



# Klasifikasi Ulasan Aplikasi E-KTP Menggunakan Bidirectional Encoder Respresentations from Transformers

Alfida Hari Rusdi<sup>1</sup>, Meilany Nonsi Tentua<sup>2\*</sup>

<sup>1,2</sup>Program Studi Informatika, Fakultas Sains dan Teknologi Universitas PGRI Yogyakarta, Yogyakarta, Indonesia

## Article Info

### Article history:

Received Januari 8, 2026  
Accepted Februari 9, 2026  
Published Februari 13, 2026

### Keywords:

BERT  
Mbert  
Review classification  
E-KTP application

## ABSTRACT

The Digital Population Identity application (E-KTP Digital) is part of e-government development aimed at improving the quality of public services. However, user reviews on the Google Play Store are still grouped based on star ratings, so the level of user satisfaction is not yet described in depth. This study aims to classify the sentiment of user reviews of the E-KTP Digital application using the Bidirectional Encoder Representations from Transformers (BERT) method with the Multilingual BERT (mBERT) model. A total of 15,000 reviews were collected from July 3, 2023 to May 31, 2025 and filtered into 1,750 reviews through data cleaning and manual labeling processes. The dataset is divided into training and testing data with ratios of 60:40, 70:30, and 80:20. The training process is conducted using the AdamW optimizer for 4 epochs with a batch size of 16. Model evaluation is planned using accuracy, precision, recall, and F1-score metrics to measure the performance of user review sentiment classification.



## Corresponding Author:

Meilany Nonsi Tentua,  
Program Studi Informatika, Fakultas Sains dan Teknologi  
Universitas PGRI Yogyakarta,  
Jl. IKIP PGRI I no. 117, Bantul, Yogyakarta, Indonesia.  
Email: \*meilany@upy.ac.id

## 1. PENGANTAR

Transformasi Kartu Tanda Penduduk (KTP) Elektronik menjadi KTP Digital merupakan bagian dari pengembangan *e-government* untuk meningkatkan kualitas layanan administrasi kependudukan secara digital [1]. Kementerian Dalam Negeri RI menargetkan 25% dari total 275.361.267 jiwa telah menggunakan Identitas Kependudukan Digital (IKD), dan seluruh Dinas Dukcapil Kabupaten/Kota diwajibkan mendorong masyarakat untuk bertransformasi [2]. Melalui digitalisasi ini, akses data kependudukan diharapkan dapat terintegrasi dengan berbagai sektor, seperti pendidikan, kesehatan, transportasi, dan logistik.

Meskipun demikian, penyelenggaraan layanan melalui Aplikasi IKD/E-KTP belum optimal. Penilaian pengguna di Playstore dan App Store menunjukkan rating 3,1 dari 49.381 vote dan 3,0 dari 1.413 vote, sementara jumlah pengguna tercatat 7.286.092 orang pada tahun 2024 [3]. Ulasan pengguna di Google Play Store masih hanya dikelompokkan berdasarkan rating bintang tanpa klasifikasi mendalam terhadap isi ulasan [4]. Hal ini membatasi informasi yang diperoleh pengembang dalam menilai kepuasan pengguna secara spesifik.

Beberapa penelitian klasifikasi ulasan menggunakan machine learning telah dilakukan. Penelitian Putri Nuriza dkk [5] menggunakan Naïve Bayes untuk klasifikasi ulasan *e-commerce* dengan akurasi 83%, namun distribusi data tidak seimbang. Penelitian Khoirunnisaa dkk [6] membandingkan Naïve Bayes dan Support Vector

