

ASPECT BASED SENTIMENT ANALYSIS MENGGUNAKAN INDOBERT MODEL TERHADAP REVIEW PENGUNJUNG OBJEK WISATA BATURRADEN

Dany Candra Febrianto¹, Maulida Ayu Fitriani², Mahazam Afrad³, Mutiara Auliya Khadija⁴

¹Teknik Informatika (Kampus Kabupaten Banyumas) Universitas Telkom, danycandra@telkomuniversity.ac.id

²Teknik Informatika Universitas Muhammadiyah Purwokerto, maulidaayuf@ump.ac.id

³Sistem Informasi (Kampus Kabupaten Banyumas) Universitas Telkom, mahazamafrad@telkomuniversity.ac.id

⁴Sekolah Vokasi, Universitas Sebelas Maret, mutiaraauliya@staff.uns.ac.id

*)Korespondensi : danycandra@telkomuniversity.ac.id

Abstrak

Pariwisata memiliki peran penting dalam mendorong perekonomian daerah dan meningkatkan kesejahteraan masyarakat. Objek wisata Baturraden, yang terletak di Kabupaten Banyumas, Jawa Tengah, merupakan salah satu destinasi yang memiliki daya tarik utama adalah wisata alam. Data perkembangan pengunjung objek wisata Baturraden selama beberapa tahun terakhir menunjukkan tren yang baik. Namun, untuk memastikan keberlanjutan dan peningkatan kualitas layanan, penting untuk memahami persepsi dan pengalaman pengunjung. Penelitian ini menggunakan pendekatan *Aspect-Based Sentiment Analysis* (ABSA) untuk menganalisis ulasan pengunjung terhadap Baturraden dengan memanfaatkan model IndoBERT, sebuah model *deep learning* berbasis *Bidirectional Encoder Representations from Transformers* (BERT) yang dirancang khusus untuk bahasa Indonesia. Analisis difokuskan pada empat aspek utama: *Attraction*, *Accessibility*, *Amenities*, dan *Ancillary Services*. Selanjutnya dilakukan proses *pre-processing* yang mencakup *Case Folding*, *Cleansing*, *Tokenizing*, *Normalisasi*, *Stemming* dan *Stopword*. Tahap Evaluasi model dilakukan menggunakan *confusion matrix* untuk menghitung nilai akurasi, *precision*, *recall*, dan *F1-score*. Hasil penelitian ini menunjukkan nilai akurasi senilai 94,61%, *precision* 83,22%, *recall* 96% dan *F1-score* 88,11%. Hal tersebut menunjukkan bahwa model dapat melakukan klasifikasi ulasan kedalam aspek – aspek yang diperlukan. Tantangan utama terletak pada analisis ulasan yang beragam dari segi bahasa dan dialek, serta jumlah data yang tidakimbang pada masing-masing aspek.

Kata Kunci: ABSA, Baturraden, BERT, IndoBERT, Pariwisata.

Abstract

Tourism significantly contributes to regional economic growth and enhances public welfare. Baturraden tourist attraction, located in Banyumas Regency, Central Java, is one of the destinations whose main attraction is nature tourism. Data on visitors to Baturraden tourist attraction over the past few years shows a good trend. To ensure long-term sustainability and enhance service quality, understanding visitor perceptions and experiences is crucial. This study employs Aspect-Based Sentiment Analysis (ABSA) to analyze visitor reviews of Baturraden. Utilizing the IndoBERT model, a deep learning architecture based on Bidirectional Encoder Representations from Transformers (BERT) specifically tailored for the Indonesian language, the research focuses on four key aspects: Attraction, Accessibility, Amenities, and Ancillary Services. Next stage, a pre-processing process is carried out which includes Case Folding, Cleansing, Tokenizing, Normalization, Stemming and Stopword. Model evaluation is conducted using a confusion matrix, assessing accuracy (94.61%), precision (83.22%), recall (96%), and F1-score (88.11). These results demonstrate the model's can classif reviews into the required aspects. A primary challenge encountered in this research involves analyzing reviews exhibiting diverse linguistic styles, including variations in language and dialect, as well as addressing the issue of imbalanced data distribution across the different aspects.

Keywords: ABSA, Baturraden, BERT, IndoBERT, Tourism.

I. PENDAHULUAN

Pariwisata merupakan salah satu penggerak sektor ekonomi yang berperan penting dalam meningkatkan pendapatan daerah dan kesejahteraan masyarakat setempat. Objek wisata Baturraden, yang terletak di Kabupaten Banyumas, Jawa Tengah, merupakan salah satu destinasi yang memiliki daya tarik utama

yaitu wisata alam. Data perkembangan pengunjung objek wisata Baturraden selama beberapa tahun terakhir menunjukkan tren yang baik. Khusus untuk Baturraden, data menunjukkan jumlah pengunjung mencapai 372.951 orang pada tahun 2023.[1]

Namun, untuk memastikan keberlanjutan dan peningkatan kualitas layanan, penting untuk



memahami persepsi dan pengalaman pengunjung[2]. Terdapat terdapat 4 Aspek yang mempengaruhi Objek Wisata yaitu: *Attraction*, *Accessibility*, *Amenities*, *Ancillary Service*. *Attraction* (Daya Tarik): Aspek ini mencakup ulasan terkait daya tarik utama dari objek wisata, seperti keindahan alam, atraksi budaya, dan kegiatan yang ditawarkan. *Accessibility* (Aksesibilitas): Aspek ini mencakup ulasan tentang kemudahan akses menuju objek wisata, termasuk kondisi jalan dan transportasi yang tersedia. *Amenities* (Fasilitas): Aspek ini mencakup ulasan tentang fasilitas yang tersedia di objek wisata, seperti tempat parkir, area bermain, restoran, dan fasilitas pendukung lainnya. *Ancillary Service* (Layanan Pendukung): Aspek ini mencakup ulasan tentang layanan tambahan yang disediakan oleh objek wisata, seperti pelayanan oleh staf, informasi wisata, dan layanan keamanan [3].

