JEMSI: Jurnal Ekonomi Manajemen Sistem Informasi
E-ISSN: 2686-5238
P-ISSN: 2686-4916
https://dinastirev.org/JEMSI dinasti.info@gmail.com +62 811 7404 455
DOI: https://doi.org/10.38035/jemsi.v7i3
https://creativecommons.org/licenses/by/4.0/
Received: 02 Januari 2026 | **Revised:** 20 Januari 2026 | **Accepted:** 19 Februari 2026

Analisis Sentimen berbasis Aspek pada Ulasan Hotel menggunakan Metode BERT untuk Rekomendasi Layanan

Muhammad Tamim Shidqi[1], Yulistia Yulistia[2]
[1]Universitas Multi Data Palembang, Sumatera Selatan, Indonesia, mtamimshidqi27@gmail.com
[1]Universitas Multi Data Palembang, Sumatera Selatan, Indonesia, yulistia@mdp.ac.id

Corresponding Author: mtamimshidqi27@gmail.com[1]

Abstract: *The hospitality sector is confronted with major obstacles in overseeing its online reputation across multiple Online Travel Agent (OTA) platforms. The subject of this research is challenged by the manual examination of countless guest reviews, a task recognized as inefficient and susceptible to bias. This study seeks to implement an automated aspect-based sentiment analysis utilizing Deep Learning techniques. We utilized the BERT (Bidirectional Encoder Representations from Transformers) algorithm with the xlm-roberta-base-sentiment pre-trained model. The evaluation results indicate that the model achieves an accuracy rate of 57% along with a negative sentiment recall of 76%. This figure demonstrates a strong capability in identifying guest grievances. The aspect-based assessment highlights that "Bathroom" and "Facilities" receive the most significant negative sentiment. A root cause analysis points to keywords such as "smell" (bau), "dirty" (kotor), and "hot" (panas) as the primary factors leading to complaints. These insights provide a foundation for strategic recommendations for management to enhance operations.*

Keywords: *Sentiment Analysis, Aspect-Based, BERT, Model Evaluation, Service Recommendation.*

Abstrak: Industri perhotelan menghadapi tantangan besar dalam mengelola reputasi di berbagai platform Online Travel Agent (OTA). Objek penelitian ini menghadapi kendala dalam menganalisis ribuan ulasan tamu yang terus bertambah secara manual. Proses tersebut dinilai tidak efisien dan rentan terhadap subjektivitas. Penelitian ini bertujuan melakukan analisis sentimen otomatis berbasis aspek menggunakan metode Deep Learning. Penulis menerapkan algoritma BERT (Bidirectional Encoder Representations from Transformers) dengan model pre-trained xlm-roberta-base-sentiment. Hasil evaluasi menunjukkan model mencapai akurasi 57% dengan nilai Recall sentimen negatif sebesar 76%. Nilai ini mengindikasikan bahwa model memiliki sensitivitas tinggi dalam mendeteksi keluhan tamu. Analisis aspek mengungkapkan bahwa "Kamar Mandi" dan "Fasilitas" mendapat sentimen negatif tertinggi. Analisis akar masalah menemukan kata kunci "bau", "kotor", dan "panas" sebagai pemicu utama keluhan. Temuan ini menjadi dasar rekomendasi strategis bagi manajemen untuk perbaikan operasional.

Kata Kunci: Analisis Sentimen, Berbasis Aspek, BERT, Evaluasi Model, Rekomendasi Layanan.

PENDAHULUAN

Perkembangan teknologi informasi dan internet telah membawa transformasi signifikan pada industri perhotelan, di mana reputasi bisnis kini sangat bergantung pada jejak digital. Ulasan *online* atau *online reviews* pada platform seperti Google Reviews, Traveloka, dan Agoda telah menjadi bentuk *electronic word-of-mouth* (e-WOM) yang memiliki pengaruh vital terhadap keputusan calon pelanggan (Danuwangsa dkk., 2025). Sebuah studi (Ratu Bilqis dkk., 2024) menemukan bahwa ulasan daring memiliki pengaruh positif dan signifikan terhadap minat beli hotel, di mana mayoritas konsumen menjadikannya referensi utama sebelum melakukan pemesanan. Bagi manajemen hotel, ulasan ini merupakan *feedback* berharga untuk mengevaluasi kualitas layanan. Namun, *volume* ulasan yang terus meningkat menciptakan tantangan tersendiri. Membaca dan menganalisis ribuan ulasan secara manual membutuhkan sumber daya yang besar dan rentan terhadap subjektivitas serta inkonsistensi (Bayu Baskoro dkk., 2021).

