



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN
TEKNOLOGI

UNIVERSITAS SUMATERA UTARA
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

PROGRAM STUDI S1 TEKNOLOGI INFORMASI

Jalan Alumni No. 3 Gedung C, Kampus USU Padang Bulan, Medan 20155
Telepon/Fax: 061-8210077 | Email: tek.informasi@usu.ac.id | Laman: http://it.usu.ac.id

FORM PENGAJUAN JUDUL



Nama : Gideon Togap Hata Oloan Manurung

NIM : 211402147

Judul diajukan oleh* : ☐ Dosen
☒ Mahasiswa

Bidang Ilmu (tuliskan dua bidang) :

1. Data Science and Intelligent System Computer
2. Graphics and Vision

Uji Kelayakan Judul** : ☐ Diterima ☐ Ditolak

Hasil Uji Kelayakan Judul : \

Calon Dosen Pembimbing I: Sarah Purnamawati, S.T., M.Sc.

(Jika judul dari dosen maka dosen tersebut berhak menjadi pembimbing I)

Calon Dosen Pembimbing II: Ade Sarah Huzaifah, S.Kom., M.Kom

Paraf Calon Dosen Pembimbing I

Medan, Maret 2025

Ka. Laboratorium Penelitian,

* Centang salah satu atau keduanya

* Pilih salah satu

(Fanindia Purnamasari, S.TI., M.IT)

NIP. 198908172019032023



RINGKASAN JUDUL YANG DIAJUKAN

Judul / Topik Skripsi	Implementasi IndoBERT dan Transformer Encoder untuk Deteksi Ulasan Palsu pada E-Commerce Tokopedia Berbasis Analisis Tekstual dan Perilaku Reviewer.
Latar Belakang dan Penelitian Terdahulu	<p>Latar Belakang</p> <p>Dalam beberapa tahun terakhir, sektor <i>e-commerce</i> di Indonesia telah mengalami pertumbuhan yang pesat. Faktor utama yang mendorong pertumbuhan ini adalah meningkatnya penetrasi internet, perubahan pola konsumsi masyarakat, serta kemudahan akses terhadap layanan digital (Berijalan, 2025). Selain itu, strategi pemasaran seperti pemberian diskon dan program gratis ongkir juga berperan signifikan dalam meningkatkan loyalitas pelanggan (Manalu, et al., 2025). Platform <i>e-commerce</i> kini tidak hanya menjadi tempat transaksi jual beli, tetapi juga menghadirkan fitur inovatif seperti live shopping, sistem rekomendasi berbasis kecerdasan buatan (AI), serta integrasi dengan layanan logistik yang lebih efisien. Transformasi ini tidak hanya menguntungkan konsumen, tetapi juga membuka peluang bagi Usaha Mikro, Kecil, dan Menengah (UMKM) untuk menjangkau pasar yang lebih luas melalui ekosistem <i>e-commerce</i> yang mendukung pemasaran digital, transaksi daring, dan distribusi yang lebih efisien.</p> <p>Shopee, Tokopedia, dan TikTok Shop merupakan tiga platform <i>e-commerce</i> yang paling banyak digunakan di Indonesia, dengan jumlah pengguna dan transaksi yang terus meningkat setiap tahunnya. Data dari Databoks (2024a) menunjukkan bahwa Shopee menjadi platform dengan jumlah kunjungan tertinggi, mencapai 237 juta kunjungan per bulan pada September 2023, meningkat sekitar 38% dibandingkan awal tahun. Selain itu, pada tahun 2022, Shopee membukukan Gross Merchandise Value (GMV) sebesar Rp718 triliun, menjadikannya salah satu platform dengan nilai transaksi terbesar di Indonesia (Databoks, 2024c). Sementara itu, Tokopedia, yang telah berintegrasi dengan TikTok Shop, mencatat 88,9 juta kunjungan bulanan pada bulan yang sama. TikTok Shop, sebagai platform social commerce, memiliki 125 juta pengguna aktif bulanan di Indonesia setelah merger dengan Tokopedia, dengan total 21 juta penjual yang beroperasi di dalamnya (Databoks, 2024c).</p> <p>Pertumbuhan pesat platform <i>e-commerce</i> sejalan dengan perubahan perilaku konsumen dalam berbelanja online. Banyaknya pilihan di <i>marketplace</i> digital, konsumen menjadi lebih selektif, dan ulasan pembeli menjadi faktor utama dalam mengambil keputusan pembelian. Data SOCO (2024) menunjukkan bahwa 77% konsumen Indonesia membaca ulasan sebelum membeli. Namun, tingginya pengaruh ulasan dimanfaatkan oleh oknum penjual yang membayar pihak tertentu untuk membuat ulasan palsu guna meningkatkan rating produk mereka atau menjatuhkan pesaing. Di Indonesia, terdapat banyak layanan berbayar yang menawarkan jasa ulasan palsu, layanan ini tersedia dengan harga bervariasi, mulai dari</p>



