

Komparasi Kinerja Model Naive Bayes, SVM, dan BERT dalam Klasifikasi Sentimen Ulasan Pada Aplikasi YUMMY

¹⁾Sabrina Amanda Salsabila, ²⁾Bayu Priyatna, ³⁾Agustia Hananto, ⁴⁾Tukino

^{1,2,3,4)}Sistem Informasi, Fakultas Ilmu Komputer, Universitas Buana Perjuangan Karawang

¹⁾si21.sabrinasalsabila@mhs.ubpkarawang.ac.id, ²⁾bayu.priyatna@ubpkarawang.ac.id,

³⁾agustia.hananto@ubpkarawang.ac.id, ⁴⁾tukino@ubpkarawang.ac.id

INFO ARTIKEL

Riwayat Artikel :

Diterima : 26 April 2025

Disetujui : 11 Mei 2025

Kata Kunci :

Analisis sentiment, Klasifikasi Teks , naïve bayes, SVM, BERT.

ABSTRAK

Ulasan pada Google Play Store memainkan peran penting dalam membentuk persepsi publik terhadap sebuah aplikasi. Penelitian ini bertujuan untuk melakukan sentimen ulasan dengan mengklasifikasikan mejadi kategori positif dan negatif di aplikasi Yummy, serta membandingkan performa tiga metode klasifikasi teks: *Naive Bayes*, *Support Vector Machine* (SVM), dan *Bidirectional Encoder Representations from Transformers* (BERT). Sebanyak 6.773 data ulasan didapatkan melalui scraping dan diproses dengan tahapan pra-pemrosesan seperti pembersihan teks, *tokenisasi*, *stopword removal*, dan *stemming*. Representasi fitur dilakukan menggunakan TF-IDF untuk *Naive Bayes* dan SVM, sedangkan BERT memanfaatkan token embedding dari model pra-latih. Hasil eksperimen menunjukkan bahwa SVM menghasilkan performa sangat efektif dengan akurasi 94%, diikuti oleh *Naive Bayes* dan BERT yang keduanya mencapai akurasi 90%. SVM juga menunjukkan keseimbangan yang lebih baik dalam mengklasifikasikan teks sentimen positif dan negatif. Temuan ini menegaskan SVM direkomendasikan sebagai solusi yang stabil dan efisien dalam analisis sentimen teks berbahasa Indonesia.

ARTICLE INFO

Article History :

Received : Apr 26, 2025

Accepted : Mei 11, 2025

Keywords:

Analysis sentiment, Classification text , naïve bayes, SVM, BERT

ABSTRACT

Reviews on the Google Play Store play an important role in shaping public perception of an application. This study aims to conduct review sentiment by classifying it into positive and negative categories in the Yummy application, and to compare the performance of three text classification methods: Naive Bayes, Support Vector Machine (SVM), and Bidirectional Encoder Representations from Transformers (BERT). A total of 6,773 review data were obtained through scraping and processed with pre-processing stages such as text cleaning, tokenization, stopwords removal, and stemming. Feature representation is done using TF-IDF for Naive Bayes and SVM, while BERT utilizes token embedding from the pre-trained model. The experimental results show that SVM produces very effective performance with an accuracy of 94%, followed by Naive Bayes and BERT which both achieve an accuracy of 90%. SVM also shows a better balance in classifying positive and negative sentiment texts. These findings confirm that SVM is recommended as a stable and efficient solution in sentiment analysis of Indonesian language texts.

1. PENDAHULUAN

Di era modern yang selalu berinovasi dan perkembangan yang pesat di berbagai sektor kehidupan menjadikan manusia sangat kompetitif sehingga lowongan pekerjaan formal yang menjamin kesejahteraan sulit didapatkan. Dari fenomena ini banyak orang yang memilih untuk menjalani pekerjaan non formal yaitu berwirausaha sebagai mata pencahariannya, seperti membuat usaha makanan dan minuman karena bidang ini sangat tinggi konsumennya. Meskipun ramai konsumen, bidang ini juga banyak persaingan terutama untuk pebisnis mikro yang biasa menggunakan resep dari situs maupun aplikasi online. Meskipun demikian tak semua resep yang ditemukan di situs dan aplikasi sesuai dengan apa yang diharapkan meskipun telah mengikuti tiap proses sesuai ketentuan yang ada, disisi lain pengguna sosial media dalam memberi ulasan online memainkan peran penting dalam memengaruhi keputusan individu maupun organisasi(Cahyo Saputra, Fauzan and Carol Aldosion, 2025).

