BAB I
PENDAHULUAN

1.1 Latar Belakang Masalah

E-commerce telah menjadi salah satu sektor yang berkembang pesat di Indonesia, dengan Tokopedia sebagai salah satu platform terbesar yang melayani jutaan pengguna setiap harinya. Menurut data terbaru, Tokopedia memiliki lebih dari 100 juta pengguna aktif dan menjadi salah satu marketplace paling populer di Indonesia Rahadian (2021). Review pengguna pada aplikasi e-commerce seperti Tokopedia merupakan sumber informasi yang sangat berharga untuk memahami kepuasan pelanggan, mengidentifikasi area perbaikan, dan meningkatkan kualitas layanan.

Analisis sentimen telah menjadi topik penelitian yang signifikan dalam bidang Natural Language Processing (NLP), khususnya untuk mengekstraksi opini dan sentimen dari teks berbahasa Indonesia. Penelitian terbaru menunjukkan bahwa model berbasis transformer seperti BERT telah memberikan peningkatan performa yang signifikan dalam tugas-tugas NLP dibandingkan dengan metode tradisional Ahmadian et al. (2024) IndoBERT, sebagai varian BERT yang dilatih khusus untuk bahasa Indonesia, telah terbukti efektif dalam berbagai tugas pemrosesan bahasa Indonesia dengan mencapai F1-score hingga 79.1% pada tugas klasifikasi emosi (Shaw, 2023) .

Meskipun IndoBERT menunjukkan performa yang baik, penelitian menunjukkan bahwa pendekatan hybrid yang menggabungkan model deep learning dengan algoritma machine learning tradisional dapat meningkatkan akurasi

Dipindai dengan CamScanner

klasifikasi. XGBoost (Extreme Gradient Boosting) telah terbukti menjadi salah satu algoritma ensemble learning yang paling efektif untuk berbagai tugas klasifikasi, termasuk analisis sentimen Chandrasekaran et al. (2024). Penelitian oleh Karabila et al. (2024) menunjukkan bahwa kombinasi BERT dengan XGBoost dapat meningkatkan akurasi klasifikasi hingga 97.18% pada dataset multimodal.

Penelitian-penelitian sebelumnya tentang analisis sentimen pada review Tokopedia umumnya menggunakan metode machine learning tradisional seperti Naive Bayes, Random Forest, dan SVM dengan tingkat akurasi berkisar antara 84-89% Idris, Rifai, & Tania (2025). Namun, belum banyak penelitian yang mengeksplorasi pendekatan hybrid antara IndoBERT dengan XGBoost untuk meningkatkan performa klasifikasi sentimen pada review berbahasa Indonesia, khususnya pada platform Tokopedia.

Berdasarkan masalah tersebut penelitian ini untuk mengimplementasikan dan mengevaluasi model hybrid IndoBERT-XGBoost untuk klasifikasi sentimen review aplikasi Tokopedia. Pendekatan ini diharapkan dapat menggabungkan kekuatan IndoBERT dalam memahami konteks semantik bahasa Indonesia dengan kemampuan XGBoost dalam mengoptimalkan proses klasifikasi, sehingga menghasilkan performa yang lebih baik dibandingkan metode tunggal.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, rumusan masalah dalam penelitian ini adalah:

1. Bagaimana mengimplementasikan model hybrid IndoBERT dan XGBoost untuk klasifikasi sentimen review aplikasi Tokopedia?

Dipindai dengan CamScanner

2. Bagaimana performa model hybrid IndoBERT-XGBoost dibandingkan dengan model IndoBERT tunggal dan XGBoost tunggal dalam klasifikasi sentimen review Tokopedia?
3. Faktor-faktor apa saja yang mempengaruhi akurasi klasifikasi sentimen pada review aplikasi Tokopedia menggunakan model hybrid IndoBERT-XGBoost?

1.3 Batasan Masalah

Untuk memfokuskan penelitian dan membatasi ruang lingkupnya, penelitian ini memiliki beberapa batasan sebagai berikut: Keterbatasan Data Hasil Scraping Google Play Store: Penelitian ini menggunakan data ulasan yang dikumpulkan secara langsung melalui teknik web scraping dari Google Play Store menggunakan library google-play-scraper berbasis Python. Proses scraping dilakukan pada periode tertentu sehingga data yang terkumpul hanya mencerminkan ulasan pada rentang waktu tersebut. Selain itu, ketersediaan ulasan bergantung pada kebijakan Google Play Store yang dapat berubah sewaktu-waktu dan tidak mencakup seluruh histori ulasan aplikasi Tokopedia.

