Jurnal Komputer, Informasi dan Teknologi Vol: 5, No 1, 2025, Page: 1-15

Pengukuran Tingkat Akurasi Pada Ulasan *E-Commerce* Menggunakan Metode *INDOBERT* Dengan *Optimizer Adam*

Agung Pratama Putra*, Ismarmiaty, Apriani

Program Studi Teknologi Informasi, Fakultas Teknik, Universitas Bumigora, Mataram

Abstrak: Pesatnya perkembangan e-commerce di Indonesia telah menyebabkan meningkatnya volume ulasan pengguna, yang kini menjadi sumber informasi penting dalam memahami kepuasan dan pengalaman pelanggan. Ulasan tersebut mengandung sentimen yang dapat dimanfaatkan untuk pengambilan keputusan bisnis, khususnya melalui pendekatan analisis sentimen. Penelitian ini bertujuan untuk mengukur tingkat akurasi model IndoBERT yang dioptimasi menggunakan algoritma Adam dalam mengklasifikasikan sentimen ulasan e-commerce berbahasa Indonesia. Dataset yang digunakan berasal dari platform Kaggle, terdiri atas 11.606 ulasan yang telah melalui tahap pra-pemrosesan seperti pembersihan teks, tokenisasi, dan pelabelan. Data kemudian dibagi menjadi 80% data latih dan 20% data uji. Model yang digunakan merupakan IndoBERT pre-trained yang dioptimasi dengan konfigurasi Adam: learning rate 2e-5, batch size 16, dan 5 epoch pelatihan. Evaluasi model dilakukan dengan menggunakan metrik akurasi, precision, recall, dan F1-score. Hasil evaluasi menunjukkan bahwa model mencapai akurasi sebesar 92,16%, precision 93,16%, recall 91,09%, dan F1-score 92,11%. Confusion matrix juga menunjukkan distribusi prediksi yang seimbang untuk sentimen positif dan negatif. Temuan ini membuktikan bahwa model IndoBERT yang dioptimasi dengan Adam memiliki kinerja yang sangat baik dalam tugas klasifikasi sentimen, serta berpotensi untuk diterapkan dalam sistem analisis sentimen otomatis pada platform e-commerce guna mendukung pengambilan keputusan berbasis data.

Kata Kunci: *IndoBERT, Optimizer Adam, E-Commerce,* Analisis Sentimen, *NLP,* Akurasi

DOI: https://doi.org/10.53697/jkomitek.v5i1.2448
*Correspondence: Agung Pratama Putra
Email: agungpratamaptr@gmail.com

Received: 22-04-2025
Accepted: 22-05-2025
Published: 22-06-2025

Copyright: © 2025 by the authors. Submitted for open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).

Abstract: *The rapid growth of e-commerce in Indonesia has led to a significant increase in the volume of user reviews, which have become an essential source of information for understanding customer satisfaction and experiences. These reviews contain sentiments that can be utilized for business decision-making, particularly through sentiment analysis approaches. This study aims to measure the accuracy level of the IndoBERT model optimized using the Adam algorithm in classifying sentiments of Indonesian-language e-commerce reviews. The dataset used was obtained from the Kaggle platform, consisting of 11,606 reviews that underwent preprocessing stages such as text cleaning, tokenization, and labeling. The data was then split into 80% training data and 20% testing data. The model used was a pre-trained IndoBERT, optimized using the Adam configuration: learning rate of 2e-5, batch size of 16, and 5 training epochs. The model was evaluated using accuracy, precision, recall, and F1-score metrics. The evaluation results showed that the model achieved an accuracy of 92.16%, precision of 93.16%, recall of 91.09%, and an F1-score of 92.11%. The confusion matrix also indicated a balanced distribution of predictions for both positive and negative sentiments. These findings demonstrate that the IndoBERT model optimized with Adam performs very well in sentiment classification tasks and has the potential to be implemented in automated sentiment analysis systems on e-commerce platforms to support data-driven decision-making.*
Keywords: *IndoBERT, Adam Optimizer, E-Commerce, Sentiment Analysis, NLP, Accuracy*

Pendahuluan

Perkembangan pesat platform E-Commerce di Indonesia tidak terlepas dari perubahan perilaku konsumen yang semakin terbiasa melakukan transaksi secara daring (Nasution, 2024). Pertumbuhan ini didukung oleh tingginya penetrasi internet, dengan jumlah pengguna mencapai 212,9 juta orang atau sekitar 77% dari total populasi. Selain itu, perangkat mobile yang terhubung ke internet tercatat sebanyak 353,8 juta unit, mengindikasikan bahwa satu individu umumnya menggunakan lebih dari satu perangkat untuk mengakses layanan digital (We Are Social, 2024).

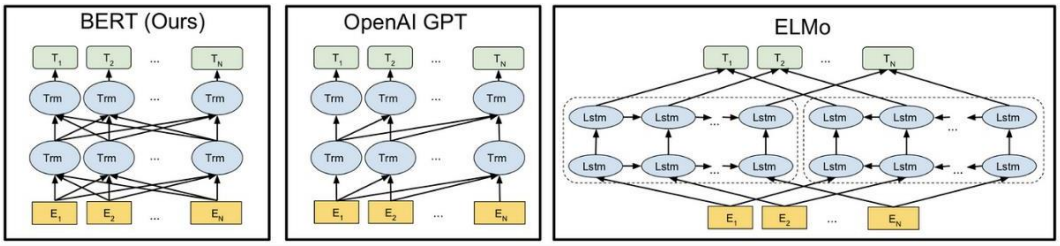
Di tengah kondisi ini, menjaga loyalitas pengguna dan meningkatkan kualitas layanan menjadi kunci utama keberhasilan platform E-Commerce. Beberapa platform populer seperti Shopee, Tokopedia, Lazada, Bukalapak, dan Blibli sering dijadikan acuan dalam mengukur kepuasan pengguna. Oleh karena itu, analisis sentimen menjadi metode yang penting untuk memahami tingkat kepuasan dan kepercayaan pengguna terhadap layanan yang mereka gunakan (Garamba & Nur, 2024). Namun, kompleksitas ulasan yang beragam dalam jumlah dan gaya bahasa menjadi tantangan tersendiri, terutama dalam konteks Indonesia yang memiliki keragaman linguistik. Teknologi Natural Language Processing (NLP) dapat dimanfaatkan untuk membantu dalam memahami, mengelompokkan, dan menganalisis teks ulasan dengan lebih akurat (Ramdhan, 2024a).