Baturraden, sebagai salah satu destinasi wisata populer di Jawa Tengah, Indonesia, menerima ulasan dari pengunjung yang membagikan pengalaman mereka melalui Google Review. Ulasan dan opini dari pengunjung objek wisata menjadi sumber informasi yang sangat berharga bagi pengelola wisata dan calon pengunjung[4]. Ulasan ini tidak hanya mencerminkan pengalaman individu, tetapi juga memberikan wawasan tentang aspek-aspek tertentu dari layanan dan fasilitas yang disediakan [5]. Namun, data ulasan yang terkumpul dari review pengunjung sangat sulit untuk dianalisis karena terdapat berbagai macam kata tidak baku serta berbagai kosa kata menggunakan dialek kedaerahan. Untuk mengatasi permasalahan kompleksitas bahasa dan memahami ulasan diperlukan sebuah teknik berbasis kecerdasan buatan. *Aspect-Based Sentiment Analysis* (ABSA) adalah teknik analisis teks yang dapat mengkategorikan data berdasarkan aspek dan mengidentifikasi sentimen yang terkait dengan masing-masing aspek tersebut [6]. Tujuan dari ABSA adalah untuk mendeteksi istilah-istilah dalam teks yang mungkin muncul secara eksplisit sehingga setiap kata memiliki makna yang berbeda ketika digabungkan dengan kata-kata lainnya [7].

Penelitian terkait ABSA menggunakan ulasan terkait destinasi wisata sebelumnya telah dilakukan pada penelitian [7] Penelitian ini menguji kinerja lima algoritme *machine learning* dan delapan skenario untuk mengklasifikasikan sentimen ulasan pengguna *Google Maps* di Candi Borobudur dan Prambanan. Algoritme *machine learning* yang digunakan adalah *Random Forest* (RF), *Naïve Bayes* (NB), *Logistic Regression* (LR), *Decision Tree* (DT), dan *Extra Tree* (ET). Aspek yang digunakan adalah *Attractions* (Att), *Amenities* (Amn), *Accessibility* (Acc), *Image* (Img), *Price* (Prc), dan *Human Resources* (HR). Kinerja algoritme dilatih dengan 10 *folds cross-validation*, dan hasilnya menunjukkan bahwa LR memperoleh skor tertinggi untuk Att (74,4%) dan Amn (81%), DT

memperoleh skor tertinggi untuk Acc (88,5%) dan HR (91,5%). Sementara ET memperoleh skor tertinggi untuk aspek Img (57,2%) dan Prc (88,1%).

ABSA merupakan tugas dalam bidang *Natural Language Processing* yang membutuhkan algoritme pemrosesan dan analisis teks yang optimal [8]. Menurut penelitian [9] algoritme dengan arsitektur *Bidirectional Encoder Representations from Transformers* (BERT) memiliki nilai akurasi tertinggi dibandingkan dengan model arsitektur lain seperti *SentWordNet*, *Logistic Regression*, dan *Long Short-Term Memory*. BERT adalah teknik berbasis *transformer* untuk NLP yang pertama kali dikembangkan oleh Jacob Devlin, et al [10] di Google dan dipublikasikan pada tahun 2018. BERT menggunakan teknik *transformer* untuk menganalisis hubungan kontekstual antara satu kata dengan kata lain dalam sebuah kalimat [11]. BERT menggunakan *masked language model* (MLM) yang memungkinkan model memprediksi kata yang hilang dalam kalimat dengan melihat konteks dari kedua arah secara bersamaan. BERT dilatih dengan dua tugas utama: MLM dan *Next Sentence Prediction* (NSP), yang memungkinkan model memahami hubungan antara kalimat. Inti dari arsitektur BERT adalah *encoder*, yang terdiri dari beberapa lapisan *self-attention* yang mampu menangkap interaksi kata-kata dalam kalimat tanpa memperhatikan urutan kata secara eksplisit [12].

Penelitian terkait ABSA menggunakan algoritme BERT dilakukan oleh Bimaputra et al [13]. Penelitian ini melakukan kajian terhadap review pengunjung pada hotel di Bali. Aspek yang dikaji dalam penelitian ini adalah aspek keamanan, kebersihan kenyamanan dan layanan hotel. Tahapan penelitian yang dilakukan adalah pengumpulan data, data *pre-processing* dan implementasi model. Data yang digunakan adalah data yang berasal dari ulasan pada *website* Tripadvisor kemudian dilakukan pelabelan sesuai dengan aspek yang dikerjakan. Evaluasi kinerja dilakukan menggunakan *confusion matrix*. Hasil penelitian ini mendapatkan nilai akurasi dengan rata-rata 95% pada setiap aspek yang dievaluasi.

IndoBERT adalah model bahasa berbasis BERT yang dikembangkan khusus untuk bahasa Indonesia [14]. Model ini dilatih menggunakan lebih dari 220 juta kata yang berasal dari tiga sumber utama: Wikipedia bahasa Indonesia, artikel berita (seperti Kompas, Tempo, dan Liputan6), serta *Indonesian Web Corpus*. IndoBERT mengikuti arsitektur BERT dan dilatih menggunakan dua tugas utama: *masked language modeling* (MLM) dan *next sentence prediction* (NSP). Dengan total pelatihan selama 2,4 juta langkah (180 *epoch*), IndoBERT mencapai tingkat *perplexity* yang sebanding dengan BERT versi bahasa Inggris. Arsitektur IndoBERT mengikuti desain BERT *base*, dengan 12 lapisan *transformer*,

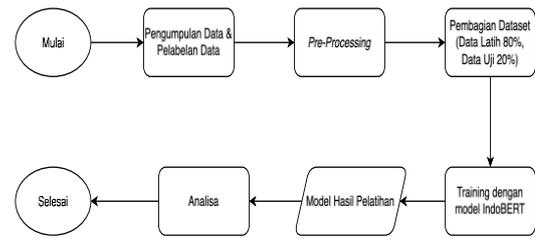
768 ukuran *hidden layer*, dan 12 *attention heads* [15]. Dalam evaluasi pada dataset benchmark IndoLEM, yang mencakup tujuh tugas NLP seperti POS tagging, named entity recognition, dan parsing dependensi, IndoBERT menunjukkan kinerja terbaik dibandingkan model lain pada penelitian [16]

Penelitian [17] menggunakan model IndoBERT untuk mendeteksi berita hoaks di Indonesia, dengan hasil: akurasi 93,2%, presisi 92%, recall 89,7%, dan F1-score 90,8%. Model ini menggabungkan IndoBERT, yang berfungsi sebagai ekstraktor fitur berbasis bahasa Indonesia, dengan *Long Short-Term Memory* (LSTM) yang mampu mengenali pola temporal dalam data berurutan. Data yang digunakan terdiri dari 1.998 artikel faktual dan 3.878 artikel hoaks, yang diperoleh dari turnbackhoax.id dan CNN Indonesia. Studi ini menyoroti pentingnya adaptabilitas model dalam lanskap berita yang dinamis, terutama menjelang peristiwa politik yang cenderung meningkatkan penyebaran misinformasi di media sosial.