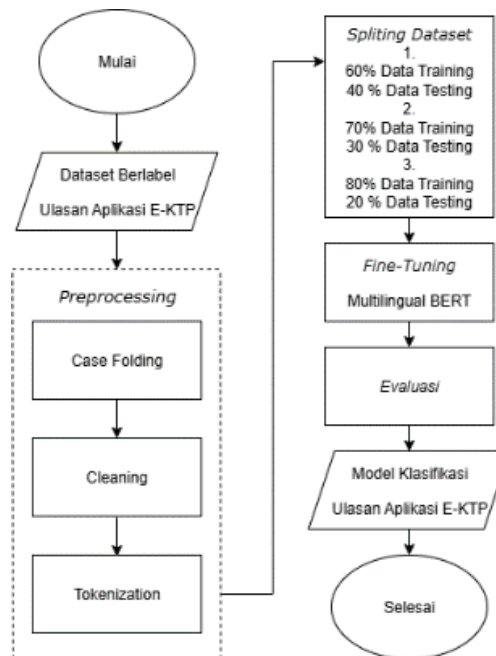
Machine (SVM) pada klasifikasi ulasan aplikasi Netflix, di mana SVM mencapai akurasi 85% dan Naïve Bayes 82%, namun keduanya belum mencapai 90%.

*Bidirectional Encoder Representations from Transformers* (BERT) adalah model representasi bahasa berbasis *Transformer* yang memproses teks secara bidirectional dan menggunakan mekanisme *self-attention* [7]. Salah satu pengembangan BERT adalah *Multilingual BERT* (mBERT), yang dilatih menggunakan korpus Wikipedia dari 104 bahasa, termasuk Bahasa Indonesia. mBERT mampu membentuk representasi semantik lintas bahasa, sehingga dapat memahami teks multibahasa secara akurat, termasuk kata atau istilah berbahasa Inggris [8][9]. Kemampuan ini diharapkan meningkatkan akurasi klasifikasi ulasan dan analisis persepsi pengguna aplikasi.

Pendekatan BERT dengan model mBERT diharapkan dapat memberikan dasar untuk mengklasifikasikan ulasan aplikasi E-KTP secara lebih mendalam, sehingga mampu memberikan informasi yang lebih akurat tentang kepuasan pengguna dan mendukung pengembangan kualitas layanan aplikasi secara berkelanjutan.

## 2. METODOLOGI PENELITIAN

Penelitian ini menggunakan pendekatan BERT untuk melakukan klasifikasi ulasan pengguna aplikasi E-KTP berdasarkan tingkat kepuasan. Model yang digunakan adalah *fine-tuning Multilingual BERT* (mBERT). Tahapan penelitian secara keseluruhan ditunjukkan pada Gambar 1, yang meliputi pengumpulan data, pelabelan, *preprocessing*, pemodelan, dan evaluasi model.



Gambar 1. Flowchart Tahapan Penelitian

### 2.1. Pengumpulan Data

Data penelitian berupa ulasan pengguna aplikasi E-KTP yang diperoleh dari Google Play Store. Pengambilan data dilakukan menggunakan metode *scraping* berdasarkan ID aplikasi. Data dikumpulkan pada rentang waktu 03 Juli 2023 hingga 31 Mei 2025 dengan total 15.000 ulasan. Setelah melalui proses *cleaning* awal dan pelabelan, jumlah data yang digunakan dalam penelitian ini adalah sebanyak 1.750 ulasan. Data tersebut kemudian digunakan sebagai dataset utama dalam proses pelatihan dan pengujian model. Contoh objek penelitian ditunjukkan pada Gambar 2.



Gambar 2. Objek Penelitian

Pelabelan data dilakukan secara manual dengan membagi ulasan ke dalam tiga kategori, yaitu puas, netral, dan tidak puas. Pedoman pelabelan masing-masing kategori disajikan pada Tabel 1, sedangkan contoh ulasan yang telah diberi label ditampilkan pada Tabel 2. Pelabelan ini bertujuan untuk menyediakan data berlabel yang representatif bagi proses pemodelan.

