

Optimasi Analisis Sentimen Ulasan Platform Pendidikan Daring Menggunakan Arsitektur ALBERT dan Teknik Augmentasi Kontekstual

Sintia Darma Pamuja^{1*}, Noviyanto²

¹Program Studi Teknologi Informasi, Fakultas Kesehatan dan Teknologi, Universitas Muhammadiyah Klaten, Klaten

²Program Studi Teknologi Informasi, Fakultas Kesehatan dan Teknologi, Universitas Muhammadiyah Klaten,

Email: ¹sintia@umkla.ac.id, ²noviyanto@umkla.ac.id

*Penulis Korespondensi

ABSTRACT — Online learning through global platforms like Coursera generates a massive volume of user reviews, which serve as vital information for educational quality improvement. However, these reviews often exhibit imbalanced label distributions, where positive sentiments significantly dominate negative and neutral ones, hindering traditional classification models. Advanced language models such as ALBERT offer parameter efficiency through cross-layer parameter sharing while maintaining high performance in complex text understanding. This study aims to evaluate the ALBERT model's performance in classifying Coursera user reviews and addressing data imbalance using Contextual Word Embedding augmentation. The methodology involves collecting 10,000 reviews followed by preprocessing steps including case folding, punctuation removal, and tokenization. The augmentation technique utilizes language models to replace words based on context to balance minority classes. The results show that ALBERT provides highly consistent performance, achieving an F1-score of 0.9710 with the contextual augmentation scenario. The model proves effective in capturing linguistic variations and remains computationally efficient. In conclusion, the ALBERT model is highly effective for sentiment analysis on the Coursera dataset, where contextual augmentation significantly enhances the model's ability to recognize minority classes that were previously difficult to identify.

KEYWORDS — Sentiment_Analysis; ALBERT; Coursera; Data_Imbalance; Contextual_Word_Embedding.

INTISARI — Pembelajaran daring melalui platform global seperti Coursera menghasilkan ulasan pengguna dalam jumlah besar yang merupakan sumber informasi vital untuk pengembangan kualitas pendidikan. Ulasan pada platform sering kali memiliki distribusi label yang tidak seimbang, sentimen positif jauh lebih mendominasi dibandingkan sentimen negatif dan netral, sehingga menyulitkan model klasifikasi tradisional untuk bekerja secara akurat. Penggunaan model bahasa tingkat lanjut seperti ALBERT menjadi relevan karena menawarkan efisiensi parameter melalui teknik *parameter sharing* namun tetap mempertahankan performa tinggi dalam memahami konteks teks yang kompleks. Penelitian ini bertujuan untuk mengevaluasi performa model ALBERT dalam mengklasifikasikan sentimen ulasan pengguna Coursera serta menangani masalah ketidakseimbangan data melalui berbagai teknik augmentasi. Metodologi penelitian dimulai dengan pengumpulan 10.000 ulasan yang kemudian diproses melalui tahap *preprocessing* berupa *case folding*, penghapusan tanda baca, dan tokenisasi. Teknik augmentasi yang diterapkan *Contextual Word Embedding* untuk menyeimbangkan jumlah data minoritas. Hasil penelitian menunjukkan bahwa ALBERT memberikan performa yang sangat konsisten dengan F1-score tertinggi mencapai 0,9799 pada skenario augmentasi. Penggunaan teknik *Contextual Word Embedding* secara spesifik menghasilkan F1-score sebesar 0,9710, yang membuktikan efektivitas model dalam menangkap variasi linguistik. Selain unggul dalam akurasi, ALBERT terbukti efisien secara komputasi dengan waktu pelatihan tercepat yakni 973,78 detik. Kesimpulannya, model ALBERT sangat efektif untuk analisis sentimen pada dataset Coursera, di mana teknik augmentasi terbukti signifikan dalam meningkatkan kemampuan generalisasi model terhadap data yang tidak seimbang.

