

Analisis Sentimen Destinasi Wisata Kabupaten Karangasem pada Ulasan Digital Menggunakan Model IndoBERT

SENTIMENT ANALYSIS OF KARANGASEM REGENCY TOURIST DESTINATIONS IN DIGITAL REVIEWS USING THE INDOBERT MODEL

Elma Regina Nababan^{*1}, I Made Dendi Maysanjaya², Gede Surya Mahendra³

^{1,2,3} Universitas Pendidikan Ganesha: Jl. Udayana No. 11, Singaraja 81116, Indonesia

^{1,2,3} Jurusan Teknik Informatika Universitas Pendidikan Ganesha

e-mail: ^{*1}elma.regina@undiksha.ac.id, ²dendi.ms@undiksha.ac.id, ³gmahendra@undiksha.ac.id

Abstrak

Kabupaten Karangasem merupakan salah satu sektor strategis pariwisata di Bali, namun data terbaru menunjukkan tren penurunan jumlah kunjungan dari sekitar 1,3 juta pada tahun 2023 menjadi 1,2 juta pada tahun 2024. Kondisi ini memerlukan evaluasi pengalaman pengunjung berbasis data untuk mengoptimalkan peningkatan kualitas dan daya saing destinasi secara responsif. Analisis sentimen diperlukan karena volume ulasan digital yang besar membuat analisis manual menjadi tidak efisien. Penelitian ini menggunakan pendekatan model *deep learning* IndoBERT melalui tahapan *pre-processing* dan penerapan *Easy Data Augmentation* (EDA) serta *Downsampling* untuk menangani ketidakseimbangan kelas. Berdasarkan hasil pengujian menggunakan *Stratified 10-Fold Cross Validation*, model berhasil mencapai akurasi sebesar 83% dengan nilai *weighted F1-score* sebesar 0,81, *precision* 0,81, dan *recall* 0,81. Hasil klasifikasi menunjukkan dominasi sentimen positif sebanyak 14.686 ulasan, sementara sentimen negatif (3.880 ulasan) dan netral (2.431 ulasan) memberikan informasi krusial sebagai dasar perbaikan layanan pariwisata di Kabupaten Karangasem.

Kata kunci — Analisis Sentimen, IndoBERT, Karangasem, Pariwisata

Abstrak

Karangasem Regency is one of Bali's strategic tourism sectors, but recent data shows a downward trend in visitor numbers from around 1.3 million in 2023 to 1.2 million in 2024. This situation requires an evaluation of visitor experiences based on data in order to optimize the quality and competitiveness of the destination in a responsive manner. Sentiment analysis is necessary because the large volume of digital reviews makes manual analysis inefficient. This study uses the IndoBERT deep learning model approach through pre-processing and the application of Easy Data Augmentation (EDA) and Downsampling to overcome class imbalance. Based on test results using Stratified 10-Fold Cross Validation, the model achieved 83% accuracy with a weighted F1 score of 0.81, precision of 0.81, and recall of 0.81. The classification results show a dominance of positive sentiment in 14,686 reviews, while negative sentiment (3,880 reviews) and neutral sentiment (2,431 reviews) provide crucial information as a basis for improving tourism services in Karangasem Regency.

Kata kunci — Sentiment Analysis, IndoBERT, Karangasem, Tourism

1. PENDAHULUAN

Pariwisata merupakan sektor strategis bagi perekonomian di Bali, termasuk di Kabupaten Karangasem yang memiliki beragam daya tarik wisata budaya dan alam [1]. Optimalisasi sektor ini menjadi krusial mengingat Karangasem merupakan daerah dengan angka kemiskinan tertinggi di Bali, yakni mencapai 29,45 ribu jiwa pada tahun 2022 [2]. Sehingga peningkatan kualitas dan daya saing destinasi menjadi kebutuhan penting sektor pariwisata

memiliki potensi besar untuk memberikan dampak ekonomi positif yang lebih luas jika dikembangkan secara optimal[3]. Namun, data terbaru menunjukkan tren penurunan jumlah kunjungan dari sekitar 1,3 juta pada tahun 2023 menjadi 1,2 juta pada tahun 2024. Kondisi ini mengindikasikan adanya tantangan dalam pengelolaan destinasi yang memerlukan evaluasi komprehensif terhadap pengalaman wisatawan. Besarnya daya tarik wisata Karangasem dapat dilihat dari data kunjungan dalam lima tahun terakhir, sebagaimana tersaji pada Gambar 1.



Gambar 1. Data Jumlah Wisatan Kab.Karangasem

Evaluasi pengalaman wisatawan dapat dilakukan melalui analisis ulasan digital pada platform seperti Google Maps dan TripAdvisor. Ulasan tersebut merepresentasikan persepsi, kepuasan, serta kritik wisatawan terhadap kualitas layanan dan fasilitas destinasi. Akan tetapi, volume data yang besar menjadikan analisis manual tidak efisien dan berpotensi subjektif, sehingga diperlukan pendekatan otomatis berbasis teknologi.