Yummy merupakan salah satu aplikasi resep tutorial memasak dan resep yang mengeksplorasi kuliner di Indonesia, aplikasi ini merupakan milik anak perusahaan dari IDN yang menerima ulasan beragam dari para pengguna di Google Play Store. dari permasalahan tersebut penelitian ini akan melakukan klasifikasi pada ulasan aplikasi yummy menjadi positif dan negatif lalu melakukan komparasi tiga metode menggunakan *naive bayes*, *Support vector machine (SVM)*, *Bidirectional Encoder Representations from Transformers (BERT)*.

Naive Bayes dikenal sederhana dan efisien untuk data bersih dan terstruktur, namun sering kali gagal memahami konteks kompleks ini dibuktikan pada penelitian yang dilakukan di balai pelatihan kerja pondok pesantren Al-Ikhwan, metode *naive bayes* mampu mengklasifikasikan data pendaftar yang lulus dengan akurasi sebesar 90% (Hananto dkk, 2024). Pada analisis sentimen yang dilakukan melalui aplikasi twitter untuk menilai komentar positif dan negatif pada proyek kereta cepat pun menghasilkan akurasi sebesar 81% yang menghasilkan suara terbanyak menyetujui adanya proyek itu(Sanjaya, Priyatna and Hilabi, 2024). Sementara hasil dari penelitian (Pratama,

2022)mengenai prediksi persediaan stok, metode *naive bayes* mendapatkan akurasi 91,20% sedangkan C4.5 mendapat akurasi 96,80% hal ini menunjukkan *naive bayes* sedikit lemah dari metode C4.5 untuk memprediksi stok.

SVM unggul dalam menangani data linier maupun non-linier dan menghasilkan klasifikasi yang lebih stabil, hal ini dibuktikan pada analisis sentimen ulasan di aplikasi ovo dan dana yang memiliki akurasi 90% dan 92% dalam menilai ulasan positif(Anshari, Alam and Hafid, 2023). Sedangkan SVM yang menggunakan linear kernel untuk melakukan sentimen pada ulasan aplikasi Primaku menghasilkan akurasi 97,5%(Tinaliah and Elizabeth, 2022). Pada sentimen opini film horor menggunakan TF-IDF SVM dapat menghasilkan nilai analisis sebesar 82,51% yang dapat menyatakan film horor diterima di indonesia(Br Sinulingga and Sitorus, 2024).

BERT merupakan *deep learning* dengan arsitektur transformer, memiliki kemampuan unggul dalam memahami konteks dua arah dan menangkap makna mendalam dalam teks ulasan. Penelitian sebelumnya yang menggunakan *epoch* 10, 20 dan 30 untuk menganalisis aplikasi *e-commerce* menghasilkan data *Promotions & Offers, dan Services* dengan tingkat akurasi masing-masing sebesar 95%, 97%, dan 99%(Ghufran dkk, 2024) Adapun data dari hasil ulasan film berbahasa inggris yang dianalisis menggunakan *fine-tuning* dengan menghasilkan akurasi 73%(Alifia Putri and Al Faraby, 2020).