Tabel 1. Pedoman Pelabelan Data

Kategori	Penjelasan
Puas	Komentar yang menyatakan kepuasan, pujian, atau pengalaman positif saat menggunakan aplikasi.
Netral	Komentar yang bersifat campuran antara puas dan tidak puas, atau tidak secara jelas menunjukkan kepuasan maupun ketidakpuasan.
Tidak Puas	Komentar yang menyatakan ketidakpuasan, keluhan, masalah teknis dan kesulitan penggunaan

Tabel 2. Contoh Ulasan Aplikasi

Kategori	Ulasan
Puas	sangat membantu dan mudah di akses
Netral	dulu itu lumayan bagus kita bisa lihat keterangan keterangan seperti kis bpjs npwp sekarang apa ga guna pengen lihat kaya hal gitu aja sekarang ga bisa
Tidak Puas	aplikasi tidak berfungsi, bukan mempermudah malah bikin emosi

## 2.2 Pre-processing

Tahap *preprocessing* dilakukan untuk meningkatkan kualitas data teks sebelum diproses oleh model mBERT. Rangkaian tahapan preprocessing ditunjukkan pada Tabel 3, yang meliputi *case folding*, *cleaning*, dan *tokenization*.

Tabel 3. Tahapan pre-processing data

Tahapan	Hasil
Sebelum preprocessing	Aplikasi digital, tapi daftar offline, gunanya apa? 🤔🤔
Case Folding	aplikasi digital, tapi daftar offline, gunanya apa? 🤔🤔
Cleaning	aplikasi digital, tapi daftar offline, gunanya apa?
Tokenization	['aplikasi', 'digital', ',', 'tapi', 'daftar', 'off', '##line', ',', 'gun', '##anya', 'apa', '?']

*Case folding* dilakukan dengan mengubah seluruh huruf menjadi huruf kecil untuk menjaga konsistensi teks[10]. Selanjutnya, proses *cleaning* bertujuan untuk menghapus karakter yang tidak relevan seperti emoji, simbol, URL, serta spasi berlebih[11]. Tahap ini juga mencakup penghapusan data duplikat agar dataset lebih bersih dan representatif.

Tokenisasi dilakukan menggunakan *tokenizer* WordPiece dari model mBERT. Pada tahap ini, teks dipecah menjadi token atau sub-kata, seperti kata “offline” yang diuraikan menjadi “off” dan “##line”. Token-token yang dihasilkan kemudian dikonversi menjadi token IDs, segment IDs, dan attention mask untuk menandai token asli dan padding. Setiap *token ID* direpresentasikan sebagai *input word vector*  $t_i$  yang kemudian

dikombinasikan dengan *absolute position vector*. Perhitungan *positional encoding*  $p_i$  dilakukan menggunakan Persamaan (1) dan (2), kemudian dijumlahkan sebagaimana ditunjukkan pada Persamaan (3) [12][13].

$$PE_{(pos,2i)} = \text{Sin}(pos/10000^{2i/d}) \tag{1}$$

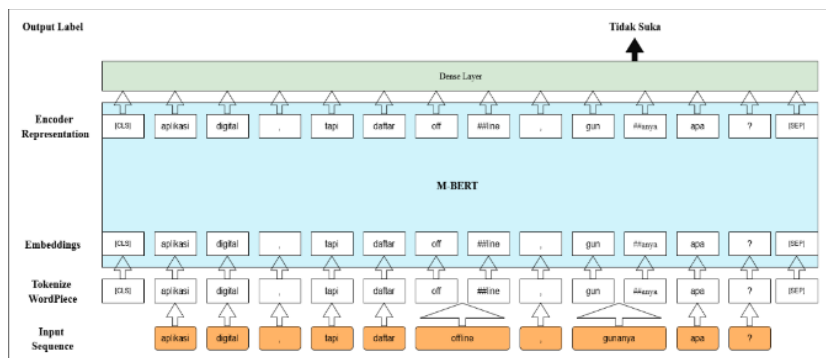
$$PE_{(pos,2i+1)} = \text{Cos}(pos/10000^{2i/d}) \tag{2}$$

$$x_i = t_i + p_i \tag{3}$$

### 2.3 Pemodelan

Sebelum pelatihan model dataset dibagi menjadi data latih dan data uji menggunakan tiga skenario rasio, yaitu 60:40, 70:30, dan 80:20 [14]. Model dilatih menggunakan optimizer AdamW dengan *learning rate* sebesar  $2e-5$ , jumlah *epoch* sebanyak 4, dan ukuran *batch* 16 [15]. Pendekatan ini bertujuan untuk mengetahui pengaruh variasi rasio data terhadap performa model.