Berbagai penelitian sebelumnya telah menerapkan metode *machine learning* seperti Naïve Bayes, Support Vector Machine, dan Long Short-Term Memory (LSTM) untuk analisis sentimen pada sektor pariwisata. Sebagian besar penelitian analisis sentimen pariwisata umumnya dilakukan skala nasional atau destinasi wisata yang sangat populer dengan cakupan data luas. Namun, analisis pada tingkat kabupaten memiliki karakteristik yang berbeda karena konteks destinasi, kondisi geografis, serta pola kunjungan wisatawan yang lebih spesifik pada wilayah tertentu. Informasi sentimen pada tingkat lokal menjadi lebih relevan bagi pemerintah daerah karena dapat memberikan gambaran yang lebih detail mengenai persepsi wisatawan terhadap kualitas layanan, fasilitas, juga pengelolaan destinasi wisata secara langsung.

Penelitian sebelumnya terlihat bahwa pendekatan berbasis transformer, seperti IndoBERT, mampu menghasilkan performa yang lebih baik dalam analisis sentimen berbahasa Indonesia dibandingkan beberapa metode *machine learning* [4]. Namun, penerapan pendekatan tersebut pada konteks destinasi wisata tingkat kabupaten masih relatif terbatas. Selain itu, permasalahan ketidakseimbangan kelas pada data ulasan sering kali belum ditangani secara optimal.

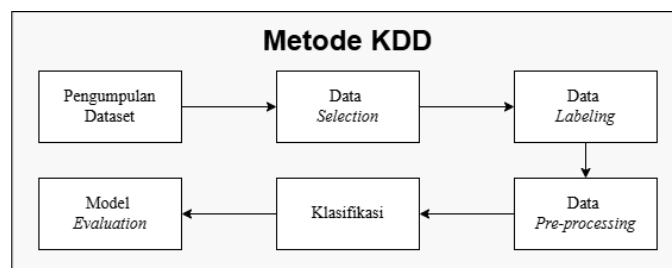
Untuk mengatasi permasalahan tersebut, penelitian ini mengimplementasikan model *deep learning* IndoBERT untuk melakukan klasifikasi sentimen ulasan wisatawan di Kabupaten Karangasem. Model berbasis transformer ini dipilih karena kemampuannya memahami konteks bahasa Indonesia secara mendalam, termasuk variasi bahasa tidak baku dalam ulasan digital [5].

Mengingat dominasi sentimen positif pada data ulasan, teknik *Easy Data Augmentation* (EDA) dan *Downsampling* diterapkan untuk meningkatkan representasi kelas minoritas.

Integrasi IndoBERT dengan EDA dan *Downsampling* diharapkan mampu menghasilkan pemetaan sentimen yang lebih presisi dan seimbang sebagai dasar evaluasi berbasis data dalam pengembangan pariwisata di Kabupaten Karangasem. Dengan demikian, penelitian ini tidak hanya berkontribusi pada penerapan model transformer dalam analisis sentimen pariwisata, tetapi juga menawarkan pendekatan penanganan ketidakseimbangan kelas yang relevan pada data ulasan digital.

2. METODE PENELITIAN

Penelitian ini menggunakan kerangka kerja *Knowledge Discovery in Databases* (KDD) sebagai pendekatan sistematis dalam mengekstraksi informasi bermakna dari kumpulan data berukuran besar. KDD merupakan proses terstruktur yang meliputi tahapan seleksi data, transformasi, pemodelan, dan evaluasi untuk menghasilkan pola atau pengetahuan yang valid dan dapat diinterpretasikan [6]. Alur penelitian yang diterapkan ditunjukkan pada Gambar 2.



Gambar 2. Alur Penelitian Berbasis Metode KDD

Pengumpulan Dataset

Pengumpulan data dilakukan untuk memperoleh ulasan wisatawan yang akan dianalisis dalam penelitian. Sumber data berasal dari dua platform ulasan digital, yaitu Google Maps dan TripAdvisor, yang dipilih karena ketersediaan ulasan berbahasa Indonesia dalam jumlah besar. Proses pengambilan data menggunakan teknik *web scraping* melalui layanan Apify untuk mengekstrak ulasan secara otomatis. Data yang dikumpulkan dibatasi pada ulasan berbahasa Indonesia yang berkaitan dengan destinasi wisata di Kabupaten Karangasem.

Data Selection

Seleksi data merupakan tahap untuk menyaring data mentah hasil *scraping* agar hanya menyisakan data yang relevan dan berkualitas [7]. Proses ini bertujuan untuk menghilangkan data yang tidak diperlukan, seperti ulasan duplikat, ulasan kosong, atau ulasan yang tidak berkaitan dengan pengalaman wisata di destinasi yang diteliti. Seleksi data dilakukan secara otomatis maupun manual untuk memastikan bahwa data yang digunakan dalam tahap selanjutnya benar-benar merepresentasikan opini wisatawan terhadap destinasi wisata di Kabupaten Karangasem.

Data Labeling

Pelabelan data merupakan proses pemberian kategori sentimen pada setiap ulasan yang telah terseleksi. Penelitian ini menggunakan tiga kategori sentimen, yaitu positif, negatif, dan netral. Sentimen positif diberikan pada ulasan yang menunjukkan kepuasan atau apresiasi terhadap

destinasi. Sentimen negatif diberikan pada ulasan yang berisi kritik, keluhan, atau ketidakpuasan. Sentimen netral diberikan pada ulasan yang bersifat deskriptif tanpa muatan emosi yang jelas. Proses pelabelan dilakukan secara manual oleh validator ahli bahasa Indonesia untuk menjamin akurasi dan konsistensi label sebagai *ground truth* dalam klasifikasi.

