Analisis Sentimen Berbasis Aspek dan Emosi pada Ulasan Hotel

Desi Natalia[1)], Viktor Handrianus Pranatawijaya[2)], Nova Noor Kamala Sari[3)]

Jurusan Teknik Informatika, Fakultas Teknik, Universitas Palangka Raya
Kampus Tunjung Nyaho Jl. Yos Sudarso Palangka Raya 73112
[1)]desiinatalia21@mhs.eng.upr.ac.id
[2)]viktorhp@it.upr.ac.id
[3)]novanoorks@it.upr.ac.id

Abstrak

Analisis sentimen terhadap ulasan pelanggan hotel merupakan pendekatan yang penting dalam mengevaluasi kualitas layanan dan memahami persepsi pengguna. Penelitian ini bertujuan untuk mengembangkan sistem klasifikasi yang mampu menganalisis ulasan hotel secara menyeluruh berdasarkan tiga dimensi utama: aspek, sentimen, dan emosi. Dua kategori aspek utama berhasil diidentifikasi melalui metode topik modeling menggunakan *Latent Dirichlet Allocation* (LDA) dan disempurnakan menggunakan *Generative AI*, yaitu aspek *Fasilitas & Lokasi* serta *Kenyamanan & Pelayanan*.Penelitian ini menggunakan tiga arsitektur model pembelajaran mendalam, yaitu *BiLSTM*, *BiLSTM + Attention*, dan *IndoBERT*. Hasil evaluasi menunjukkan bahwa model *IndoBERT* memiliki performa terbaik pada klasifikasi aspek dan sentimen, dengan *accuracy* mencapai 96% untuk aspek dan 98% untuk sentimen, serta *F1-score* masing-masing sebesar 0.96 dan 0.98. Model ini unggul dalam mengenali polaritas dan aspek ulasan, bahkan dalam kondisi distribusi data yang tidak seimbang.
Untuk klasifikasi emosi, model *BiLSTM + Attention* menunjukkan performa paling seimbang, dengan *accuracy* mencapai 91%, serta *precision*, *recall* , dan *F1-score* yang konsisten pada nilai 90–91%. Model ini mampu menangani variasi emosi dengan lebih baik, khususnya dalam mengenali kelas minoritas seperti *percaya* dan *sedih*, meskipun masih menghadapi tantangan pada emosi seperti *takut* yang memiliki jumlah data sangat sedikit.

Kata kunci: Analisis Sentimen, Emosi, Aspek, Ulasan Hotel, *IndoBERT*, *BiLSTM*, *Attention*.

Abstract

Sentiment analysis of hotel customer reviews is an essential approach to evaluating service quality and understanding user perceptions. This study aims to develop a classification system capable of analyzing hotel reviews from three perspectives: aspect, sentiment, and emotion. Two primary aspect categories were extracted using topic modeling with Latent Dirichlet Allocation (LDA) and refined with the help of Generative AI, *resulting in the categories* Facilities & Location *and* Comfort & Service.
Three deep learning architectures were employed in this research: BiLSTM, BiLSTM + Attention, and IndoBERT. Evaluation results showed that IndoBERT outperformed the other models in aspect and sentiment classification tasks, achieving 96% accuracy for aspect and 98% for sentiment, with F1-scores of 0.96 and 0.98, respectively. IndoBERT demonstrated strong capabilities in recognizing review polarity and aspect orientation, even under imbalanced data conditions.

DOI:https://doi.org/10.47111/jointecoms.v6i3 *Received: 01-09-2026*
Accepted: 10-09-2026

In the emotion classification task, the BiLSTM + Attention model delivered the most balanced performance, with an overall accuracy of 91%, and precision, recall , and F1-score values consistently ranging between 90–91%. This model effectively handled a wide range of emotional expressions, especially for minority classes such as trust *and* sadness*, although it still faced challenges in identifying emotions with very limited data, such as* fear.*The integration of aspect- and emotion-based sentiment analysis provides a more in-depth understanding of customer experiences. The proposed system not only detects general sentiment but also links it to specific service aspects and emotional expressions, offering valuable insights for context-aware service improvements in the hospitality industry.*

Keywords: *Sentiment Analysis, Emotion, Aspect, Hotel Reviews, IndoBERT, BiLSTM, Attention*

1. PENDAHULUAN

Perhotelan berperan penting dalam pertumbuhan ekonomi, terutama di kota dengan sektor pariwisata yang berkembang [13]. Perkembangan teknologi digital mengubah cara pelanggan memilih hotel, dengan ulasan daring dari platform seperti Traveloka dan Tiket.com menjadi faktor utama dalam pengambilan keputusan [17].