KATA KUNCI — Analisis_Sentimen; ALBERT; Coursera; Ketidakseimbangan_Data; Contextual_Word_Embedding.

I. PENDAHULUAN

Pembelajaran daring melalui platform *Massive Open Online Courses* (MOOC) telah menjadi komponen krusial dalam sistem pendidikan dan pengembangan karier global [1]. Platform Coursera merupakan penyedia layanan pendidikan representatif

yang memiliki jutaan pengguna aktif dari berbagai penjuru dunia. Ulasan yang diberikan oleh pengguna pada platform ini menjadi sumber informasi vital untuk membantu calon peserta dalam pengambilan keputusan. Analisis sentimen diperlukan sebagai teknik otomatis untuk mengidentifikasi serta mengklasifikasikan opini pengguna ke dalam kategori positif,

netral, atau negatif [2]. Penelitian terdahulu mengenai ulasan Coursera menggunakan model Naive Bayes menunjukkan keterbatasan dengan tingkat akurasi yang hanya mencapai 75% [3]. Model tradisional tersebut memiliki kelemahan dalam memproses struktur kalimat yang kompleks. Dataset ulasan Coursera memiliki karakteristik distribusi label yang sangat tidak seimbang, jumlah ulasan bersentimen positif mendominasi secara signifikan dibandingkan dengan ulasan negatif dan netral [2],[4]. Ketidakseimbangan data ini berpotensi menyebabkan model mengalami bias terhadap kelas mayoritas dan menurunkan akurasi pada kelas minoritas.

Pengembangan model bahasa berbasis *transformer* seperti BERT (*Bidirectional Encoder Representations from Transformers*) telah meningkatkan performa analisis sentimen secara signifikan [5]. Model ALBERT (*A Lite BERT*) hadir sebagai versi yang lebih teroptimasi dengan jumlah parameter yang lebih sedikit namun tetap mempertahankan kinerja tinggi. Arsitektur ALBERT menggunakan teknik *cross-layer parameter sharing* dan *factorized embedding parameterization* untuk meningkatkan efisiensi memori serta kecepatan pelatihan [6]. Penggunaan ALBERT menjadi solusi strategis untuk mengatasi hambatan komputasi pada pemrosesan data skala besar,.

Teknik augmentasi data diusulkan dalam penelitian ini untuk menangani masalah ketidakseimbangan dataset pada ulasan Coursera. Metode *Contextual Word Embedding* diterapkan untuk menambah variasi data pada kelas minoritas tanpa mengubah makna asli ulasan [7], [8]. Integrasi antara model ALBERT dan teknik augmentasi diharapkan dapat meningkatkan kemampuan generalisasi model dalam mengenali berbagai pola sentimen pengguna secara akurat dan efisien

II. TINJAUAN PUSTAKA

Penelitian mengenai analisis sentimen telah berkembang pesat dengan pemanfaatan model bahasa berbasis *transformer*. Bagian ini menguraikan penelitian terdahulu yang relevan dengan penggunaan model ALBERT, analisis pada dataset Coursera, serta penerapan teknik augmentasi untuk menangani ketidakseimbangan data. Penelitian terhadap ulasan kursus pada platform edukasi seperti Coursera telah dilakukan dengan berbagai pendekatan. Analisis sentimen menggunakan model Naive Bayes dan mencoba menangani ketidakseimbangan data dengan ulasan sintesis dari GPT-3 telah dilakukan, namun akurasi yang dihasilkan hanya sebesar 75% [3]. Keterbatasan model tradisional dalam memproses kalimat yang kompleks mendorong penggunaan model yang lebih mutakhir. Model BERT secara signifikan lebih efektif dibandingkan metode konvensional dalam mengklasifikasikan sentimen siswa di platform e-learning dengan akurasi mencapai 82% [8]. Model ALBERT dikembangkan untuk mengatasi kendala besarnya parameter dan kebutuhan memori pada model pendahulunya [6]. ALBERT menggunakan teknik *cross-layer parameter sharing* dan *factorized embedding parameterization* yang menjadikannya lebih ringan tanpa mengorbankan performa [1]. Dalam penelitian komparatif, ALBERT terbukti mampu memberikan hasil superior pada dataset yang memerlukan pemahaman konteks tinggi mencapai akurasi tertinggi sebesar 86% pada dataset SQuAD v2 dibandingkan dengan variasi BERT lainnya [9], [10], [11]. Kemampuan ALBERT dalam menangkap hubungan antar token melalui