Penelitian ini, akan melakukan *Aspect-Based Sentiment Analysis* (ABSA) menggunakan model IndoBERT untuk menganalisis ulasan pengunjung terhadap objek wisata Baturraden. Tujuan penelitian ini adalah untuk mengidentifikasi aspek-aspek spesifik yang sering dibahas dalam ulasan, yaitu *Attraction*, *Accessibility*, *Amenities*, dan *Ancillary Service*, serta memahami sentimen yang terkait dengan setiap aspek tersebut. Dengan menggunakan IndoBERT, yang merupakan model bahasa terkini untuk Bahasa Indonesia. Hasil analisis menggunakan model ini kemudian akan dievaluasi menggunakan *confusion matrix* untuk menghitung nilai akurasi, *precision*, *recall*, dan F1-Score.

II. METODE

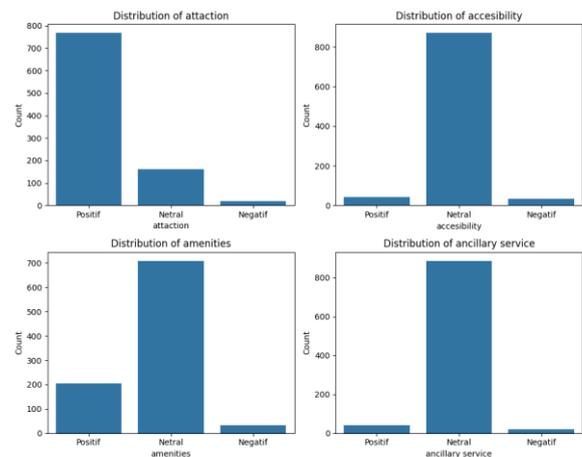
Alur penelitian pada terdiri dari beberapa tahap, dimulai dengan tahap pengumpulan data, di mana data yang relevan dikumpulkan dari ulasan wisata baturraden dari halaman *Google Maps* dan diberikan label sesuai aspek *Attraction*, *Accessibility*, *Amenities*, serta *Ancillary Service*. Selanjutnya proses *pre-processing* yang mencakup *Case Folding*, *Cleansing*, *Tokenizing*, *Normalisasi*, *Stemming* dan *Stopword*. Dataset kemudian dibagi menjadi dua bagian: data latih (80%) untuk melatih model dan data uji (20%) untuk mengevaluasi performanya. Pada tahap evaluasi, model yang telah dilatih kemudian dievaluasi menggunakan matrik akurasi, presisi, *recall*, dan *F1-Score* untuk mengukur kinerja yang dihasilkan. Pengembangan model menggunakan Bahasa Pemrograman Python dan Library PyTorch. *Base Model* IndoBERT yang digunakan dapat diakses melalui portal huggingface. Gambar 1 menunjukkan diagrama alir untuk alur penelitian.



Gambar 1. Alur Penelitian

2.1 Pengumpulan Data & Pelabelan Data

Pengumpulan data merupakan langkah awal yang dilakukan dalam penelitian ini. Data diperoleh dari ulasan *Google Maps* tentang Objek Wisata Baturraden. Teknik web scraping menggunakan bantuan *library SerpAPI* untuk mengumpulkan ulasan tersebut. Sebanyak 946 ulasan berhasil dikumpulkan. Data dalam penelitian ini dikumpulkan pada bulan Juli 2024, yaitu dari tahun 2020 hingga Juli 2024. Data yang dikumpulkan berupa data teks berbahasa Indonesia. Setelah data terkumpul, setiap ulasan diberi label secara manual berdasarkan aspek sentimennya.



Gambar 2. Distribusi Dataset per Aspek

Ulasan-ulasan tersebut dikelompokkan menjadi aspek *Attraction*, *Accessibility*, *Amenities*, dan *Ancillary Service*. Masing-masing aspek dikelompokkan menjadi tiga kategori sentimen: positif, netral, dan negatif. Contoh pelabelan data ulasan: “Tempat wisata favorit dengan pemandangan indah di Banyumas untuk menenangkan diri. Selain lingkungannya asri, udaranya sejuk. Aksesnya mudah dan ramah” {“Attraction”: “Positif”, “Accessibility”: “Positive”, “Amenities”: “Neutral”, “Ancillary Service”: “Positive”}. Gambar 2 menunjukkan distribusi data per masing-masing aspek.

2.2 Pre-Processing

Pada tahap *pre-processing* dilakukan untuk mengolah data yang telah dikumpulkan agar bisa digunakan dan memiliki nilai dalam klasifikasi. Secara umum data teks memiliki pola yang tidak

terstruktur dan multidimensi, sehingga memerlukan proses pembersihan data [18]. Tahap *preprocessing* yang digunakan dalam penelitian ini yaitu *case folding*, *text cleaning*, tokenisasi, *stopword removal*, *stemming*.

1) Case Folding

Case folding adalah salah satu langkah penting dalam bidang NLP yang bertujuan untuk menyeragamkan teks dengan mengubah semua huruf menjadi huruf kecil. Proses ini esensial dalam memastikan bahwa analisis teks tidak terganggu oleh perbedaan kapitalisasi yang tidak relevan secara semantik. Contoh dari kata tahap *case folding* adalah 'Tempat' menjadi 'tempat'.

2) Text Cleaning

Text cleaning bertujuan untuk meningkatkan kualitas data teks sebelum dianalisis lebih lanjut. Proses *text cleaning* melibatkan beberapa langkah sistematis. Pertama, simbol dan karakter khusus seperti &, %, dan * dihapus karena tidak memiliki kontribusi semantik. Selanjutnya, alamat URL dan hashtag dihapus untuk menghindari gangguan dalam analisis, terutama dalam konteks media sosial di mana elemen-elemen ini sering muncul.