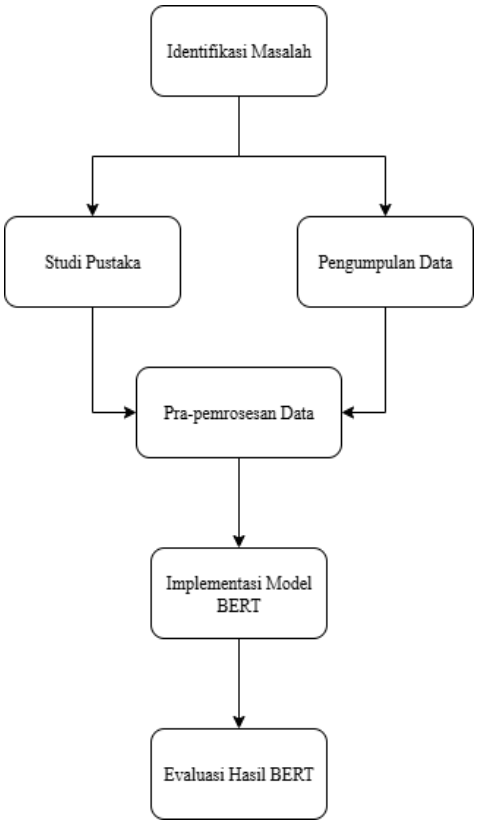
Permasalahan ini juga dihadapi oleh objek penelitian ini, Hotel XYZ, yang mengalami kesulitan dalam mengelola ulasan tamu yang terus bertambah. Analisis manual yang dilakukan saat ini dinilai tidak efisien dan sering kali gagal menangkap konteks spesifik dari keluhan tamu. Selain itu, indikator rating bintang sering kali bersifat ambigu dan tidak cukup untuk merepresentasikan pengalaman tamu secara utuh. Sebuah penilaian bintang tiga dapat menyembunyikan keluhan spesifik mengenai kebersihan kamar yang tertutup oleh kepuasan terhadap lokasi strategis. Oleh karena itu, diperlukan metode analisis sentimen otomatis yang mampu membedah opini pelanggan hingga ke level aspek layanan.

Analisis sentimen, atau *opinion mining*, adalah cabang dari *Natural Language Processing* (NLP) yang bertujuan untuk mengidentifikasi dan mengekstraksi informasi subjektif dari teks. Dalam beberapa tahun terakhir, pendekatan *Deep Learning* telah menunjukkan performa *superior* dalam tugas ini. Analisis sentimen dilakukan untuk melihat pendapat atau kecenderungan opini publik terhadap sebuah isu atau objek, apakah pandangan tersebut cenderung positif atau negatif (Yunitasari Yessi, 2023). Salah satu terobosan terpenting adalah metode BERT (*Bidirectional Encoder Representations from Transformers*) yang diperkenalkan oleh Google dan resmi dirilis secara global pada tahun 2018 (Alfarobby & Irawan, 2024). Berbeda dengan model konvensional yang memproses teks satu arah, BERT menggunakan mekanisme *attention* untuk memahami konteks kata dari dua arah (kiri-ke-kanan dan kanan-ke-kiri) secara bersamaan (Ravichandiran, 2021). Keunggulan ini membuat BERT mampu memahami nuansa bahasa yang kompleks, termasuk bahasa informal yang sering digunakan dalam ulasan media sosial. Penelitian sebelumnya menunjukkan bahwa BERT mampu mengungguli model *Long Short-Term Memory* (LSTM) dalam klasifikasi teks dengan akurasi yang signifikan (Aji Riyantoko dkk., 2022).