mekanisme *self-attention* yang efisien menjadikannya solusi ideal untuk analisis teks skala besar [6], [12].

Ketidakeimbangan distribusi label merupakan tantangan umum dalam dataset ulasan Coursera, di mana ulasan positif jauh lebih banyak dibandingkan ulasan negatif dan netral [13], [14]. Teknik augmentasi data diusulkan sebagai solusi untuk memperkaya variasi pada kelas minoritas tanpa mengubah makna asli teks [15]. Penggunaan teknik augmentasi seperti *Synonym Replacement*, *Contextual Word Embedding*, dan *Back Translation* dapat meningkatkan performa model secara signifikan hingga mencapai akurasi 99,3% pada tugas klasifikasi teks [16], [17]. Teknik *Back Translation*, khususnya, sangat berguna untuk menghasilkan variasi frasa baru namun tetap mempertahankan inti pesan dari ulasan pengguna [18]. Berdasarkan tinjauan pustaka di atas, penelitian ini memfokuskan penggunaan model ALBERT secara spesifik pada dataset ulasan Coursera. Berbeda dengan penelitian sebelumnya yang menggunakan model Naive Bayes atau BERT konvensional, penelitian ini mengintegrasikan efisiensi parameter ALBERT dengan tiga teknik augmentasi sekaligus untuk mengatasi bias data,. Pendekatan ini diharapkan dapat menghasilkan performa klasifikasi yang lebih konsisten dan efisien secara komputasi dalam mengenali sentimen pengguna di platform pendidikan daring.

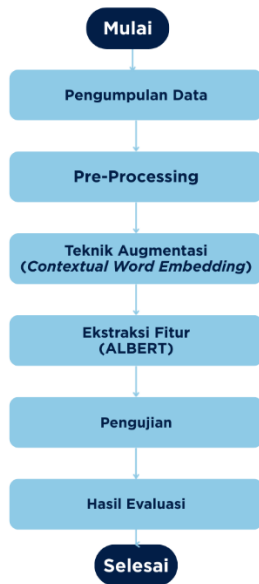
III. METODOLOGI

A. DESKRIPSI SISTEM

Penelitian ini mengevaluasi sistem analisis sentimen yang dirancang untuk mengklasifikasikan ulasan pengguna platform Coursera ke dalam tiga kategori: Positif, Netral, dan Negatif. Dataset utama yang digunakan bersifat publik dan bersumber dari Kaggle, yang kemudian diproses melalui tahapan *preprocessing* dan teknik augmentasi untuk menangani masalah ketidakseimbangan data yang signifikan. Model bahasa yang menjadi fokus utama dalam arsitektur klasifikasi ini adalah ALBERT yang bertindak sebagai ekstraktor fitur sebelum dilakukan tahap prediksi dan evaluasi.

B. TAHAPAN PENELITIAN

Secara sistematis, tahapan penelitian dimulai dari pengumpulan data hingga evaluasi akhir guna memastikan proses berjalan sesuai dengan tujuan penelitian. Tahapan Penelitian ditunjukkan pada Gambar 1.



Gambar 1. Tahapan Penelitian

1) PENGUMPULAN DATA

Data yang digunakan berjumlah 10.000 ulasan Coursera yang terdiri dari ID, ulasan teks, dan label. Karakteristik data menunjukkan dominasi ulasan positif (7.134 ulasan) dibandingkan ulasan negatif (396) dan netral (470).