3) Tokenisasi

Tokenisasi berfungsi untuk memecah teks menjadi unit-unit yang lebih kecil, seperti kata-kata atau token. Proses ini memungkinkan analisis teks dilakukan pada tingkat yang lebih granular, sehingga memudahkan pemahaman dan manipulasi data teks. Dengan memecah kalimat menjadi token, hal ini memungkinkan untuk melakukan analisis lebih lanjut secara individual, yang berguna dalam berbagai aplikasi NLP seperti analisis frekuensi kata, pengindeksan, dan pengenalan entitas. Misalnya, kalimat "tempat wisata yang sejuk dan menyenangkan" dapat dipecah menjadi token: 'tempat', 'wisata', 'yang', 'sejuk', 'dan', 'menyenangkan'. Proses ini memungkinkan algoritma untuk memproses dan memahami makna dari setiap kata dalam konteks yang lebih luas.

4) Normalisasi

Setiap ekspresi yang disajikan di media sosial sering kali mengandung penggunaan bahasa non-standar atau informal, yang dapat bervariasi antara satu pengguna dengan yang lainnya. Oleh karena itu, diperlukan proses standarisasi untuk menormalisasi kata-kata tersebut agar lebih konsisten dan mudah diproses dalam analisis teks.

5) Stopword Removal

Pada tahap *stopword removal* akan menghilangkan kata-kata umum yang dianggap tidak mempunyai arti dan biasanya muncul dalam jumlah besar dalam suatu dokumen teks. Dengan menghilangkan kata-kata umum yang tidak signifikan, proses ini memastikan bahwa analisis difokuskan pada elemen yang lebih relevan dan bermakna. Contoh kata yang dihapus seperti 'dan', 'atau', 'yang' dan sejenisnya sering muncul dalam jumlah besar dalam dokumen teks, namun tidak memberikan kontribusi berarti terhadap analisis semantik.

6) Stemming

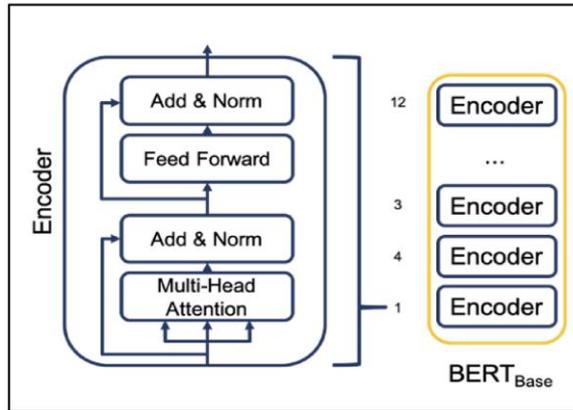
Stemming digunakan untuk mengubah kata menjadi kata dasar [19]. Pada penelitian ini menggunakan *library* Sastrawi yang dimana mempunyai kamus kata dasar Bahasa Indonesia. Proses *stemming* bertujuan untuk menyederhanakan analisis teks dengan mengurangi variasi kata yang memiliki makna dasar yang sama. Sebagai contoh, kata 'menyenangkan' diubah menjadi kata dasar 'senang'. Dengan demikian, *stemming* membantu dalam menyatukan berbagai bentuk kata yang berbeda menjadi satu entitas yang seragam.

2.3 Pembagian Dataset

Pembagian dataset dalam penelitian ini dilakukan dengan rasio 80:20, di mana 80% dari data digunakan sebagai data latih (*training data*), dan 20% sisanya digunakan sebagai data uji (*test data*). Data latih berfungsi untuk melatih model agar mampu mengenali pola dan relasi dalam data yang tersedia, memungkinkan model belajar menghasilkan prediksi yang akurat. Sementara itu, data uji digunakan untuk mengevaluasi performa model yang telah dilatih, memastikan kemampuan generalisasi model terhadap data baru yang tidak dilibatkan dalam proses pelatihan.

2.4 Pelatihan

Proses pelatihan model IndoBERT dalam penelitian ini mengikuti arsitektur yang diperkenalkan dalam penelitian [15]. Model IndoBERT yang digunakan merupakan varian dari BERT yang dioptimalkan untuk bahasa Indonesia, yang memanfaatkan model *transformer* dengan 12 lapisan encoder dan mekanisme *self-attention* untuk menangkap hubungan antar kata dalam sebuah kalimat secara bidirectional. Gambar 3 menunjukkan arsitektur BERT[17].



Gambar 3. Arsitektur BERT [17]

Setiap *encoder* pada dasarnya terdiri dari dua komponen utama: *Multi-Head Attention* dan *Feed Forward Neural Network*, yang dilengkapi dengan mekanisme *Add & Norm* [10]. *Multi-Head Attention* memungkinkan model untuk memperhatikan berbagai bagian dari teks secara bersamaan, sehingga dapat menangkap hubungan antar kata dengan lebih baik. Kedua, lapisan *Add & Norm* berfungsi untuk menstabilkan pelatihan dengan menambahkan *residual connection* dan melakukan normalisasi data, memastikan konsistensi distribusi data. Selanjutnya, *Feed Forward Neural Network* memproses informasi secara non-linear melalui dua lapisan *dense*, memungkinkan model untuk menangkap pola yang lebih kompleks. Terakhir, lapisan *Add & Norm* kedua memastikan hasil dari *Feed Forward Neural Network* distabilkan sebelum diteruskan ke lapisan berikutnya. Kombinasi dari komponen-komponen ini memungkinkan BERT untuk membangun representasi teks yang kaya dan kontekstual, menjadikannya efektif untuk berbagai aplikasi pemrosesan bahasa alami.

Data latih yang digunakan telah melalui proses pra-pemrosesan, termasuk tokenisasi dan konversi ke dalam format numerik yang sesuai, seperti token ID, segment ID, dan positional embedding. Setelah diproses, data ini dimasukkan ke dalam model melalui lapisan embedding yang meliputi *word embeddings*, *position embeddings*, dan *token type embeddings* [20].