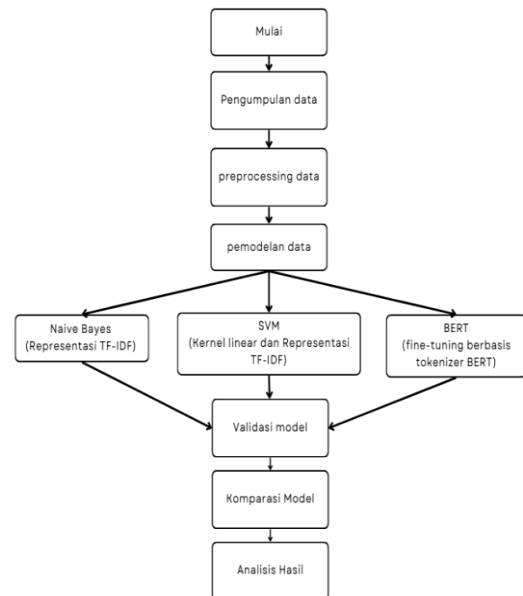
Penelitian terdahulu menunjukkan kinerja algoritma SVM secara konsisten memberikan hasil klasifikasi yang lebih akurat dibandingkan Naïve Bayes, baik dalam domain analisis sentimen *marketplace* di media sosial Twitter (Kurniawan dkk, 2023). Maupun pada ulasan aplikasi *mobile banking* BRImo di Google Play Store (Astuti, Alam and Jaelani, 2022).Kedua studi ini menggunakan data yang telah diproses melalui tahapan *preprocessing, feature extraction*, dan validasi berlapis, serta menunjukkan bahwa SVM memiliki keunggulan akurasi dalam pengolahan teks dan klasifikasi opini pengguna di berbagai sektor. Disisi lain Hasil penelitian yang dilakukan oleh (Arthansa dkk, 2024)dan (Muttaqin and Kharisudin, 2021) menunjukkan bahwa algoritma SVM pun memiliki keunggulan dalam klasifikasi sentimen

berbasis teks dibandingkan metode lain seperti *Naive Bayes* dan *K-Nearest Neighbor* Dalam konteks ulasan film *Avengers: Endgame* dari IMDb, SVM memiliki akurasi lebih tinggi (83%) dibandingkan *Naive Bayes* (82%), sedangkan pada ulasan aplikasi Gojek, SVM dengan kernel linear mencapai akurasi dan presisi yang lebih baik (87,98% dan 88,55%) dibandingkan KNN. Temuan ini menegaskan efektivitas SVM dalam mengenali pola sentimen secara konsisten di berbagai domain analisis teks. Penelitian yang dilakukan untuk mengklasifikasikan email *spam* menggunakan bert dinyatakan lebih baik dari LSTM dibuktikan dengan dengan nilai evaluasinya (Hutagaol and Arifin, 2024). Berbagai penelitian telah membuktikan efektivitas model BERT dan variannya dalam tugas klasifikasi teks berbahasa Indonesia. (Imron, Setiawan and Santoso, 2023) menunjukkan bahwa kombinasi IndoBERT dan CNN berhasil meningkatkan akurasi deteksi aspek dalam ulasan *e-commerce* hingga 94,86%. (Amin dkk, 2024) turut membuktikan keunggulan IndoBERT dalam klasifikasi *spam* SMS dan email dengan tingkat akurasi mencapai 98%, melampaui model multibahasa seperti *MultilingualBERT*. Sementara itu, (Chandradev, Suarjaya and Bayupati, 2023) mengungkap bahwa model *SmallBERT* yang telah melalui proses *fine-tuning* mampu mengklasifikasikan sentimen ulasan hotel secara akurat dengan capaian akurasi 91,40%. Temuan-temuan ini menguatkan peran penting model BERT dalam meningkatkan kinerja klasifikasi teks di berbagai domain bahasa Indonesia.

Dengan melakukan analisis sentimen tiga metode ini diharapkan dapat mengklasifikasikan data ulasan aplikasi yummy menjadi positif dan negatif serta melakukan komparasi metode yang paling optimal jika menilai ulasan berbahasa Indonesia khususnya pada platform digital seperti Google Play Store

2. METODE

Penelitian ini menerapkan pendekatan kuantitatif dengan metode eksperimen komparatif dalam mengevaluasi kinerja tiga algoritma klasifikasi sentimen, yaitu *Naive Bayes*, SVM, dan BERT.