2) PRE-PROCESSING

Tahapan ini bertujuan untuk membersihkan teks dan mengurangi dimensi input. *Case Folding* bertujuan mengubah seluruh teks menjadi huruf kecil untuk konsistensi. *Remove Punctuation* bertujuan menghapus tanda baca, angka, dan simbol yang tidak relevan. *Tokenization* bertujuan memecah string menjadi unit kata atau token [19].

3) TEKNIK AUGMENTASI

Teknik Augmentasi *Contextual Word Embedding* digunakan untuk menambah jumlah ulasan pada kelas negatif dan netral. Berikut merupakan contoh kalimat yang sudah dilakukan Augmentasi pada Tabel 1.

TABEL 1. KALIMAT AUGMENTASI

Kalimat Asli (Sebelum Augmentasi)	Hasil <i>Contextual Word Embedding</i> (Sesudah)
<i>the platform is very confusing and makes it hard to follow the lessons</i>	<i>the platform is quite confusing and makes it difficult to follow the lessons</i>
<i>the course content is okay but it requires a lot of time to complete</i>	<i>the course material is acceptable but it requires a significant amount of time to finish</i>
<i>i learned many useful skills from this course</i>	<i>i gained many valuable skills from this course</i>

Berbeda dengan metode statis, teknik ini mengubah kata-kata dalam ulasan berdasarkan konteks di sekitarnya, sehingga menghasilkan representasi yang lebih dinamis dan halus [16].

4) ARSITEKTUR MODEL ALBERT

Model ALBERT dipilih karena efisiensinya dalam penggunaan parameter melalui teknik *cross-layer parameter sharing* dan *factorized embedding parameterization*. Arsitektur ini terdiri dari 12 lapisan Transformer namun hanya memiliki sekitar 12 juta parameter, jauh lebih ringan dibanding BERT konvensional [1]. Proses ekstraksi fitur pada ALBERT melibatkan penggunaan token [CLS] sebagai representasi semantik keseluruhan teks, yang kemudian diteruskan ke lapisan *fully connected* dan fungsi aktivasi *softmax* untuk menentukan probabilitas sentimen akhir

5) SKENARIO EKSPERIMEN DAN EVALUASI

Eksperimen dilakukan dengan melatih model ALBERT pada data yang telah diaugmentasi menggunakan *Contextual Word Embedding* dan data yang tidak dilakkan augmentasi. Pelatihan menggunakan optimizer AdamW dengan learning rate sebesar 2×10^{-5} selama 3 epoch. Performa model diukur menggunakan metrik evaluasi yang diturunkan dari *Confusion Matrix* [20], yaitu:

- *Accuracy*: Mengukur seberapa akurat model mengklasifikasikan sentimen secara keseluruhan.
- *Precision*: Rasio prediksi positif yang benar dibandingkan dengan total hasil prediksi positif.
- *Recall*: Kemampuan model dalam menemukan kembali seluruh label kelas positif yang benar.
- *F1-Score*: Nilai rata-rata harmonik antara precision dan recall untuk memberikan gambaran performa yang seimbang.

IV. HASIL DAN PEMBAHASAN

A. HASIL PENGUMPULAN DAN DISTRIBUSI DATA

Pada tahap awal, dataset ulasan Coursera yang dikumpulkan berjumlah 10.000 data dengan distribusi label yang sangat tidak seimbang. Terdapat 7.134 ulasan positif, 470 ulasan netral, dan 396 ulasan negatif. Ketidakseimbangan ini diatasi dengan menggunakan teknik augmentasi *Contextual Word Embedding* untuk memperbanyak data pada kelas minoritas (netral dan negatif) agar sebanding dengan kelas mayoritas. Pembagian distribusi label ditunjukkan pada Tabel 2.