Model yang digunakan dalam penelitian ini adalah Multilingual BERT (*bert-base-multilingual-uncased*) yang telah dilatih sebelumnya. Arsitektur pemodelan fine-tuning mBERT ditunjukkan pada Gambar 4.



Gambar 4. Arsitektur Pemodelan *Fine-Tuning* Multilingual BERT

Proses pemodelan dimulai dari teks masukan yang telah melalui tahap *preprocessing*. Token khusus [CLS] ditambahkan di awal kalimat sebagai representasi keseluruhan konteks teks, sedangkan token [SEP] digunakan sebagai penanda akhir. Setiap token kemudian dipetakan ke dalam *embedding* yang merupakan gabungan dari token *embedding*, *segment embedding*, dan *position embedding*. Vektor gabungan tersebut diteruskan ke dalam lapisan *Transformer* mBERT yang menggunakan mekanisme *multi-head self-attention* secara *bidirectional* untuk memahami konteks kata dari keseluruhan kalimat. Representasi akhir dari token [CLS] digunakan sebagai masukan ke lapisan klasifikasi untuk memprediksi label ulasan, yaitu puas, netral, atau tidak puas [16][17][18].

### 2.4. Evaluasi Model

Hasil evaluasi model klasifikasi ditampilkan menggunakan *classification report* dan *confusion matrix*. *Classification report* menyajikan nilai *accuracy*, *precision*, *recall*, dan *F1-score* untuk setiap kelas, sedangkan *confusion matrix* ditampilkan dalam bentuk tabel yang menggambarkan distribusi prediksi model terhadap data aktual. *Confusion matrix* berfungsi sebagai alat evaluasi performa klasifikasi machine learning dengan menunjukkan empat kemungkinan hasil prediksi, yaitu true positive (TP), false positive (FP), false negative (FN), dan true negative (TN). Nilai-nilai tersebut digunakan sebagai dasar perhitungan metrik evaluasi untuk menilai kinerja model secara menyeluruh, di mana F1-score digunakan sebagai metrik gabungan yang merepresentasikan keseimbangan antara *precision* dan *recall*. Dengan menggunakan *confusion matrix*, kemampuan model dalam memprediksi data secara benar serta pola kesalahan klasifikasi dapat dianalisis secara sistematis. Seluruh metrik evaluasi dihitung menggunakan Persamaan (4), (5), (6), dan (7) sebagai parameter pengujian kinerja model [19][20][21].

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \tag{4}$$

$$Recall = \frac{TP}{TP+FN} \tag{5}$$

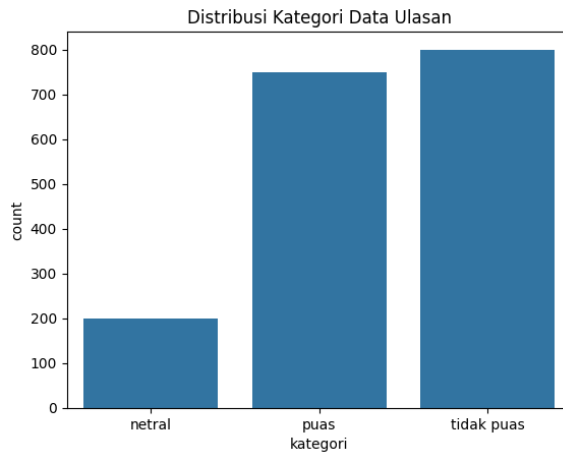
$$Precision = \frac{TP}{TP+FP} \tag{6}$$