Pelatihan dilakukan selama 10 *epoch* dengan menggunakan optimizer Adam dan learning rate yang disetel pada 0.00001 [21], yang dikombinasikan dengan *learning rate scheduling* untuk mengatur pembaruan bobot model selama proses pelatihan. Proses pelatihan dilakukan secara iteratif, di mana model dilatih dengan data latih yang mencakup 80% dari dataset dan hasilnya dievaluasi menggunakan data uji yang mencakup 20% dataset. Untuk memantau performa model, matrik evaluasi seperti akurasi, *precision*, *recall*, dan F1-Score dihitung setiap *epoch*, yang digunakan untuk mengevaluasi

kemampuan model dalam menggeneralisasi data yang tidak terlihat selama pelatihan. Selama proses ini, teknik regularisasi seperti dropout dengan nilai 0.1 diterapkan untuk mencegah *overfitting* dan meningkatkan kemampuan generalisasi model[22].

2.5 Evaluasi Hasil

Proses evaluasi hasil kinerja model klasifikasi menggunakan *confusion matrix*. Setelah model dilatih dan diuji menggunakan, hasil prediksi dibandingkan dengan nilai sebenarnya menggunakan *confusion matrix*. *Confusion matrix* memberikan gambaran rinci tentang jumlah prediksi yang benar dan salah, yang kemudian digunakan untuk menghitung matrik evaluasi kinerja model seperti akurasi, presisi, recall, dan F1-score.

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

$$\text{Presisi} = \frac{TP}{TP+FP} \quad (2)$$

$$\text{Recall} = \frac{TP}{TP+FN} \quad (3)$$

$$\text{F1 - Score} = 2 \times \frac{\text{Presisi} \times \text{Recall}}{\text{Presisi} + \text{Recall}} \quad (4)$$

Komponen Utama *Confusion Matrix*:

1. *True Positive* (TP): Prediksi benar untuk kelas positif.
2. *True Negative* (TN): Prediksi benar untuk kelas negatif.
3. *False Positive* (FP): Prediksi salah, model memprediksi positif padahal seharusnya negatif (Type I error).
4. *False Negative* (FN): Prediksi salah, model memprediksi negatif padahal seharusnya positif (Type II error).

III. HASIL DAN PEMBAHASAN

Penelitian ini, kami menggunakan model IndoBert untuk menganalisis sentimen ulasan pengunjung pada objek wisata Baturaden. Tabel 1 menunjukkan contoh hasil tahap pre-processing. Pelatihan dilakukan selama 10 *epoch* dengan menggunakan optimizer Adam dan learning rate 0.00001, yang dikombinasikan dengan *learning rate scheduling* untuk mengatur pembaruan bobot model selama proses pelatihan. Model dilatih dengan data latih yang mencakup 80% dari dataset dan hasilnya dievaluasi menggunakan data uji yang mencakup 20% dataset Berikut adalah hasil dan pembahasan untuk masing-masing model.

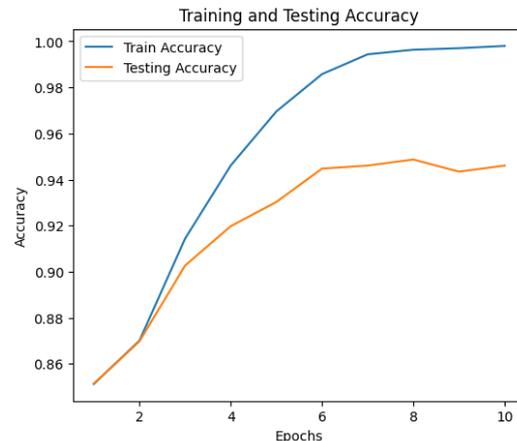
Tabel 1. Contoh Hasil Pre-Processing

Proses	Hasil
Review	Tempatnya sangat luas dengan indahnya air terjun hawanya juga sejuk,banyak pedagang juga dan fasilitas lainnya apalagi ada 2 kolam berenang ini cocok buat tempat liburan keluarga 😊👍🌿
Casefolding	tempatnya sangat luas dengan indahnya air terjun hawanya juga sejuk,banyak pedagang juga dan fasilitas lainnya apalagi ada 2 kolam berenang ini cocok buat tempat liburan keluarga 😊👍🌿
Text Cleaning	tempatnya sangat luas dengan indahnya air terjun hawanya juga sejuk banyak pedagang juga dan fasilitas lainnya apalagi ada kolam berenang ini cocok buat tempat liburan keluarga
Tokenisasi	['tempatnya', 'sangat', 'luas', 'dengan', 'indahny', 'air', 'terjun', 'hawanya', 'juga', 'sejuk', 'banyak', 'pedagang', 'juga', 'dan', 'fasilitas', 'lainnya', 'apalagi', 'ada', 'kolam', 'berenang', 'ini', 'cocok', 'buat', 'tempat', 'liburan', 'keluarga']
Normalisasi	['tempatnya', 'sangat', 'luas', 'dengan', 'indahny', 'air', 'terjun', 'hawanya', 'juga', 'sejuk', 'banyak', 'pedagang', 'juga', 'dan', 'fasilitas', 'lainnya', 'apalagi', 'ada', 'kolam', 'berenang', 'ini', 'cocok', 'buat', 'tempat', 'liburan', 'keluarga']
Stemming	['tempat', 'sangat', 'luas', 'dengan', 'indah', 'air', 'terjun', 'hawa', 'juga', 'sejuk', 'banyak', 'dagang', 'juga', 'dan', 'fasilitas', 'lain', 'apalagi', 'ada', 'kolam', 'renang', 'ini', 'cocok', 'buat', 'tempat', 'libur', 'keluarga']
Stopword Removal	['luas', 'indah', 'air', 'terjun', 'hawa', 'sejuk', 'dagang', 'fasilitas', 'kolam', 'renang', 'cocok', 'libur', 'keluarga']

Gambar 4 menunjukkan akurasi model pada saat pelatihan dan evaluasi menggunakan data uji. Hasil pelatihan menunjukkan nilai akurasi antara data latih dan data uji tidak berbeda jauh, hal ini mengartikan bahwa model dapat belajar dan memahami objek dengan cukup baik. Tabel 2 menunjukkan hasil kinerja model menggunakan matriks akurasi, precision, recall, dan F1-score.