Data Pre-processing

Pre-processing data bertujuan untuk membersihkan dan mentransformasi teks ulasan mentah menjadi data terstruktur dan digunakan dalam pemodelan[8]. Tahapan *pre-processing* terdiri dari:

- Cleaning*: menghapus karakter tidak relevan seperti tanda baca, angka, simbol, dan emoji [9].
- Case folding*: mengubah seluruh huruf dalam teks menjadi huruf kecil (*lowercase*) [10].
- Tokenizing*: memecah teks ulasan menjadi unit-unit kata tunggal [10].
- Normalization*: mengubah kata tidak baku, *slang*, atau *typo* menjadi bentuk baku [11].
- Stopword removal*: menghapus kata umum yang tidak memiliki makna signifikan [12].
- Stemming*: mengembalikan kata berimbuhan jadi kata dasar dengan *library* Sastrawi [13].

Penanganan Imbalance Data

Ketidakseimbangan kelas (*class imbalance*) merupakan kondisi di mana jumlah data pada suatu kelas jauh lebih banyak dibandingkan kelas lainnya, yang berpotensi menyebabkan model menjadi bias terhadap kelas mayoritas dan kurang sensitif terhadap kelas minoritas. Untuk mengatasi hal tersebut, penelitian ini menerapkan dua strategi penyeimbangan data secara simultan. Pertama, *downsampling* dilakukan pada kelas mayoritas dengan mengurangi jumlah sampel secara acak, yang dipilih karena jumlah data pada kelas mayoritas yang besar memungkinkan reduksi tanpa kehilangan representasi yang bermakna[14]. Kedua, *Easy Data Augmentation* (EDA) diterapkan pada kelas minoritas melalui empat teknik augmentasi teks, yaitu *synonym replacement*, *random insertion*, *random swap*, dan *random deletion*. Kombinasi kedua teknik ini bertujuan untuk menciptakan distribusi data yang lebih proporsional sebelum proses pelatihan model.

Klasifikasi dengan Model IndoBERT

Tahap klasifikasi dilakukan menggunakan model IndoBERT, arsitektur Transformer yang telah melalui *pre-training* pada korpus besar berbahasa Indonesia sehingga unggul dalam memahami konteks semantik dan nuansa linguistik secara mendalam [15]. Proses *fine-tuning* dilakukan dengan konfigurasi *learning rate* $3e-5$, *batch size* 16, dan 3 *epoch* menggunakan *optimizer* AdamW. Dataset diproses dengan mekanisme *subword tokenization* dan batas maksimum 384 token. Untuk memastikan stabilitas dan objektivitas performa, model dievaluasi menggunakan skema *Stratified 10-Fold Cross Validation* yang menjaga proporsi setiap kelas di setiap *fold* pengujian.

Spesifikasi Lingkungan Komputasi

Proses *fine-tuning* model IndoBERT pada penelitian ini dilakukan menggunakan lingkungan komputasi berbasis GPU pada platform Google Colaboratory Pro. Lingkungan komputasi ini digunakan untuk menjalankan proses pelatihan model serta evaluasi menggunakan skema *10-fold cross validation*. Spesifikasi lingkungan komputasi yang digunakan dirincikan pada Tabel 1.

Tabel 1. Spesifikasi Lingkungan Komputasi

Komponen	Spesifikasi
CPU	Intel Xeon 2.30 GHz (2 Core)
RAM	12.7 GB
GPU	NVIDIA Tesla T4 (VRAM 15 GB)
Sistem Operasi	Ubuntu 20.04
Python	3.10.12
PyTorch	2.5.1
Transformers	4.46.1
CUDA	13.0
Lingkungan Eksperimen	Google Colaboratory Pro

Model Evaluation

Model evaluasi bertujuan mengukur tingkat keberhasilan dan keandalan model IndoBERT dalam mengklasifikasikan sentimen ulasan wisatawan. Mengingat ukuran dataset yang diolah cukup besar, penelitian ini mengimplementasikan skema pengujian *10-Fold Cross Validation* untuk memastikan objektivitas dan stabilitas performa model. Melalui skema ini, dataset dibagi secara acak ke dalam 10 bagian (*fold*) yang berukuran sama. Pada setiap iterasinya, satu bagian digunakan sebagai data uji (*testing data*) dan sembilan bagian lainnya digunakan sebagai data latih (*training data*). Pendekatan ini secara efektif meminimalisir risiko *overfitting* dan memastikan bahwa seluruh data telah melalui proses pengujian secara komprehensif.

Kinerja model selanjutnya dievaluasi secara kuantitatif menggunakan metrik standar klasifikasi *machine learning*, yang meliputi Akurasi (*Accuracy*), Presisi (*Precision*), *Recall*, dan *F1-Score*. Hasil prediksi dari model terhadap kelas aktual sentimen (positif, negatif, dan netral) akan direpresentasikan ke dalam bentuk *Confusion Matrix* multi-kelas. Metrik-metrik tersebut dihitung secara berbobot (*weighted average*) untuk mengevaluasi kemampuan model dalam mendeteksi dan mengklasifikasikan setiap sentimen, sehingga kelemahan model pada kelas minoritas dapat diidentifikasi secara transparan.

