasi. Analisis ini dengan jelas menunjukkan adanya *trade-off* (keseimbangan) antara kinerja prediktif, terutama pada keseimbangan F1-Score, dengan efisiensi waktu komputasi antar metode.

Metode GRU mencapai F1-Score tertinggi (0.8082), dengan nilai presisi (0.8093) dan recall (0.8072) yang paling seimbang. Hal ini mengindikasikan bahwa kemampuan GRU dalam memahami konteks kalimat efektif untuk klasifikasi ulasan. Namun, keunggulan kinerja ini membutuhkan biaya komputasi yang paling lambat, dengan waktu pelatihan 877.72 detik dan waktu inferensi 17.19 detik.

Metode *Naïve Bayes* menunjukkan akurasi tertinggi (0.8841) dan presisi yang juga sangat tinggi (0.8646). Namun, nilai recall-nya (0.7420) adalah yang terendah. Ini menunjukkan bahwa *Naïve Bayes* sangat baik dalam memastikan ulasan yang diprediksi positif, tetapi cenderung kurang sensitif ke beberapa kasus positif yang sebenarnya. Keunggulan utamanya adalah efisiensi komputasi yang cepat, dengan waktu pelatihan hanya 0.027 detik dan inferensi 0.0031 detik.

Terakhir, Metode LightGBM menghasilkan F1-Score terendah (0.7597) di antara ketiganya, meskipun menggunakan *boosting*, reduksi dimensi dengan SVD mungkin sedikit mengurangi performa F1-Score pada dataset ini.

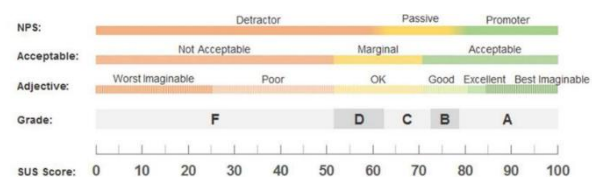
3.2.2. Pembahasan *Black Box Testing*

Hasil pengujian menunjukkan bahwa seluruh skenario uji coba yang dijalankan pada fitur-fitur utama aplikasi menghasilkan status "valid". Ini mengindikasikan bahwa prototipe aplikasi telah berhasil diimplementasikan sesuai dengan kebutuhan fungsional yang diharapkan. Aplikasi mampu menerima input ulasan, menjalankan analisis sentimen menggunakan ketiga model, menampilkan hasil prediksi dan terdapat perbandingan kinerja pada halaman yang sesuai, serta menyediakan fungsi unduh data tanpa ditemukan adanya *error* atau *bug* fungsional selama pengujian. Keberhasilan pengujian *Black Box* ini menandakan bahwa sistem secara fungsional telah siap dan berjalan sesuai rancangan.

3.2.3. Pembahasan *System Usability Scale* (SUS)

Pengujian *System Usability Scale* (SUS) telah dilakukan dengan menyebarkan kuisisioner secara *online* selama periode 26 - 28 Oktober 2025. Kuisisioner ini diisi oleh 30 responden setelah mereka mencoba menggunakan aplikasi beserta seluruh fitur utama-nya.

Dari hasil perhitungan skor ke-30 responden, diperoleh skor rata-rata SUS sebesar 93.84. Berdasarkan standar interpretasi skor SUS, nilai ini termasuk dalam kategori "Excellent" atau Sangat Baik (Grade A+) yang ditunjukkan pada **Gambar 3**.



Gambar 3 Hasil Interpretasi Skor SUS [17]

Hasil ini menunjukkan bahwa sistem yang dikembangkan dinilai memiliki tingkat kegunaan yang sangat tinggi oleh pengguna, serta tampilan antarmuka, kemudahan navigasi, dan kejelasan fiturnya sudah berjalan dengan sangat baik.

4. Kesimpulan

Berdasarkan hasil perancangan dan pengujian sistem klasifikasi sentimen ulasan *e-commerce*, dapat ditarik beberapa kesimpulan sebagai berikut:

1. Prototipe aplikasi web yang dirancang telah berfungsi penuh sesuai dengan rancangan. Hal ini dibuktikan dengan hasil "valid" pada seluruh skenario pengujian *Black Box Testing*, yang mencakup fungsionalitas dari fitur - fitur utama.
2. Sistem yang dikembangkan memiliki tingkat kegunaan (*usability*) yang tinggi. Berdasarkan pengujian *System Usability Scale* (SUS) terhadap 30 responden, sistem memperoleh

- skor rata-rata 93.84, yang termasuk dalam kategori "Excellent" (Grade A+)
3. Dalam perbandingan kinerja model, ditemukan adanya *trade-off* (keseimbangan) yang jelas antara performa dan efisiensi komputasi. Model GRU unggul dalam keseimbangan kinerja (F1-Score 0.8082), namun memiliki kelemahan pada waktu komputasi yang paling lambat (877.72 detik). Sebaliknya, model *Naïve Bayes* unggul dalam efisiensi (waktu latih 0.027 detik) dan Akurasi (0.8841), namun memiliki nilai Recall terendah (0.7420) yang mengindikasikan sensitivitas yang lebih rendah terhadap kasus positif.