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN
TEKNOLOGI

UNIVERSITAS SUMATERA UTARA
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

PROGRAM STUDI S1 TEKNOLOGI INFORMASI

Jalan Alumni No. 3 Gedung C, Kampus USU Padang Bulan, Medan 20155
Telepon/Fax: 061-8210077 | Email: tek.informasi@usu.ac.id | Laman: <http://it.usu.ac.id>

Rp25.000 untuk 5 produk hingga Rp66.000 untuk 15 produk (Septiani, 2023). Prosesnya melibatkan komunikasi, kesepakatan konten ulasan, pembayaran, hingga unggahan ulasan palsu oleh penyedia jasa.

Ulasan palsu atau fiktif dapat dihasilkan melalui dua metode utama. Metode pertama melibatkan perekrutan individu untuk menulis ulasan berdasarkan skrip tertentu. Dalam kasus ini, para penulis ulasan sering kali tidak pernah melihat produk aslinya, tetapi tetap memberikan ulasan berbayar. Sementara itu, metode kedua menggunakan teknologi berbasis komputer, seperti otomatisasi teks, untuk menghasilkan ulasan palsu secara otomatis (Mohawes et al., 2021). Secara konvensional, jumlah ulasan berbayar yang ditulis oleh manusia terbatas, dan pemilik perusahaan hanya dapat mempekerjakan sejumlah kecil penyebar spam untuk memposting ulasan palsu. Namun, kemajuan pesat dalam pembelajaran mesin (*machine learning*) dan pemrosesan bahasa alami (*natural language processing* atau NLP) telah memungkinkan pengembangan algoritma canggih yang dapat menghasilkan ulasan palsu dalam jumlah tak terbatas dengan kecepatan jauh lebih tinggi dibandingkan manusia (Yin et al., 2021).

Ulasan palsu telah menjadi masalah serius di era digital. Keberadaannya tidak hanya mengancam keaslian lingkungan belanja online, tetapi juga menyesatkan pengguna serta berdampak negatif terhadap reputasi penjual dan platform *e-commerce* (Sun et al., 2024). Selain itu, ulasan palsu dapat merusak kepercayaan konsumen terhadap sistem rating dan rekomendasi produk, menyebabkan skeptisisme yang lebih tinggi dalam proses pengambilan keputusan pembelian (Peña-García et al., 2024). Dampak lainnya adalah terdistorsinya peringkat produk akibat manipulasi algoritma pencarian di platform *e-commerce*, yang dapat secara tidak adil menguntungkan atau merugikan suatu bisnis (Golrezaei et al., 2020). Secara ekonomi, ulasan palsu juga dapat menyebabkan penurunan pendapatan bagi perusahaan, di mana penelitian menunjukkan bahwa penurunan satu bintang dalam peringkat suatu merek dapat mengakibatkan penurunan pendapatan sebesar 5-9% (TUFAIL et al., 2022).

Seriusnya masalah ulasan palsu di *e-commerce* menuntut adanya sistem deteksi yang efektif untuk melindungi pelanggan dari pemborosan uang dan waktu. Penelitian terkait deteksi ulasan palsu telah berkembang pesat dengan berbagai pendekatan, termasuk analisis berbasis *text features*, analisis perilaku pengulas (*reviewer behavioral*), serta kombinasi keduanya dengan teknik deep learning. Deteksi ulasan palsu berbasis *text features* telah banyak dilakukan dalam berbagai penelitian. (Shan et al., 2021) mengusulkan metode analisis inkonsistensi ulasan, yang mengintegrasikan ketidaksesuaian antara rating dan sentimen, inkonsistensi bahasa, serta fitur sentimen untuk mendeteksi ulasan palsu. Studi ini menunjukkan bahwa inkonsistensi linguistik lebih sering ditemukan dalam ulasan palsu dibandingkan ulasan asli, sehingga meningkatkan akurasi deteksi. (Goyal et al., 2023) menggunakan XGBoost dan Extra Trees Classifier, yang mampu mendeteksi ulasan palsu dengan akurasi 96% pada dataset Yelp. (Nguyen, 2023) menerapkan Support Vector Classifier (SVC) dengan TF-IDF, yang menunjukkan keandalan tinggi dalam membedakan ulasan palsu dan asli. Studi ini menyoroti efektivitas term weighting dan pendeteksian opinion spamming dalam mengidentifikasi ulasan yang tidak autentik.