3. HASIL DAN PEMBAHASAN

Lingkungan Eksperimen

Proses pelatihan (*training*) dan evaluasi model IndoBERT pada penelitian ini dilakukan menggunakan platform *Google Colaboratory Pro* dengan dukungan komputasi berbasis GPU. Berdasarkan pengujian yang dilakukan, rata-rata waktu pelatihan untuk setiap *fold* pada skema *10-fold cross validation* berkisar antara 15–20 menit. Informasi spesifikasi GPU selama proses eksperimen ditunjukkan pada Gambar 3.

NVIDIA-SMI 580.82.07			Driver Version: 580.82.07			CUDA Version: 13.0		
GPU	Name		Persistence-M	Bus-Id	Disp.A	Volatile Uncorr. ECC		
Fan	Temp	Perf	Pwr:Usage/Cap		Memory-Usage	GPU-Util	Compute M.	
							MIG M.	
0	Tesla T4		Off	00000000:00:04.0	Off		0	
N/A	44C	P8	9W / 70W		3MiB / 15360MiB	0%	Default	N/A
Processes:								
GPU	GI	CI	PID	Type	Process name		GPU Memory	
ID	ID	ID					Usage	
No running processes found								

Gambar 3. Informasi Spesifikasi GPU NVIDIA

Gambar 3 menunjukkan spesifikasi GPU yang digunakan selama eksperimen. Berdasarkan informasi tersebut, lingkungan komputasi menggunakan NVIDIA Tesla T4 dengan kapasitas memori 15360 MiB (≈ 15 GB VRAM) dan dukungan CUDA 13.0. Pada saat pengambilan gambar, utilisasi GPU tercatat 0% dengan penggunaan memori sebesar 3 MiB dari total 15360 MiB, serta tidak terdapat proses komputasi lain yang berjalan pada GPU.

Pengumpulan Dataset

Proses *web scraping* yang dilakukan pada 28 destinasi wisata di Kabupaten Karangasem menghasilkan total 22.835 ulasan mentah dari platform Google Maps dan TripAdvisor. Pengumpulan data difokuskan pada ulasan berbahasa Indonesia dalam rentang waktu tahun 2020 hingga 2025. Data mentah ini masih berupa teks ulasan lengkap yang disertai dengan informasi tambahan seperti nama pengulas, rating, dan tanggal ulasan.

Data Selection

Tahapan seleksi dilakukan untuk memastikan kualitas dataset sebelum memasuki proses pelabelan dan pemodelan. Dari total 22.835 ulasan awal, tahap seleksi ini menghasilkan dataset sebanyak 20.997 ulasan yang telah tervalidasi dan siap untuk diproses lebih lanjut. Ulasan yang berhasil dikumpulkan terlihat pada Tabel 2.

Tabel 2. Hasil Data *Selection*

No.	Ulasan
1	Indah, tetapi panas karena terbuka
	...
20997	Pantai terpencil yang indah. Harga tiket masuk 10.000, turunannya kecil dan relatif mudah, kami sampai di ombak besar (Tahun Baru), berayun di atasnya sungguh luar biasa! Ada 4 warung di pantai. 2 kursi berjemur - 100.000 untuk berdua.

Data Labeling

Hasil pelabelan manual oleh validator ahli bahasa Indonesia menunjukkan distribusi sentimen yang timpang. Dari 20.997 ulasan, sentimen positif mendominasi dengan jumlah 14.686 ulasan

(69,94%), diikuti sentimen negatif sebanyak 3.880 ulasan (18,48%), dan sentimen netral sebanyak 2.431 ulasan (11,58%). Distribusi lengkap disajikan pada Tabel 3.

Tabel 3. Jumlah Label Pada Masing-Masing Kelas

Kelas	Jumlah	Persentase
Positif	14686	69,94%
Netral	2431	11,58%
Negatif	3880	18,48%
Total Dataset	20997	100%

Meskipun jumlahnya lebih sedikit, ulasan negatif mengandung informasi krusial terkait keluhan fasilitas maupun pelayanan wisata. Ketimpangan data yang signifikan ini menjadi dasar dilakukannya penanganan ketidakseimbangan kelas sebelum proses pelatihan model.

Pre-processing Data

Tahapan pra-pemrosesan mengubah teks ulasan mentah menjadi data terstruktur melalui enam tahap: *cleaning*, *case folding*, *tokenizing*, *normalization*, *stopword removal*, dan *stemming*. Proses transformasi teks pada setiap tahap menggunakan satu contoh ulasan ditampilkan pada Tabel 4.