TABEL 2. DISTRIBUSI DATA

Kategori Sentimen	Sebelum Augmentasi	Sesudah Augmentasi (<i>Contextual Embedding</i>)
Positif	7.134	7.134
Netral	470	7.134
Negatif	396	7.134
Total Data	10.000	21.402

Jumlah dataset meningkat setelah proses augmentasi menjadi 21.402 ulasan, di mana masing-masing kategori memiliki 7.134 data. Data tersebut kemudian dibagi menjadi 80% untuk pelatihan, 10% untuk validasi, dan 10% untuk pengujian guna memastikan model dapat dievaluasi pada distribusi kelas yang seimbang ditunjukkan pada Tabel 3.

TABEL I. DISTRIBUSI KELAS

Jenis Dataset	Persentase	Jumlah Data (Per Kelas)	Total Data
Data Pelatihan (Training)	80%	5.707	17.121
Data Validasi (Validation)	10%	714	2.142
Data Pengujian (Testing)	10%	714	2.142

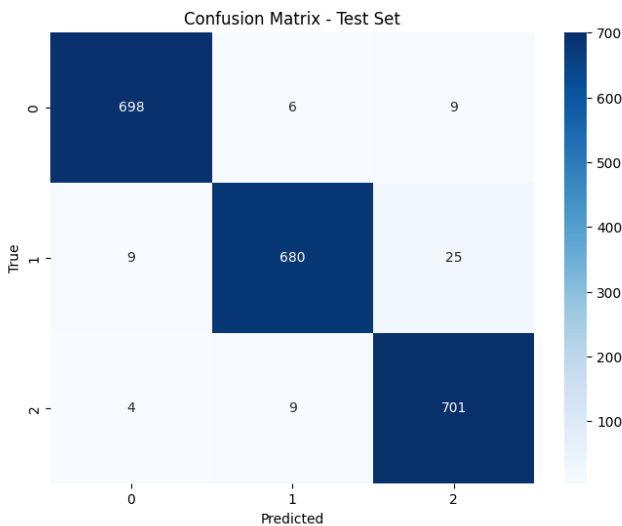
B. HASIL PENGUJIAN MODEL ALBERT

Pengujian model ALBERT pada dataset Coursera dilakukan melalui dua skenario utama, yaitu menggunakan teknik augmentasi dan tanpa augmentasi, untuk melihat efektivitas penanganan data tidak seimbang ditunjukkan pada Tabel 4.

TABEL II. HASIL EKSPERIMEN

Skenario Eksperimen	F1-Score	Keterangan
ALBERT Tanpa Augmentasi	0,9170	Mengalami kesulitan memprediksi kelas minoritas.
ALBERT + Contextual Embedding	0,9710	Peningkatan signifikan dalam mengenali variasi sentimen.

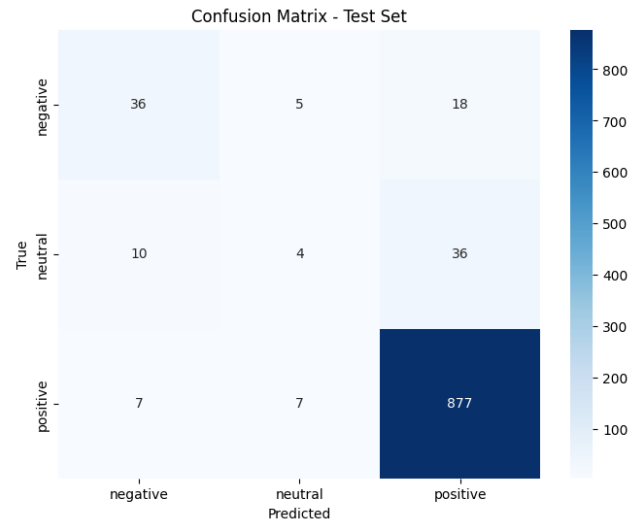
• ALBERT dengan *Contextual Word Embedding*: Hasil eksperimen menunjukkan performa yang sangat tinggi dengan perolehan F1-Score sebesar 0,9710. Penggunaan *Contextual Word Embedding* terbukti efektif dalam menangkap variasi linguistik yang kompleks tanpa mengubah makna asli ulasan pengguna. *Confusion Matrix* ALBERT dengan Augmentasi ditunjukkan pada Gambar 2.