Tabel 2. Rata- rata kinerja prediksi model

Model	Akurasi	Precision	Recall	F1-Score
IndoBert	94,61%	83,22 %	96 %	88,11 %



Gambar 4 Akurasi model pada proses pelatihan dan pengujian

Model yang digunakan memiliki akurasi sebesar 94,61%, F1-Score sebesar 88,22%, precision 83,22%, dan recall 96%. Semakin tinggi tingkat akurasi menunjukkan kemampuan model dalam mengklasifikasikan data secara keseluruhan dengan semakin baik. F1-Score yang mencapai 88,22% mencerminkan keseimbangan antara precision dan recall, yang sangat penting dalam menangani dataset tidak seimbang. Precision sebesar 83,22% menunjukkan bahwa 83% prediksi positif yang dibuat oleh model benar-benar positif, meskipun masih terdapat peluang untuk mengurangi kesalahan pada prediksi positif. Sementara itu, recall yang mencapai 96% menandakan kemampuan model dalam mengidentifikasi hampir seluruh contoh positif dengan akurat, dengan hanya sedikit data positif yang terlewat.

Hasil evaluasi ini dirangkum dalam Tabel 3, yang menggambarkan kinerja model pada analisis sentimen untuk setiap aspek yang diuji. Dalam analisis sentimen berbasis aspek terhadap objek wisata Baturraden, hasil menunjukkan variasi kinerja model pada setiap aspek. Untuk aspek Attraction, sentimen negatif memiliki precision 98%, recall 80%, dan F1-score 88%. Sentimen netral menunjukkan precision 97,10%, recall 97,81%, dan F1-score 97,45%, sedangkan sentimen positif memiliki nilai precision, recall, dan F1-score sebesar 93,75%. Pada aspek Accessibility, sentimen negatif memiliki precision 98,9%, recall 75%, dan F1-score 85,71%. Sentimen netral mencapai precision 96,35%, recall 97%, dan F1-score 97,06%, sementara sentimen positif menunjukkan precision 94%, recall 92%, dan F1-score 93,07%. Untuk aspek Amenities, sentimen negatif memiliki precision 98,80%, recall 62,50%, dan F1-score 76,92%. Sentimen netral menunjukkan precision 93,30%, recall 96,21%, dan F1-score 94,78%, serta sentimen positif dengan precision 87,76%, recall 86%, dan F1-score 86,87%. Terakhir, pada aspek Ancillary Services, sentimen negatif

memiliki *precision* 98,80%, *recall* 25%, dan *F1-score* 40%. Sentimen netral mencapai *precision* 93,48%, *recall* 98%, dan *F1-score* 95%, sedangkan sentimen positif menunjukkan *precision* 96,08%, *recall* 88,09%, dan *F1-score* 92%.

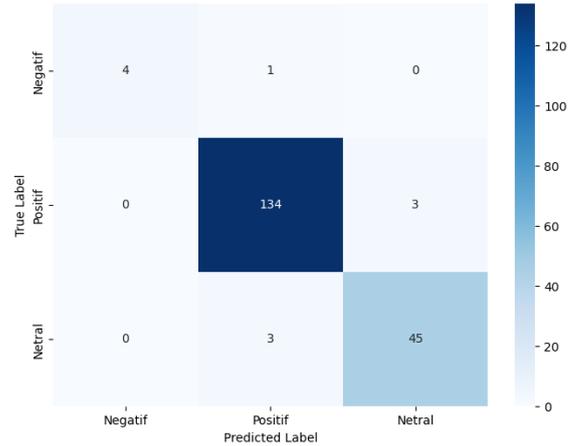
Tabel 3. *Confusion matrix* pada setiap aspek

Aspek	Sentiment	Precision	Recall	F1-Score
Attraction	Negatif	98%	80%	88%
	Netral	97,10%	97,81%	97,45%
	Positif	93,75%	93,75%	93,75%
Accessibility	Negatif	98,9%	75%	85,71%
	Netral	96,35%	97%	97,06%
	Positif	94%	92%	93,07%
Amenities	Negatif	98,80%	62,50%	76,92%
	Netral	93,30%	96,21%	94,78%
	Positif	87,76%	86,00%	86,87%
Ancillary Services	Negatif	98,80%	25%	40%
	Netral	93,48%	98,00%	95%
	Positif	96,08%	88,09%	92%

Jumlah data yang kurang seimbang terlihat pada Gambar 2, mengakibatkan hasil prediksi pada beberapa point aspek mendapat kinerja yang kurang baik. Terlihat pada Tabel 3 aspek *ancillary services* pada sentimen negatif mendapat nilai 25% pada matriks *recall* dan pada matriks *F1-score* mendapatkan 40%. Tidak seimbangnya jumlah data latih mengakibatkan hasil prediksi data menjadi kurang optimal.

3.1 Hasil Evaluasi Model pada Aspek Attraction

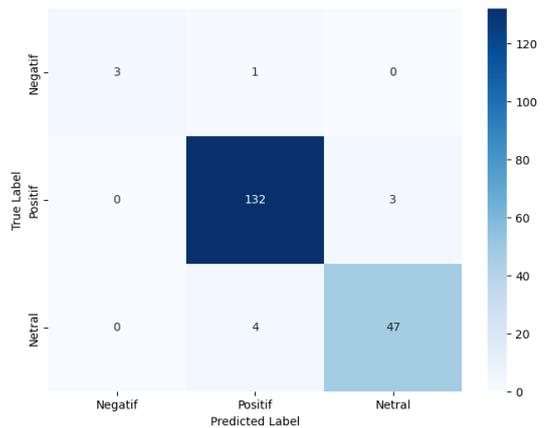
Gambar 5 menunjukkan hasil kinerja model dalam memprediksi sentimen ulasan terkait aspek *attraction* (daya tarik wisata). Akurasi kinerja model mencapai 96,32% mencerminkan kinerja yang baik dalam mengklasifikasikan sentimen pengunjung terhadap daya tarik wisata. Kinerja rata-rata *precision* sebesar 96,95% menunjukkan bahwa model dapat secara akurat mengidentifikasi ulasan yang relevan dengan sentimen yang tepat. Selain itu rata-rata *recall* sebesar 90,52% menandakan bahwa model memiliki kemampuan yang baik dalam mendeteksi semua ulasan yang relevan, meskipun terdapat beberapa ulasan yang terlewatkan. Rata-rata *F1-Score* sebesar 93,36% mengindikasikan keseimbangan yang baik antara *precision* dan *recall*, mendndakan kinerja model dalam menangani variasi sentimen.