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN
TEKNOLOGI

UNIVERSITAS SUMATERA UTARA
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

PROGRAM STUDI S1 TEKNOLOGI INFORMASI

Jalan Alumni No. 3 Gedung C, Kampus USU Padang Bulan, Medan 20155
Telepon/Fax: 061-8210077 | Email: tek.informasi@usu.ac.id | Laman: <http://it.usu.ac.id>

Selain analisis berbasis *text features*, beberapa peneliti juga memanfaatkan pendekatan berbasis *graph* dan jaringan untuk menganalisis metadata serta pola perilaku pengguna. Yu et al. (2021) menggunakan Graph Neural Networks (GNNs) yang dikombinasikan dengan algoritma HITS untuk mengidentifikasi pola *co-review* antar pengguna yang mencurigakan, dengan tingkat akurasi mencapai 93,5%. Sementara itu, He et al. (2022) menerapkan model *network structure detection* untuk menganalisis histori ulasan serta anomali dalam timestamp, yang menghasilkan tingkat akurasi sebesar 89,5%. Selain itu, Manaskasemsak et al. (2023) menerapkan pendekatan Graph Partitioning untuk mendeteksi jaringan reviewer palsu menggunakan metode BeGPX, yang mencapai akurasi 92,1% dalam mendeteksi hubungan mencurigakan antar pengguna.

Pendekatan berbasis *deep learning* juga telah digunakan untuk meningkatkan akurasi deteksi ulasan palsu. Seiring dengan perkembangan teknologi, metode berbasis *neural network* secara bertahap menjadi standar dalam deteksi ulasan palsu, dengan berbagai model *sequence-based* seperti Recurrent Neural Network (RNN), Long Short-Term Memory (LSTM), Bidirectional LSTM (BiLSTM), serta model Text-CNN yang banyak digunakan dalam pemrosesan konten teks. Salah satu penelitian yang dilakukan oleh Qayyum et al. (2023) mengembangkan FRD-LSTM (Fake Review Detection menggunakan BiLSTM) dengan Deep Contextualized Word Representation (DCWR), yang berhasil mencapai akurasi 97,2%, menjadikannya salah satu metode dengan performa terbaik di bidang ini.

Namun, model berbasis *sequence-based* memiliki beberapa keterbatasan, terutama dalam menangani ketergantungan jangka panjang antar kata serta dalam aspek paralelisasi selama proses pelatihan. Oleh karena itu, dalam beberapa tahun terakhir, penelitian telah banyak mengadopsi model BERT (Bidirectional Encoder Representations from Transformers) untuk mengatasi permasalahan tersebut. Model BERT menggunakan encoder berbasis transformer untuk mengekstraksi word embeddings yang bersifat kontekstual dengan representasi yang lebih akurat. Selain itu, mekanisme self-attention dalam BERT memungkinkan proses pelatihan yang lebih efisien secara paralel dibandingkan dengan pendekatan berbasis RNN atau LSTM.

Salah satu studi yang dilakukan oleh Alsaad & Joshi (2024) membandingkan berbagai model transformer dan menemukan bahwa RoBERTa memiliki performa lebih unggul dibandingkan dengan model lain seperti BERT dan DistilBERT, dengan akurasi mencapai 97,1% dalam klasifikasi ulasan palsu. Namun, penelitian tersebut tidak mempertimbangkan review behavior sebagai faktor yang dapat digunakan dalam deteksi ulasan palsu. Untuk mengatasi keterbatasan ini, Sun et al. (2024) mengusulkan metode yang menggabungkan analisis teks dan perilaku reviewer guna meningkatkan akurasi deteksi ulasan palsu. Studi tersebut menunjukkan bahwa beberapa faktor perilaku seperti review burst (jumlah ulasan yang tinggi dalam waktu singkat), rating inconsistency, serta pola pengulangan komentar positif/negatif merupakan indikator kuat yang dapat digunakan dalam mendeteksi ulasan palsu. Metode yang digunakan dalam



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN
TEKNOLOGI

UNIVERSITAS SUMATERA UTARA
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

PROGRAM STUDI S1 TEKNOLOGI INFORMASI

Jalan Alumni No. 3 Gedung C, Kampus USU Padang Bulan, Medan 20155
Telepon/Fax: 061-8210077 | Email: tek.informasi@usu.ac.id | Laman: <http://it.usu.ac.id>

penelitian tersebut adalah model Fake Review Detection dengan pendekatan pengolahan teks menggunakan BERT dan pengolahan review behavior menggunakan Transformer Encoder.