Kota Palangka Raya sebagai ibu kota Kalimantan Tengah menunjukkan peningkatan sektor pariwisata dan bisnis perhotelan [5]. Namun, masih sedikit penelitian yang secara khusus menganalisis ulasan pelanggan hotel di wilayah ini menggunakan pendekatan analisis sentimen berbasis aspek dan emosi [9].

Ulasan yang beragam sering kali mengandung opini dengan sentimen dan emosi berbeda, sehingga sulit dianalisis secara manual. Oleh karena itu, diperlukan metode otomatis seperti *Aspect-Based Sentiment Analysis (ABSA***)** dan deteksi emosi untuk memahami kepuasan pelanggan secara efektif [7] [11].

Penelitian ini mengombinasikan tiga model: *IndoBERT, Bi-LSTM, dan Bi-LSTM + Attention*, untuk memahami hubungan antar kata dan mengidentifikasi sentimen serta emosi pada aspek-aspek seperti fasilitas, layanan, dan kondisi kamar [3]. Studi ini melengkapi penelitian sebelumnya dan berkontribusi dalam pengembangan teknik NLP untuk industri perhotelan, khususnya di kota Palangka Raya yang belum banyak diteliti dalam konteks ini [15] [14].

2. TINJAUAN PUSTAKA

2.1 Analisis Sentimen Berbasi Aspek

Aspect-Based Sentiment Analysis (ABSA) adalah pendekatan yang lebih mendalam dalam analisis sentimen, yang tidak hanya mengklasifikasikan sentimen secara umum, tetapi juga mengidentifikasi sentimen terhadap aspek-aspek spesifik dari layanan atau produk yang dibahas. Dengan demikian, *ABSA* memberikan wawasan yang lebih terperinci tentang aspek-aspek tertentu yang memengaruhi opini pengguna [16].

2.2 Deteksi Emosi

Teori-teori terbaru mengenai emosi yang dikembangkan oleh Cowen dan Keltner [2] mendukung pemilihan emosi ini, menunjukkan bahwa emosi-emosi tersebut lebih relevan dalam konteks ulasan hotel, sementara terkejut dan antisipasi lebih jarang diungkapkan dalam ulasan yang fokus pada pengalaman yang telah terjadi.

2.3 *Deep Learning*

Deep Learning adalah model komputasi untuk mempelajari representasi data yang baik. Model komputasi biasanya berupa jaringan saraf. Baru-baru ini, pembelajaran mendalam telah meraih keberhasilan luar biasa dalam menjalankan tugas-tugas seperti pengenalan gambar dan ucapan, serta berpotensi memengaruhi setiap bidang sains. Artikel ini bertujuan untuk membuat pembelajaran mendalam dapat diakses oleh khalayak yang lebih luas, termasuk orang-orang yang berpikiran teknis tetapi bukan ahli dalam bidang pemrograman atau kurang paham matematika [11].

2.4 Attention Mechanism

Mekanisme *Attention* adalah konsep dalam kecerdasan buatan yang memungkinkan model untuk fokus pada bagian paling relevan atau penting dari input saat memproses data sekuensial. Mekanisme ini pertama kali diperkenalkan dalam konteks penerjemahan mesin untuk mengatasi keterbatasan model *encoder-decoder* tradisional yang sulit mempertahankan informasi penting dalam urutan panjang. Proses ini memungkinkan jaringan saraf untuk menangkap hubungan antar kata dalam kalimat yang panjang, meningkatkan akurasi dalam berbagai tugas pemrosesan bahasa alami (NLP), seperti penerjemahan mesin, analisis sentimen, dan rangkuman teks [4].

2.5 Topic Modelling Latent Dirichlet Allocation (LDA)

Latent Dirichlet Allocation (LDA) adalah model probabilistik generatif yang digunakan dalam pemodelan topik untuk menemukan pola tersembunyi dalam kumpulan dokumen teks. LDA membentuk model topik, di mana setiap ulasan dipetakan ke dalam beberapa kategori berdasarkan distribusi probabilitas kata.