Gambar 1. Diagram alur penelitian

Data yang digunakan berasal dari ulasan pengguna aplikasi Yummy yang diambil melalui proses web scraping dari platform Google Play Store menggunakan skrip Python. Total ulasan yang berhasil dikumpulkan sebanyak 6.773 entri, yang kemudian dikategorikan ke dalam dua jenis sentimen, yakni positif dan negatif

Tahapan *preprocessing* meliputi pembersihan teks dari simbol, angka, *URL*, dan tanda baca, *casefolding*, tokenisasi, penghapusan *stopword*, serta *stemming*. Untuk fitur representasi, digunakan teknik TF-IDF pada model *Naive Bayes* dan SVM, sedangkan pada model BERT digunakan *fine tuning*.

Data dibedakan menjadi dua bagian, yaitu 80% data latih dan 20% data uji. Evaluasi kinerja dilakukan menggunakan metrik akurasi, *precision*, *recall*, dan *F1-score* untuk menilai dari efektivitas masing-masing algoritma dalam mengklasifikasikan sentimen ulasan. Komparasi dilakukan untuk menentukan model yang paling optimal dalam konteks klasifikasi sentimen teks berbahasa Indonesia.

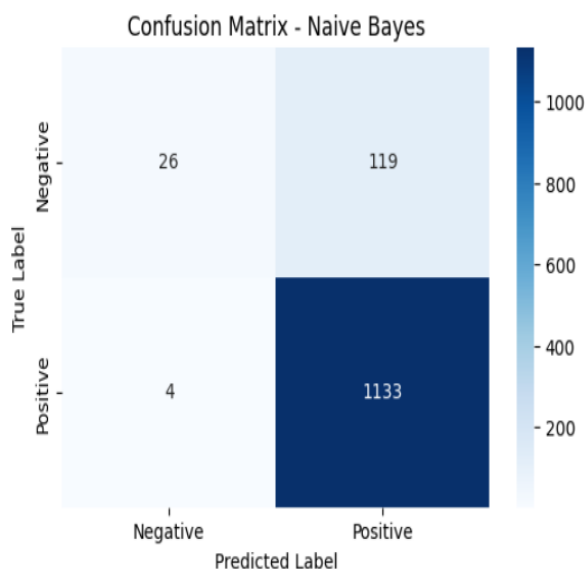
3. HASIL DAN PEMBAHASAN

Setelah dilakukan pelatihan dan pengujian model klasifikasi sentimen terhadap 6.773 data ulasan aplikasi Yummy, diperoleh hasil evaluasi performa dari tiga metode klasifikasi, yaitu *Naive Bayes*, *Support Vector Machine* (SVM), dan BERT. Evaluasi dilakukan menggunakan metrik akurasi, *precision*, *recall*, dan *F1-score*, sebagai berikut:

Tabel 1. Evaluasi Performa model

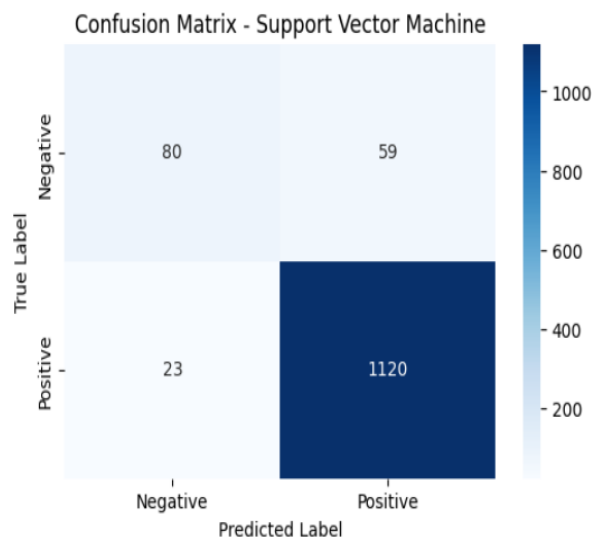
Model	Akurasi	Precision	Recall	F1-Score
<i>Naive bayes</i>	90%	90%	100%	95%
SVM	94%	95%	98%	96%
BERT	90%	90%	100%	95%

Model *naive bayes* menghasilkan nilai akurasi dan *precision* yang besar, *recall* 100% dan *F1-score* 95%. Tingginya nilai *recall* ini menunjukkan bahwa model ini sangat baik dalam menangkap seluruh data ulasan positif, namun *precision* yang sedikit lebih rendah menandakan adanya beberapa kesalahan dalam mengklasifikasikan ulasan negatif menjadi positif