Gambar 5. Hasil prediksi model pada aspek *attraction*

3.2 Hasil Evaluasi Model pada Aspek Accessibility

Kinerja model dalam memprediksi sentimen ulasan terkait *accessibility* berhasil mendapat kinerja akurasi senilai 95,79% rata-rata *precision* sebesar 96,78%, rata-rata *recall* sebesar 88,31%, sedangkan untuk rata-rata *F1-score* sebesar 91,95%. Gambar 6 menunjukkan hasil prediksi pada aspek *accessibility*. *Confusion matrix* yang ditampilkan dalam gambar 6 memberikan gambaran tentang distribusi prediksi dan label sebenarnya.

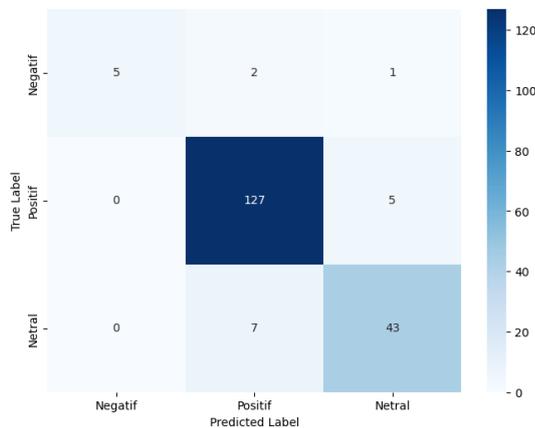


Gambar 6. Hasil prediksi model pada aspek *accessibility*

3.3 Hasil Evaluasi Model pada Aspek Amenities

Kinerja model pada *amenities* mampu mendapatkan hasil 92,11% pada matriks akurasi. Kinerja pada rata-rata *precision* sebesar 93,71%, rata-rata *recall* sebesar 81,57% dan rata-rata *F1-Score* sebesar 86,19%. *Recall* lebih rendah karena kinerja pada kelas negatif yang terbatas. Gambar 7 menunjukkan hasil prediksi pada aspek *amenities*. *Confusion matrix* yang ditampilkan dalam gambar 7

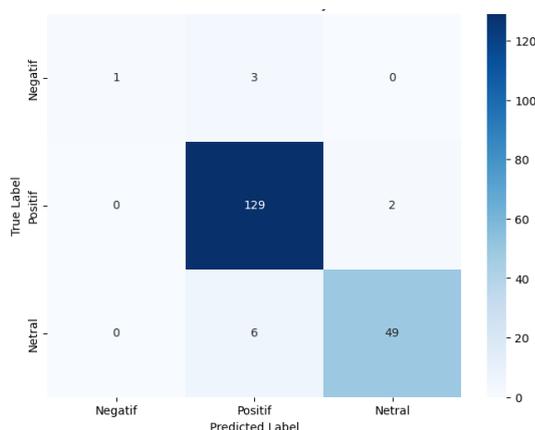
memberikan gambaran tentang distribusi prediksi dan label sebenarnya



Gambar 7. Hasil prediksi model pada aspek amenities

3.4 Hasil Evaluasi Model pada Aspek Ancillary Services

Gambar 8 menunjukkan hasil kinerja model dalam memprediksi sentimen ulasan terkait aspek ancillary service. Confusion matrix yang ditampilkan dalam gambar 8 memberikan gambaran tentang distribusi prediksi dan label sebenarnya. Akurasi kinerja model mencapai 95,79% mencerminkan kinerja yang baik dalam mengklasifikasikan sentimen pengunjung terhadap aspek ancillary service. Kinerja rata-rata precision sebesar 96,78%, rata-rata recall sebesar 88,31% dan rata-rata F1-Score sebesar 95,79%.



Gambar 8. Hasil prediksi model pada aspek ancillary service

IV. PENUTUP

4.1. Kesimpulan

Penelitian ini menghasilkan model berbasis IndoBERT yang efektif untuk melakukan analisis sentimen berbasis aspek terhadap ulasan pengunjung objek wisata Baturraden, difokuskan pada empat aspek utama: *Attraction*, *Accessibility*, *Amenities*, dan

Ancillary Services. Model ini menunjukkan kinerja yang mengesankan dengan akurasi 94,61%, precision 83,22%, recall 96%, dan F1-score 88,11%. Hasil ini menunjukkan model IndoBERT dapat digunakan dalam menganalisis sentimen berdasarkan aspek tertentu. Namun, ketidakseimbangan data, terutama pada aspek dengan sentimen negatif, mengindikasikan bahwa masih ada ruang untuk peningkatan kinerja dalam penelitian mendatang. Tantangan utama yang dihadapi adalah keragaman ulasan dari segi bahasa dan dialek, serta ketidakseimbangan jumlah data pada masing-masing aspek.

4.2. Saran

Penelitian selanjutnya diharapkan dapat menambahkan jumlah variasi data yang digunakan atau mengimplementasikan algoritme yang mampu mengatasi masalah ketidakseimbangan kelas data. Pendekatan ini penting untuk meningkatkan hasil akurasi dan generalisasi model dalam melakukan analisis sentimen. Penggunaan teknik seperti oversampling, undersampling, atau algoritme yang dirancang khusus untuk menangani ketidakseimbangan data, seperti SMOTE atau algoritme ensemble, dapat menjadi solusi untuk penelitian selanjutnya.

DAFTAR PUSTAKA

- [1] BPS Kabupaten Banyumas, "Jumlah Pengunjung Obyek Wisata yang Dikelola oleh Pemerintah Kabupaten Banyumas 2023," <https://banyumaskab.bps.go.id/id/statistics-table/1/NjY5IzE=/jumlah-pengunjung-obyek-wisata-yang-dikelola-oleh-pemerintah-kabupaten-banyumas-2023.html>.
- [2] Ni Luh Putu Merawati, Amrullah Ahmad Zuli, and Ismarmiaty, "Analisis Sentimen dan Pemodelan Topik Pariwisata Lombok Menggunakan Algoritma Naive Bayes dan Latent Dirichlet Allocation," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 5, no. 1, pp. 123–131, 2021, doi: 10.29207/resti.v5i1.2587.
- [3] N. Kresna Diwangkara1, S. R. Sari2, and R. S. Rukayah3, "PENGEMBANGAN PARIWISATA KAWASAN BATURRADEN," *Jurnal Arsitektur ARCADE*, vol. 4, no. 2, pp. 120–128, Jul. 2020, doi: 10.31848/arcade.v4i2.431.
- [4] B. Mathayomchan and K. Sripanidkulchai, "Utilizing Google Translated Reviews from Google Maps in Sentiment Analysis for Phuket Tourist Attractions," in *2019 16th International Joint Conference on Computer Science and Software Engineering (JCSSE)*,