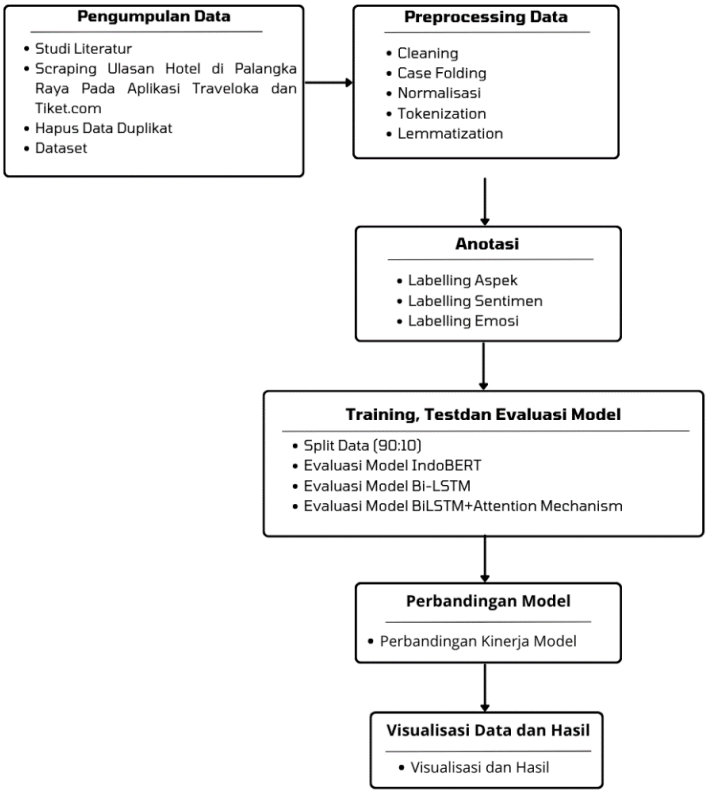
Keunggulan utama LDA adalah kemampuannya untuk melakukan pelabelan secara otomatis tanpa memerlukan anotasi manual, serta dapat menangkap hubungan semantik antar kata meskipun tidak identik [6].

Dengan cara ini, *LDA* membantu mengidentifikasi pola dan hubungan tersembunyi antara kata dan topik dalam kumpulan dokumen [1].

2.6 *Web Scraper*

Web Scraper adalah ekstensi *Chrome* populer untuk membuat peta situs untuk situs web dan mengekstrak data terstruktur . Fitur: Antarmuka yang mudah digunakan, mendukung pagination, mengekspor data dalam format CSV [8].

3. METODE PENELITIAN



Gambar 1. Metodologi Penelitian

Pada gambar diatas ini, menunjukkan alur penelitian yang akan dilakukan. Proses penelitian terdiri dari beberapa tahap utama yang dimulai dari pengumpulan data, diikuti oleh *preprocessing* data untuk memastikan data siap digunakan. Selanjutnya, dilakukan proses klasifikasi untuk mengidentifikasi aspek, sentimen, dan emosi dalam ulasan. Setelah itu, model dievaluasi menggunakan berbagai metode untuk menentukan performa terbaik. Terakhir, hasil analisis divisualisasikan untuk menyajikan temuan penelitian secara informatif.

4. PEMBAHASAN

4.1 Pengumpulan Data

Proses pengumpulan data yang dilakukan, diperoleh data ulasan hotel di Kota Palangka Raya melalui teknik *web scraping* dengan menggunakan ekstensi *Web Scraper* pada peramban *Google Chrome*. Sumber data berasal dari dua platform pemesanan hotel yaitu Traveloka dan Tiket.com. Proses scraping dilakukan pada 30 hotel yang tersedia di kedua platform tersebut, 30 hotel dari Traveloka dan 29 hotel dari Tiket.com. Secara keseluruhan, jumlah data yang berhasil dikumpulkan mencapai 16.000 ulasan, yang terdiri dari 9.934 ulasan dari Traveloka dan 6.066 ulasan dari Tiket.com. Rentang waktu pengambilan data mencakup ulasan yang dipublikasikan mulai tahun Januari 2017 hingga Maret 2025.