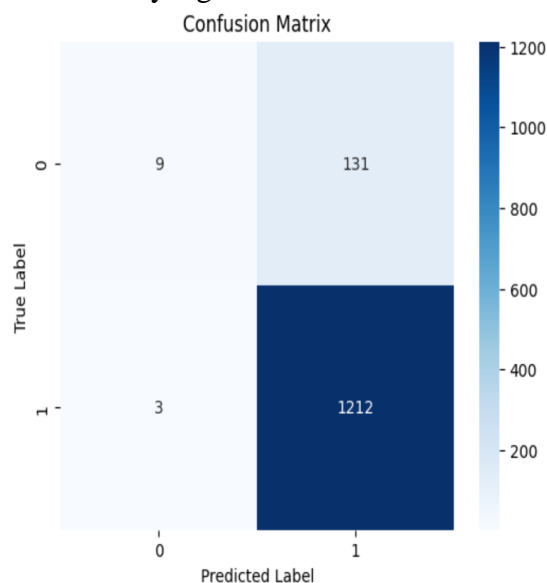
Gambar 2. Confussion Matriks Naive bayes

Model SVM memberikan hasil terbaik di antara ketiga model dengan akurasi 94%, *precision* 95%, *recall* 98%, dan *F1-score* 96%. Ini menunjukkan bahwa SVM mampu mengklasifikasikan ulasan dengan keseimbangan tinggi antara akurasi dan ketepatan identifikasi ulasan positif maupun negatif. Hasil ini memperkuat temuan di penelitian sebelumnya bahwa SVM konsisten memberikan performa tinggi dalam klasifikasi sentimen berbasis teks.



Gambar 3. Confussion matriks SVM

Pada model BERT menghasilkan akurasi 90%, dengan *precision* 90%, *recall* 100%, dan *F1-score* 95%. Performanya identik dengan Naive Bayes dari segi angka metrik, namun secara umum model BERT diharapkan mampu memahami konteks kalimat lebih dalam. Meski dalam penelitian ini tidak jauh berbeda dengan *Naive Bayes*, kekuatan BERT biasanya lebih terasa pada dataset yang lebih kompleks atau nuansa kalimat yang lebih halus



Gambar 4. Confussion matriks BERT

Pada metode Naive Bayes dan SVM, ulasan teks data yang telah melalui tahapan pra-pemrosesan (pembersihan teks, tokenisasi, *stopword removal*, dan *stemming*) konversi menjadi representasi numerik menggunakan teknik TF-IDF. TF-IDF mengubah teks menjadi vektor berdasarkan bobot kemunculan kata, sehingga memungkinkan model untuk

membedakan kata-kata yang relevan dengan masing-masing kelas sentimen. Naive Bayes kemudian mengklasifikasikan teks berdasarkan perhitungan probabilitas dari setiap kata terhadap label sentimen, sedangkan SVM membangun hyperplane optimal untuk memisahkan ulasan positif dan negatif dalam ruang vektor fitur.

Pada model BERT, teks ulasan ditokenisasi menggunakan *tokenizer* IndoBERT, menghasilkan input berbentuk token ID dan masker perhatian. Data ini kemudian diproses melalui lapisan Transformer, yang memahami konteks kata dalam dua arah (*bidirect*) sebelum menghasilkan prediksi kelas melalui lapisan klasifikasi. Proses ini memungkinkan BERT untuk menangkap makna kontekstual yang lebih dalam dibandingkan metode berbasis TF-IDF.

Model SVM, dengan akurasi dan kestabilannya yang tinggi, sangat direkomendasikan untuk aplikasi nyata seperti sistem monitoring ulasan aplikasi, *feedback* pelanggan pada platform *e-commerce*, serta deteksi opini dalam media sosial. Model *Naive Bayes*, dengan kecepatan latih dan inferensi yang tinggi, sesuai untuk kebutuhan klasifikasi *real-time* dalam aplikasi dengan keterbatasan sumber daya komputasi. Sementara itu, BERT lebih cocok diterapkan pada analisis opini tingkat lanjut, pengembangan chatbot berbasis sentimen, atau sistem rekomendasi berbasis ulasan, terutama jika data yang dianalisis memiliki struktur bahasa kompleks dan konteks yang beragam.