Tabel 4. Hasil Transformasi *Pre-processing*

Tahapan	Hasil Teks
Data Mentah	Pantai terpencil yang indah. Harga tiket masuk 10.000, turunannya kecil dan relatif mudah, kami sampai di ombak besar (Tahun Baru), berayun di atasnya sungguh luar biasa! Ada 4 warung di pantai. 2 kursi berjemur - 100.000 untuk berdua.
<i>Cleaning</i>	Pantai terpencil yang indah Harga tiket masuk turunannya kecil dan relatif mudah kami sampai di ombak besar Tahun Baru berayun di atasnya sungguh luar biasa Ada warung di pantai kursi berjemur untuk berdua
<i>Case Folding</i>	pantai terpencil yang indah harga tiket masuk turunannya kecil dan relatif mudah kami sampai di ombak besar tahun baru berayun di atasnya sungguh luar biasa ada warung di pantai kursi berjemur untuk berdua
<i>Tokenizing</i>	[pantai, terpencil, yang, indah, harga, tiket, masuk, turunannya, kecil, dan, relatif, mudah, kami, sampai, di, ombak, besar, tahun, baru, berayun, di, atasnya, sungguh, luar, biasa, ada, warung, di, pantai, kursi, berjemur, untuk, berdua]
<i>Normalization</i>	[pantai, terpencil, yang, indah, harga, tiket, masuk, turunannya, kecil, dan, relatif, mudah , kami, sampai, di, ombak, besar, tahun, baru, berayun, di, atasnya, sungguh, luar, biasa, ada, warung, di, pantai, kursi, berjemur, untuk, berdua]
<i>Stopword Removal</i>	[pantai, terpencil, indah, harga, tiket, masuk, turunannya, kecil, relatif, mudah, sampai, ombak, besar, tahun, baru, berayun, atasnya, sungguh, luar, biasa, warung, pantai, kursi, berjemur, berdua]
<i>Stemming</i>	[pantai, terpencil, indah, harga, tiket, masuk, turun , kecil, relatif, mudah, sampai, ombak, besar, tahun, baru, ayun , atas , sungguh, luar, biasa, warung, pantai, kursi, jemur , dua]

Penanganan *Imbalance Data*

Berdasarkan distribusi label pada Tabel 3, ditemukan ketidakseimbangan kelas yang signifikan dengan sentimen positif mendominasi (14.686 ulasan, 69,94%), sementara negatif (3.880 ulasan, 18,48%) dan netral (2.431 ulasan, 11,58%) jauh lebih sedikit. Untuk mengatasi hal ini, dilakukan penanganan ketidakseimbangan data dengan strategi kombinasi *downsampling* dan *Easy Data Augmentation* (EDA).

Downsampling diterapkan pada kelas positif dengan mengurangi jumlah sampel secara acak dari 14.686 menjadi 7.806 ulasan. Sementara itu, EDA diterapkan pada kelas negatif dan netral melalui empat teknik: *synonym replacement*, *random insertion*, *random swap*, dan *random deletion* untuk memperkaya variasi data minoritas. Hasil penerapan strategi tersebut menghasilkan distribusi data baru yang lebih proporsional dengan total 16.234 ulasan, terdiri dari positif 7.806 ulasan, negatif 5.204 ulasan, dan netral 3.224 ulasan.

Hasil Evaluasi Model Klasifikasi

Evaluasi keandalan model dilakukan menggunakan skema *10-Fold Cross Validation* untuk memastikan bahwa performa klasifikasi stabil dan teruji pada berbagai variasi partisi data. Metrik yang digunakan meliputi akurasi, presisi, *recall*, dan *F1-score* yang dihitung secara *macro* dan *weighted*. Berdasarkan hasil pengujian secara keseluruhan, model IndoBERT berhasil mencapai tingkat akurasi sebesar 83%. Model juga mencatatkan nilai *macro F1-score* sebesar 0.81 dan *weighted F1-score* sebesar 0.83. Rincian metrik evaluasi untuk masing-masing kelas sentimen disajikan pada Tabel 5.

Tabel 5. Klasifikasi Gabungan Seluruh *Fold*

Kelas	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>	<i>Support</i>
Negatif (N)	0,85	0,86	0,86	5204
Netral (O)	0,68	0,68	0,68	3224
Positif (P)	0,89	0,88	0,88	7806
Akurasi			0,83	16234
Macro Avg	0,81	0,81	0,81	16234
Weighted Avg	0,83	0,83	0,83	16234

Berdasarkan Tabel 3 terkait jumlah label, performa klasifikasi terbaik dicapai pada kelas sentimen positif dengan nilai *F1-Score* tertinggi sebesar 0.88. Tingginya nilai presisi (0.89) pada kelas ini mengindikasikan bahwa model sangat akurat dalam mengenali pola linguistik yang merepresentasikan kepuasan wisatawan.

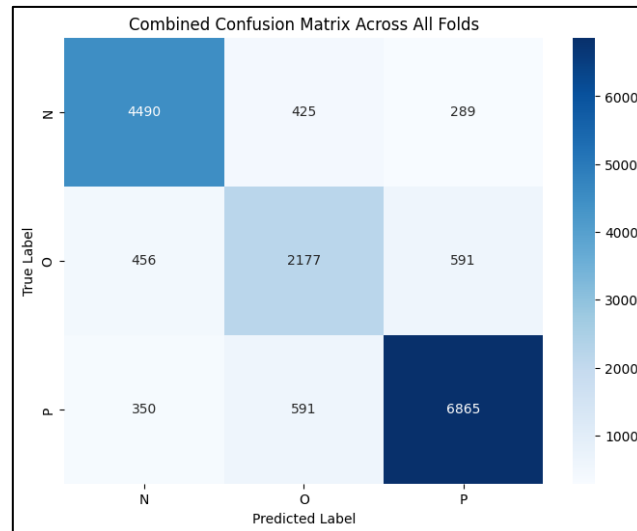
Kelas Negatif juga mencatatkan performa yang sangat baik dengan *F1-Score* 0,86. Nilai *recall* yang tinggi (0,86) pada kelas ini membuktikan bahwa model memiliki sensitivitas yang baik dalam mendeteksi ulasan yang mengandung kritik atau ketidakpuasan wisatawan. Keberhasilan ini tidak lepas dari penerapan strategi penyeimbangan data berupa kombinasi EDA pada kelas minoritas (negatif dan netral) dan *downsampling* pada kelas mayoritas (positif).