Salah satu model NLP yang unggul adalah INDOBERT (Bidirectional Encoder Representations From Transformers), yang terbukti efektif dalam memahami makna dan konteks dalam teks berbahasa Indonesia (Kusuma & Cahyono, 2023). Dalam penelitian ini, INDOBERT diaplikasikan pada dataset yang diperoleh dari platform Kaggle, yang mencakup berbagai ulasan konsumen terkait aspek produk, pengiriman, harga, dan layanan pelanggan. Tujuan dari penelitian ini adalah untuk mengidentifikasi pola sentimen dan opini secara menyeluruh, serta menyajikan hasil analisis yang dapat memberikan wawasan strategis bagi platform E-Commerce dalam meningkatkan layanan dan memenuhi ekspektasi pengguna (Priambodo & Zuliarso, 2024).

Dengan mengoptimalkan model INDOBERT menggunakan algoritma Adam, penelitian ini juga bertujuan untuk mengukur tingkat akurasi model INDOBERT dalam menganalisis ulasan pengguna platform E-Commerce yang ditulis dalam bahasa Indonesia. Selain itu, penelitian ini juga difokuskan pada penerapan algoritma optimasi Adam dalam proses pelatihan model, untuk mengevaluasi sejauh mana algoritma tersebut dapat meningkatkan akurasi serta mempercepat konvergensi model dalam tugas klasifikasi sentimen. Dengan menggabungkan pendekatan model bahasa berbasis transformer yang telah dilatih sebelumnya dengan teknik optimasi yang efisien, diharapkan penelitian ini mampu memberikan kontribusi terhadap pengembangan sistem analisis sentimen otomatis yang lebih akurat dan responsif terhadap dinamika data pengguna di sektor E-Commerce (Darmawan, 2023).

Metodologi

Penelitian ini menggunakan pendekatan kuantitatif dengan metode pemrosesan bahasa alami *Natural Language Processing (NLP*) yang fokus pada penerapan model INDOBERT untuk analisis sentimen ulasan pengguna *E-Commerce*. Model *BERT (Bidirectional Encoder Representations from Transformers),* yang dikembangkan oleh *Google*, diterapkan untuk menganalisis teks secara *Bidirectional*, memungkinkan pemahaman konteks yang lebih dalam dibandingkan model bahasa satu arah (Computer, 2023). Dengan menggunakan arsitektur Transformer multi-layer, BERT menganalisis teks input secara simultan dari kedua arah, kiri dan kanan. Perbedaan utama antara BERT dan model lain seperti OpenAIGPT dan ELMo terletak pada kemampuan BERT untuk memahami konteks secara lebih holistik dengan memperhatikan informasi dari kedua arah.



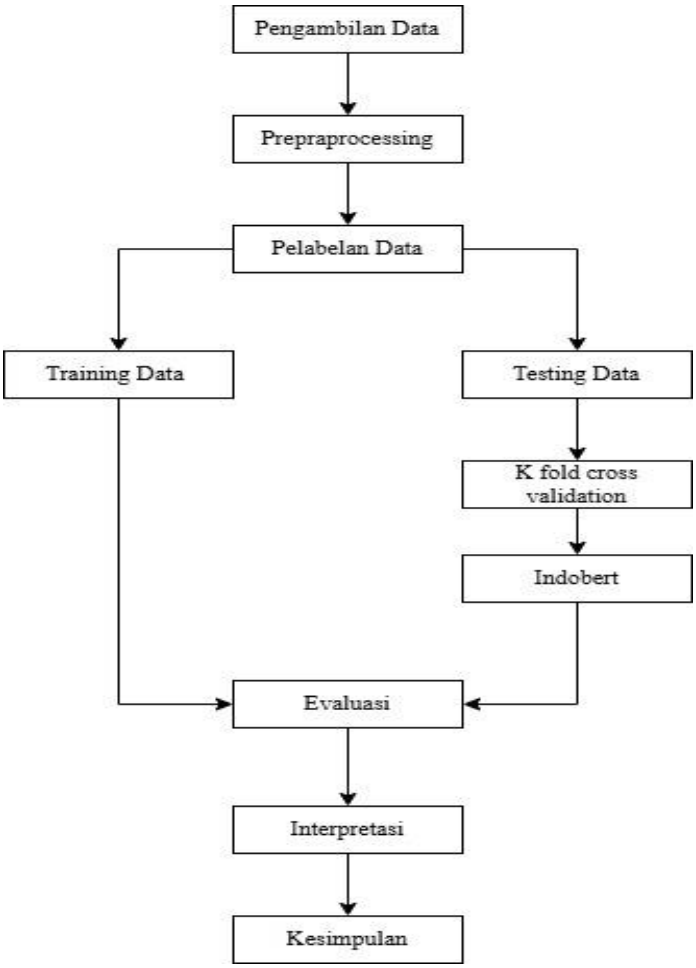
Gambar 1. Perbedaan Arsitektur

Perbedaan tersebut memberikan BERT kemampuan unik dalam memahami teks secara lebih menyeluruh, karena dapat menangkap ketergantungan antar kata dengan lebih baik. Sebaliknya, model dengan arsitektur satu arah hanya mempertimbangkan konteks dari satu arah, yang dapat membatasi pemahaman dalam beberapa tugas pemrosesan bahasa. Dengan pendekatan Bidirectional, BERT mampu mengenali hubungan dan konteks yang lebih kompleks, yang telah terbukti efektif dalam berbagai tugas pemrosesan bahasa alami dan telah merevolusi paradigma NLP secara keseluruhan (Ramdhan, 2024b).

Pada tahap pelatihan, terdapat dua proses utama yang dilakukan untuk memahami konteks dalam teks, yaitu Pre-Training dan Fine-Tuning. Pre-Training adalah tahap di mana BERT mempelajari bahasa input dan konteksnya menggunakan data dalam jumlah besar. Selama Pre-Training, BERT melakukan dua tugas utama secara bersamaan, yaitu Masked Language Modeling (MLM), yang berfokus pada prediksi kata yang hilang dalam kalimat, dan Next Sentence Prediction (NSP), yang memprediksi hubungan antara dua kalimat yang berurutan (Imaduddin, 2023).