4.2 *Preprocessing* Data

4.2.1 Data *Cleaning*

Data Cleaning atau pembersihan data dimaksudkan untuk menerjemahkan teks mentah ke dalam format bersih dan menganalisisnya. Fungsi utama yang digunakan, yaitu "clean_text

(konten)," dirancang untuk mengambil entri dalam bentuk teks (*string*) dan kembali dalam bentuk teks yang dibersihkan.

4.2.2 *Case Folding*

Case folding dilakukan dengan mengubah semua huruf dalam teks menjadi huruf kecil, misalnya kata "Hotel" dan "hotel" akan dianggap sebagai satu kata yang sama.

4.2.3 *Normalization*

Proses *Normalization* (Normalisasi) teks untuk mengimbangi bentuk kata-kata dalam teks sesuai dengan format standar, terutama dengan mengganti kata atau singkatan non-standar dengan kamus bahasa gaul, sesuai dengan format standar.

4.2.4 *Tokenization*

Tokenization (tokenisasi) adalah proses memecah teks menjadi unit kata atau frasa yang lebih kecil. Dalam algoritma ini, digunakan model bahasa spaCy `xx_ent_wiki_sm` yang mampu mengenali dan memproses teks dalam Bahasa Indonesia. Hasil dari proses ini adalah kolom baru bernama `tokens_content` yang berisi daftar kata (token) dari setiap teks, yang siap digunakan untuk tahap analisis lanjutan seperti klasifikasi, pencocokan, atau pembentukan fitur.

4.2.5 *Lemmatization*

Lemmatization adalah proses linguistik yang bertujuan mengubah setiap kata dalam teks ke bentuk dasarnya agar konsisten dan lebih mudah dianalisis, misalnya kata “kebersihan”, “membersihkan” akan dikembalikan ke bentuk dasar “bersih”. Hasil akhirnya adalah DataFrame yang memiliki dua kolom tambahan: `lemmatized_content` berupa daftar kata dasar, dan `ulasan` berupa teks lengkap yang telah terlepas dari bentuk kata terikat atau turunan.

4.3 Anotasi Aspek

4.3.1 *Bag of Words*

Bag of Words (BoW) digunakan untuk mengubah teks menjadi bentuk angka menggunakan metode frekuensi kata. Proses ini dilakukan agar komputer bisa memahami teks dan mengolahnya.

```
[['hotel', 'nyaman', 'dan', 'bersih', 'sayang', 'makan', 'di', 'resto', 'kurang', 'banyak', 'pilih']]
[[(0, 1), (1, 1), (2, 1), (3, 1), (4, 1), (5, 1), (6, 1), (7, 1), (8, 1), (9, 1), (10, 1)]]
```

Gambar 2. Contoh Hasil BoW dalam 1 kalimat teks ulasan

4.3.2 Pemodelan *Latent Dirichlet Allocation (LDA)*

Latent Dirichlet Allocation (LDA) adalah proses untuk menentukan jumlah topik terbaik. Proses ini menggunakan metrik *coherence score* untuk menilai kualitas topik yang dihasilkan. Coherence score mengukur seberapa erat hubungan antar kata dalam satu topik semakin tinggi nilainya, semakin relevan kata-kata tersebut dalam satu konteks.

4.3.3 Visualisasi dan Interpretasi Topik

Langkah selanjutnya, memvisualisasikan hasil pemodelan topik LDA (*Latent Dirichlet Allocation*) menggunakan pustaka *pyLDAvis*, yang sangat membantu dalam memahami struktur topik yang dihasilkan. Pertama, dilakukan impor pustaka yang dibutuhkan untuk visualisasi interaktif. Kemudian, model LDA dengan nilai *coherence* tertinggi dipilih dari daftar model yang telah dibuat sebelumnya. Pemilihan ini dilakukan dengan mencari indeks dari nilai *coherence* tertinggi, lalu mengambil model yang sesuai dari daftar model. Setelah itu, visualisasi topik dilakukan menggunakan fungsi prepare() dari modul *gensimvis*, yang menggabungkan model LDA, korpus dalam bentuk bag-of-words, dan kamus kata. Hasilnya adalah representasi visual interaktif dari topik-topik yang dihasilkan, yang dapat digunakan untuk eksplorasi lebih lanjut, seperti melihat distribusi kata dalam tiap topik atau hubungan antar topik.