Meskipun berbagai penelitian telah berhasil mengembangkan metode deteksi ulasan palsu, pendekatan berbasis machine learning dan deep learning tidak cukup hanya mengandalkan textual features untuk mendeteksi apakah sebuah ulasan bersifat manipulatif atau tidak. Penting untuk juga menganalisis bagaimana perilaku reviewer dan merchant dalam sistem e-commerce. Ulasan palsu sering kali dibuat oleh reviewer yang memiliki pola aktivitas mencurigakan, seperti memberikan banyak ulasan dalam waktu singkat, menggunakan kata-kata hiperbolik, serta memiliki hubungan erat dengan merchant tertentu. Beberapa studi juga menunjukkan bahwa pola lain seperti harga produk, durasi produk di pasaran, serta interaksi pengguna dengan sistem e-commerce dapat berkontribusi terhadap deteksi ulasan palsu.

Dalam penelitian ini, dataset yang digunakan berbasis bahasa Indonesia, sehingga diperlukan pendekatan yang dapat menangkap struktur linguistik dan pola perilaku pengguna dalam ekosistem e-commerce lokal. Oleh karena itu, penelitian ini mengusulkan pendekatan yang mengombinasikan Deep Learning dengan IndoBERT serta Transformer Encoder untuk mendeteksi ulasan palsu dalam dataset berbahasa Indonesia dari platform Tokopedia. Model ini dirancang untuk tidak hanya mendeteksi ulasan palsu berdasarkan teks, tetapi juga untuk menganalisis perilaku reviewer, termasuk hubungan dengan merchant, serta pola interaksi dalam jaringan ulasan e-commerce.

Sebagai pembeda dari penelitian sebelumnya, penelitian ini juga menerapkan Aspect-Based Sentiment Analysis (ABSA) untuk memahami opini pengguna terhadap aspek-aspek spesifik dari produk. Dengan ABSA, sistem dapat mengidentifikasi ketidaksesuaian antara sentimen dalam teks ulasan dan rating yang diberikan, yang merupakan indikator kuat dalam mendeteksi ulasan palsu yang dibuat secara manipulatif. Kombinasi dari pendekatan berbasis IndoBERT, Transformer Encoder, dan ABSA diharapkan dapat menghasilkan model deteksi ulasan palsu yang lebih akurat dan komprehensif dibandingkan metode sebelumnya.



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN
TEKNOLOGI

UNIVERSITAS SUMATERA UTARA
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

PROGRAM STUDI S1 TEKNOLOGI INFORMASI

Jalan Alumni No. 3 Gedung C, Kampus USU Padang Bulan, Medan 20155
Telepon/Fax: 061-8210077 | Email: tek.informasi@usu.ac.id | Laman: <http://it.usu.ac.id>

Penelitian Terdahulu

No.	Penulis	Judul	Tahun
1	Sun, P., Bi, W., Zhang, Y., Wang, Q., Kou, F., Lu, T., & Chen, J.	Review Detection Model Based on Comment Content and Review Behavior.	2024
2	Shan, G., Zhou, L., & Zhang, D.	From conflicts and confusion to doubts: Examining review inconsistency for fake review detection	2021
3	Nguyen, L. N.	An Empirical Study on Fake Review Detection	2023
4	Yu, S., Ren, J., Li, S., Naseriparsa, M., & Xia, F.	Graph learning for fake review detection	2021
5	He, S., Hollenbeck, B., Overgoor, G., Proserpio, D., & Tosyali, A.	Detecting fake-review buyers using network structure: Direct evidence from Amazon	2022



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN
TEKNOLOGI

UNIVERSITAS SUMATERA UTARA
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

PROGRAM STUDI S1 TEKNOLOGI INFORMASI

Jalan Alumni No. 3 Gedung C, Kampus USU Padang Bulan, Medan 20155
Telepon/Fax: 061-8210077 | Email: tek.informasi@usu.ac.id | Laman: <http://it.usu.ac.id>

	6	Manaskasemsak, B., Tantisuwankul, J., & Rungsawang, A.	Fake review and reviewer detection through behavioral graph partitioning integrating deep neural network.	2021
	7	Goyal, N., & Goyal, D	Fake Review Detection Using Ensemble Techniques by the Fusion of Chronology, Aspect and Sentiment of Reviews and Oversampling by Smote	2023
	8	Alsaad, M. M. B., & Joshi, H.	Transformer-Based Language Deep Learning Detection of Fake Reviews on Online Products	2024
	9	Yin, C., Cuan, H., Zhu, Y., & Yin, Z.	Improved Fake Reviews Detection Model Based on Vertical Ensemble Tri-Training and Active Learning	2021
Rumusan Masalah	Perkembangan pesat industri <i>e-commerce</i> di Indonesia telah mengubah pola perilaku konsumen dalam mengambil keputusan pembelian. Salah satu faktor yang berperan penting dalam proses ini adalah ulasan produk yang ditulis oleh pengguna lain, yang menjadi sumber informasi utama bagi calon pembeli. Namun, meningkatnya ketergantungan terhadap ulasan ini juga diikuti dengan tantangan baru, yaitu maraknya ulasan palsu yang dibuat untuk memanipulasi persepsi konsumen. Ulasan palsu sering kali dibuat untuk meningkatkan popularitas suatu produk secara tidak wajar atau menjatuhkan reputasi pesaing, sehingga menyebabkan ketidakseimbangan dalam persaingan pasar. Keberadaan ulasan yang tidak autentik ini tidak hanya merugikan konsumen, tetapi juga menurunkan kepercayaan terhadap platform <i>e-commerce</i> itu sendiri. Oleh karena itu, dibutuhkan sistem yang dapat secara efektif mendeteksi ulasan palsu dengan mempertimbangkan pola bahasa yang digunakan serta perilaku <i>reviewer</i> dalam memberikan ulasan.			



