



Eksplorasi Sentimen Aplikasi BCA Mobile Pada Google Playstore Menggunakan Metode *Bidirectional Encoder Representations from Transformers* (BERT)

Elisheva Immanuela Leihitu^{1*}, Cahyawati Diah Kusumarini²

^{1,2}Universitas Gunadarma

Email: elishevaimmanuela@gmail.com

Received: 13-06-2025

Revised: 25-06-2025

Accepted : 27-06-2025

Published : 03-07-2025

Abstrak

Penelitian ini bertujuan untuk menganalisis sentiment pengguna terhadap aplikasi BCA Mobile berdasarkan ulasan di Google Play Store. Dengan memahami persepsi pengguna, penelitian ini diharapkan dapat membantu pengembang aplikasi dalam meningkatkan kualitas layanan. Sebanyak 5000 ulasan dikumpulkan melalui proses scraping dan dianalisis menggunakan metode BERT, khususnya model IndoBERT-base. Dataset melalui tahap preprocessing, kemudian dibagi menjadi data pelatihan, validasi, dan pengujian. Model dilatih menggunakan parameter batch size 32, learning rate $2e-5$ dan 5 epoch. Hasil penelitian menunjukkan bahwa model berhasil mengenali sentimen positif dengan baik (F1-score 0.82 pada data testing), tetapi kesulitan dalam mengenali sentiment netral (F1-score 0.23). akurasi keseluruhan mencapai 73% pada data testing, dengan indikasi overfitting setelah epoch, ke-3, di mana akurasi validasi stagnan meskipun akurasi pelatihan meningkat. Penelitian ini memberikan wawasan tentang persepsi pengguna terhadap aplikasi BCA Mobile dan menyoroti pentingnya distribusi data yang seimbang untuk meningkatkan kemampuan model dalam mengenali semua kategori sentimen. Selain itu, disarankan untuk menambah jumlah data dan mengeksplorasi model dengan kapasitas yang lebih besar untuk meningkatkan hasil analisis, sehingga hasil temuan ini dapat dijadikan sebagai dasar pengambilan keputusan dalam strategi pengembangan fitur dan layanan aplikasi perbankan digital secara lebih responsif terhadap kebutuhan pengguna.

Kata Kunci : Analisis Sentimen, BCA Mobile, Metode BERT

Abstract

This study aims to analyse user sentiment towards the BCA Mobile app based on reviews from Google Play Store. By understanding user perceptions, this research is expected to assist app developers in improving service quality. A total of 5000 reviews were collected through a scraping process and analyzed using the BERT method, specifically the IndoBERT-base model. The dataset underwent preprocessing, followed by splitting into training, validation, and testing sets. The model was trained using batch size 32, learning rate of $2e-5$, and 5 epochs. The results show that the model successfully identified positive sentiment with a high F1-score of 0.82 on the test data, but struggled to recognize neutral sentiment (F1-score 0.23). The overall accuracy reached 73% on the test data, with signs of overfitting after the third epoch, where validation accuracy stagnated despite an increase in training accuracy. This study provides insights into user perceptions of the BCA Mobile app and highlights the importance of balanced data distribution to improve the model's ability to recognize all sentiment categories. Additionally, it is recommended to increase the dataset size and explore models with larger capacities to improve the analysis results, so that the findings can be used as a basis for decision making in the development strategy of digital banking application features and services that are more responsive to user needs.

Keywords: Sentiment Analysis, BCA Mobile, BERT Method



Pendahuluan

Dalam era digital, perbankan telah bertransformasi signifikan, terutama dengan munculnya aplikasi mobile banking seperti BCA Mobile. Aplikasi ini, dirilis pada 2011 oleh Bank BCA, telah menjadi pilihan utama dengan lebih dari 50 juta unduhan di Google Playstore. Dengan banyaknya ulasan pengguna, analisis sentimen menjadi penting untuk memahami persepsi mereka. Meskipun banyak penelitian telah dilakukan, penggunaan metode BERT dalam analisis sentimen BCA Mobile masih terbatas. BERT unggul dalam pemrosesan bahasa alami, memungkinkan analisis sentimen yang lebih akurat. Penelitian ini menggunakan data ulasan BCA Mobile dari Google Playstore, yang akan dianalisis setelah praproses data untuk mengidentifikasi sentimen positif, negatif, atau netral. Namun demikian, belum banyak studi yang secara khusus menerapkan model IndoBERT untuk mengkaji persepsi pengguna aplikasi perbankan lokal seperti BCA Mobile, sehingga diperlukan penelitian lebih lanjut untuk menjembatani kekosongan tersebut. Berdasarkan latar belakang tersebut, maka rumusan masalah dalam penelitian ini adalah: bagaimana kinerja model IndoBERT dalam mengklasifikasikan sentimen pengguna terhadap aplikasi BCA Mobile berdasarkan ulasan di Google Playstore?