- IEEE, Jul. 2019, pp. 260–265. doi: 10.1109/JCSSE.2019.8864150.
- [5] M. Afrad, C. Febrianto, S. Wijayanto, and Y. Fathoni, “Analysis of Visitor Reviews on Baturaden Tourist Attraction Using Machine Learning Methods,” *Edu Komputika*, vol. 11, no. 1, pp. 57–64, 2024, doi: doi.org/10.15294/edukom.v11i1.10561.
- [6] M. A. Khadija, I. S. D. Jayanti, and F. U. Nimah, “Towards Smart City: Aspect Based Sentiment Analysis of Indonesian Public Aspiration Complaints Data Using Machine Learning,” in *Proceedings - International Conference on Informatics and Computational Sciences*, Institute of Electrical and Electronics Engineers Inc., 2024, pp. 215–220. doi: 10.1109/ICICoS62600.2024.10636859.
- [7] D. Arianto and I. Budi, “Aspect-based Sentiment Analysis on Indonesia’s Tourism Destinations Based on Google Maps User Code-Mixed Reviews (Study Case: Borobudur and Prambanan Temples),” in *Proceedings of the 34th Pacific Asia Conference A Language, Information and Computation*, Oct. 2020, pp. 359–367. [Online]. Available: <https://aclanthology.org/2020.pacific-1.41.pdf>
- [8] U. Nuha and C. H. Lin, “Aspect-Based Sentiment Analysis with Semi-Supervised Approach on Taiwan Social Distancing App User Reviews,” in *5th International Conference on Artificial Intelligence in Information and Communication, ICAIIC 2023*, Institute of Electrical and Electronics Engineers Inc., 2023, pp. 444–447. doi: 10.1109/ICAIIC57133.2023.10067048.
- [9] D. A. K. Khotimah and R. Sarno, “Sentiment analysis of hotel aspect using probabilistic latent semantic analysis, word embedding and LSTM,” *International Journal of Intelligent Engineering and Systems*, vol. 12, no. 4, pp. 275–290, 2019, doi: 10.22266/ijies2019.0831.26.
- [10] J. Devlin, M.-W. Chang, K. Lee, K. T. Google, and A. I. Language, “BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding.” [Online]. Available: <https://github.com/tensorflow/tensor2tensor>
- [11] X. Li, L. Bing, W. Zhang, and W. Lam, “Exploiting BERT for End-to-End Aspect-based Sentiment Analysis,” Oct. 2019, [Online]. Available: <http://arxiv.org/abs/1910.00883>
- [12] A. Aziz, M. A. Khadija, W. Nurharjadmo, D. C. Febrianto, M. R. u. Firmansyah, and R. A. Pratama, “Enhancing Freelancer Project Matching with a BERT-Powered Deep Learning Indonesian Chatbot,” in *International Conference on Computer, Control, Informatics and its Applications, IC3INA*, Institute of Electrical and Electronics Engineers Inc., 2024, pp. 213–218. doi: 10.1109/IC3INA64086.2024.10732347.
- [13] D. S. Bimaputra and E. Sutoyo, “Aspect-Based Sentiment Analysis of Hotels in Bali on Tripadvisor Using BERT Algorithm,” *International Journal of Advances in Data and Information Systems*, vol. 4, no. 2, Apr. 2023, doi: 10.25008/ijadis.v4i2.1284.
- [14] H. Santosa, F. Rachman, S. A. Austen, Christianto, and A. S. Girsang, “IndoBERT for classifying hate speech in Twitter,” in *AIP Conference Proceedings*, American Institute of Physics, Mar. 2024. doi: 10.1063/5.0199750.
- [15] B. Wilie *et al.*, “IndoNLU: Benchmark and Resources for Evaluating Indonesian Natural Language Understanding.” [Online]. Available: <https://github.com/annisanurulazhar/absa-playground>
- [16] F. Koto, A. Rahimi, J. H. Lau, and T. Baldwin, “IndoLEM and IndoBERT: A Benchmark Dataset and Pre-trained Language Model for Indonesian NLP,” Nov. 2020, [Online]. Available: <http://arxiv.org/abs/2011.00677>
- [17] D. Y. Yefferson, V. Lawijaya, and A. S. Girsang, “Hybrid model: IndoBERT and long short-term memory for detecting Indonesian hoax news,” *IAES International Journal of Artificial Intelligence*, vol. 13, no. 2, pp. 1911–1922, Jun. 2024, doi: 10.11591/ijai.v13.i2.pp1913-1924.
- [18] I. R. Hidayat and W. Maharani, “General Depression Detection Analysis Using IndoBERT Method,” *International Journal on Information and Communication Technology (IJoICT)*, vol. 8, no. 1, pp. 41–51, Aug. 2022, doi: 10.21108/ijoiect.v8i1.634.
- [19] N. Yusliani, R. Primartha, and M. Diana, “Multiprocessing Stemming: A Case Study of Indonesian Stemming,” *Int J Comput Appl*, vol. 182, pp. 15–19, Dec. 2019, doi: 10.5120/ijca2019918476.
- [20] E. Yulianti and N. K. Nissa, “ABSA of Indonesian customer reviews using IndoBERT: single-sentence and sentence-pair classification approaches,” *Bulletin of Electrical Engineering and Informatics*, vol. 13, no. 5, pp. 3579–3589, Oct. 2024, doi: 10.11591/eei.v13i5.8032.

- [21] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," Dec. 2014, [Online]. Available: <http://arxiv.org/abs/1412.6980>
- [22] J. H. Computer, S. M. Honova, V. P. Computer, C. A. Setiawan, I. H. Parmonangan, and Diana, "Sentiment Analysis of Skincare Product Reviews in Indonesian Language using IndoBERT and LSTM," in *Proceeding - IEEE 9th Information Technology International Seminar, ITIS 2023*, Institute of Electrical and Electronics Engineers Inc., 2023. doi: 10.1109/ITIS59651.2023.10420222.