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN TEKNOLOGI

UNIVERSITAS SUMATERA UTARA
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

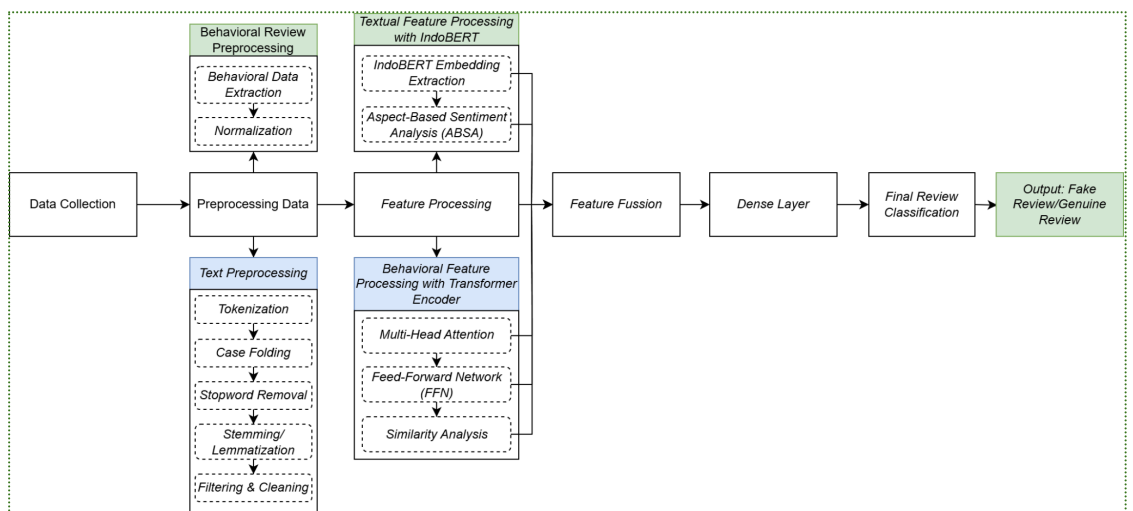
PROGRAM STUDI S1 TEKNOLOGI INFORMASI

Jalan Alumni No. 3 Gedung C, Kampus USU Padang Bulan, Medan 20155
Telepon/Fax: 061-8210077 | Email: tek.informasi@usu.ac.id | Laman: http://it.usu.ac.id

Dalam upaya mengatasi masalah ini, berbagai pendekatan telah dikembangkan, mulai dari analisis berbasis teks hingga analisis perilaku *reviewer* (*reviewer behavioral analysis*). Meskipun pendekatan berbasis teks mampu mengidentifikasi pola linguistik dalam ulasan palsu, metode ini masih memiliki keterbatasan dalam menangkap pola interaksi antar-pengulas dan penjual yang dapat mengindikasikan adanya manipulasi ulasan. Oleh karena itu, teknologi berbasis *deep learning*. Penelitian ini bertujuan untuk mengembangkan model deteksi ulasan palsu dengan menggabungkan Indo-BERT untuk analisis *text features* dan Transformer Encoder untuk analisis pola perilaku *reviewer*. Dengan pendekatan ini, sistem yang diusulkan tidak hanya dapat mendeteksi ulasan palsu berdasarkan isi teks, tetapi juga mampu menganalisis hubungan dan pola aktivitas *reviewer* dalam ekosistem *e-commerce*, sehingga memberikan solusi yang lebih akurat dalam menangani permasalahan ulasan manipulatif di platform seperti Tokopedia.