Penelitian terkait mengenai analisis sentimen pada aplikasi di Google Playstore dilakukan penelitian terdahulu. Penelitian (Saraswati dkk., 2023) menggunakan model BERT untuk menganalisis sentimen terhadap perundangan siber pada twitter menunjukkan bahwa metode BERT sangat efektif dalam analisis sentimen, dengan akurasi 99% dan f1-score 98,9%, mengklasifikasikan 5.254 dari 5.437 komentar sebagai positif. Penelitian (Senjaya, 2022) menggunakan model BERT untuk menganalisis sentimen aplikasi video berdasarkan review pada Google Playstore, 6.000 komentar dan mencapai akurasi 76%, lebih baik dari metode Support Vector Machine (SVM) yang hanya mencapai 62%, menunjukkan keunggulan metode yang digunakan dalam pengolahan dan analisis data. Penelitian (Putri dkk., 2020) menggunakan metode BERT untuk menganalisis sentimen Review Film Berbahasa Inggris, menggunakan dataset movie review polarity dari conelldu yang terdiri dari 2000 ulasan film (1000 positif dan 1000 negatif). Model Bert-base mencapai akurasi 73,3% meskipun hanya menggunakan 128 karakter terakhir dari dokumen. Berbeda dari studi-studi sebelumnya yang lebih banyak berfokus pada domain hiburan, media sosial, dan film, penelitian ini berupaya mengisi kekosongan dengan menerapkan BERT secara spesifik dalam konteks aplikasi layanan keuangan di Indonesia, khususnya mobile banking BCA Mobile, yang hingga saat ini belum banyak dieksplorasi secara mendalam.

Metode

Metode penelitian yang digunakan dalam penelitian ini dibagi menjadi beberapa tahap, yaitu:

1. Metode Pengumpulan Data

Data untuk penelitian ini dikumpulkan dengan menggunakan teknik web scraping melalui google colab dan Python dari Google Playstore. Dataset yang digunakan mencakup 5000 data ulasan yang akan diuji dalam studi ini.

2. Metode Perancangan

- a) Preprocessing

Pada penelitian ini, tahap preprocessing data ulasan aplikasi BCA Mobile dilakukan dalam beberapa langkah untuk mempersiapkan data bagi model BERT. Pertama, data dibersihkan dan dinormalisasi dengan menghapus karakter-karakter yang tidak



relevan serta mengubah semua huruf menjadi huruf kecil. Selanjutnya, teks dipecah menjadi token-token (baik kata maupun sub-kata) dan dikonversi menjadi ID token yang sesuai dengan kamus BERT. Untuk memastikan panjang kalimat yang konsisten, dilakukan padding dengan menambahkan token khusus. Teknik Masked Language Modeling juga diterapkan untuk memprediksi kata yang hilang, guna meningkatkan kemampuan model dalam memahami konteks. Terakhir, data dibagi menjadi tiga set: pelatihan, validasi, dan pengujian, untuk melatih serta mengevaluasi model dengan efektif.

b) Modeling

Dalam tahap ini, terdapat dua pendekatan utama dalam menerapkan model BERT untuk analisis sentimen: menggunakan model BERT yang telah dilatih sebelumnya dan melatih model dari awal. Pertama, kita dapat memanfaatkan model BERT yang telah terlatih pada dataset besar dan beragam, seperti BERT-base atau BERT-large, yang sudah disesuaikan untuk berbagai tugas, termasuk analisis sentimen. Pendekatan ini sangat praktis karena tidak memerlukan pelatihan model dari awal. Namun, jika kita memiliki dataset yang sangat khusus atau ingin meningkatkan kinerja model pada domain tertentu, kita bisa melakukan fine-tuning pada model BERT yang ada dengan dataset kita sendiri. Fine-tuning melibatkan pelatihan ulang pada beberapa lapisan terakhir model BERT dengan data baru. Setelah model disesuaikan, kita dapat mengevaluasi performanya menggunakan metrik seperti akurasi, precision, recall, dan F1-score, biasanya dilakukan dengan menggunakan confusion matrix untuk menunjukkan jumlah prediksi yang benar dan salah. Dalam penelitian ini, dipilih model IndoBERT-base karena dirancang khusus untuk Bahasa Indonesia dan terbukti unggul dalam berbagai tugas NLP berbahasa lokal. Parameter pelatihan yang digunakan adalah batch size 32, learning rate $2e-5$, dan 5 epoch, yang dipilih berdasarkan studi terdahulu dan pertimbangan kestabilan pelatihan pada dataset berukuran sedang.