Berdasarkan urgensi tersebut, penelitian ini bertujuan untuk mengimplementasikan metode BERT guna melakukan analisis sentimen berbasis aspek pada data ulasan Hotel XYZ. Penelitian ini berfokus untuk menjawab pertanyaan utama: bagaimana mengklasifikasikan sentimen dan mengidentifikasi aspek layanan spesifik dari ulasan tamu menggunakan algoritma BERT untuk menghasilkan rekomendasi peningkatan pelayanan yang akurat? Dengan memanfaatkan kemampuan model *pre-trained* xlm-roberta-base-sentiment, penelitian ini diharapkan dapat memberikan wawasan mendalam mengenai kekuatan dan kelemahan operasional hotel, serta menyediakan dasar pengambilan keputusan berbasis data bagi manajemen untuk meningkatkan kepuasan pelanggan secara berkelanjutan.

METODE

Desain dan Kerangka Penelitian ini menerapkan pendekatan kerangka sistematis yang menggambarkan tahapan-tahapan yang dilakukan dalam penelitian implementasi metode BERT untuk analisis sentimen pada data *review*. Fokus utama penelitian adalah pengembangan model analisis sentimen otomatis berbasis *Deep Learning* menggunakan algoritma BERT (*Bidirectional Encoder Representations from Transformers*) untuk mengolah data teks ulasan yang tidak terstruktur menjadi wawasan strategis bagi manajemen hotel.



Gambar 1. Metode Penelitian

Identifikasi Masalah

Tahap pertama dimulai dengan identifikasi masalah, di mana penulis menelaah kendala operasional yang dihadapi oleh Hotel XYZ terkait ketidakefisienan dalam memproses ribuan ulasan tamu secara manual.

Studi Pustaka

Berdasarkan masalah tersebut, dilakukan tahap studi pustaka untuk mengkaji teori-teori relevan mengenai analisis sentimen, *Natural Language Processing* (NLP), dan keunggulan algoritma *Deep Learning* khususnya arsitektur BERT dalam menangani data teks yang kompleks. Kajian ini menjadi landasan teoretis dalam merancang solusi sistem yang diusulkan.

Pengumpulan Data

Data dikumpulkan dari tiga *platform Online Travel Agent* (OTA) utama yang mendominasi pasar, yaitu Google Reviews, Traveloka, dan Agoda. Proses pengumpulan data dilakukan menggunakan teknik *web scraping* pada tanggal 30 September 2025. *Web scraping* didefinisikan sebagai suatu teknik untuk melakukan ekstraksi sejumlah besar data yang terdapat pada halaman website tertentu secara otomatis (Dwicahyo & Indah Ratnasari, 2023). Pengambilan data Google Reviews dilakukan menggunakan *browser automation* dengan *Selenium*, sedangkan data Traveloka dan Agoda diambil menggunakan teknik API *Scraping*.

Rentang waktu data ulasan yang diambil adalah periode tahun 2022 hingga 2025. Dari proses ini, terkumpul total data mentah sebanyak 605 ulasan, yang terdiri dari 495 ulasan Google Reviews, 87 ulasan Agoda, dan 23 ulasan Traveloka.

Pra-pemrosesan Data

Prosedur Pra-pemrosesan Data Sebelum memasuki tahap pemodelan, data mentah melalui serangkaian tahapan pra-pemrosesan untuk membersihkan *noise* dan menstandarisasi format teks. Tahapan ini meliputi pembersihan data (*data cleaning*) dari emoji dan simbol, pengubahan huruf menjadi kecil (*case folding*), dan *stemming*. Mengingat lokasi objek penelitian, ditemukan banyak penggunaan bahasa daerah dan istilah tidak baku. Oleh karena itu, penelitian ini menerapkan normalisasi bahasa berbasis kamus (*dictionary-based normalization*) untuk memetakan istilah lokal (seperti "lemak" menjadi "enak" atau "dak" menjadi "tidak") ke dalam bahasa Indonesia baku. Proses dilanjutkan dengan tokenisasi dan penghapusan kata sambung (*stopword removal*). Setelah melalui tahap pembersihan, jumlah data yang siap digunakan untuk pemodelan adalah sebanyak 394 ulasan.