Penelitian ini menggunakan data yang diperoleh dari platform Kaggle, yang terdiri dari 16.648 entri data. Dari jumlah tersebut, hanya dua kolom yang digunakan untuk analisis lebih lanjut, yaitu kolom "review" yang berisi data teks (*string*) dan kolom "sentimen" yang berisi data numerik *(integer*). Data yang telah dipilih dan disaring ini kemudian disimpan dalam format CSV dengan nama file *ecommerce-product-reviews-Sentiment.csv*. Data ini berisi ulasan produk dari berbagai pengguna serta penilaian sentimen mereka terhadap produk tersebut, dengan nilai 0 menunjukkan sentimen negatif dan nilai 1 menunjukkan sentimen positif. Sebagai contoh, beberapa entri data awal yang telah disimpan terlihat dalam tabel, di mana ulasan pengguna dan terjemahannya tercatat

bersama dengan nilai sentimen yang sesuai, menggambarkan tanggapan pengguna terhadap produk yang mereka beli. Tahapan alur penelitian sebagai berikut:



Gambar 2. Gambar Alur Penelitian

Penelitian ini terdiri dari beberapa tahapan yang sistematis. Pertama, pengambilan data dilakukan dengan mengambil 11.606 ulasan produk dari platform Kaggle. Data ini mencakup dua jenis sentimen, yaitu positif dan negatif, yang terdiri dari kolom "review" (isi ulasan), "sentimen" (label sentimen, 0 untuk negatif dan 1 untuk positif), dan "translate" (terjemahan ulasan dalam bahasa Inggris). Proses pengambilan data dilakukan menggunakan Python, dan data disimpan dalam format CSV untuk analisis selanjutnya. Setelah data terkumpul, tahap berikutnya adalah preprocessing, di mana data ulasan yang telah disimpan dalam format CSV dibersihkan dari karakter yang tidak berguna seperti tanda baca dan kata-kata yang tidak relevan. Proses ini dilakukan menggunakan Python untuk memastikan data siap dianalisis (Mei, 2024).

Selanjutnya, dilakukan sampling dan pelabelan data. Proses ini melibatkan pemilihan data secara acak dan pelabelan berdasarkan konteks dan kata-kata dalam ulasan untuk menentukan apakah sentimen yang terkandung bersifat positif atau negatif. Hasil dari pelabelan ini menghasilkan data dengan proporsi sentimen positif dan negatif yang seimbang dengan rasio 1:1. Setelah pelabelan, data dibagi menjadi dua set: data latih dan data uji. Data latih digunakan untuk melatih model, sementara data uji digunakan untuk

mengevaluasi kinerja model. Pembagian data dilakukan dengan proporsi 80% untuk data latih (9.284 data) dan 20% untuk data uji (2.322 data).

Tahap berikutnya adalah pemodelan dengan menggunakan INDOBERT, model berbasis BERT yang sudah dilatih dengan dataset bahasa Indonesia. Model yang digunakan adalah INDOBERTClassifier dengan konfigurasi optimizer Adam, learning rate 2e-5, fungsi loss CrossEntropyLoss, batch size 16, dan jumlah epoch sebanyak 5. Terakhir, evaluasi dilakukan dengan menggunakan metrik seperti akurasi, precision, recall, dan F1-score. Jika metrik evaluasi tidak memenuhi ekspektasi, dilakukan tuning terhadap hyperparameter model, seperti jumlah epoch dan ukuran batch (Perwira, 2025). Pada percobaan ini, model berhasil memberikan hasil yang optimal dengan metrik evaluasi yang memadai. Pada tahap Preprocessing, data yang diperoleh dari Kaggle melalui penelitian ini akan dipersiapkan untuk analisis lebih lanjut. Tahap ini terdiri dari beberapa proses untuk membersihkan dan mengolah data menjadi format yang lebih siap pakai, yang meliputi dua tahapan utama berikut:

Menghapus Karakter yang Tidak Memiliki Makna

Proses pertama adalah menghapus karakter-karakter yang tidak memberikan makna atau relevansi terhadap kalimat. Karakter seperti tanda baca, angka yang tidak penting, dan kata-kata yang tidak memiliki makna akan dihilangkan. Teknik yang digunakan dalam proses ini adalah *regular expression (regex),* yang memungkinkan pencarian dan penghapusan pola karakter tertentu dalam teks. Sebagai contoh, kalimat seperti "kondisi dus robek2, tolong di perhatikan lagi" akan diproses menjadi "kondisi dus robek tolong di perhatikan lagi" setelah karakter yang tidak penting (seperti angka) dihapus. Hasil dari tahap ini dapat dilihat pada tabel berikut:

Tabel 1. Contoh Hasil *Preprocessing*

KALIMAT	kondisi dus robek2, tolong di perhatikan lagi
HASIL	kondisi dus robek tolong di perhatikan lagi

Tokenisasi

Selanjutnya, pada tahap Tokenisasi, setiap kalimat akan dipisahkan menjadi kata-kata yang lebih kecil atau yang disebut token. Proses ini bertujuan untuk memecah teks menjadi unit-unit yang lebih kecil agar lebih mudah dianalisis. Tokenisasi dilakukan dengan menggunakan Natural Language Toolkit (NLTK), sebuah pustaka Python yang umum digunakan untuk memproses teks dalam Bahasa Indonesia. Hasil dari proses ini adalah token-token yang mewakili kata-kata dalam kalimat. Sebagai contoh, kalimat "Suara dan mic puas. Hanya saya pesan yg nyala dan sudah sy pastikan lg klik yg nyala tp yg Datang ga nyala" akan diproses menjadi token-token yang terpisah (Febryanti, 2024). Hasil tokenisasi untuk kalimat tersebut menunjukkan bahwa setiap kata atau unit penting dalam kalimat dipisahkan dan dikodekan menjadi ID tertentu, seperti yang terlihat pada tabel berikut:

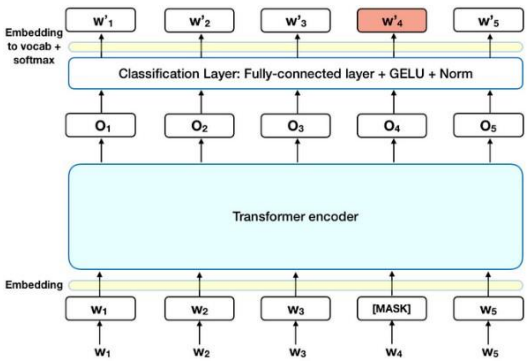
Tabel 2. Contoh Hasil Tokenisasi

KALIMAT	Suara dan mic puas. Hanya saya pesan yg nyala dan sudah sy pastikan lg klik yg nyala tp yg *Data*ng ga nyala
HASIL	*Review* 1: suara dan mic puas hanya saya pesan yg nyala dan sudah sy pastikan lg klik yg nyala tp yg *Data*ng ga nyala *Token* IDs: [2, 1699, 41, 3197, 4416, 344, 209, 2054, 741, 14906]...

Proses preprocessing ini bertujuan untuk menyiapkan data agar model dapat menganalisis teks dengan lebih efisien dan akurat dalam tahapan berikutnya.

Masked Language Modeling (MLM)

Pada tahap Masked Language Modeling (MLM), BERT dilatih untuk memahami konteks dan makna kata dalam sebuah kalimat dengan cara memprediksi kata yang hilang atau disamarkan dalam teks. Selama fase Pre-Training, sebagian token dalam teks input digantikan secara acak dengan token [MASK], dan model dilatih untuk memprediksi kata yang tepat berdasarkan konteks kata-kata di sekitarnya. Proses ini membantu BERT memahami hubungan antar kata dan konteksnya secara lebih mendalam, yang meningkatkan pemahaman semantik bahasa. Sekitar 15% dari token dalam teks dipilih untuk disamarkan, dengan rincian: 80% diganti dengan token [MASK], 10% diganti dengan token acak, dan 10% dibiarkan tidak berubah.