1. Pelabelan Otomatis Berdasarkan Rating (Proxy Labeling): Pelabelan sentimen dilakukan secara otomatis dan deterministik berdasarkan peringkat bintang, di mana rating 4-5 = Positif, rating 3 = Netral, dan rating 1-2 = Negatif. Pendekatan ini memiliki keterbatasan karena terdapat kemungkinan ketidaksesuaian (mismatch) antara nada teks ulasan dan rating numerik (misalnya, pengguna menulis keluhan panjang tetapi memberi rating 4 karena alasan tertentu). Validasi manual pada sampel acak

Dipindai dengan CamScanner

akan dilakukan untuk mengestimasi tingkat akurasi proxy labeling ini, namun tidak untuk seluruh dataset.

2. Bahasa dan Konteks pada Ulasan Aplikasi: Model hanya akan dikembangkan dan dievaluasi pada teks berbahasa Indonesia yang ditemukan dalam ulasan aplikasi seluler. Meskipun IndoBERT relatif kuat, performa model terhadap bahasa gaul yang sangat baru, campuran kode dengan bahasa daerah, atau istilah teknis khusus produk yang tidak muncul dalam data pelatihan pre-trained IndoBERT mungkin terbatas.
3. Klasifikasi Tingkat Ulasan (Document-Level): Penelitian berfokus pada klasifikasi sentimen tingkat dokumen, yaitu memberikan satu label sentimen untuk setiap ulasan secara utuh. Analisis yang lebih detail seperti analisis sentimen berbasis aspek (Aspect-Based Sentiment Analysis), yang mengidentifikasi sentimen terhadap entitas atau fitur spesifik (contoh: "pengiriman cepat", "UI membingungkan"), tidak dilakukan dalam cakupan penelitian ini.
4. Arsitektur Hybrid Spesifik dan Model Baseline: Eksperimen hybrid terbatas pada kombinasi IndoBERT-base (sebagai feature extractor) dengan XGBoost (sebagai classifier). Variasi arsitektur lain seperti penggunaan model transformer lainnya (IndoBERT-Large, mBERT), atau penggabungan dengan algoritma ensemble selain XGBoost (seperti LightGBM, CatBoost) tidak dieksplorasi.

1.4 Tujuan Penelitian

Tujuan dari penelitian ini adalah:

Dipindai dengan CamScanner

1. Mengimplementasikan model hybrid IndoBERT dan XGBoost untuk klasifikasi sentimen review aplikasi Tokopedia.
2. Menganalisis dan membandingkan performa model hybrid IndoBERT-XGBoost dengan model IndoBERT tunggal dan XGBoost tunggal dalam klasifikasi sentimen review Tokopedia.
3. Mengidentifikasi faktor-faktor yang mempengaruhi akurasi klasifikasi sentimen menggunakan model hybrid IndoBERT-XGBoost.

1.5 Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan manfaat sebagai berikut:

1. Manfaat Teoritis

   Penelitian ini diharapkan dapat memberikan kontribusi pada pengembangan ilmu pengetahuan, khususnya dalam bidang Natural Language Processing dan sentiment analysis untuk bahasa Indonesia. Selain itu, penelitian ini juga dapat menjadi referensi bagi penelitian selanjutnya tentang penggunaan model hybrid dalam analisis sentimen berbahasa Indonesia.

2. Manfaat Praktis

   Manfaat praktis dari penelitian ini meliputi:

   a. Bagi Platform E-commerce: Menyediakan solusi otomatis untuk menganalisis feedback pelanggan secara real-time, sehingga dapat membantu dalam pengambilan keputusan strategis dan peningkatan kualitas layanan.

Dipindai dengan CamScanner

b. Bagi Peneliti: Memberikan metodologi dan framework yang dapat digunakan untuk pengembangan sistem analisis sentimen pada platform e-commerce lainnya.
c. Bagi Pengembang: Menyediakan model yang dapat diintegrasikan ke dalam sistem monitoring kepuasan pelanggan pada aplikasi e-commerce.