Sebaliknya, kelas netral mencatatkan performa terendah dengan *F1-Score* 0,68. Hal ini mengindikasikan adanya tantangan dalam mengklasifikasikan sentimen netral secara akurat karena batas semantik antara opini netral dan opini evaluatif sering kali tidak tegas. Meskipun

telah dilakukan augmentasi pada kelas netral, ambiguitas linguistik yang melekat pada teks deskriptif masih menyulitkan model dalam membedakannya dari teks evaluatif.

Analisis *Confusion Matrix*

Untuk mengidentifikasi pola kesalahan klasifikasi yang dilakukan model, terutama pada kelas netral, dilakukan analisis menggunakan matriks kebingungan (*confusion matrix*). Hasil visualisasinya dapat dilihat pada Gambar 4.



Gambar 4. *Confusion Matrix* IndoBERT

Berdasarkan Gambar 4, model menunjukkan kinerja terbaik pada kelas positif dengan 6.865 ulasan berhasil diklasifikasikan dengan benar. Hal ini menunjukkan bahwa pola linguistik yang merepresentasikan sentimen positif relatif mudah dikenali oleh model, terutama pada ulasan yang mengandung kata-kata evaluatif seperti “indah”, “bagus”, atau “rekomendasi”. Selain itu, model juga mampu mengklasifikasikan 4.490 ulasan negatif secara tepat, yang menunjukkan bahwa ekspresi ketidakpuasan wisatawan dapat dikenali dengan cukup baik oleh model.

Meskipun demikian, masih terdapat beberapa kesalahan klasifikasi. Sebanyak 425 ulasan negatif diprediksi sebagai netral. Kesalahan ini umumnya terjadi pada ulasan yang mengandung kritik secara tidak langsung atau tidak menggunakan kata-kata negatif yang eksplisit, sehingga model cenderung menginterpretasikan kalimat tersebut sebagai deskripsi biasa. Selain itu, beberapa ulasan juga mengandung campuran informasi deskriptif dan evaluatif sehingga batas antara sentimen negatif dan netral menjadi kurang jelas bagi model.

Kesalahan klasifikasi juga terlihat pada kelas Netral, di mana 591 ulasan netral diprediksi sebagai positif. Hal ini menunjukkan bahwa model masih mengalami kesulitan membedakan ulasan yang bersifat deskriptif dengan ulasan yang benar-benar mengandung evaluasi positif. Sebagai contoh, ulasan seperti “Makanan warung lokal OK, porsi kecil untuk harga segitu di Amed” secara konteks bersifat informatif dan tidak sepenuhnya menunjukkan evaluasi positif, namun diprediksi sebagai sentimen positif karena adanya kata “OK” yang sering diasosiasikan dengan pengalaman yang baik. Contoh lainnya adalah ulasan “Jalan menuju Tirta Gangga mulus tapi parkir susah, pengunjung banyak” yang seharusnya termasuk kategori netral karena memuat

Analisis *Word cloud*

a. *Word cloud* Positif

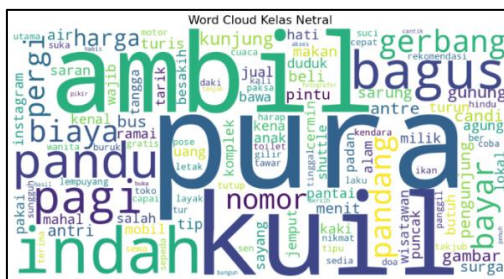


Pada Gambar 5, *word cloud* kelas positif didominasi kata evaluatif seperti “*indah*”, “*bagus*”, “*pantai*”, dan “*kunjung*”. Dominasi kata adjektiva bermuatan apresiatif menunjukkan bahwa persepsi wisatawan cenderung menekankan aspek keindahan dan kualitas destinasi. Pola leksikal yang eksplisit ini memudahkan model dalam mengidentifikasi sentimen positif, sehingga berkontribusi pada tingginya performa klasifikasi pada kelas ini.

[illegible]

Pada Gambar 6, *word cloud* pada kelas negatif, kata-kata seperti “bayar”, “mahal”, “biaya”, “tipu”, dan “antri” muncul dominan. Hal ini menunjukkan bahwa faktor harga, biaya tambahan, dan pengalaman layanan menjadi sumber utama ketidakpuasan wisatawan. Berbeda dengan kelas positif, pola bahasa pada kelas negatif lebih kontekstual dan beragam, sehingga memerlukan pemahaman konteks yang lebih dalam oleh model.