Gambar 3. Masked Language Modelling (Euna, 2020)

Setelah proses prediksi kata dilakukan melalui **Encoder**, lapisan klasifikasi digunakan untuk menghitung probabilitas setiap kata dalam vocabulary dengan menggunakan fungsi **softmax**. Hal ini memungkinkan BERT untuk memilih kata yang paling sesuai dalam konteks kalimat tersebut. Dengan cara ini, MLM berperan penting dalam meningkatkan kemampuan BERT untuk memahami konteks kalimat secara lebih menyeluruh, yang pada akhirnya memperbaiki kinerjanya dalam berbagai tugas pemrosesan bahasa alami (Sinapoy, 2023).

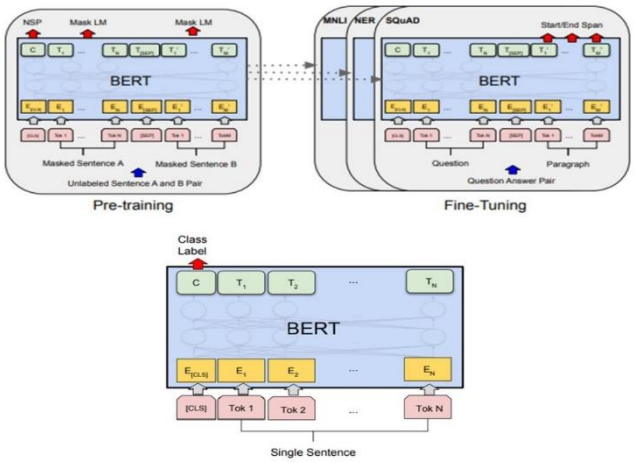
Next Sentence Prediction (NSP)

Pada tahap *Next Sentence Prediction (NSP)*, BERT dilatih untuk memprediksi apakah dua kalimat yang diberikan merupakan kalimat yang berurutan atau tidak. NSP bertujuan untuk membantu BERT memahami hubungan antar kalimat dan menangkap

ketergantungan jangka panjang antara kalimat-kalimat tersebut. Selama proses pelatihan, model diberikan pasangan kalimat sebagai input, di mana 50% dari pasangan tersebut adalah kalimat yang berurutan secara logis, sementara 50% lainnya adalah kalimat acak yang dipilih dari seluruh korpus (Hulliyah, 2022).

Untuk merepresentasikan input, BERT menggunakan tiga jenis embedding, yaitu:

1. Token Embeddings: Menggambarkan setiap kata atau token dalam teks.
2. Segment Embeddings: Menandai apakah token tersebut berasal dari kalimat pertama atau kedua, yang membantu BERT membedakan antara kedua kalimat tersebut.
3. Positional Embeddings: Menyediakan informasi tentang posisi setiap token dalam urutan teks, sehingga model dapat memahami urutan kata dalam kalimat.



Gambar 4. Proses Fine-Tuning BERT (Euna, 2020)

Kombinasi dari ketiga embedding ini memungkinkan BERT untuk memahami konteks yang lebih kaya dan kompleks, baik pada level kata maupun antar kalimat, yang memperkaya pemahamannya terhadap struktur teks secara keseluruhan.

Representasi Input dan Output INDOBERT

Proses representasi input dalam model INDOBERT terdiri dari beberapa tahap yang memastikan teks dapat diproses secara efektif oleh model. Berikut adalah langkah-langkah utama dalam representasi input tersebut:

1. Tokenisasi dengan WordPiece: Data teks akan dipecah menjadi token menggunakan metode WordPiece Tokenization. Setiap kata diperiksa apakah terdapat dalam vocabulary INDOBERT. Jika tidak, kata tersebut diubah menjadi token [UNK] atau dipisah menjadi sub-kata dengan sufiks yang dimulai dengan tanda "##". Ini memastikan model tetap dapat memproses kata-kata yang tidak ada dalam vocabulary.
2. Penyisipan Token Khusus: Model INDOBERT menyisipkan dua token khusus pada setiap kalimat, yaitu [CLS] di awal dan [SEP] di akhir kalimat. Token [CLS] mewakili seluruh kalimat, sedangkan [SEP] digunakan sebagai pemisah antara dua kalimat jika input berupa pasangan kalimat.
3. Penyeragaman Panjang Kalimat: Untuk memastikan bahwa semua kalimat memiliki panjang yang sama, token [PAD] ditambahkan hingga mencapai panjang maksimum

yang ditentukan. Token [PAD] tidak memiliki makna semantik dan hanya digunakan untuk padding.

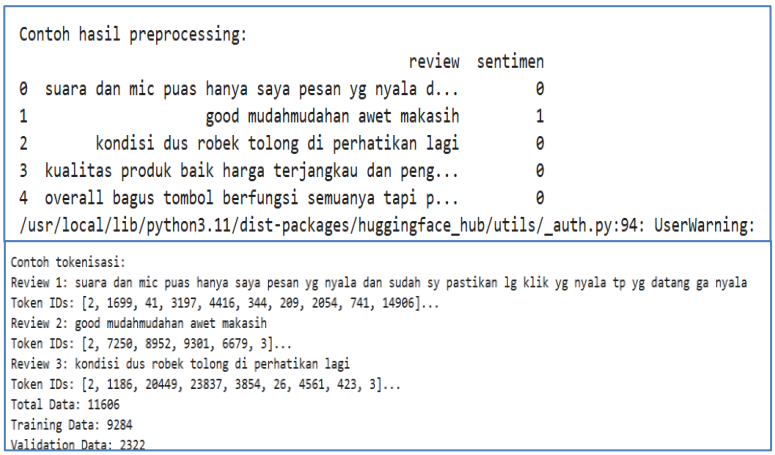
4. Konversi ke ID Token: Setiap token, termasuk token khusus, akan diubah menjadi ID numerik berdasarkan indeksnya dalam vocabulary. Model INDOBERT hanya menerima input dalam bentuk ID token.
5. Sentence Embedding: Setiap token selain [PAD] diberi nilai 1, sedangkan [PAD] diberi nilai 0. Ini membantu model mengenali token yang relevan dan mengabaikan [PAD] dalam proses pelatihan atau inferensi.
6. Positional Embedding: Setiap token juga diberikan nilai positional embedding yang menunjukkan posisi token dalam urutan kalimat. Ini memungkinkan model memahami urutan kata dalam kalimat dan konteksnya (Anderies, 2021).
7. Proses di Encoder: Setelah representasi input terbentuk, data tersebut diproses melalui tumpukan Encoder. Proses ini terdiri dari dua komponen utama: Self-attention untuk memahami hubungan antar token dan Feed-forward network yang memperkuat representasi token.
8. Output Vektor Hidden Layer: Setelah diproses melalui Encoder, setiap token menghasilkan vektor hidden layer dengan ukuran 768 dimensi, yang merepresentasikan informasi kontekstual dari setiap token dalam kalimat.
9. Prediksi Label (Klasifikasi): Dalam tugas klasifikasi, vektor dari token [CLS] akan digunakan sebagai representasi dari seluruh kalimat dan diteruskan ke lapisan klasifikasi untuk menghasilkan output akhir, seperti prediksi label.