4.3.4 Identifikasi Topik Dominan Pada Ulasan

Proses identifikasi ini bertujuan untuk menggabungkan hasil ekstraksi topik dari model LDA (Latent Dirichlet Allocation) dengan data ulasan hotel yang tersedia. Data ulasan mencakup informasi seperti teks ulasan asli, versi teks yang sudah dilemmatized (dinormalisasi), nama hotel, dan waktu ulasan. Setelah data dimuat, setiap ulasan dianalisis untuk mengetahui topik apa yang paling dominan di dalamnya.

4.3.5 Identifikasi Aspek Berdasarkan Topik Pada Ulasan

Setelah proses ekstraksi topik menggunakan model LDA, dilakukan interpretasi topik untuk menentukan aspek yang sesuai. Dengan bantuan model bahasa Gemini, setiap topik dianalisis berdasarkan kata kunci dominan dan dikelompokkan ke dalam dua aspek utama, yaitu "Fasilitas & Layanan" dan "Kondisi Kamar".

4.4 Anotasi Sentimen

Analisis sentimen menggunakan model pre-trained *IndoBERT* yang dirancang khusus untuk Bahasa Indonesia. Proses analisis sentimen dimulai dengan tokenisasi teks menggunakan tokenizer *IndoBERT*, yang mengubah teks menjadi token yang dapat diproses oleh model. Selanjutnya, model *IndoBERT* melakukan prediksi sentimen, mengklasifikasikan teks ke dalam dua kategori: Positif atau Negatif.

4.5 Anotasi Emosi

Proses analisis emosi dilakukan dengan memanfaatkan NRC Emotion Lexicon versi Bahasa Indonesia untuk mengidentifikasi emosi yang terkandung dalam teks ulasan pelanggan. Setiap teks dianalisis berdasarkan sentimen yang telah ditentukan sebelumnya (positif atau negatif), kemudian dicocokkan dengan kata-kata yang terdapat dalam leksikon emosi. Emosi yang dianalisis meliputi bahagia, percaya, marah, jijik, sedih, dan takut, di mana setiap kata yang cocok akan meningkatkan skor untuk emosi terkait.

4.5.1 Analisis Ulasan ber Emosi Negatif Per Aspek Untuk Mendapatkan *Insight*



Gambar 3. *Wordcloud* Aspek Fasilitas Layanan Pada Ulasan Ber Emosi Negatif

```
10 kata yang paling sering muncul di Fasilitas & Layanan:
hotel: 216
kamar: 153
layan: 141
makan: 86
tempat: 61
fasilitas: 61
harga: 60
parkir: 59
jauh: 48
lokasi: 44

Masalah umum yang perlu diperbaiki: ['layan', 'makan']
```

Gambar 4. Kata Yang Sering Muncul Pada Aspek Fasilitas & Layanan

Tabel 1. Contoh Ulasan Untuk Masalah Yang Perlu Diperbaiki

Contoh Ulasan	Hotel	Aspek	Sentimen	Emosi
Pelayanan dari petugas kurang nyaman	M Bahalap Hotel Palangka Raya	Fasilitas & Layanan	Negatif	Marah

Tempat makan menu makan kurang variative kurang enak	Luwansa Hotel Palangka Raya	Fasilitas & Layanan	Negatif	Marah

4.6 Pemodelan dan Evaluasi

4.6.1 *Split* Data

Pembagian data dengan rasio 90:10 diterapkan untuk ketiga model yang digunakan dalam penelitian ini, yaitu *IndoBERT, BiLSTM, dan BiLSTM+Attention.* Proses split ini dilakukan secara terpisah untuk masing-masing tugas anotasi Aspek, Sentimen, dan Emosi dengan tujuan agar proses pelatihan dan evaluasi berjalan secara optimal dan adil. Sebanyak 90% data digunakan sebagai data latih untuk membangun dan mengoptimalkan model, sementara 10% sisanya sebagai data uji untuk mengukur performa model terhadap data yang belum pernah dipelajari sebelumnya.

4.6.2 Pemodelan *IndoBERT*

Secara keseluruhan, meskipun model tampil sangat baik untuk kelas yang dominan, performanya masih perlu ditingkatkan untuk emosi yang jarang muncul.