Implementasi Model BERT

Pada tahap ini, model *pre-trained* xlm-roberta-base-sentiment diterapkan untuk melakukan dua tugas utama yaitu mengklasifikasikan sentimen ulasan ke dalam kategori positif, negatif, atau netral; serta mendeteksi aspek layanan yang dibahas (seperti Kamar, Pelayanan, atau Fasilitas). Model ini juga diintegrasikan dengan logika *Root Cause Analysis* untuk mengekstraksi kata kunci pemicu keluhan. Hasil pemodelan kemudian dibangun ke dalam sebuah aplikasi berbasis web menggunakan kerangka kerja Flask untuk visualisasi.

Evaluasi Hasil BERT

Tahap terakhir adalah evaluasi hasil BERT dilakukan menggunakan metode *Confusion Matrix* untuk mengukur kinerja model berdasarkan metrik Akurasi (*Accuracy*), Presisi (*Precision*), *Recall*, dan *F1-Score*, guna memastikan keandalan sistem dalam mendeteksi keluhan tamu secara akurat.

HASIL DAN PEMBAHASAN

Pra-pemrosesan Data

Preprocessing adalah serangkaian langkah yang bertujuan untuk membersihkan, mentransformasi, dan menstandarisasi data teks mentah menjadi format yang lebih bersih, terstruktur, dan siap untuk diproses lebih lanjut oleh model analisis sentimen (Amna dkk., 2023). Sebelum dilakukan pemodelan menggunakan algoritma BERT, data ulasan mentah yang telah dikumpulkan dari *platform Online Travel Agent* (OTA) harus melalui serangkaian tahapan pra-pemrosesan. Tahapan ini krusial mengingat data *User-Generated Content* (UGC) seringkali bersifat tidak terstruktur dan mengandung banyak *noise* yang dapat menurunkan akurasi model. Hasil akhir yang diharapkan dari tahap preprocessing adalah kumpulan data teks yang terdiri dari kata-kata dasar atau token yang bermakna dan siap untuk diolah (Rizal Chandra Rivaldi & T.D. Wismarini, 2024).

Pembersihan Data

Pada tahap ini, sistem menghapus elemen-elemen yang tidak relevan dan dapat mengganggu proses analisis sentimen. Elemen tersebut meliputi penghapusan tanda baca yang berlebihan, angka yang tidak memiliki makna konteks, simbol mata uang, *emoticon* atau emoji, serta karakter non-alfanumerik lainnya. Penghapusan ini bertujuan agar model fokus hanya pada kata-kata yang mengandung makna sentimen (Prasetyo dkk., 2024).

Case Folding

Setelah data dibersihkan dari karakter gangguan, dilakukan penyeragaman format teks dengan mengubah seluruh huruf menjadi huruf kecil (*lowercase*) (Ariansyah, 2024). Tahap ini penting untuk memastikan konsistensi data, sehingga kata yang ditulis dengan variasi huruf kapital (misalnya "BAGUS", "Bagus", atau "bagus") dianggap sebagai token yang sama oleh mesin.

Tabel 1. Hasil Pembersihan dan *Case Folding*

No	Teks Review Asli	Hasil Pembersihan & Case Folding
1	Kamarnya KOTOR banget!!! Banyak debu di meja	kamarnya kotor banget banyak debu di meja
2	Pelayanannya RAMAH, tapi check-in lama...	pelayanannya ramah tapi checkin lama
3	AC nya panas & tidak dingin sama sekali.	ac nya panas tidak dingin sama sekali

Stopword Removal

Tahap ini bertujuan untuk menyaring kata-kata yang memiliki frekuensi kemunculan tinggi namun memiliki makna sentimen yang rendah (kata-kata netral) (T Angel & Pranatawijaya, 2024). Seperti kata sambung "dan", "yang", "di", "ke", atau "dari". Dengan menghapus kata-kata ini, model dapat lebih fokus pada kata-kata kunci yang membawa muatan sentimen atau opini.