REFERENSI

- [1] R. Rachmiani, N. Kintan Oktadinna, and T. Rachmat Fauzan, "The Impact of Online Reviews and Ratings on Consumer Purchasing Decisions on E-commerce Platforms," *International Journal of Management Science and Information Technology*, vol. 4, no. 2, pp. 504–515, Dec. 2024, doi: 10.35870/ijmsit.v4i2.3373.
- [2] N. M. Y. D. A. Ni Made Yulia Dewati Ayu and Jakaria, "Pengaruh E-Commerce Terhadap Ekonomi Indonesia," *Jurnal Ekonomi Trisakti*, vol. 3, no. 2, pp. 2891–2900, Aug. 2023, doi: 10.25105/jet.v3i2.17499.
- [3] K. Taha, P. D. Yoo, C. Yeun, D. Homouz, and A. Taha, "A comprehensive survey of text classification techniques and their research applications: Observational and experimental insights," *Comput Sci Rev*, vol. 54, p. 100664, Nov. 2024, doi: 10.1016/j.cosrev.2024.100664.
- [4] B. I. Hasbi and I. S. Putro, "Perbandingan Kinerja Algoritma Naive Bayes Dan Support Vector Machine untuk Analisis Sentimen Ulasan Produk E-Commerce Menggunakan Pembobotan TF-IDF," vol. 9, no. 2, p. 123, 2025, doi: 10.37817/ikraith-informatika.v9i2.
- [5] D. Wardana, L. Muflikhah, and R. Perdana, "Analisis Sentimen Ulasan Pengguna Aplikasi MyPertamina Pada Google Play Store Menggunakan Metode Xgboost dan Word2Vec Embedding," *J-PTIHK: Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 9, no. 8, Aug. 2025.
- [6] R. Merdiansah, S. Siska, and A. Ali Ridha, "Analisis Sentimen Pengguna X Indonesia Terkait Kendaraan Listrik Menggunakan IndoBERT," *Jurnal Ilmu Komputer dan Sistem Informasi (JIKOMSI)*, vol. 7, no. 1, pp. 221–228, Mar. 2024, doi: 10.55338/jikoms.v7i1.2895.
- [7] J. A. Wibowo, V. C. Mawardi, and T. Sutrisno, "Penerapan Support Vector Machine untuk Analisis Sentimen Fitur Layanan pada Ulasan Gojek," *Jurnal Ilmu Komputer dan Sistem Informasi*, vol. 12, no. 1, Jan. 2024, doi: 10.24912/jiksi.v12i1.28211.
- [8] P.-Ning. Tan, Michael. Steinbach, Anuj. Karpatne, and Vipin. Kumar, *Introduction to data mining*. Pearson Education, Inc., 2019.
- [9] Pristiyono, M. Ritonga, M. A. Al Ihsan, A. Anjar, and F. H. Rambe, "Sentiment analysis of COVID-19 vaccine in Indonesia using Naïve Bayes Algorithm," *IOP Conf Ser Mater Sci Eng*, vol. 1088, no. 1, p. 012045, Feb. 2021, doi: 10.1088/1757-899X/1088/1/012045.
- [10] G. Ke *et al.*, "LightGBM: A Highly Efficient Gradient Boosting Decision Tree." [Online]. Available: <https://github.com/Microsoft/LightGBM>.
- [11] Ekka Pujo Ariesanto Akhmad, "Enhancing the Accuracy of Airline Review Classification Using SMOTE and Grid Search with Cross Validation for Hyperparameter Tuning," *Journal of Information Systems Engineering and Management*, vol. 10, no. 25s, pp. 850–861, Mar. 2025, doi: 10.52783/jisem.v10i25s.4174.
- [12] K. Cho *et al.*, "Learning Phrase Representations using RNN Encoder–Decoder for Statistical Machine Translation," in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Stroudsburg, PA, USA: Association for Computational Linguistics, 2014, pp. 1724–1734. doi: 10.3115/v1/D14-1179.
- [13] A. Géron, *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: concepts, tools, and techniques to build intelligent systems*. O'Reilly Media, Inc., 2019.
- [14] D. A. Mohammed and K. A. Patel, "An Improved GRU Based On Recurrent Attention Unit and Self-Attention Technique for Text Sentiment Analysis," *ICTACT Journal on Soft Computing*, vol. 15, no. 4, pp. 3737–3745, Jan. 2025, doi: 10.21917/ijsc.2025.0518.
- [15] A. P. Putra, F. Andriyanto, K. Karisman, T. D. M. Harti, and W. P. Sari, "Pengujian Aplikasi Point of Sale Menggunakan Blackbox Testing," *Jurnal Bina Komputer*, vol. 2, no. 1, pp. 74–78, Feb. 2020, doi: 10.33557/binakomputer.v2i1.757.
- [16] S. Aisyah, E. Saputra, N. Rozanda, and T. AhSyar, "Evaluasi Usability Website Dinas Pendidikan Provinsi Riau Menggunakan Metode System Usability Scale," *Jurnal Ilmiah Rekayasa dan Manajemen Sistem Informasi*, vol. 7, no. 2, pp. 125–132, Aug. 2021, Accessed: Oct. 26, 2025. [Online]. Available: <http://dx.doi.org/10.24014/rmsi.v7i2.13066>
- [17] M. Rafid Pratama, J. Umam, and R. Yakok, "Usability Testing pada Aplikasi iJateng Menggunakan Metode System Usability Scale," *Jurnal Sistem Informasi, Manajemen dan Teknologi Informasi*, vol. 2, no. 1, pp. 15–23, Jan. 2024, doi: 10.33020/jsimtek.v2i1.556.

Aulia Dwi Yulianti, saat ini sebagai Mahasiswa Program Studi Teknik Informatika, Fakultas Teknologi Informasi Universitas Tarumanagara.

Tony, memperoleh gelar S.Kom. pada tahun 2005 dari Universitas Tarumanagara, M.Kom. pada tahun 2010 dari Universitas Indonesia, dan Ph.D. pada tahun 2021 dari Curtin University. Saat ini sebagai staff pengajar di Fakultas Teknologi Informasi, Universitas Tarumanagara.