Gambar 2. Confusion Matrix ALBERT-Augmentasi

Penerapan augmentasi data berbasis *contextual word embedding* menggunakan ALBERT terbukti secara signifikan meningkatkan performa klasifikasi pada kelas minoritas, khususnya kelas neutral. *Confusion matrix* menunjukkan distribusi kesalahan yang jauh lebih seimbang serta peningkatan kemampuan generalisasi model. Hasil ini mengindikasikan bahwa augmentasi kontekstual efektif dalam mempertahankan makna semantik sekaligus memperkaya representasi data latih.

• ALBERT tanpa Augmentasi: Sebagai pembandingan, pengujian tanpa teknik augmentasi menghasilkan *F1-Score* sebesar 0,9170. Meskipun masih kompetitif, model mengalami kesulitan signifikan dalam memprediksi kelas minoritas, di mana hanya 4 sampel netral yang berhasil diprediksi dengan benar, sementara sebagian besar lainnya salah diklasifikasikan ke label lain.



Gambar 3. Confusion Matrix ALBERT- Tanpa Augmentasi

Hasil evaluasi menunjukkan bahwa model ALBERT tanpa augmentasi mampu mencapai performa yang sangat baik pada kelas positif, namun menunjukkan kelemahan signifikan dalam mengklasifikasikan kelas negative dan terutama neutral. *Confusion matrix* mengindikasikan adanya bias kuat terhadap kelas mayoritas, yang menyebabkan rendahnya recall pada kelas neutral. Hal ini menegaskan bahwa meskipun akurasi keseluruhan tinggi, model belum memiliki kemampuan generalisasi yang seimbang antar kelas. Strategi penyeimbang data augmentasi teks, *class weighting*, atau *resampling* untuk meningkatkan performa pada kelas minoritas lebih diprioritaskan.

V. KESIMPULAN

Implementasi model ALBERT pada ulasan pengguna Coursera terbukti memberikan performa klasifikasi sentimen yang sangat akurat melalui pendekatan penanganan data tidak seimbang. Penerapan teknik augmentasi *Contextual Word Embedding* secara signifikan meningkatkan kemampuan model dalam mengenali sentimen netral dan negatif yang sebelumnya sulit teridentifikasi akibat dominasi ulasan positif. Hasil pengujian menunjukkan peningkatan nilai F1-Score yang substansial, yakni dari 0,9170 pada data asli menjadi 0,9710 setelah dataset diseimbangkan secara proporsional. Kemampuan ALBERT dalam menangkap variasi linguistik yang kompleks tanpa mengubah makna asli ulasan menjadikannya solusi efektif untuk memitigasi bias terhadap kelas mayoritas. Penelitian ini menegaskan bahwa penggunaan model bahasa berbasis *transformer* yang didukung oleh data seimbang mampu menghasilkan generalisasi performa yang superior pada platform pendidikan daring.

REFERENSI

[1] C. Wang, S. Huang, and Y. Zhou, "Sentiment analysis of ALBERT-BiLSTM model MOOC reviews via," vol. 05008, pp. 1–8, 2021.