Dengan tahapan-tahapan ini, model INDOBERT mampu memproses dan memahami hubungan antar kata dan kalimat secara mendalam, serta menghasilkan prediksi atau output yang lebih akurat.

Hasil dan Pembahasan

Preprocessing Data

Data yang digunakan dalam penelitian ini diperoleh dari platform Kaggle, yang mencakup 11.606 ulasan produk dengan sentimen positif dan negatif. Data ini terdiri dari beberapa kolom, antara lain kolom "review" yang berisi isi ulasan dari pengguna, kolom "sentimen" yang berisi label sentimen (0 untuk negatif dan 1 untuk positif), serta kolom "translate" yang memuat terjemahan ulasan dalam bahasa Inggris (Fernandez, 2022). Proses preprocessing dilakukan menggunakan Python dengan tujuan untuk menghilangkan karakter-karakter yang tidak relevan, seperti tanda baca, serta kata-kata atau kalimat yang tidak memiliki makna. Hasil dari proses preprocessing dapat dilihat pada gambar berikut:



Gambar 5. Hasil *Preprocessing*

Pelabelan Data

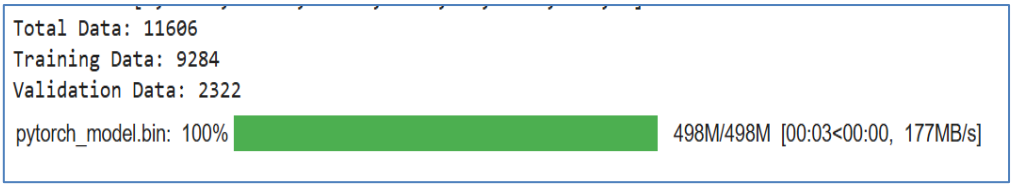
Setelah tahap pembersihan data, pelabelan dilakukan pada data yang dipilih secara acak. Pelabelan ini dilakukan dengan mempertimbangkan konteks dan kata-kata dalam setiap ulasan untuk menentukan apakah sentimen yang terkandung dalam ulasan tersebut bersifat positif atau negatif. Data yang telah dilabeli terdiri dari beberapa kolom, seperti kolom "review" yang berisi isi ulasan dari pengguna, dan kolom "sentimen" yang memuat label sentimen (0 untuk negatif dan 1 untuk positif). Contoh hasil pelabelan dapat dilihat pada tabel berikut:

Tabel 3. Hasil Pelabelan Data

Review	Sentimen
Suara dan mic puas. Hanya saya pesan yg nyala ...	0
Good. Mudah-mudahan awet. Makasih	1
Kondisi dus robek-robek, tolong di perhatikan lagi	0
Kualitas produk baik. Harga terjangkau dan pengiriman cepat	0
Overall bagus, tombol berfungsi semuanya, tapi suara tidak terlalu kencang	0

Pembagian Data Latih dan Uji

Setelah data diberi label, data tersebut dibagi menjadi dua bagian: Data latih dan Data uji. Data latih digunakan untuk melatih model, sedangkan Data uji digunakan untuk mengevaluasi kinerja model. Pembagian ini dilakukan dengan proporsi 80% untuk Data latih dan 20% untuk Data uji, yang terdiri dari 9.284 data latih dan 2.322 data uji. Data yang digunakan untuk pembagian ini adalah data teks yang telah melewati tahap preprocessing. Pembagian data ini bertujuan agar model dapat mengenali pola dan beradaptasi dengan data. Hasil pembagian data dapat dilihat pada gambar berikut:



Gambar 6. Hasil Pembagian Data

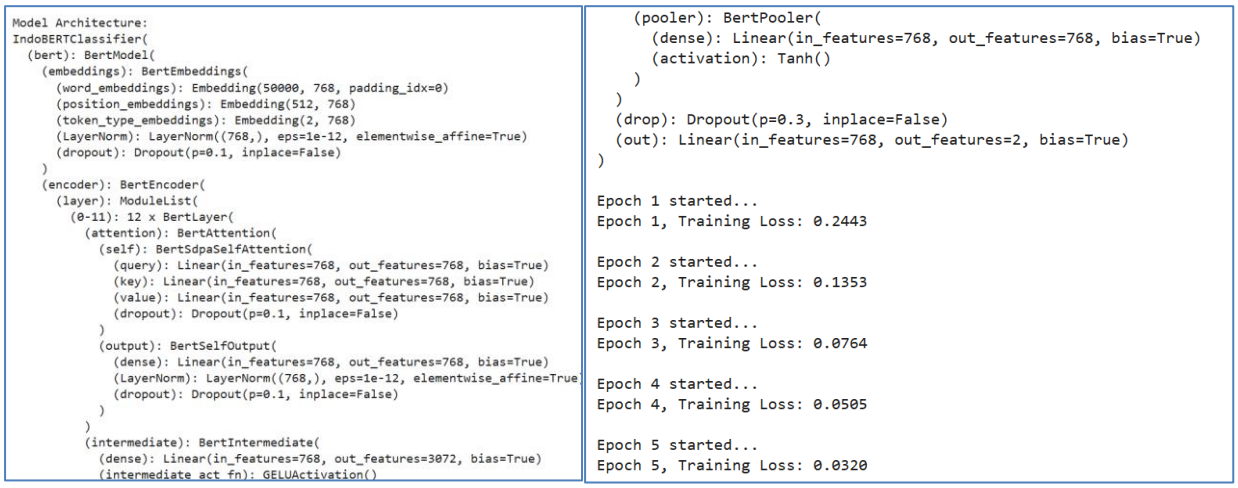
Pelatihan Model

Pelatihan model dilakukan menggunakan INDOBERT dengan konfigurasi yang telah ditentukan untuk mencapai prediksi sentimen yang lebih akurat. Model yang digunakan adalah INDOBERT (indobenchmark/INDOBERT-base-p2), dengan optimizer Adam, fungsi loss CrossEntropyLoss, batch size 16, dan jumlah epoch sebanyak 5. Proses pelatihan ini juga memanfaatkan perangkat GPU jika tersedia, atau CPU, dengan dropout 0.3 untuk mencegah overfitting (Yulianti, 2024).