Secara keseluruhan, SVM menunjukkan performa terbaik di semua aspek evaluasi dibandingkan *Naive Bayes* dan BERT dalam klasifikasi sentimen ulasan aplikasi Yummy. Temuan ini mengindikasikan bahwa untuk data ulasan berbahasa Indonesia dari platform digital seperti Google Play Store, SVM merupakan metode yang paling optimal untuk diterapkan dalam tugas klasifikasi sentimen teks.

4. PENUTUP

4.1. Kesimpulan

Dari hasil evaluasi terhadap tiga metode klasifikasi sentimen, yakni Naive Bayes, SVM, dan BERT, dapat diketahui bahwa seluruh model mampu menjalankan tugas klasifikasi dengan performa yang baik, ditunjukkan oleh nilai

akurasi yang berada di atas 90%. Namun, dari algoritma ketiga yang diuji, metode SVM menunjukkan performa paling unggul dalam mengklasifikasikan sentimen ulasan aplikasi Yummy.

Model SVM berhasil mencapai nilai akurasi sebesar 94%, dengan *precision* 95%, *recall* 98%, dan *F1-score* 96%. Keseimbangan antara presisi dan recall yang tinggi ini mencerminkan kemampuan SVM dalam mengidentifikasi kedua kelas sentimen (positif dan negatif) secara lebih akurat dan konsisten. Sementara itu, meskipun model Naive Bayes dan BERT menunjukkan hasil performa yang sama dalam hal akurasi dan *F1-score* (masing-masing 90% dan 95%), keduanya masih belum melampaui keunggulan SVM dari aspek presisi dan stabilitas klasifikasi. Dengan pengujian ini SVM dapat diandalkan dalam konteks pemrosesan teks berbahasa Indonesia, khususnya untuk klasifikasi sentimen menggunakan data ulasan aplikasi di platform digital seperti Google Play Store.

4.2. Saran

Disarankan untuk memperluas jumlah serta variasi data ulasan agar model seperti BERT dapat mengoptimalkan kemampuannya dalam memahami konteks kalimat yang lebih kompleks. Selain itu, penelitian mendatang dapat mengeksplorasi pengaturan hyperparameter pada proses fine-tuning, seperti variasi jumlah epoch dan learning rate, untuk meningkatkan performa model berbasis deep learning. Temuan ini memperkuat bukti bahwa metode SVM dapat diandalkan dalam konteks pemrosesan teks berbahasa Indonesia, khususnya untuk klasifikasi sentimen pada ulasan aplikasi di platform digital seperti Google Play Store.

5. DAFTAR PUSTAKA

- Alifia Putri, C. and Al Faraby, S. (2020) 'Analisis Sentimen Review Film Berbahasa Inggris Dengan Pendekatan Bidirectional Encoder Representations from Transformers', *Jurnal Teknik Informatika dan Sistem Informasi*, 6(2), pp. 181–193. Available at: <http://jurnal.mdp.ac.id>.
- Amin, M.B.M. *et al.* (2024) 'Deteksi Spam Berbahasa Indonesia Berbasis Teks Menggunakan Model Bert', *Jurnal*