Metodologi

Dalam mendeteksi ulasan palsu pada *e-commerce*, peneliti mengombinasikan dua pendekatan utama: analisis fitur teks menggunakan Indo-RoBERTa dan analisis perilaku pengulas menggunakan BeGP (Behavioral Graph Partitioning). Kombinasi ini memungkinkan sistem untuk mengidentifikasi ulasan palsu berdasarkan pola bahasa serta hubungan dan aktivitas pengguna. Proses deteksi ini mencakup tahapan sebagai berikut:



1. Data Collection

Pada tahap ini dilakukan pengumpulan data ulasan yang diambil dari e-commerce Tokopedia. Fitur data yang dikumpulkan meliputi:

- a. **Teks Ulasan**, ulasan yang diberikan oleh pengguna mengenai pengalaman mereka terhadap suatu produk. Data ini digunakan untuk analisis berbasis teks guna mendeteksi pola linguistik yang mencurigakan.



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN
TEKNOLOGI

UNIVERSITAS SUMATERA UTARA
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

PROGRAM STUDI S1 TEKNOLOGI INFORMASI

Jalan Alumni No. 3 Gedung C, Kampus USU Padang Bulan, Medan 20155
Telepon/Fax: 061-8210077 | Email: tek.informasi@usu.ac.id | Laman: http://it.usu.ac.id

b. Rating Produk, skor yang diberikan pengguna dalam bentuk bintang (1–5). Data ini dianalisis untuk menemukan ketidaksesuaian antara rating dan isi ulasan, misalnya ulasan negatif dengan rating tinggi atau sebaliknya.

c. ID Pengguna dan ID Merchant, informasi mengenai identitas pengguna dan merchant digunakan untuk menganalisis hubungan antara *reviewer* dan *merchant*. Data ini berguna dalam mendeteksi kemungkinan adanya aktivitas ulasan palsu yang dilakukan secara sistematis oleh sekelompok pengguna.

d. Timestamp Ulasan, waktu ketika ulasan diberikan dikumpulkan untuk mendeteksi lonjakan aktivitas ulasan dalam waktu singkat. Adanya lonjakan ulasan dalam periode tertentu dapat menjadi indikasi manipulasi ulasan yang disengaja.

e. Frekuensi Ulasan oleh Reviewer, jumlah ulasan yang diberikan oleh reviewer dalam periode waktu tertentu digunakan sebagai indikator dalam mendeteksi pola aktivitas mencurigakan, seperti reviewer yang memberikan banyak ulasan dalam waktu singkat atau menunjukkan pola ulasan yang tidak alami.

2. ***Preprocessing Data***

Pada tahap ini, data yang dikumpulkan dibersihkan dan dipersiapkan agar sesuai untuk pemrosesan lebih lanjut. Proses ini terdiri dari dua bagian utama: Text Preprocessing dan Behavioral Review Preprocessing.

A. Text Preprocessing

Tahapan dalam *text processing* mencakup:

a. Tokenization, memecah teks ulasan menjadi unit kata atau token agar lebih mudah dianalisis.

b. Case Folding, mengonversi semua huruf menjadi huruf kecil untuk menjaga konsistensi data.

c. Stopword Removal, menghapus kata-kata umum, seperti "dan", "atau", "dengan".

d. Stemming/Lemmatization, mengubah kata ke dalam bentuk dasarnya untuk menghindari perbedaan makna akibat variasi kata.

e. Filtering & Cleaning, menghapus tanda baca, angka, serta karakter khusus yang tidak relevan dalam analisis teks.

B. Behavioral Preprocessing

Behavioral preprocessing dilakukan untuk menganalisis pola interaksi reviewer berdasarkan aktivitas mereka di platform. Langkah-langkahnya meliputi:

a. Behavioral Data Extraction, mengambil fitur perilaku reviewer seperti jumlah ulasan dalam waktu tertentu, pola rating, dan hubungan dengan merchant.



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN
TEKNOLOGI

UNIVERSITAS SUMATERA UTARA
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

PROGRAM STUDI S1 TEKNOLOGI INFORMASI

Jalan Alumni No. 3 Gedung C, Kampus USU Padang Bulan, Medan 20155
Telepon/Fax: 061-8210077 | Email: tek.informasi@usu.ac.id | Laman: http://it.usu.ac.id

b. Normalization, menstandarisasi data perilaku agar lebih mudah dianalisis oleh Transformer Encoder.

3. Textual Feature Processing menggunakan IndoBERT

Teks ulasan diproses menggunakan IndoBERT, yang menghasilkan representasi semantik dalam bentuk word embeddings. Model ini memahami konteks kata dalam ulasan dan mampu mengenali pola linguistik yang mencurigakan.

a. Aspect-Based Sentiment Analysis (ABSA)

Sebagai pembeda utama dalam penelitian ini, **ABSA** diterapkan untuk menganalisis sentimen berdasarkan aspek tertentu dari produk (misalnya, **kualitas produk, harga, atau layanan**). Dengan ABSA, sistem dapat mengidentifikasi apakah ada ketidaksesuaian antara sentimen yang diungkapkan dalam teks ulasan dan rating yang diberikan oleh pengguna.