$$F1 - score = 2x \frac{Precision \times Recall}{Precision + Recall} \tag{7}$$

### 3. HASIL DAN PEMBAHASAN

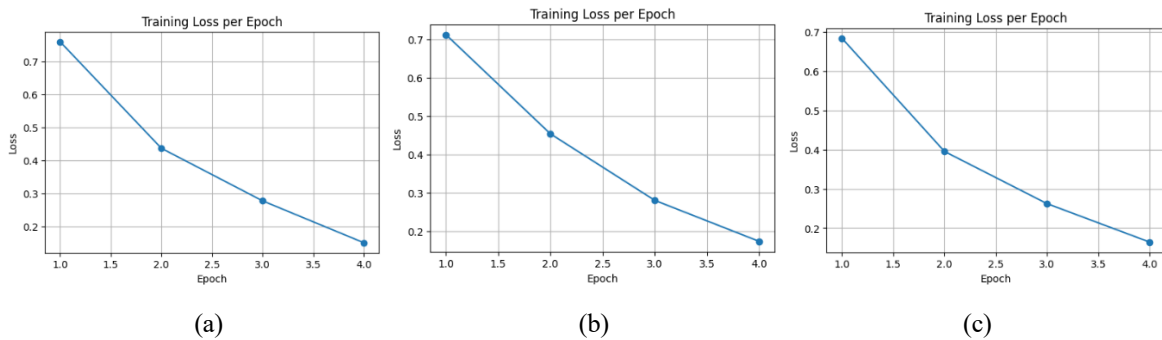
#### 3.1. Hasil

Data ulasan pengguna aplikasi E-KTP Digital dikumpulkan dari Google Play Store pada periode 3 Juli 2023 hingga 31 Mei 2025 dengan total 15.000 ulasan. Setelah melalui proses *cleaning* awal dan pelabelan manual, diperoleh 1.750 ulasan yang digunakan sebagai dataset penelitian dengan tiga kategori sentimen, yaitu puas, netral, dan tidak puas.



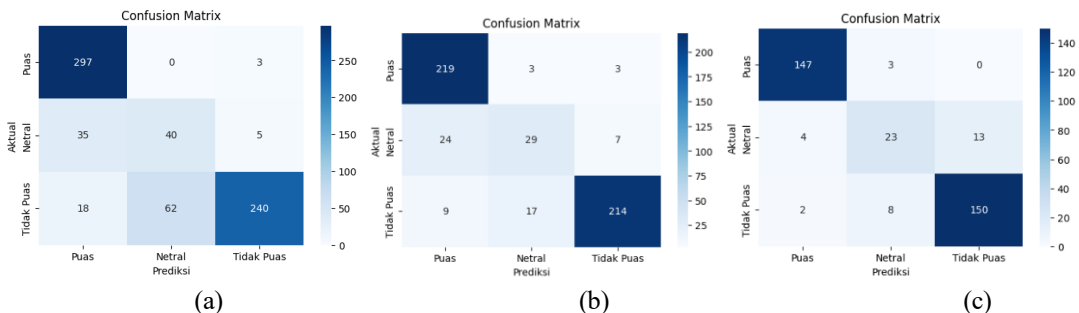
Gambar 5. Distribusi kategori ulasan

Grafik Training Loss untuk setiap skenario rasio pembagian data 60:40, 70:30 dan 80:20 disajikan pada Gambar 4a-c.. Visualisasi ini digunakan untuk melihat perkembangan performa model selama proses pelatihan pada masing-masing rasio.



Gambar 4. Grafik Training Loss (a) 60:40, (b)70:30 dan (c)80:20

Performa model dapat dievaluasi secara lebih rinci, termasuk kesalahan prediksi antar kelas menggunakan Confusion Matrix seperti yang terlihat pada Gambar 5.



Gambar 5. Hasil Confusion Matrix (a) 60:40, (b)70:30 dan (c)80:20

Hasil perhitungan precision, recall, F1-Score dan akurasi dari confusion matriks disajikan pada Tabel 3.