c) Implementasi

Tahap implementasi melibatkan pelatihan model BERT dengan menggunakan data ulasan yang telah diproses sebelumnya. Proses ini melibatkan iterasi melalui data pelatihan untuk mengoptimalkan parameter model, dengan tujuan menghasilkan prediksi sentimen yang lebih akurat. Setelah model berhasil dilatih, langkah berikutnya adalah menerapkannya pada ulasan baru untuk menganalisis sentimen. Hasil analisis sentimen dapat divisualisasikan dalam format histogram atau diagram batang untuk mengeksplorasi distribusi sentimen dari ulasan aplikasi BCA Mobile di Google Play Store. Metode ini mencakup seluruh proses dari pelatihan model hingga analisis hasil akhir.

Hasil dan Pembahasan

Hasil Web Scraping

Web scraping digunakan untuk mengumpulkan 5000 ulasan aplikasi BCA Mobile dari Google Playstore, yang disimpan ke dalam format CSV. Proses ini melibatkan pengambilan data ulasan library google-play-scaper dan Google Colab.

Hasil Labelisasi Dataset

Dataset di labelisasi menjadi tiga kategori positif, negatif, dan netral berdasarkan rating bintang. Presentase ulasan adalah 62% positif, 13.3% negatif, dan 24.7% netral. Distribusi yang tidak seimbang ini mempengaruhi hasil pelatihan model, terutama dalam mengenali sentimen netral yang cenderung lebih ambigu dan kurang representatif dalam dataset.

Hasil PreProcessing Dataset

Tahapan ini bertujuan untuk mengubah dataset yang tidak terstruktur menjadi terstruktur, sehingga analisis sentimen dapat dilakukan dengan mudah, yang meliputi:

- a. Case folding: Mengubah huruf kapital menjadi huruf kecil.
- b. Data cleaning: Menghapus elemen yang tidak relevan.
- c. Tokenizing: memisahkan kalimat menjadi kata-kata
- d. Stopwords Removal: menghapus kata tidak signifikan
- e. Stemming: mengubah kata ke bentuk dasarnya.
- f. Normalize: mengganti kata tidak baku dengan kata baku.

Hasil Splitting Dataset

Proses pemisahan *dataset* diperlukan untuk mengevaluasi kinerja model yang digunakan. *Dataset* dibagi menjadi tiga bagian utama yaitu *dataset training* untuk melatih model, *dataset testing* untuk menguji akurasi model setelah dilatih, dan *dataset validasi* yang membantu mengurangi risiko *overfitting*. Dari 5000 dataset yang tersedia, 3500 data digunakan untuk *dataset training*, 495 data digunakan untuk *dataset testing*, dan 1005 data akan digunakan untuk *dataset validasi*. Proporsi ini mengacu pada praktik standar dalam deep learning yang juga diterapkan pada studi Saraswati et al. (2023), yang menunjukkan bahwa validasi yang cukup besar diperlukan untuk menjaga kestabilan model selama proses fine-tuning. Hasil pembagian *dataset* ini dapat dilihat pada Gambar 1.

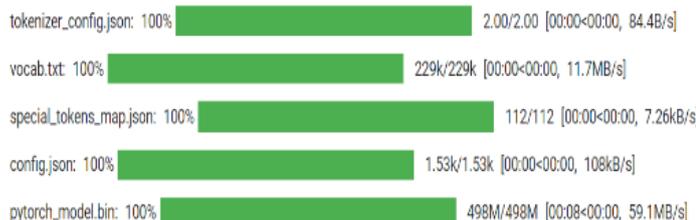
Train shape: (3500, 2)
 Val shape: (1005, 2)
 Test shape: (495, 2)

Gambar 1 Hasil Splitting Dataset

Hasil Implementasi BERT

- a. Load Model

Load model adalah langkah awal dalam menggunakan BERT, dan model BERT yang dimuat ini bisa dijadikan dasar untuk berbagai tugas pemrosesan Bahasa alami. Pada tahap ini, digunakan *pre-trained* model indobert-base-pd dari IndoBERT untuk proses load model. Fungsi dari *load model* ini adalah untuk melatih dataset yang sudah dibagi sebelumnya, dengan tujuan memastikan penggunaan memori yang efisien. Dengan demikian, seluruh dataset dapat dimuat ke dalam memori secara bersamaan. Berikut adalah tahapan dalam proses load model.