4. Behavioral Feature Processing dengan Transformer Encoder

Pola perilaku reviewer diproses menggunakan Transformer Encoder, yang terdiri dari beberapa komponen utama:

a. Multi-Head Attention, memungkinkan model untuk fokus pada berbagai aspek perilaku reviewer dalam ulasan.

b. Feed-Forward Network (FFN), memproses fitur perilaku yang telah diperoleh untuk menghasilkan representasi yang lebih abstrak.

c. Similarity Analysis, menganalisis kesamaan pola ulasan antar-reviewer guna mengidentifikasi kemungkinan kolusi atau sindikat ulasan palsu.

5. Feature Fusion

Setelah kedua fitur (tekstual dan perilaku) diproses, dilakukan Feature Fusion, yaitu proses penggabungan informasi dari IndoBERT dan Transformer Encoder. Tujuan dari tahap ini adalah untuk menggabungkan wawasan dari konten teks dan perilaku pengguna guna meningkatkan akurasi model dalam membedakan ulasan asli dan palsu.

6. Final Classification

Pada tahap ini, hasil dari Indo-BERT (berbasis teks) dan Transformer Encoder (berbasis perilaku reviewer) digabungkan untuk meningkatkan akurasi deteksi ulasan palsu.

Output dari klasifikasi akhir adalah:

- **Fake Review** → Jika sistem mendeteksi pola anomali dalam teks ulasan atau perilaku reviewer.

- **Genuine Review** → Jika ulasan tidak memiliki indikasi kepalsuan.



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN
TEKNOLOGI

UNIVERSITAS SUMATERA UTARA
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

PROGRAM STUDI S1 TEKNOLOGI INFORMASI

Jalan Alumni No. 3 Gedung C, Kampus USU Padang Bulan, Medan 20155
Telepon/Fax: 061-8210077 | Email: tek.informasi@usu.ac.id | Laman: <http://it.usu.ac.id>

Referensi	<p>Alsaad, M. M. B., & Joshi, H. (2024). Transformer-Based Language Deep Learning Detection of Fake Reviews on Online Products. <i>Journal of Electrical Systems</i>, 20(3), 2368-2378.</p> <p>Asaad, W. H., Allami, R., & Ali, Y. H. (2023, Oktober). Fake Review Detection Using Machine Learning. <i>Revue d'Intelligence Artificielle</i>, 37(5). 10.18280/ria.370507</p> <p>Bashir, U., & Statista. (2024, October 23). <i>Most popular categories for online purchases in Indonesia 2024</i>. Statista. Retrieved March 10, 2025, from https://www.statista.com/forecasts/823409/most-popular-categories-for-online-purchases-in-indonesia</p> <p>Berijalan. (2025, Maret 4). <i>Tren Belanja Online 2025: Teknologi dan Perilaku Konsumen</i>. Tren Belanja Online 2025: Teknologi dan Perilaku Konsumen. Retrieved Maret 14, 2025, from https://berijalan.co.id/article-detail/tren-belanja-online-2025-teknologi-dan-perilaku-konsumen</p> <p>Databoks. (2024, January 10). <i>5 E-Commerce dengan Pengunjung Terbanyak Sepanjang 2023</i> Databoks. Databoks. Retrieved March 8, 2025, from https://databoks.katadata.co.id/teknologi-telekomunikasi/statistik/3c9132bd3836eff/5-e-commerce-dengan-pengunjung-terbanyak-sepanjang-2023</p> <p>Databoks. (2024, March 19). <i>TikTok Shop Gabung Tokopedia, Ini Jumlah Penggunanya</i> Databoks. Databoks. Retrieved March 8, 2025, from https://databoks.katadata.co.id/teknologi-telekomunikasi/statistik/a6be2d006205170/tiktok-shop-gabung-tokopedia-ini-jumlah-penggunanya</p> <p>Databoks. (2024, July 18). <i>Shopee Dominasi Pasar E-Commerce Asia Tenggara pada 2023</i> Databoks. Databoks. Retrieved March 8, 2025, from https://databoks.katadata.co.id/teknologi-telekomunikasi/statistik/66989de7b7168/shopee-dominasi-pasar-e-commerce-asia-tenggara-pada-2023</p>
-----------	--



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN
TEKNOLOGI

UNIVERSITAS SUMATERA UTARA
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

PROGRAM STUDI S1 TEKNOLOGI INFORMASI

Jalan Alumni No. 3 Gedung C, Kampus USU Padang Bulan, Medan 20155
Telepon/Fax: 061-8210077 | Email: tek.informasi@usu.ac.id | Laman: <http://it.usu.ac.id>