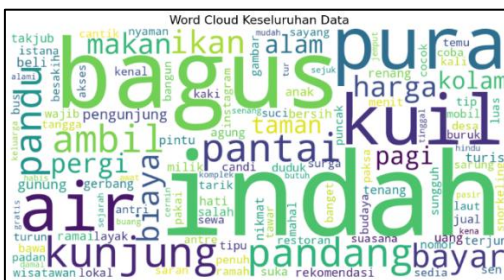
c. *Word Cloud Netral*



Gambar 7. *Word Cloud Netral*

Pada Gambar 7, *Word cloud* sentimen netral didominasi kata-kata seperti "pura", "kuil", "harga", "bayar", dan "gerbang". Hal ini menunjukkan bahwa ulasan netral cenderung membahas informasi faktual mengenai destinasi, tiket masuk, dan akses. Kata "bagus" dan "indah" juga muncul meskipun dengan frekuensi lebih kecil, mengindikasikan adanya tumpang tindih kosakata dengan kelas positif. Dominasi nomina deskriptif dibandingkan adjektiva evaluatif menjadi ciri khas kelas netral, yang menjelaskan mengapa model kesulitan mengklasifikasikannya secara akurat karena tidak memiliki penanda sentimen yang eksplisit.

d. *Word cloud Keseluruhan*



Gambar 8. *Word Cloud Keseluruhan Data*

Pada Gambar 8, dapat dilihat secara keseluruhan, *word cloud* menunjukkan dominasi kata "indah", "bagus", dan "kunjung", yang mengindikasikan kecenderungan sentimen positif dalam dataset. Namun, kemunculan kata seperti "bayar" dan "harga" menegaskan bahwa aspek biaya tetap menjadi perhatian utama wisatawan. Distribusi ini mendukung hasil klasifikasi yang menunjukkan dominasi sentimen positif dengan beberapa isu spesifik terkait biaya dan layanan.

4. KESIMPULAN

Penelitian ini berhasil mengimplementasikan model IndoBERT untuk analisis sentimen ulasan wisatawan di Kabupaten Karangasem. Model mencapai akurasi sebesar 83% dengan *weighted F1-score* 0,83 melalui pengujian *10-fold cross validation*. Penerapan teknik *downsampling* dan *Easy Data Augmentation* (EDA) terbukti efektif dalam mengatasi ketidakseimbangan kelas, yang ditunjukkan dengan nilai *recall* kelas negatif mencapai 0,86. Meskipun demikian, model masih menghadapi tantangan dalam mengklasifikasikan sentimen netral akibat ambiguitas semantik yang tinggi, dengan *F1-score* hanya 0,68.

Secara umum, destinasi wisata di Kabupaten Karangasem memiliki citra yang sangat baik di mata wisatawan, dibuktikan dengan dominasi ulasan bersentimen positif mencapai 69,94% dari

total 20.997 ulasan yang terkumpul. Namun, deteksi terhadap 18,48% ulasan bersentimen negatif memberikan bahan evaluasi krusial bagi pemerintah daerah. Berdasarkan analisis *word cloud*, keluhan wisatawan didominasi oleh kata-kata seperti "bayar", "mahal", dan "tipu" yang mengindikasikan permasalahan pada aspek biaya dan transparansi harga, serta "antri" yang menunjukkan perlunya peningkatan kualitas pelayanan. Dengan kemampuan mengidentifikasi pola kepuasan dan keluhan wisatawan secara sistematis, model ini berpotensi diintegrasikan sebagai instrumen analitik otomatis untuk mendukung evaluasi dan pengambilan keputusan berbasis data dalam pengelolaan destinasi wisata di Kabupaten Karangasem.

Berdasarkan keterbatasan yang ditemukan dalam penelitian ini, penelitian selanjutnya disarankan untuk meningkatkan performa klasifikasi pada sentimen netral yang masih memiliki *F1-score* 0,68 akibat ambiguitas semantik antara teks deskriptif dan evaluatif. Hal ini dapat dilakukan melalui eksplorasi metode *ensemble learning* atau penyesuaian *threshold* probabilitas agar model lebih mampu membedakan ulasan yang bersifat faktual tanpa muatan emosi. Selain itu, mengingat cakupan penelitian ini hanya terbatas pada 28 destinasi wisata di Kabupaten Karangasem, penelitian mendatang perlu menguji model pada cakupan wilayah yang lebih luas dengan karakteristik wisatawan yang berbeda untuk mengetahui tingkat generalisasi model yang dikembangkan.