Gambar 2 Proses Load Model

b. Inisiasi Kelas Dokumen Sentimen dan Dokumen Sentimen Data Loader

Pada tahap ini, penulis menginisialisasi kelas `Document Sentiment Dataset` dan `DocumentSentiment Data Loader` untuk menyiapkan dan memproses data analisis sentimen. Document Sentiment Dataset` menyediakan data, sedangkan `Document Sentiment Data Loader` memuat data dalam batch ke model BERT. Inisialisasi melibatkan pengaturan parameter seperti sumber data, panjang token maksimum, ukuran batch, jumlah proses, dan metode pengambilan data.

```
train_dataset_path = '/content/train_set.tsv'
valid_dataset_path = '/content/val_set.tsv'
test_dataset_path = '/content/test_set.tsv'

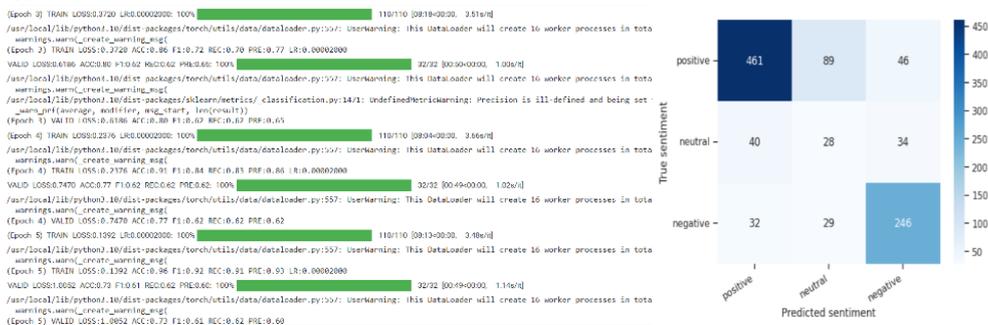
# fungsi dataset loader dari utils IndoNLU
train_dataset = DocumentSentimentDataset(train_dataset_path, tokenizer, lowercase=True)
valid_dataset = DocumentSentimentDataset(valid_dataset_path, tokenizer, lowercase=True)
test_dataset = DocumentSentimentDataset(test_dataset_path, tokenizer, lowercase=True)

train_loader = DocumentSentimentDataLoader(dataset=train_dataset, max_seq_len=512, batch_size=32, num_workers=16, shuffle=True)
valid_loader = DocumentSentimentDataLoader(dataset=valid_dataset, max_seq_len=512, batch_size=32, num_workers=16, shuffle=False)
test_loader = DocumentSentimentDataLoader(dataset=test_dataset, max_seq_len=512, batch_size=32, num_workers=16, shuffle=False)
```

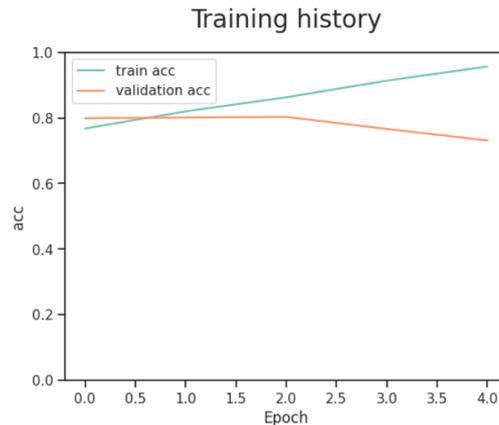
Gambar 3 Tahap Inisiasi Kelas

Hasil Training Dataset

Pada tahap ini, fine-tuning dilakukan dengan learning rate $2e-5$. Lapisan dropout diaktifkan dengan `model.train()`, dan komputasi diaktifkan dengan `torch.set_grad_enabled(True)`. Data batch diproses dengan `forward_sequence_classification`, gradien dihitung menggunakan `loss.backward()`, dan parameter diperbarui dengan `optimizer.step()`. Proses ini diulang untuk setiap batch selama pelatihan model BERT dengan 5 epoch, menggunakan dataset pelatihan dan validasi. Hasil akurasi dari pelatihan dan validasi ditampilkan pada Gambar 4.