Stemming

Untuk meminimalisir variasi morfologis kata, dilakukan proses *stemming* yang mengubah kata-kata berimbuhan yang menyebabkan ambigu menjadi bentuk kata dasarnya (Ilham & Maharani, 2022). Sebagai contoh, kata "membersihkan", "dibersihkan", atau "kebersihan" akan direduksi menjadi kata dasar "bersih". Proses ini membantu sistem dalam mengelompokkan keluhan atau pujian yang memiliki inti makna yang sama, yang sangat berguna pada fitur *Root Cause Analysis*.

Tabel 2. Hasil *Stopword Removal* dan *Stemming*

No	Hasil Pembersihan & *Case Folding*	Hasil *Stopword Removal*	Hasil *Stemming*
1	kamarnya kotor banget banyak debu di meja	kamarnya kotor banget debu meja	kamar kotor banget debu meja
2	pelayanannya ramah tapi checkin lama	pelayanannya ramah checkin lama	layan ramah checkin lama
3	ac nya panas tidak dingin sama sekali	ac panas dingin	ac panas dingin

Normalisasi Bahasa Daerah

Mengingat objek penelitian berlokasi di Kota Palembang, ditemukan bahwa sebagian data ulasan mengandung terminologi lokal (Bahasa Palembang) serta bahasa tidak baku (*slang*). Karena model BERT (xlm-roberta-base-sentiment) dilatih menggunakan korpus bahasa standar, keberadaan kata-kata daerah ini dapat menimbulkan *noise* atau ambiguitas yang berpotensi menurunkan akurasi klasifikasi sentimen. Untuk mengatasi hal tersebut, penelitian ini menerapkan tahapan Normalisasi Berbasis Kamus.

Tabel 3. Daftar Kata Bahasa Daerah

No	Kata Asli	Kata Hasil Normalisasi
1	lemak	enak
2	dak/idak/ndak	tidak
3	basing	terserah
4	nian	sangat
5	ringam	terganggu

6	igo	terlalu
7	wong	orang

Tokenization

Tokenization merupakan proses untuk memecah kalimat menjadi unit-unit kata atau sub-kata (*tokens*) yang dapat dipahami oleh arsitektur transformer (Wafda, 2025), meskipun tidak ditampilkan secara visual pada antarmuka pengguna.

Tabel 4. Hasil *Tokenization*

No	Teks Hasil Preprocessing	*Tokenization*
1	kamar kotor banget debu meja	['kamar', 'kotor', 'banget', 'debu', 'meja']
2	layan ramah checkin lama	['layan', 'ramah', 'check', 'lama']
3	ac panas dingin	['ac', 'panas', 'dingin']

Penentuan Kategori Aspek dan Kamus Kata Kunci Untuk mendukung analisis yang lebih mendalam, penelitian ini menetapkan kategori aspek layanan sebagai acuan klasifikasi. Penentuan aspek ini didasarkan pada observasi awal terhadap frekuensi kata yang muncul dalam dataset serta merujuk pada dimensi kualitas layanan industri perhotelan dari penelitian terdahulu (Astriningsih, 2023). Berdasarkan referensi tersebut, aspek penilaian dibagi menjadi enam kategori utama, yaitu: Fasilitas, Harga, Kamar, Kamar Mandi, Lokasi, Pelayanan, dan Sarapan. Setiap aspek didefinisikan dengan kumpulan kata kunci terkait untuk mempermudah identifikasi topik dalam ulasan, sebagaimana dirincikan pada Tabel 5.