- Fil'ilmi, P. R., Manalu, D. A., Tara, I. N. A., Barnanti, R., & Asral. (2025, Januari).
PENGARUH STRATEGI DISKON TERHADAP LOYALITAS PELANGGAN DI
PLATFORM SHOPEE. *Neraca Manajemen, Ekonomi*, 14(4). 10.8734/mnmae.v1i2.359
- Golrezaei, N., Manshadi, V., Schneider, J., & Sekar, S. (2020). Learning Product Rankings
Robust to Fake Users. *arXiv*. <https://doi.org/10.48550/arXiv.2009.05138>
- Goyal, N., & Goyal, D. (2023). Fake Review Detection Using Ensemble Techniques by the
Fusion of Chronology, Aspect and Sentiment of Reviews and Oversampling by Smote.
Journal of Harbin Engineering University, 44(6).
<https://www.researchgate.net/publication/372138938>
- He, S., Hollenbeck, B., Overgoor, G., Proserpio, D., & Tosyali, A. (2022). Detecting fake-review
buyers using network structure: Direct evidence from Amazon. *Proceedings of the
National Academy of Sciences*, 119(47). 10.1073/pnas.2211932119
- Manaskasemsak, B., Tantisuwankul, J., & Rungsawang, A. (2021). Fake review and reviewer
detection through behavioral graph partitioning integrating deep neural network. *Neural
Computing and Applications*, 35(2), 1169-1182. 10.1007/s00521-021-05948-1
- MOHAWESH, R., XU, S., TRAN, S. N., OLLINGTON, R., SPRINGER, M., JARARWEH, Y.,
& MAQSOOD, S. (2021, April). Fake Reviews Detection: A Survey. *IEEE Access*.
10.1109/ACCESS.2021.3075573
- Nguyen, L. N. (2023). An Empirical Study on Fake Review Detection. *TRA VINH UNIVERSITY
JOURNAL OF SCIENCE*. 10.35382/tvujs.13.6.2023.2107
- Peña-García, N., Losada-Otálora, M., Auza, D. P., & Cruz, M. P. (2024). Reviews, trust, and
customer experience in online marketplaces: the case of Mercado Libre Colombia.
Frontiers, 9, 1460321. <https://doi.org/10.3389/fcomm.2024.1460321>
- Qayyum, H., Ali, F., Nawaz, M., & Nazir, T. (2023). FRD-LSTM: A novel technique for fake
reviews detection using DCWR with the Bi-LSTM Method. *Multimedia Tools and
Applications*, 82(20), 31505-31519. 10.1007/s11042-023-15098-2



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN
TEKNOLOGI

UNIVERSITAS SUMATERA UTARA
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

PROGRAM STUDI S1 TEKNOLOGI INFORMASI

Jalan Alumni No. 3 Gedung C, Kampus USU Padang Bulan, Medan 20155
Telepon/Fax: 061-8210077 | Email: tek.informasi@usu.ac.id | Laman: http://it.usu.ac.id

- Septiani, L. (2023, Desember 21). *Tarif Jasa Review Produk Palsu dan Cara Mengenalinya*.
Katadata. Retrieved Maret 12, 2025, from
<https://katadata.co.id/digital/e-commerce/6584029461099/tarif-jasa-review-produk-palsu-dan-cara-mengenalinya>
- Shan, G., Zhou, L., & Zhang, D. (2021). From conflicts and confusion to doubts: Examining review inconsistency for fake review detection. *Elsevier*, 144, 113513.
10.1016/j.dss.2021.113513
- Sun, P., Bi, W., Zhang, Y., Wang, Q., Kou, F., Lu, T., & Chen, J. (2024, November 4). Fake Review Detection Model Based on Comment Content and Review Behavior. *Electronics*, 13, 4322. <https://doi.org/10.3390>
- TUFAIL, H., ASHRAF, M. U., ALSUBHI, K., & ALJAHDALI, H. M. (2022). The Effect of Fake Reviews on e-Commerce During and After Covid-19 Pandemic: SKL-Based Fake Reviews Detection. *IEEE Access*, 10. 10.1109/ACCESS.2022.3152806
- Yin, C., Cuan, H., Zhu, Y., & Yin, Z. (2021, Mei). Improved Fake Reviews Detection Model Based on Vertical Ensemble Tri-Training and Active Learning. *ACM Transactions on Intelligent Systems and Technology*, 12(4), 33. <https://doi.org/10.1145/3450285>
- Yu, S., Ren, J., Li, S., Naseriparsa, M., & Xia, F. (2022). Graph learning for fake review detection. *Frontiers in Artificial Intelligence*, 5. 10.3389/frai.2022.922589

Medan, Maret 2025

Mahasiswa yang mengajukan,

(Gideon Togap Hata Oloan Manurung)

NIM.211402147