- [2] A. Pilicita-garrido and E. Barra, "Sentiment Analysis With Transformers Applied to Education : Systematic Review," vol. 9, no. 2, pp. 72–83, 2025, doi: 10.9781/ijimai.2025.02.008.
- [3] C. Suhaeni and H.-S. Yong, "Mitigating Class Imbalance in Sentiment Analysis through GPT-3-Generated Synthetic Sentences," *Appl. Sci.*, vol. 13, no. 17, p. 9766, 2023, doi: 10.3390/app13179766.
- [4] A. R. Nair, R. P. Singh, D. Gupta, and P. Kumar, "Evaluating the Impact of Text Data Augmentation on Text Classification Tasks Using DistilBERT BT - Procedia Computer Science," 2024, pp. 102–111. doi: 10.1016/j.procs.2024.04.013.
- [5] Y. Wu, Z. Jin, C. Shi, P. Liang, and T. Zhan, "Research on the Application of Deep Learning-based BERT Model in Sentiment Analysis".
- [6] Z. Lan, M. Chen, S. Goodman, K. Gimpel, P. Sharma, and R. Soicrut, "ALBERT: A Lite BERT for Self-Supervised Learning of Language Representations," *arXiv*, 2020.
- [7] A. Pingle and G. Kale, "Robust Sentiment Analysis for Low Resource languages Using Data Augmentation Approaches : A Case Study in Marathi," 2024.
- [8] X. J. Lee, T. Tzen, V. Yap, H. Ng, and V. T. Goh, "Comparison of Word Embeddings for Sentiment," vol. 1, pp. 488–502, 2022, doi: 10.2991/978-94-6463-094-7.
- [9] P. Suryavanshi and S. Gangele, "INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING Sentiment Analysis for Prediction of Brand Value Using Albert Model," 2024.
- [10] B. Ozkurt and A. Cemozkurt, "Comparative Analysis of BERT Variants on Contextual Question Answering Tasks," *Expert Syst. Appl.*, vol. 235, p. 121021, 2024, doi: 10.1016/j.eswa.2023.121021.
- [11] A. Sahoo and R. Chanda, "Comparative Analysis of BERT Models for Sentiment Analysis on Twitter Data," 2023.
- [12] V. Sugiarto, A. S. Agstringtyas, and Y. M. Cholily, "Performance Comparison of BERT , ALBERT and RoBERTa for Sentiment Analysis in Critical Pilot Communication Prior to Aviation Accident .," vol. 01, no. 01, pp. 19–32, 2025.
- [13] W. Safira, B. Prabaswara, A. S. Karnyoto, and B. Pardamean, "Leveraging ALBERT for Sentiment Classification of Long-Form ChatGPT Reviews on Twitter," vol. 17, no. 1, pp. 1–12, 2025.
- [14] D. Marutho and V. G. Utomo, "Benchmarking IndoBERT and Transformer Models for Sentiment Classification on Indonesian E-Government Service Reviews," vol. 23, no. 1, pp. 85–95, 2025.
- [15] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," *NAACL-HLT*, 2023, doi: 10.18653/v1/N19-1423.
- [16] A. Abbas, M. Lee, N. Shanavas, and V. Kovatchev, "Clinical concept annotation with contextual word embedding in active transfer learning environment," *Digit. Heal.*, vol. 10, 2024, doi: 10.1177/20552076241308987.
- [17] L. Rasmy, "Med-BERT: pretrained contextualized embeddings on large- scale structured electronic health records for disease prediction," *npj Digit. Med.*, 2021, doi: 10.1038/s41746-021-00455-y.
- [18] D. R. Beddiar, S. Jahan, and M. Oussalah, "Data expansion using back translation and paraphrasing for hate speech," *Online Soc. Networks Media*, vol. 24, no. May, p. 100153, 2021, doi: 10.1016/j.osnem.2021.100153.
- [19] A. Alkhusayni and J. Lee, "Improving Multilingual Sentiment Classification Using Back Translation," *Appl. Soft Comput.*, vol. 150, p. 110873, 2025, doi: 10.1016/j.asoc.2024.110873.
- [20] H. Murfi, T. Gowandi, G. Ardaneswari, and S. Nur-, "BERT-Based Combination of Convolutional and Recurrent Neural Network for Indonesian Sentiment Analysis," pp. 1–15.