- Teknologi Informasi dan Ilmu Komputer*, 11(6), pp. 1291–1302. Available at: <https://doi.org/10.25126/jtiik.2024118121>.
- Anshari, R. AL, Alam, S. and Hafid, M.T. (2023) ‘Komparasi Payment Digital Untuk Analisis Sentimen Berdasarkan Ulasan Di Google Playstore Menggunakan Metode Support Vector Machine’, *STORAGE – Jurnal Ilmiah Teknik dan Ilmu Komputer*, 2(3), pp. 118–128. Available at: <https://doi.org/10.55123>.
- Astuti, A.P., Alam, S. and Jaelani, I. (2022) ‘Komparasi Algoritma Support Vector Machine Dengan Naive Bayes Untuk Analisis Sentimen Pada Aplikasi BRImo’, *Bangkit Indonesia*, XI(02).
- Br Sinulingga, J.E. and Sitorus, H.C.K. (2024) ‘Analisis Sentimen Opini Masyarakat terhadap Film Horor Indonesia Menggunakan Metode SVM dan TF-IDF’, *Jurnal Manajemen Informatika (JAMIKA)*, 14(1), pp. 42–53. Available at: <https://doi.org/10.34010/jamika.v14i1.11946>.
- Cahyo Saputra, D., Fauzan, M. and Carol Aldosion, G. (2025) ‘Pengaruh Rating Dan Komentar Pengguna Di Google Playstore Terhadap Keputusan Pengguna Dalam Mengunduh Aplikasi’, *Spectrum: Multidisciplinary Journal*, 2(1).
- Chandradev, V., I Made Agus Dwi Suarjaya and I Putu Agung Bayupati (2023) ‘Analisis Sentimen Review Hotel menggunakan Metode Deep Learning BERT 107 Analisis Sentimen Review Hotel Menggunakan Metode Deep Learning BERT’, *Jurnal buana informatika*, 14(2).
- Ghufran, A. *et al.* (2024) *Comparison of Service and Ease of e-Commerce User Applications Using BERT*, *Jurnal Sistem Cerdas*.
- Hananto, A.L. *et al.* (2024) ‘Determination of Training Participants in Community Work Training Centers Using the Naïve Bayes Classifier Algorithm’, *International Journal On Informatics Visualization*, pp. 1162–1167. Available at: www.joiv.org/index.php/joiv.
- Hutagaol, Y.R. and Arifin, Y. (2024) ‘Klasifikasi Spam Email Berbasis Semantik Menggunakan Metode Bert Semantic-Based Email Spam Classification Using Bert Method’, *Journal of Information Technology and Computer Science (INTECOMS)*, 7(5).
- Imron, S., Setiawan, E.I. and Santoso, J. (2023) ‘Deteksi Aspek Review E-Commerce Menggunakan IndoBERT Embedding dan CNN’, *Journal of Intelligent System and Computation*, 5(1), pp. 10–16. Available at: <https://doi.org/10.52985/insyst.v5i1.267>.
- Kurniawan, I. *et al.* (2023) ‘Perbandingan Algoritma Naive Bayes Dan SVM Dalam Sentimen Analisis Marketplace Pada Twitter’, *Jurnal Teknik Informatika dan Sistem Informasi*, 10(1). Available at: <http://jurnal.mdp.ac.id>.
- Muttaqin, M.N. and Kharisudin, I. (2021) ‘Analisis Sentimen Pada Ulasan Aplikasi Gojek Menggunakan Metode Support Vector Machine dan K Nearest Neighbor’, *UNNES Journal of Mathematics*, 10(2), pp. 22–27. Available at: <http://journal.unnes.ac.id/sju/index.php/ujm>.
- Radendha Muhammad Arthansa, Dhea Intan Sagita and Anggraini Puspita Sari (2024) ‘Komparasi Analisis Sentimen Ulasan Film Avengers: Endgame Di Imdb Menggunakan Metode Naïve Bayes Dan Svm’, *STORAGE – Jurnal Ilmiah Teknik dan Ilmu Komputer*, 3(4).
- Rian Pratama *et al.* (2022) ‘Perbandingan Algoritma C4.5 dan Naïve Bayes dalam Menentukan Persediaan Stok’, *Metik Jurnal*, 6(2), pp. 115–122. Available at: <https://doi.org/10.47002/metik.v6i2.379>.
- Sanjaya, J., Priyatna, B. and Shofia Hilabi, S. (2024) ‘Analisis Sentimen Terhadap Opini Proyek Kereta Cepat Menggunakan Metode Naïve Bayes Classifier’, *Jurnal Fasilkom*, 14(1).
- Tinaliah and Elizabeth, T. (2022) ‘Analisis Sentimen Ulasan Aplikasi PrimaKu Menggunakan Metode Support Vector Machine’, *Jurnal Teknik Informatika dan Sistem Informasi*, 9(4). Available at: <http://jurnal.mdp.ac.id>.