DAFTAR PUSTAKA

- [1] G. A. Pradnyana and I. G. M. Darmawiguna, "Web-Based System for Bali Tourism Sentiment Analysis during The Covid-19 Pandemic using Django Web Framework and Naive Bayes Method," *Proc. 4th Int. Conf. Innov. Res. Across Discip. (ICIRAD 2021)*, vol. 613, no. Icirad, pp. 316–320, 2022, doi: 10.2991/assehr.k.211222.050.
- [2] K. E. Marsitadewi and I. W. Sudemen, "Kemiskinan Struktural di Karangasem: Menganalisis Penyebab dan Intervensi Pemerintah," *J. Good Gov.*, vol. 20, no. 1, 2024, doi: <https://doi.org/10.32834/gg.v20i1.745>.
- [3] S. G. Mahendra and I. N. I. Wiradika, "Decision Support System for Selection of Favorite Tourist Attractions Using Piprecia-Cocoso With Python," *Teknomatika*, vol. 14, no. 1, pp. 1–12, 2024, doi: <http://ojs.palcomtech.ac.id/index.php/teknomatika/article/view/616>.
- [4] C. A. Bahri and L. H. Suadaa, "Aspect-Based Sentiment Analysis in Bromo Tengger Semeru National Park Indonesia Based on Google Maps User Reviews," *IJCCS (Indonesian J. Comput. Cybern. Syst.)*, vol. 17, no. 1, p. 79, Feb. 2023, doi: 10.22146/ijccs.77354.
- [5] A. Dwiyono, M. Fachrurrozi, J. Palembang-Prabumulih, K. Ogan Ilir, and S. Selatan, "Analisis Perbandingan Klasifikasi Intent Chatbot Menggunakan Deep Learning BERT, RoBERTa, dan IndoBERT," *J. Inf. Syst. Res.*, vol. 6, no. 1, pp. 605–616, 2024, doi: 10.47065/josh.v6i1.6051.
- [6] I. P. D. W. Darmawan, G. A. Pradnyana, and I. B. N. Pascima, "Optimasi Parameter Support Vector Machine Dengan Algoritma Genetika Untuk Analisis Sentimen Pada Media Sosial Instagram," *SINTECH (Science Inf. Technol. J.)*, vol. 6, no. 1, pp. 58–67, Apr. 2023, doi: 10.31598/sintechjournal.v6i1.1245.
- [7] N. Fibriyanti Arminda, N. Sulistiyowati, and T. Nur Padilah, "Implementasi Algoritma Multinomial Naive Bayes Pada Analisis Sentimen Terhadap Ulasan Pengguna Aplikasi Brimo," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 7, no. 3, pp. 1817–1822, 2023, doi: 10.36040/jati.v7i3.7012.
- [8] A. Saputra, M. Subing, and R. Pratama, "Perbandingan Metode Naïve Bayes Classifier Dan Support Vector Machine Untuk Analisis Sentimen Pengguna Twitter Mengenai Piala

- Dunia Fifa 2022 COMPARISON OF NAIVE BAYES TAXONOMY AND SUPPORT VECTOR MACHINES FOR ANALYZING TWITTER USER SENTIMENT ON FIFA WORLD,” *Teknomatika*, vol. 13, no. 01, 2023.
- [9] I. D. G. M. Parawangsa, G. S. Mahendra, and R. N. P. Pinatih, “Implementasi Metode XLM-RoBERTa Untuk Analisis Sentimen Ulasan Pengunjung Objek Wisata Kerta Gosa Klungkung,” *RIGGS J. Artif. Intell. Digit. Bus.*, vol. 4, no. 4, pp. 9893–9900, 2026, doi: 10.31004/riggs.v4i4.5368.
- [10] G. A. Pradnyana, I. G. M. Darmawiguna, D. K. S. Suditresna Jaya, and A. Sasmita, “Performance analysis of support vector machines with polynomial kernel for sentiment polarity identification: A case study in lecturer’s performance questionnaire,” *J. Phys. Conf. Ser.*, vol. 1810, no. 1, 2021, doi: 10.1088/1742-6596/1810/1/012033.
- [11] Aprilia Monica Sari, I Made Gede Sunarya, and I Md. Dendi Maysanjaya, “Perancangan Aplikasi Analisis Sentimen Mengenai Program Petani Milenial Berbasis IndoBERT,” *Kumpul. Artik. Mhs. Pendidik. Tek. Inform.*, vol. 15, no. 1, pp. 118–126, Jan. 2026, doi: 10.23887/karmapati.v15i1.108665.
- [12] A. Y. Setiawan, I. G. M. Darmawiguna, and G. Pradnyana, “Sentiment Summarization Learning Evaluation Using LSTM (Long Short Term Memory) Algorithm,” in *Proceedings of the 4th International Conference on Vocational Education and Technology, IConVET 2021, 27 November 2021, Singaraja, Bali, Indonesia*, EAI, 2022. doi: 10.4108/eai.27-11-2021.2315533.
- [13] I. G. M. Darmawiguna, G. A. Pradnyana, and I. B. Jyotisananda, “Indonesian sentiment summarization for lecturer learning evaluation by using textrank algorithm,” *J. Phys. Conf. Ser.*, vol. 1810, no. 1, 2021, doi: 10.1088/1742-6596/1810/1/012024.
- [14] S. Pratama *et al.*, “Penerapan NLP untuk Menganalisis Komentar Publik pada Pelantikan Menteri di Kanal Sekretariat Presiden,” *Teknomatika*, vol. 15, no. 02, pp. 1–10, 2025.
- [15] D. A. Agustina, S. Subanti, E. Zukhronah, P. S. Statistika, and U. S. Maret, “Implementasi Text Mining Pada Analisis Sentimen Pengguna Twitter Terhadap Marketplace di Indonesia Menggunakan Algoritma Support Vector Machine,” *Indones. J. Appl. Stat.*, vol. 3, no. 2, pp. 109–122, 2021.
-