Gambar 4 Proses *Training dan Evaluasi*



Gambar 5 *Learning Curve* Hasil *Training* dan Validasi

Evaluasi

Pada tahap evaluasi, hasil pengujian model disajikan menggunakan confusion matrix untuk menilai akurasi prediksi. Selain itu, classification report menampilkan skor akurasi, presisi, recall, dan f1-score. Confusion matrix menunjukkan bahwa model cukup baik dalam memprediksi sentimen positif dan negatif, namun kesulitan dengan sentimen netral. Perhitungan metrik evaluasi dilakukan menggunakan fungsi `classification_report()` dari `sklearn`

	precision	recall	f1-score	support
positive	0.86	0.77	0.82	596
neutral	0.19	0.27	0.23	102
negative	0.75	0.80	0.78	307
accuracy			0.73	1005
macro avg	0.60	0.62	0.61	1005
weighted avg	0.76	0.73	0.74	1005

Gambar 6 *Classification Report*

Model memiliki presisi tertinggi untuk sentimen positif sebesar 0.86, dengan recall 0.77 dan F1-score 0.82, menunjukkan keseimbangan yang baik antara presisi dan recall. Untuk sentimen negatif, presisi mencapai 0.75, recall 0.80, dan F1-score 0.78. Namun, model kesulitan memprediksi sentimen netral dengan presisi 0.19, recall 0.27, dan F1-score 0.23. Secara keseluruhan, akurasi model adalah 0.73, yang berarti model mampu memprediksi dengan benar 73% dari total data. Hasil ini menunjukkan kinerja yang kompetitif dibandingkan studi Putri et al. (2020) yang mencatat akurasi 76% untuk domain aplikasi video. Performa tinggi pada sentimen positif dan negatif dapat dikaitkan dengan ekspresi linguistik yang lebih eksplisit, sedangkan lemahnya hasil pada kategori netral kemungkinan besar disebabkan oleh ketidakteraturan dalam struktur kalimat netral, yang seringkali mengandung ambiguitas semantik. Hal ini penting diperhatikan dalam pengembangan sistem serupa agar dilakukan pelabelan manual yang lebih hati-hati pada kelas netral. Secara praktis, hasil ini memberikan landasan untuk mengembangkan sistem pemantauan kepuasan pelanggan secara otomatis, yang dapat membantu tim pengembang dalam memetakan aspek layanan mana yang perlu diperbaiki berdasarkan pola sentimen dominan.



Kesimpulan

Hasil evaluasi pada penelitian dengan bahasa pemrograman python, untuk mengklasifikasikan *review* kedalam sentimen *positive*, *negative*, dan *neutral* telah berhasil dilakukan. Hasil analisis ini memberikan wawasan mengenai pandangan umum pengguna terhadap aplikasi BCA Mobile, termasuk aspek-aspek positif yang diapresiasi, masalah-masalah yang dihadapi oleh pengguna yang tercermin dalam ulasan negatif, serta saran-saran perbaikan dari pengguna. Dapat disimpulkan bahwa sebagian pengguna memiliki pandangan yang positif terhadap aplikasi BCA Mobile. Melalui penelitian ini, analisis sentimen berhasil dilakukan dengan menggunakan model IndoBERT-base melalui teknik fine-tuning. Model ini mencapai akurasi sebesar 73% dengan pemilihan hyperparameter yang meliputi batch size (32), learning rate $2e-5$, dan epoch 5. Dari penelitian ini juga dapat dilihat bahwa ulasan pengguna cenderung terpolarisasi ke arah sentimen positif dan negatif. Kontribusi utama dari penelitian ini terletak pada penerapan pendekatan pemodelan berbasis IndoBERT dalam konteks layanan keuangan digital di Indonesia, yang memberikan alternatif metodologi yang kuat dan adaptif dalam analisis sentimen berbahasa Indonesia. Secara aplikatif, temuan ini dapat menjadi dasar pengembangan sistem pemantauan kepuasan pengguna secara otomatis untuk mendukung peningkatan kualitas layanan mobile banking, khususnya dalam menangkap dinamika persepsi pengguna secara real-time.