1.6 Sistematika Penulisan

Sistematika penulisan penelitian ini disusun sebagai berikut:

BAB I PENDAHULUAN

Bab ini berisi latar belakang penelitian tentang analisis sentimen review Tokopedia, gap penelitian mengenai belum adanya kombinasi IndoBERT dengan XGBoost, rumusan masalah, batasan penelitian, tujuan, manfaat, dan ringkasan penelitian terkait.

BAB II TINJAUAN PUSTAKA

Bab ini menjelaskan teori dasar tentang analisis sentimen, cara kerja IndoBERT untuk memahami bahasa Indonesia, keunggulan XGBoost untuk klasifikasi, konsep model hybrid, dan metode evaluasi yang akan digunakan.

BAB III METODOLOGI PENELITIAN

Bab ini memaparkan langkah-langkah penelitian: pengumpulan data review melalui web scraping Google Play Store menggunakan library google-play-scraper, preprocessing teks, fine-tuning IndoBERT, training XGBoost, pembuatan model pembanding, dan evaluasi hasil.

Dipindai dengan CamScanner

BAB IV HASIL DAN PEMBAHASAN

Bab ini menyajikan hasil implementasi model, meliputi gambaran umum dataset, hasil preprocessing, penanganan ketidakseimbangan kelas dengan SMOTE, hasil fine-tuning IndoBERT,hasil model hybrid IndoBERT-XGBoost, perbandingan dengan model baseline, hasil validasi statistik(K-Fold Cross Validation dan Paired T-Test), analisis interpretabilitas model menggunakan SHAP, serta pembahasan menyeluruh atas temuan penelitian.

BAB V KESIMPULAN DAN SARAN

Bab ini berisi kesimpulan yang menjawab rumusan masalah penelitian berdasarkan hasil implementasi dan pengujian model hybrid IndoBERT-XGBoost, serta saran bagi penelitian Selanjutnya, Pengembang Platform Tokopedia, dan program studi.

1.7 Penelitian Terkait

Tabel 1.1 Penelitian Terkait

Penulis dan Tahun	Judul	Hasil / Pencapaian
Ahmadian et al. (2024)	Hybrid Models for Emotion Classification and	Penelitian ini menggabungkan IndoBERT dengan BiLSTM, BiGRU, dan attention mechanism.

Dipindai dengan CamScanner

	Sentiment Analysis	Hasil penelitian menunjukkan bahwa model hybrid IndoBERT-BiGRU dengan attention mechanism mencapai F1-score 0.791 untuk klasifikasi emosi dan akurasi 92% untuk analisis sentimen pada dataset IndoNLU. Penelitian ini menggunakan empat hidden layers terakhir dari IndoBERT sebagai input untuk neural network models, yang terbukti meningkatkan performa dibandingkan menggunakan satu hidden layer saja.
Setiawan, Iswavigra, & Anggiratih (2025)	Implementation of IndoBERT for Sentiment Analysis of the Constitutional Court's Decision Regarding the Minimum Age of	menganalisis performa IndoBERT untuk klasifikasi sentimen pada komentar YouTube berbahasa Indonesia terkait keputusan Mahkamah Konstitusi. Dataset terdiri dari 23.796 komentar yang diklasifikasikan menjadi positif, negatif, dan netral. IndoBERT mencapai akurasi 95%

Dipindai dengan CamScanner

	Vice Presidential Candidates	dibandingkan dengan BERT (92%), SVM (88%), dan Random Forest (85%). Penelitian ini menekankan kemampuan IndoBERT dalam menangkap nuansa kontekstual bahasa Indonesia yang kompleks, terutama dalam konteks politik dan hukum.
Kusoema & Ibrahim (2025)	Sentiment Analysis on PT Pertamina Corruption Case	mengintegrasikan arsitektur IndoBERT-RCNN untuk analisis sentimen pada isu sosial-politik, khususnya kasus korupsi PT Pertamina. Penelitian ini mengevaluasi berbagai kombinasi hyperparameter dengan learning rate 2e-5 dan batch size 16 menghasilkan performa optimal dengan akurasi 84% dan F1-score 83%. Model menunjukkan performa kuat dalam mengklasifikasikan komentar negatif, namun akurasi untuk