Tabel 3. Perbandingan Hasil Evaluasi Model

Rasio Data	Kelas	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>	Akurasi
60:40	Puas	0.85	0.99	0.91	82,43%
	Netral	0.39	0.50	0.44	
	Tidak Puas	0.97	0.75	0.85	
70:30	Puas	0.87	0.97	0.92	88%
	Netral	0.59	0.48	0.53	
	Tidak Puas	0.96	0.89	0.92	
80:20	Puas	0.96	0.98	0.97	91,43%
	Netral	0.68	0.57	0.62	
	Tidak Puas	0.92	0.94	0.93	

Pada rasio 60:40 dataset terdiri dari 1.050 data training dan 700 data testing dengan akurasi model sebesar 82,43%. Pada skenario ini nilai precision tertinggi diperoleh dari kelas tidak puas sebesar 0,97 sedangkan terendah terdapat pada kelas netral sebesar 0,39. Nilai recall tertinggi ditunjukkan oleh kelas puas sebesar 0,99 dan terendah kembali pada kelas netral sebesar 0,50. Untuk f1-score, nilai tertinggi diperoleh kelas puas sebesar 0,91 sementara nilai terendah terdapat pada kelas netral sebesar 0,44. Hasil ini menunjukkan bahwa model lebih efektif dalam mengenali sentimen puas dan tidak puas dibandingkan sentimen netral.

Selanjutnya, pada rasio 70:30 dataset terdiri dari 1.225 data training dan 525 data testing sehingga akurasi model meningkat menjadi 88%. Pada skenario ini, nilai precision tertinggi diperoleh oleh kelas tidak puas sebesar 0,96, sedangkan nilai terendah terdapat pada kelas netral sebesar 0,59. Nilai recall tertinggi ditunjukkan oleh kelas puas sebesar 0,97 sementara nilai terendah kembali terdapat pada kelas netral sebesar 0,48. Untuk nilai f1-score kelas puas dan tidak puas sama-sama memperoleh nilai tertinggi sebesar 0,92 sedangkan nilai terendah terdapat pada kelas netral sebesar 0,53. Peningkatan akurasi ini menunjukkan bahwa penambahan data training berdampak positif terhadap kemampuan model dalam mengenali pola sentimen.

Pada rasio 80:20 dataset terdiri dari 1.400 data training dan 350 data testing dengan akurasi tertinggi sebesar 91,43%. Pada skenario ini nilai precision tertinggi diperoleh oleh kelas puas sebesar 0,96, sedangkan nilai terendah terdapat pada kelas netral sebesar 0,68. Nilai recall tertinggi ditunjukkan oleh kelas puas sebesar 0,98 sementara nilai terendah kembali terdapat pada kelas netral sebesar 0,57. Untuk nilai f1-score, nilai tertinggi diperoleh kelas puas sebesar 0,97 sedangkan nilai terendah terdapat pada kelas netral sebesar 0,62. Hal ini menegaskan bahwa proporsi data training yang lebih besar mampu meningkatkan performa model secara keseluruhan.

Berdasarkan perbandingan ketiga skenario tersebut, rasio pembagian data 80:20 merupakan konfigurasi terbaik pada penelitian ini. Meskipun demikian, performa pada kategori netral masih lebih rendah dibandingkan kategori lainnya, yang disebabkan oleh jumlah data yang lebih sedikit dan karakteristik sentimen yang cenderung ambigu. Oleh karena itu, peningkatan jumlah data netral serta pengoptimalan preprocessing dan parameter pelatihan disarankan pada penelitian selanjutnya untuk meningkatkan keseimbangan performa model.

#### 4. KESIMPULAN

Model mBERT berhasil diterapkan untuk klasifikasi ulasan aplikasi E-KTP ke dalam tiga kategori sentimen, yaitu puas, netral, dan tidak puas. Model mampu memahami konteks ulasan yang bersifat tidak terstruktur serta mengandung variasi bahasa.

Hasil pengujian menunjukkan bahwa rasio pembagian data 80:20 memberikan performa terbaik dengan akurasi sebesar 91,43%, lebih tinggi dibandingkan rasio 70:30 sebesar 88% dan rasio 60:40 sebesar 82,43%. Hal ini menunjukkan bahwa peningkatan proporsi data training berpengaruh positif terhadap kinerja model.