4.6.3 Pemodelan *BiLSTM*

Secara keseluruhan, model BiLSTM memiliki performa yang kuat dalam mengenali emosi yang umum dan sering muncul, namun masih menghadapi tantangan besar dalam mendeteksi emosi minoritas.

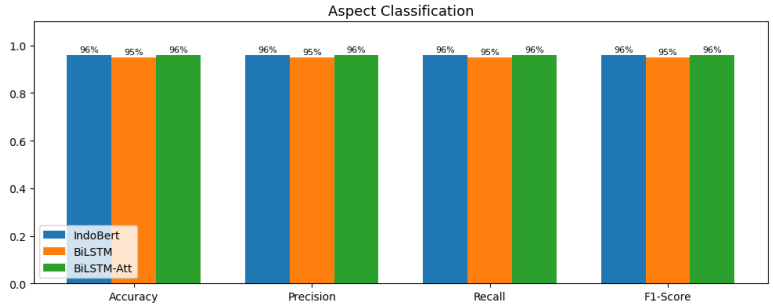
4.6.4 Pemodelan *BiLSTM+Attention*

Secara keseluruhan, meskipun model BiLSTM + Attention menunjukkan performa unggul pada emosi mayoritas, tetap dibutuhkan strategi peningkatan untuk menangani ketidakseimbangan kelas.

4.7 Perbandingan Model

Pada tahap ini, dilakukan evaluasi terhadap kinerja ketiga model untuk mengidentifikasi model dengan performa paling unggul. Evaluasi ini didasarkan pada metrik-metrik seperti *accuracy, precision, recall , dan f1-score,* dengan menggunakan rasio pembagian data 90:10 untuk masing-masing tugas klasifikasi: *aspect*, *sentiment,* dan *emotion.*

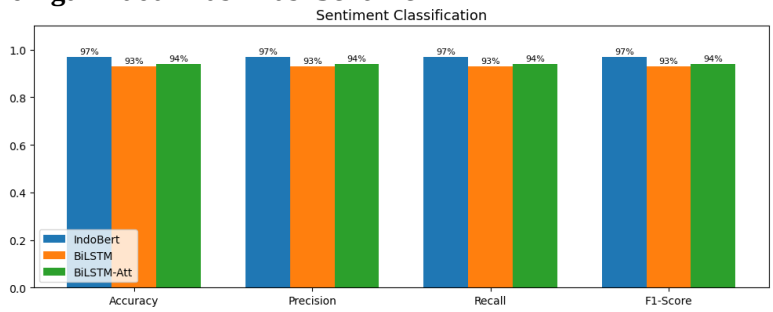
4.7.1 Perbandingan Pada Klasifikasi Aspek



Gambar 5. *Bar chart* ketiga model untuk *Aspect*

Gambar di atas menunjukkan perbandingan metrik evaluasi — *Accuracy*, *Precision*, *Recall* , dan *F1-score* — untuk tiga model berbeda: *IndoBERT, BiLSTM,* dan *BiLSTM+Attention.* Dari hasil tersebut, terlihat bahwa model *IndoBERT* unggul secara konsisten dalam semua metrik, dengan nilai 96% untuk seluruh metrik. Sementara itu, *BiLSTM* dan *BiLSTM+Attention* memiliki performa yang sama, yakni 95% di semua metrik, yang menunjukkan performa yang baik namun masih sedikit di bawah *IndoBERT.*

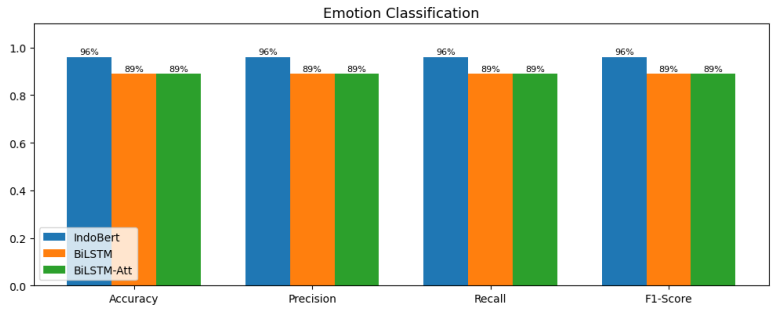
4.7.2 Perbandingan Pada Klasifikasi Sentimen