Pelatihan model bertujuan untuk meminimalkan kesalahan prediksi dengan menyesuaikan bobot parameter model. Dalam pelatihan ini, terdapat beberapa konsep penting yang diterapkan, yaitu:

1. Transfer Learning: INDOBERT, sebagai model yang telah dilatih sebelumnya dengan data besar, hanya memerlukan fine-tuning pada dataset spesifik.
2. Backpropagation: Algoritma ini digunakan untuk memperbarui bobot model berdasarkan perbedaan antara prediksi dan label sebenarnya.
3. Optimizer Adam: Digunakan untuk mempercepat konvergensi dengan menggabungkan keunggulan metode momentum dan RMSprop.
4. CrossEntropyLoss: Fungsi loss yang digunakan untuk mengukur perbedaan antara distribusi prediksi model dan distribusi label asli.
5. Regularization: Teknik dropout diterapkan untuk mencegah overfitting dan meningkatkan kemampuan generalisasi model.

Proses pelatihan dimulai dengan inisialisasi model INDOBERT, yang kemudian dipindahkan ke perangkat yang tersedia. Optimizer dan fungsi loss diatur sesuai dengan konfigurasi yang telah ditentukan. Setelah itu, training loop dimulai, dengan langkah-langkah yang dilakukan setiap epoch, yaitu forward pass untuk mengonversi data ulasan menjadi vektor representasi, prediksi keluaran untuk menentukan sentimen, menghitung loss untuk membandingkan dengan label sentimen asli, dan melakukan backpropagation untuk memperbarui bobot model. Optimisasi dilakukan dengan memperbarui parameter model menggunakan optimizer Adam. Nilai loss dicetak setelah setiap epoch untuk memantau proses pembelajaran. Setelah pelatihan selesai, model disimpan untuk digunakan dalam evaluasi. Hasil dari pemodelan INDOBERT dan training loops dapat dilihat pada gambar berikut yang menggambarkan progres pelatihan model selama Hasil pelatihan menunjukkan bahwa nilai loss menurun pada setiap epoch, yang menandakan bahwa model berhasil belajar dari data dan mengalami perbaikan dalam prediksinya (Baskara, 2023).

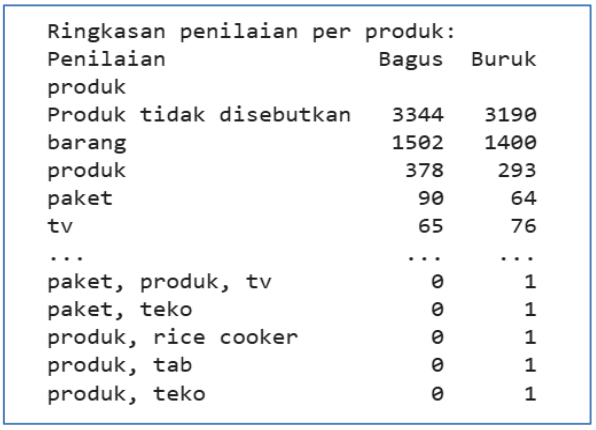


```
Model Architecture:
IndoBERTClassifier(
  (bert): BertModel(
    (embeddings): BertEmbeddings(
      (word_embeddings): Embedding(50000, 768, padding_idx=0)
      (position_embeddings): Embedding(512, 768)
      (token_type_embeddings): Embedding(2, 768)
      (LayerNorm): LayerNorm((768,), eps=1e-12, elementwise_affine=True)
      (dropout): Dropout(p=0.1, inplace=False)
    )
    (encoder): BertEncoder(
      (layer): ModuleList(
        (0-11): 12 x BertLayer(
          (attention): BertAttention(
            (self): BertSdpaSelfAttention(
              (query): Linear(in_features=768, out_features=768, bias=True)
              (key): Linear(in_features=768, out_features=768, bias=True)
              (value): Linear(in_features=768, out_features=768, bias=True)
              (dropout): Dropout(p=0.1, inplace=False)
            )
            (output): BertSelfOutput(
              (dense): Linear(in_features=768, out_features=768, bias=True)
              (LayerNorm): LayerNorm((768,), eps=1e-12, elementwise_affine=True)
              (dropout): Dropout(p=0.1, inplace=False)
            )
          )
          (intermediate): BertIntermediate(
            (dense): Linear(in_features=768, out_features=3072, bias=True)
            (intermediate_act_fn): GELUActivation()
          )
      (pooler): BertPooler(
        (dense): Linear(in_features=768, out_features=768, bias=True)
        (activation): Tanh()
      )
    )
    (drop): Dropout(p=0.3, inplace=False)
    (out): Linear(in_features=768, out_features=2, bias=True)
)

Epoch 1 started...
Epoch 1, Training Loss: 0.2443

Epoch 2 started...
Epoch 2, Training Loss: 0.1353

Epoch 3 started...
Epoch 3, Training Loss: 0.0764

Epoch 4 started...
Epoch 4, Training Loss: 0.0505

Epoch 5 started...
Epoch 5, Training Loss: 0.0320
```

Gambar 7. Hasil Pemodelan dan Training Loops Pemodelan

Ekstrasi Produk Dan Kategorisasi *Sentiment*

Hasil pengujian ekstraksi produk dan kategorisasi sentimen menunjukkan bahwa sistem berhasil mengenali produk yang disebutkan dalam ulasan dengan tingkat akurasi yang baik, serta dapat mengategorikan sentimen berdasarkan rating sesuai dengan logika yang telah ditentukan. Pengujian ini dilakukan pada dataset ecommerce-product-reviews-Sentiment.csv dan bertujuan untuk memastikan bahwa proses ekstraksi dan kategorisasi bekerja dengan efektif.



Ringkasan penilaian per produk:

Penilaian	Bagus	Buruk
produk		
Produk tidak disebutkan	3344	3190
barang	1502	1400
produk	378	293
paket	90	64
tv	65	76
...	...	...
paket, produk, tv	0	1
paket, teko	0	1
produk, rice cooker	0	1
produk, tab	0	1
produk, teko	0	1

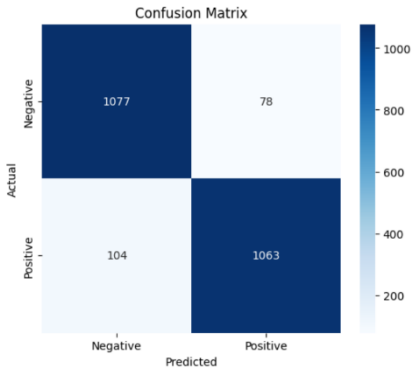
Gambar 8. Ekstrasi dan Kategorisasi Produk