Saran dari penelitian ini mencakup penambahan jumlah dataset untuk meningkatkan akurasi hasil analisis. Selain itu, disarankan untuk menggunakan model dengan parameter yang lebih besar, seperti indoBERT-large. Temuan dari analisis ini juga dapat digunakan sebagai bahan evaluasi atau acuan bagi pengembang dalam upaya meningkatkan kualitas aplikasi BCA Mobile serta pelayanan yang diberikan.

Daftar Pustaka

- Ardiani, L., Sujaini, H., & Tursina, T. (2020). Implementasi sentiment analysis tanggapan masyarakat terhadap pembangunan di Kota Pontianak. *JUSTIN (Jurnal Sistem dan Teknologi Informasi)*, 8(2), 183–190.
- Atmaja, R. M. R. W. P. K., & Yustanti, W. (2021). Analisis sentimen customer review aplikasi Ruang Guru dengan metode BERT (Bidirectional Encoder Representations from Transformers). *Journal of Emerging Information Systems and Business Intelligence*, 2(3). <https://doi.org/10.35760/ik.2023.v28i3.10090>
- Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. *NAACL HLT 2019 - Proceedings of the Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 1(Mlm), 4171–4186.
- Fauzi, A., & Yunial, H. A. (2024). *Analisis Sentimen (Sentiment Analysis): Evaluasi Sentimen Layanan Dataset Twitter US Airline*. CV Bintang Semesta Media.
<https://play.google.com/store/apps/details?id=com.bca&hl=id>
<https://www.bca.co.id/-/media/Feature/Report/File/Sustainability-Index/GRI/GRI-1011Nama-organisasi.pdf>
- Lazuardi, J. (2023). Analisis sentimen ulasan pengguna aplikasi Joox pada Android menggunakan metode Bidirectional Encoder Representation from Transformer (BERT). *Jurnal Ilmiah Informatika Komputer*, 28(3), 251–260. <https://doi.org/10.35760/ik.2023.v28i3.10090>



- Mubaraq, M. F., & Maharani, W. (2022). Sentiment analysis on Twitter social media towards climate change on Indonesia using IndoBERT model. *Jurnal Media Informatika Budidarma*.
- Mulyatun, S., Utama, H., & Mustopa, A. (2021). Pendekatan natural language processing pada aplikasi chatbot sebagai alat bantu customer service. *Journal of Information System Management*.
- Muttaqin, M. N., & Kharisudin, I. (2021). Analisis sentimen pada ulasan aplikasi Gojek menggunakan metode support vector machine dan k-nearest neighbor. *UNNES Journal of Mathematics*, 10(2), 22–27.
- Nugraha, F. A., Harani, N. H., & Habibi, R. (2020). *Analisis Sentimen terhadap Pembatasan Sosial Menggunakan Deep Learning*. Kreatif Industri Nusantara.
- Putra, H. K., Bijaksana, M. A., & Romadhony, A. (2021). Deteksi penggunaan kalimat abusive pada teks bahasa Indonesia menggunakan metode IndoBERT. *Jurnal Tugas Akhir Fakultas Informatika*.
- Putra, R. R., Isa, I. G. T., & Malyan, A. B. J. (2023). *Buku Ajar Pengantar Deep Learning dalam Pemrosesan Citra*. Penerbit NEM.
- Putri, C. A., Adiwijaya, A., & 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), 181–193.
- Runimeirati, A. M. F., & Muis. (2023). Pelatihan text mining menggunakan bahasa pemrograman Python. *Jurnal Pengabdian Kepada Masyarakat*.
- Rudawan, R. A., Sularsa, A., & Qodryantha, A. P. (2023). Pembangunan sistem pengawasan cerdas dengan visualisasi 3D. *e-Proceeding of Applied Science*.
- Saraswati, N. P. V. D., Yudistira, N., & Adikara, P. P. (2023). Analisis sentimen terhadap perundungan siber pada Twitter menggunakan algoritma Bidirectional Encoder Representations from Transformer (BERT). *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 7(2), 909–916. <http://j-ptiik.ub.ac.id>
- Sofiana, S. (2024). *Konsep BERT pada Natural Language Processing*. Eureka Media Aksara.
- Subawa, N. S., & Widhiasthini, N. W. (2020). *Waralaba 4.0: Isu, Tren, dan Evolusi Waralaba di Era Digital*. Nilacakra.