Gambar 6. *Bar chart* ketiga model untuk *Sentiment*

Berdasarkan visualisasi bar chart untuk klasifikasi sentimen, model *IndoBERT* secara konsisten menunjukkan performa terbaik dibandingkan dengan *BiLSTM* dan *BiLSTM + Attention*. *IndoBERT* mencatat skor tertinggi pada seluruh metrik evaluasi — *Accuracy*, *Precision*, *Recall* , dan *F1-score* — masing-masing sebesar 97%, menunjukkan bahwa model ini sangat andal dalam mengenali dan mengklasifikasikan polaritas sentimen dengan akurasi tinggi.

4.7.3 Perbandingan Pada Klasfikasi Emosi

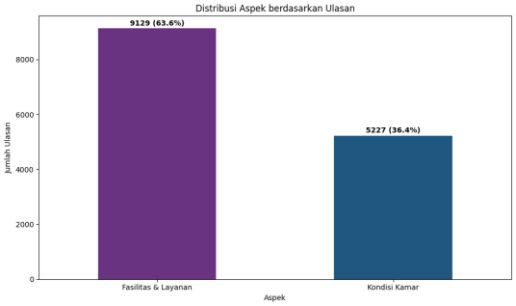


Gambar 7. *Bar chart* ketiga model untuk *Emotion*

Pada grafik bar chart klasifikasi emosi di atas, terlihat bahwa ketiga model — *IndoBERT*, *BiLSTM*, dan *BiLSTM + Attention* — memiliki performa yang relative seimbang. Metrik *accuracy*, precision, *recall* , dan *F1-score* berkisar antara 89% hingga 91%, dengan BiLSTM + Attention sedikit unggul pada sebagian besar metrik.

4.8 Visualisasi dan Hasil

4.8.1 Distribusi *Aspect*

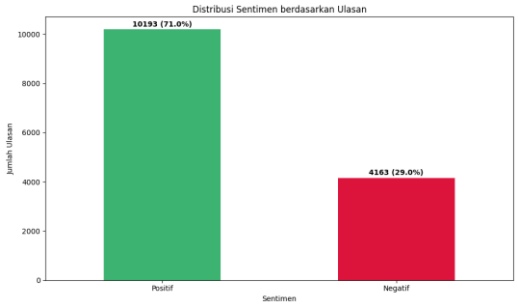


Gambar 8. Distribusi apek berdasarkan ulasan

Tabel 2. Jumlah dan persentase aspek berdasrkan ulasan

Aspect	Jumlah Ulasan	Persentase
Fasilitas & layanan	9129	63.6%
Kondisi Kamar	5227	36.4%

4.8.2 Distribusi *Sentiment*

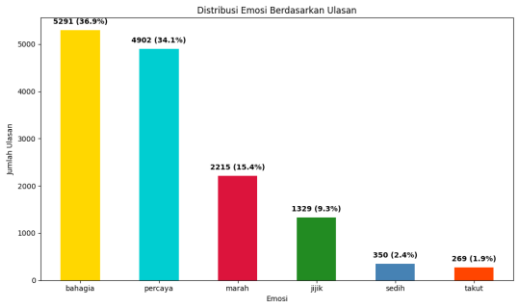


Gambar 8. Distribusi sentiment berdasarkan ulasan

Tabel 3. Jumlah dan persentase sentimen per aspek

Sentiment	**Jumlah Ulasan**	**Persentase**
Positif	10193	71.0%
Negatif	4163	29.0%

4.8.3 Distribusi *Emotion*



Gambar 9. Distribusi berdasarkan ulasan

Tabel 4. Jumlah dan persentase emosi berdasarkan ulasan

Emotion	**Jumlah Ulasan**	**Persentase**
Bahagia	5291	39.9%
Percaya	4902	34.1%
Marah	2215	15.4%
Jijik	1329	9.3%
Sedih	350	2.4%
Takut	269	1.9%

4.8.4 Visualisasi Mengginalan *Streamlit*

Dalam penelitian ini, untuk memudahkan penyampaian hasil analisis secara interaktif, dibangun sebuah dashboard visualisasi menggunakan *Streamlit*, yaitu sebuah *framework* berbasis *Python* yang memungkinkan pembuatan aplikasi web dengan cepat dan sederhana.