Dari hasil analisis yang ditampilkan terlihat bahwa sebagian besar ulasan pelanggan tidak menyebutkan produk secara spesifik, dengan kategori "Produk tidak disebutkan" memiliki jumlah ulasan terbanyak, baik positif (3.488) maupun negatif (3.414). Hal ini menunjukkan bahwa banyak pelanggan memberikan review tanpa mencantumkan nama produk yang jelas. Selain itu, kategori "barang" juga mendapatkan banyak ulasan, dengan 1.529 ulasan positif dan 1.443 ulasan negatif, yang menunjukkan distribusi sentimen yang hampir seimbang. Kategori "produk" memiliki lebih banyak ulasan positif (385) dibandingkan ulasan negatif (309), yang mengindikasikan bahwa pengalaman pelanggan dengan produk tersebut cenderung lebih baik. Kategori seperti "paket" memiliki jumlah ulasan yang lebih sedikit, dengan 94 ulasan positif dan 72 ulasan negatif. Beberapa kombinasi produk, seperti "barang, produk" atau "produk, teko," juga menunjukkan jumlah

ulasan yang sangat terbatas, dengan mayoritas ulasan bernada negatif. Secara keseluruhan, distribusi sentimen dalam hampir semua kategori terlihat cukup seimbang, menandakan bahwa pengalaman pelanggan sangat bervariasi, tergantung pada jenis produk dan individu yang memberikan ulasan. Hal ini menunjukkan pentingnya mempertimbangkan variasi dalam ulasan pelanggan ketika menganalisis sentimen produk di platform e-commerce.

Evaluasi Model

Hasil evaluasi model klasifikasi menunjukkan kinerja yang cukup baik dengan akurasi mencapai 92%. Berdasarkan Confusion Matrix yang ditampilkan pada gambar berikut



Gambar 9. Hasil Confusion Matrix

Model berhasil mengklasifikasikan 1.077 sampel negatif dengan benar (True Negative - TN) dan 1.063 sampel positif dengan benar (True Positive - TP). Namun, terdapat beberapa kesalahan prediksi, yaitu 78 sampel yang seharusnya negatif tetapi diprediksi sebagai positif (False Positive - FP) dan 104 sampel yang seharusnya positif tetapi diprediksi sebagai negatif (False Negative - FN). Metrik evaluasi model menunjukkan nilai Precision sebesar 93%, yang berarti 93% dari semua prediksi positif benar-benar positif. Sementara itu, Recall model mencapai 91%, menunjukkan bahwa 91% dari seluruh data yang seharusnya positif berhasil diklasifikasikan dengan benar. Dengan demikian, model memiliki kemampuan yang cukup baik dalam mengidentifikasi data positif dan menghindari kesalahan klasifikasi. F1-score, yang merupakan gabungan dari precision dan recall, menghasilkan nilai 92%, menegaskan bahwa model memiliki keseimbangan yang baik antara ketepatan dan sensitivitas. Secara keseluruhan, evaluasi ini menunjukkan bahwa model klasifikasi yang dibangun memiliki performa yang sangat baik dalam melakukan analisis sentimen dengan tingkat akurasi dan presisi yang tinggi.

Pseudocode Penelitian

Penelitian ini dimulai dengan memuat dan memproses dataset untuk analisis sentimen, di mana teks dalam kolom 'review' dibersihkan dengan cara mengubahnya menjadi huruf kecil (lowercase), menghapus tanda baca, angka, dan ruang ekstra. Teks yang telah diproses kemudian di-tokenisasi menggunakan model tokenizer INDOBERT. Selanjutnya, dataset dibagi menjadi dua bagian: Data latih (80%) dan Data uji (20%), dan dimasukkan ke dalam DataLoader untuk memudahkan proses batching.

```
# 1. Preprocessing Function
def preprocess_text(text):
    # Lowercase
    text = text.lower()
    # Remove punctuation
    text = re.sub(r'[^\w\s]', '', text)
    # Remove numbers
    text = re.sub(r'\d+', '', text)
    # Remove extra spaces
    text = re.sub(r'\s+', ' ', text).strip()
    return text

# 2. Dataset Class
class ReviewDataset(Dataset):
    def __init__(self, reviews, sentiments, tokenizer, max_len):
        self.reviews = reviews
        self.sentiments = sentiments
        self.tokenizer = tokenizer
        self.max_len = max_len

    def __len__(self):
        return len(self.reviews)

    def __getitem__(self, idx):
        review = self.reviews[idx]
        sentiment = self.sentiments[idx]
        encoding = self.tokenizer.encode_plus(
            review,
            max_length=self.max_len,
            add_special_tokens=True,
            padding="max_length",
            truncation=True,
            return_tensors="pt",
        )
        return {
            'input_ids': encoding['input_ids'].flatten(),
            'attention_mask': encoding['attention_mask'].flatten(),
            'labels': torch.tensor(sentiment, dtype=torch.long)
        }

# 3. Main Script
def main():
    # Load dataset
    df = pd.read_csv('/content/ecommerce-product-reviews-sentiment.csv')  # Dataset dengan kolom 'review' dan 'sentimen'
    print(f"Dataset berhasil dimuat. Kolom tersedia:\n{df.columns}")

    df['review'] = df['review'].apply(preprocess_text)  # Apply preprocessing
    print("\nContoh hasil preprocessing:")
    print(df[['review', 'sentimen']].head())

    reviews = df['review'].values
    sentiments = df['sentimen'].values  # nilai sentimen berupa angka (0/1)
```

Gambar 10. Kodingan Sistem

Model klasifikasi dibangun menggunakan pre-trained INDOBERT, yang dilengkapi dengan lapisan dropout dan linear layer untuk klasifikasi sentimen. Model ini dilatih selama beberapa epoch, di mana Loss dihitung dan dilakukan optimasi menggunakan backpropagation. Setelah pelatihan selesai, model diuji menggunakan Data uji untuk mengevaluasi performa model. Evaluasi dilakukan dengan mengukur akurasi dan menghasilkan laporan klasifikasi yang mencakup metrik precision, recall, dan F1-score. Hasil evaluasi ini kemudian ditampilkan untuk memberikan gambaran mengenai kinerja model dalam menganalisis sentimen.

Simpulan

Berdasarkan hasil penelitian yang telah dilakukan, model INDOBERT Classifier menunjukkan kinerja yang sangat baik dalam tugas klasifikasi sentimen ulasan pelanggan. Proses preprocessing telah dilaksanakan dengan benar, termasuk tokenisasi menggunakan INDOBERT, yang sangat membantu dalam pemrosesan teks. Dataset yang digunakan terdiri dari 11.606 sampel, dengan 80% untuk pelatihan dan 20% untuk validasi. Selama pelatihan, nilai Loss menunjukkan penurunan yang signifikan dari epoch pertama hingga kelima, yang menandakan bahwa model dapat belajar dengan efektif.