Model mBERT menunjukkan performa yang cukup baik pada kategori puas dan tidak puas dengan nilai precision, recall, dan F1-score yang tinggi, sedangkan kategori netral masih memiliki performa yang lebih rendah akibat keterbatasan jumlah data dan karakteristik sentimen yang cenderung ambigu.

Secara keseluruhan, penerapan mBERT mampu menghasilkan akurasi yang lebih baik dibandingkan metode klasifikasi konvensional sebelumnya dalam memahami makna kontekstual ulasan pengguna. Sehingga hasil yang diperoleh dapat memberikan gambaran kepuasan pengguna aplikasi E-KTP secara lebih objektif dan mendalam dibandingkan pengelompokan berdasarkan rating bintang saja.

#### DAFTAR PUSTAKA

- [1] L. S. Wati and G. W. Pradana, "Kajian Aplikasi E-Government Dalam Penerapan Identitas Kependudukan Digital di Kota Surabaya (Studi Kasus Kecamatan Semampir)," *Publika*, pp. 563–572, Jul. 2024, doi: 10.26740/PUBLIKA.V12N2.P563-572.
- [2] M. Alfarizi, "Digitalization of Indonesian Identity Card and Millennial Participation: Investigation of Receiving Digital Transformation in Indonesian Civil Registry Policy," *J. Stud. Kebijak. Publik*, vol. 2, no. 1, pp. 41–54, May 2023, doi: 10.21787/JSKP.2.2023.41-54.
- [3] Y. F. Setyawan and L. Rudita, "Performance Evaluation of the Digital Population Identity Application Services," *J. Kebijak. Publik*, vol. 15, no. 3, pp. 301–307, 2024.
- [4] C. Hayati, S. Shofiah Hilabi, A. Lia Hananto, S. Informasi, and U. Buana Perjuangan Karawang, "Klasifikasi dan Prediksi Ulasan Aplikasi Dana pada Google Play Store Menggunakan Algoritma Naive Bayes," *J. Inform. Teknol. dan Sains*, vol. 7, no. 2, pp. 596–605, May 2025, doi: 10.51401/JINTEKS.V7I2.5691.
- [5] A. Putri Nuriza, E. Novalia, B. Priyatna, F. Ilmu Komputer, and U. Buana Perjuangan Karawang, "Klasifikasi dan Prediksi Ulasan E-Commerce Menggunakan Algoritma Naive Bayes," *JOISIE (Journal Inf. Syst. Informatics Eng.)*, vol. 9, no. 1, pp. 207–217, Jul. 2025, doi: 10.35145/JOISIE.V9I1.4993.
- [6] N. Khoirunnisaa, K. Nabila, N. Kesuma, S. Setiawan, A. Yunizar, and P. Yusuf, "Klasifikasi Teks Ulasan Aplikasi Netflix Pada Google Play Store Menggunakan Algoritma Naive Bayes dan Svm," *SKANIKA Sist. Komput. dan Tek. Inform.*, vol. 7, no. 1, pp. 64–73, Jan. 2024, doi: 10.36080/SKANIKA.V7I1.3138.
- [7] N. M. Gardazi, A. Daud, M. K. Malik, A. Bukhari, T. Alsahfi, and B. Alshemaimri, "BERT applications in natural language processing: a review," *Artif. Intell. Rev.*, vol. 58, no. 6, pp. 1–49, Jun. 2025, doi: 10.1007/S10462-025-11162-5/TABLES/17.
- [8] I. F. Putra, A. Purwarianti, and U.-C. Ai-Vlb, "Improving Indonesian Text Classification Using Multilingual Language Model," Sep. 2020.
- [9] L. W. Hao and R. K. Liu, "Transfer Learning Approach for Sentiment Analysis in Low-Resource Austronesian Languages Using Multilingual BERT," *J. Technol. Informatics Eng.*, vol. 4, no. 1, pp. 75–94, 2025, doi: 10.51903/jtie.v4i1.276.