5. KESIMPULAN

Penelitian ini menunjukkan bahwa model IndoBERT unggul dibandingkan BiLSTM dan BiLSTM + Attention dalam klasifikasi aspek, sentimen, dan emosi, dengan metrik evaluasi tinggi, terutama pada klasifikasi aspek dan sentimen yang mencapai akurasi hingga 97%. Hasil topik modeling LDA dan pemetaan Generative AI mengelompokkan aspek ulasan ke dalam dua kategori utama, yaitu Fasilitas & Layanan serta Kondisi Kamar. IndoBERT mampu mengidentifikasi polaritas ulasan dengan akurasi tinggi meski data tidak seimbang, serta memberikan hasil paling seimbang dalam klasifikasi emosi, khususnya untuk kelas minoritas seperti sedih, meski tantangan tetap ada pada emosi dengan data rendah seperti takut. Pendekatan

gabungan berbasis aspek dan emosi memberikan wawasan mendalam dan kontekstual, menggambarkan baik perasaan pelanggan maupun aspek layanan yang menjadi sasaran emosi tersebut.

6. DAFTAR PUSTAKA

[1] Analytics Vidhya. (2021). *Step-by-Step Guide to Master NLP: Topic Modelling using LDA (Probabilistic Approach)*.

[2] Cowen, A. S., & Keltner, D. (2023). Semantic Space Theory: A New Approach to Emotions. *Journal of Emotion Studies*, 22(3), 231-245.

[3] Devlin, J., Chang, M-W., Lee, K., & Toutanova, K. (2020). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. Proceedings of NAACL, 4171-4186.

[4] GeeksforGeeks, "Attention mechanism," GeeksforGeeks, 2021. \[Online]. Available: [https://www.geeksforgeeks.org/attention-mechanism/] (https://www.geeksforgeeks.org/attention-mechanism/). \[Accessed: Aug. 11, 2025].

[5] Kusuma, A., & Wijaya, R. (2022). Hotel Industry Growth in Developing Cities: A Case Study of Palangka Raya. Journal of Tourism Studies, 45(2), 210-225.

[6] Lewis, C.M., Grossetti, F. (2022). A Statistical Approach for Optimal Topic Model Identification. *Journal of Machine Learning Research*, 23.

[7] Medhat, W., Hassan, A., & Korashy, H. (2024). Sentiment Analysis Algorithms and Applications: A Comprehensive Review. Artificial Intelligence Journal, 29(1), 1-23.

[8] PromptCloud. (2024, Juni 15). How to Use Web Scraper Chrome Extension to Extract Data. *PromptCloud Blog*, *1*(1), 1-5.

[9] Medhat, W., Hassan, A., & Korashy, H. (2024). Sentiment Analysis Algorithms and Applications: A Comprehensive Review. Artificial Intelligence Journal, 29(1), 1-23.

[10] Putra, T., Sari, M., & Nugroho, R. (2023). Aspect-Based Sentiment Analysis for Hotel Reviews Using Deep Learning. International Journal of Data Science, 18(3), 112-127.

[11] Sharma, D. (2020). Deep Learning without Tears: A Simple Introduction. *Resonance*, 25(1), pp. 15-32.

[12] Sharma, D., et al. (2021). *Emotion Detection in Customer Reviews Using Deep Learning*. Journal of Artificial Intelligence and Data Mining

[13] Smith, J. (2020). The Role of Hospitality Industry in Economic Growth. *Tourism Economics*, 26(4), 547-564.

[14] Sun, C., Huang, L., & Qiu, X. (2020). Utilizing BERT for Aspect-Based Sentiment Analysis via Constructing Auxiliary Sentence. *NAACL-HLT*, 380-385.

[15] Yang, L., Zhang, X., & Li, Y. (2020). *Natural Language Processing* for Hotel Review Analysis: Methods and Applications. Journal of Computational Linguistics, 35(4), 345-359.

[16] Zhang, H., Li, Y., Liu, Y., & Luo, X. (2020). The impact of online reviews on hotel performance: A meta-analysis. *Tourism Management*, 79, 104099.

[17] Zhang, Y., Wang, H., & Zhang, J. (2021). The Impact of Online Reviews on Hotel Performance. *Journal of Travel Research*, 60(5), 1024-1039.