Hasil evaluasi model menunjukkan akurasi sebesar 92,16%, dengan nilai precision 93,16%, recall 91,09%, dan F1-score 92,11%. Angka-angka ini membuktikan bahwa model dapat mengklasifikasikan sentimen dengan akurat dan seimbang dalam memprediksi sentimen positif dan negatif. Dengan demikian, model ini memiliki potensi besar untuk digunakan dalam analisis sentimen ulasan dalam bahasa Indonesia.

Selain klasifikasi sentimen, penelitian ini juga mencakup ekstraksi informasi produk dari ulasan pelanggan. Hasil analisis menunjukkan bahwa sebagian besar ulasan tidak secara eksplisit menyebutkan produk yang dibeli. Dari 100 sampel ulasan yang dianalisis, mayoritas tidak mencantumkan nama produk secara langsung. Ringkasan penilaian juga menunjukkan bahwa beberapa produk memiliki lebih banyak ulasan negatif dibandingkan positif, yang dapat mengindikasikan adanya masalah dengan kualitas atau layanan produk tersebut. Selain itu, beberapa ulasan mengacu pada lebih dari satu produk, yang menyebabkan variasi dalam pengelompokan data.

Daftar Pustaka

Anderies. (2021). Finetunning IndoBERT to Understand Indonesian Stock Trader Slang Language. *Proceedings of 2021 1st International Conference on Computer Science and Artificial Intelligence, ICCSAI 2021*, 42–46. https://doi.org/10.1109/ICCSAI53272.2021.9609746

Baskara, A. R. (2023). Classification of User Reviews for Software Maintenance in Indonesian Language Using IndoBERT-BiLSTM (Case Study: MyPertamina). *2023 8th International Conference on Informatics and Computing, ICIC 2023*. https://doi.org/10.1109/ICIC60109.2023.10381946

Computer, J. H. (2023). Sentiment Analysis of Skincare Product Reviews in Indonesian Language using IndoBERT and LSTM. *Proceeding - IEEE 9th Information Technology International Seminar, ITIS 2023*. https://doi.org/10.1109/ITIS59651.2023.10420222

Darmawan, M. A. (2023). Experiments on IndoBERT Implementation for Detecting Multi-Label Hate Speech with Data Resampling through Synonym Replacement Method. *8th International Conference on Recent Advances and Innovations in Engineering: Empowering Computing, Analytics, and Engineering Through Digital Innovation, ICRAIE 2023*. https://doi.org/10.1109/ICRAIE59459.2023.10468099

Euna, S. J. (2020, November 7). BERT: Bidirectional Encoder Representations from Transformer. *Sophie Euna Jang*. https://sophieeunajang.wordpress.com/2020/11/07/bert/

Garamba, Y., & Nur, A. (2024). Analisis Kepuasan Pelanggan pada E-Commerce. *Jurnal Ilmu Komputer dan Informatika| E-ISSN: 3063-9026*, *1*(2), 27–37.

Febryanti, D. (2024). Customer Sentiment Analysis of Local Skincare Reviews Using IndoBERT and Graph Neural Networks. *2024 International Conference on Intelligent Cybernetics Technology and Applications, ICICyTA 2024*, 320–325. https://doi.org/10.1109/ICICYTA64807.2024.10912986

Fernandez, E. (2022). Improving IndoBERT for Sentiment Analysis on Indonesian Stock Trader Slang Language. *Proceedings of the 2022 IEEE International Conference on Internet of Things and Intelligence Systems, IoTaIS 2022*, 240–244. https://doi.org/10.1109/IoTaIS56727.2022.9975975

Hulliyah, K. (2022). Development of A Chatbot for the Online Application Telegram Chat with An Approach to the Emotion Classification Text Using the Indobert-Lite Method. *2022 4th International Conference on Cybernetics and Intelligent System, ICORIS 2022*. https://doi.org/10.1109/ICORIS56080.2022.10031483

Imaduddin, H. (2023). Sentiment Analysis in Indonesian Healthcare Applications using IndoBERT Approach. *International Journal of Advanced Computer Science and Applications*, *14*(8), 113–117. https://doi.org/10.14569/IJACSA.2023.0140813

Kusuma, I. H., & Cahyono, N. (2023). Analisis Sentimen Masyarakat Terhadap Penggunaan E-Commerce Menggunakan Algoritma K-Nearest Neighbor. *Jurnal Informatika: Jurnal Pengembangan IT*, *8*(3), 302–307.

Mei, N. C. (2024). Multi-Label Aspect-Sentiment Classification on Indonesian Cosmetic Product Reviews with IndoBERT Model. *International Journal of Advanced Computer Science and Applications*, *15*(11), 712–720. https://doi.org/10.14569/IJACSA.2024.0151168

Nasution, D. (2024). Strategi Penjualan E-Commerce Untuk Meningkatkan Penjualan Handphone Di Era Digital. *Jurnal Teknologi Ilmu Komputer*, *3*(1), 155–164.

Perwira, R. I. (2025). Domain-Specific Fine-Tuning of IndoBERT for Aspect-Based Sentiment Analysis in Indonesian Travel User-Generated Content. *Journal of Information Systems Engineering and Business Intelligence*, *11*(1), 30–40. https://doi.org/10.20473/jisebi.11.1.30-40

Priambodo, W., & Zuliarso, E. (2024). Combination K-Means And Lstm For Social Media Black Campaign Detection Of Indonesia Presidential Candidates 2024. *Jurnal Teknik Informatika (JUTIF)*, *5*(2), 539–550.

Ramdhan, H. M. (2024a). Sentiment Analysis of Beauty Product Reviews Using the IndoBERT Method and Naive Bayes Classification. *2024 12th International Conference on Information and Communication Technology, ICoICT 2024*, 397–404. https://doi.org/10.1109/ICoICT61617.2024.10698198

Ramdhan, H. M. (2024b). Sentiment Analysis of Beauty Product Reviews Using the IndoBERT Method and Naive Bayes Classification. *2024 12th International Conference on Information and Communication Technology, ICoICT 2024*, 397–404. https://doi.org/10.1109/ICoICT61617.2024.10698198

Sinapoy, M. I. K. (2023). Comparison of LSTM and IndoBERT Method in Identifying Hoax on Twitter. *Jurnal RESTI*, *7*(3), 657–662. https://doi.org/10.29207/resti.v7i3.4830

We Are Social. (2024). *Laporan Digital 2024*. https://wearesocial-com.translate.goog/id/blog/2024/01/digital-2024/?_x_tr_sl=en&_x_tr_tl=id&_x_tr_hl=id&_x_tr_pto=tc

Yulianti, E. (2024). ABSA of Indonesian customer reviews using IndoBERT: single-sentence and sentence-pair classification approaches. *Bulletin of Electrical Engineering and Informatics*, *13*(5), 3579–3589. https://doi.org/10.11591/eei.v13i5.8032