Dipindai dengan CamScanner

		kelas netral dan positif relatif lebih rendah karena overlap semantik dalam ekspresi pengguna.
Chandrasekaran et al. (2024)	Multimodal Sentiment Analysis with LSTM-XGBoost	mengembangkan Hybrid LXGB Model yang menggabungkan LSTM dengan XGBoost classifier untuk analisis sentimen multimodal. Model ini mencapai akurasi exceptional 97.18% pada dataset CMU-MOSEI, melampaui classifier alternatif termasuk LSTM, CNN, DNN, dan XGBoost standalone. Penelitian ini menunjukkan bahwa kombinasi deep learning dengan gradient boosting dapat menangkap spektrum nuansa sentimen yang lebih baik pada dataset multimodal yang mencakup data tekstual, gambar, dan audio.
Perwira et al. (2025)	Domain-Specific Fine-Tuning of	melakukan fine-tuning IndoBERT untuk aspect-based sentiment analysis pada user-generated

Dipindai dengan CamScanner

	IndoBERT for Tourism	content pariwisata Indonesia dari Google Reviews. Model yang di-fine-tune pada domain spesifik pariwisata menunjukkan kemampuan superior dalam mengidentifikasi aspek kunci seperti kualitas layanan dan fasilitas. Penelitian ini menekankan pentingnya domain adaptation dalam meningkatkan performa model pre-trained untuk aplikasi spesifik.
Idris et al. (2025)	Sentiment Analysis of Tokopedia App Reviews	mengevaluasi performa tiga algoritma machine learning - SVM, Random Forest, dan Gaussian Naïve Bayes - dikombinasikan dengan Word2Vec dan FastText untuk analisis sentimen review aplikasi Tokopedia. Dataset sebanyak 59.811 review dikumpulkan dari Google Play Store. SVM mencapai performa terbaik

Dipindai dengan CamScanner

		dengan akurasi 89.06% menggunakan FastText dan 89.02% menggunakan Word2Vec. Penelitian ini menunjukkan bahwa perbedaan performa antara Word2Vec dan FastText minimal di semua algoritma, mengindikasikan efektivitas yang sebanding.
Karabila, Darraz, El-Ansari, Alami, & Mallahi (2024)	BERT-enhanced sentiment analysis for personalized e-commerce recommendations	mengintegrasikan collaborative filtering dengan sentiment analysis menggunakan BERT fine-tuned model untuk rekomendasi e-commerce yang dipersonalisasi. Pendekatan hybrid mereka menunjukkan akurasi 91% dalam klasifikasi sentimen, secara signifikan lebih baik dari model sebelumnya. Model BERT digunakan untuk ekstraksi insight dari review teks, yang kemudian diintegrasikan ke dalam sistem rekomendasi untuk

Dipindai dengan CamScanner

		meningkatkan akurasi dan personalisasi.
Ulhaq & Suprayogi (2025)	Comparing Machine Learning Models for Tokopedia Reviews	melakukan evaluasi komparatif model machine learning untuk analisis sentimen review Tokopedia dalam bahasa Indonesia. Dataset terdiri dari 10.236 entri unik setelah preprocessing. Setiap model dievaluasi menggunakan hold-out validation (80:20 split) dan 5-fold cross-validation dengan metrik akurasi, precision, recall, dan F1-score. Penelitian ini memberikan benchmark performa berbagai algoritma pada dataset review e-commerce Indonesia.
Nsaif & Abd (2022)	SentiXGboost for Enhanced Sentiment Analysis	mengusulkan metode ensemble classifier novel yang menggunakan kombinasi multiple feature sets dengan ensemble

Dipindai dengan CamScanner

		classification berbasis XGBoost. Feature sets yang digunakan meliputi Bag of Words, TF-IDF, Part of Speech, N-gram, Opinion Lexicon, dan Term Frequency. Eksperimen mengkonfirmasi bahwa metode ensemble yang diusulkan mengungguli semua classifier individual dan secara signifikan meningkatkan performa klasifikasi sentimen keseluruhan pada dataset yang sering digunakan dalam analisis sentimen, mencapai akurasi 95.16%.
Mushtaq Gardazi et al. (2025)	BERT Applications in Natural Language Processing	menunjukkan bahwa model berbasis BERT mengungguli metode lain dalam analisis sentimen e-commerce dan code-mixed language. Studi ini mendemonstrasikan efektivitas BERT pada platform e-commerce dengan hasil yang well-structured.