- [10] M. Krisna, H. Prasetyo, I. Much, and I. Subroto, "Penerapan Metode BERT ( Bidirectional Encoder Reprerentations from Transformers ) Pada Analisis Emosi Terhadap Program Kerja Laporan Mas Wapres Presiden RI Dengan Presepsi Pengguna Media Sosial X," vol. 7, no. 1, pp. 37–42, 2025.
- [11] K. C. K. and P. B. D. K. C. K. and P. B. Dhumane and ijrbat, "Review Of Data Pre-Processing Techniques In Data Mining," *Int. J. Res. Biosci. Agric. Technol.*, 2022, doi: 10.29369/IJRBT.2022.010.3.0016.
- [12] B. Nath, S. Tamang, O. Elwasila, and Y. Gulzar, "Task-Oriented Evaluation of Assamese Tokenizers Using Sentiment Classification," *Int. J. Adv. Comput. Sci. Appl.*, vol. 16, no. 9, pp. 826–836, Sep. 2025, doi: 10.14569/IJACSA.2025.0160979.
- [13] J. Acs, E. Hamerlik, R. Schwartz, N. A. Smith, and A. Kornai, "Morphosyntactic probing of multilingual BERT models," *Nat. Lang. Eng.*, vol. 30, no. 4, pp. 753–792, May 2024, doi: 10.1017/S1351324923000190.
- [14] Y. A. Prasetyo, E. Utami, and A. Yaqin, "Pengaruh Komposisi Split Data Terhadap Performa Akurasi Analisis Sentimen Algoritma Naive Bayes dan SVM," *J. Electr. Eng. Comput.*, vol. 6, no. 2, pp. 382–390, Oct. 2024, doi: 10.33650/JEECOM.V6I2.9188.
- [15] A. Salam and S. R. Sidiq, "SciBERT Optimisation for Named Entity Recognition on NCBI Disease Corpus with Hyperparameter Tuning," *J. Appl. Informatics Comput.*, vol. 9, no. 2, pp. 432–441, Mar. 2025, doi: 10.30871/JAIC.V9I2.9283.
- [16] Y. Huh and Y. Seo, "Multilingual BERT-based Classification and Recommendation Model for Supporting Innovation Finance Decisions," *ICAIF 2025 - 6th ACM Int. Conf. AI Financ.*, vol. 25, pp. 823–828, Nov. 2025, doi: 10.1145/3768292.3770384;SUBPAGE:STRING: BASIC.
- [17] K. Dedes, Fatimatuzzahra, M. Hermansyah, A. B. Setiawan, R. P. Pradana, and A. F. M. Harvyanti, "BERT Sentimen: Fine-Tuning Multibahasa untuk Ulasan Bahasa Indonesia," *J. Komput. Teknol. Inf. Sist. Inf.*, vol. 4, no. 2, pp. 1080–1084, Sep. 2025, doi: 10.62712/JUKTISI.V4I2.585.
- [18] N. M. Gardazi, A. Daud, M. K. Malik, A. Bukhari, T. Alsahfi, and B. Alshemaimri, "BERT applications in natural language processing: a review," *Artif. Intell. Rev.* 2025 586, vol. 58, no. 6, pp. 166–, Mar. 2025, doi: 10.1007/S10462-025-11162-5.
- [19] A. T. Riadi, F. Indriani, M. I. Mazdadi, M. R. Faisal, and R. Herteno, "Cross-Temporal Generalization of IndoBERT for Indonesian Hoax News Classification," *J. Tek. Inform.*, vol. 6, no. 5, pp. 5291–5304, Oct. 2025, doi: 10.52436/1.JUTIF.2025.6.5.4757.
- [20] A. Nanyonga and G. Wild, "Classification of Operational Records in Aviation Using Deep Learning Approaches," *2025 Int. Conf. Pervasive Comput. Technol. ICPCCT 2025*, pp. 997–1002, 2025, doi: 10.1109/ICPCCT64145.2025.10940469.
- [21] M. De *et al.*, "Classification of User Reports for Detection of Faulty Computer Components using NLP Models: A Case Study," Mar. 2025.