Dipindai dengan CamScanner

		Review ini juga menyoroti bahwa (Karabila, Darraz, El-Ansari, Alami, & Mallahi, 2024) mengusulkan model hybrid yang menggabungkan BERT dengan XGBoost untuk review ChatGPT.

1.8 Kesimpulan dan Relevansi Penelitian Terdahulu

Berdasarkan tinjauan penelitian terdahulu yang telah diuraikan, dapat disimpulkan bahwa pengembangan model analisis sentimen untuk bahasa Indonesia telah mengalami kemajuan signifikan melalui penggunaan model transformer berbasis BERT dan pendekatan hybrid. Penelitian-penelitian tersebut memberikan beberapa temuan penting yang relevan dengan penelitian ini:

Pertama, IndoBERT telah terbukti efektif untuk tugas analisis sentimen bahasa Indonesia. Ahmadian, Abidin, Riza, & Muchtar (2024) menunjukkan bahwa IndoBERT mencapai akurasi 92% untuk analisis sentimen, sementara Ahmadian, Abidin, Riza, & Muchtar (2025) memperoleh akurasi 88% pada klasifikasi sentimen komentar YouTube. Penelitian Kusoema & Ibrahim (2025) menggunakan pendekatan IndoBERT-RCNN dengan akurasi 84%. Temuan ini mengonfirmasi superioritas IndoBERT dibandingkan model multilingual BERT standar dalam menangkap nuansa kontekstual bahasa Indonesia.

Kedua, pendekatan hybrid yang menggabungkan deep learning dengan machine learning tradisional menunjukkan hasil yang menjanjikan. Chandrasekaran et al. (2024) mencapai akurasi 97.18% dengan model LSTM-

Dipindai dengan CamScanner

XGBoost, mendemonstrasikan bahwa kombinasi deep learning untuk ekstraksi fitur dengan gradient boosting untuk klasifikasi dapat menghasilkan performa superior. Hal serupa ditunjukkan oleh Karabilaet al. (2024) yang mengintegrasikan BERT dengan XGBoost mencapai akurasi 91% dalam klasifikasi sentimen review e-commerce. Nsaif & Abd (2022) juga membuktikan efektivitas ensemble XGBoost dengan multiple feature sets mencapai akurasi 95.16%.

Ketiga, penelitian-penelitian terdahulu pada domain e-commerce Indonesia menunjukkan tantangan khusus dalam menangani bahasa informal, slang, dan variasi dialek. Ulhaq & Suprayogi (2025) melakukan evaluasi komparatif berbagai model machine learning pada review Tokopedia dengan dataset 10.236 entri, memberikan benchmark performa yang menjadi acuan penting. Idris et.al (2025) menekankan pentingnya preprocessing yang komprehensif untuk meningkatkan akurasi hingga 5% pada teks bahasa Indonesia.

Relevansi penelitian ini dengan studi terdahulu terletak pada pengintegrasian kekuatan IndoBERT dalam pemahaman kontekstual bahasa Indonesia dengan kemampuan optimalisasi XGBoost dalam klasifikasi. Berbeda dengan penelitian Ahmadian et al. (2024) yang menggunakan kombinasi IndoBERT-BiGRU, atau Kusoema & Ibrahim (2025) yang menggunakan IndoBERT-RCNN, penelitian ini mengadopsi pendekatan hybrid IndoBERT-XGBoost yang belum dieksplorasi secara mendalam pada konteks review e-commerce Indonesia. Penelitian ini juga melengkapi studi Ulhaq & Suprayogi (2025) dengan menyediakan alternatif metodologi yang menggabungkan state-of-the-art deep learning dengan gradient boosting machine. Berbeda dengan penelitian

Dipindai dengan CamScanner

sebelumnya yang umumnya menggunakan dataset dari platform pihak ketiga, penelitian ini mengumpulkan data secara langsung dari Google Play Store melalui teknik web scraping menggunakan library google-play-scraper, sehingga berhasil mengumpulkan 35.038 ulasan mentah yang setelah preprocessing menghasilkan 27.018 ulasan bersih yang bersifat primer, aktual, dan representatif. Hal ini memungkinkan evaluasi yang lebih objektif terhadap efektivitas pendekatan hybrid IndoBERT-XGBoost yang diusulkan.

Dipindai dengan CamScanner