Tabel 5. Definisi Aspek Layanan

Aspek	Kata Kunci Terkait
Fasilitas	fasilitas, ac, wifi, parkir, kolam
Harga	harga, murah, mahal, biaya, jangkau
Kamar	kamar, ranjang, kasur, tidur
Kamar Mandi	kamar mandi, wc, toilet, air panas, shower
Lokasi	lokasi, strategis, dekat, jauh
Pelayanan	layan, staff, resepsionis, ramah, sopan, bantu, profesional
Sarapan	makan, sarapan, menu, restoran, enak, lezat, variatif

Selain penentuan aspek, penelitian ini juga menyusun sebuah kamus kata kunci (*lexicon dictionary*) yang berfungsi sebagai *ground truth* atau acuan awal dalam metode pelabelan berbasis leksikon. Kamus ini memuat daftar kata-kata yang bermuatan opini positif dan negatif di mana setiap kata memiliki sentimen tertentu (Rachmadana Ismail dkk., 2023). Daftar kata dasar (*seed words*) dikembangkan dari kosakata umum Bahasa Indonesia yang kemudian diperluas dengan istilah-istilah spesifik yang sering muncul dalam konteks ulasan perhotelan di lokasi studi kasus. Kamus kata kunci ini menjadi landasan penting dalam memvalidasi sentimen yang terdeteksi oleh model.

Tabel 6. Daftar Kata Kunci Sentimen

Kategori	Daftar Kata Kunci
Positif	puas, mantap, bagus, baik, bersih, nyaman, ramah, enak, cepat, strategis, luar biasa, rekomendasi, sempurna, terbaik, bantu, sopan, rapi, wangi, luas, dingin, lengkap, lezat, murah, jangkau, hebat, profesional, istimewa, juara, senang, keren, modern, baru, terawat, tenang, sejuk, adem, nyenyak, sigap, solutif.
Negatif	kecewa, buruk, jelek, kotor, bau, berisik, lama, lambat, sulit, susah, mahal, masalah, kurang, tidak, rusak, bocor, debu, jorok, kasar, antri, panas, pengap, parah, apek, lembab, kusam, noda, tua, usang, sempit, terbatas, macet, bising, gelap, menurun, jutek, acuh, rumit, bingung, menyesal.

Modeling

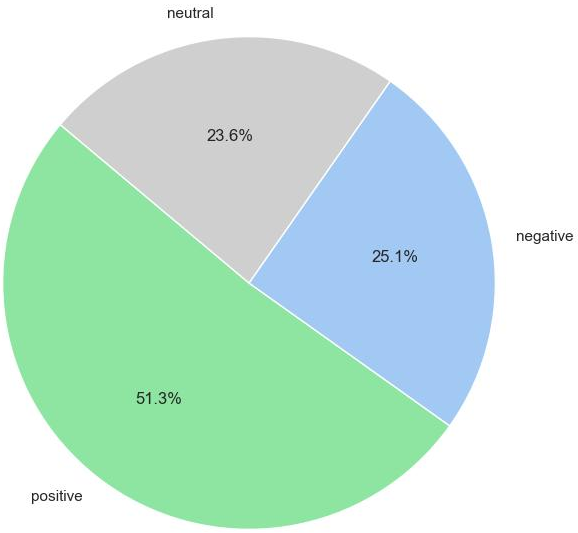
Tahap *modeling* merupakan inti dari proses penambangan data (*data mining*), di mana algoritma analisis diterapkan pada data ulasan yang telah melalui tahap pra-pemrosesan (*data preparation*). Penelitian ini menggunakan model *Deep Learning* berbasis transformer, yaitu BERT (*Bidirectional Encoder Representations from Transformers*). Secara spesifik, model yang digunakan adalah varian multibahasa cardiffnlp/twitter-xlm-roberta-base-sentiment yang dimuat menggunakan pustaka *Hugging Face*. Library ini menyediakan model *pre-trained* yang dapat digunakan untuk melakukan tugas-tugas klasifikasi (Nurfitri & Chowanda, 2024). Pemilihan model ini didasarkan pada kemampuannya menangani teks multibahasa dan konteks informal yang umum ditemukan pada media sosial atau ulasan daring.

Proses pemodelan dilakukan dengan menerapkan model pada setiap baris teks ulasan yang telah dibersihkan (*text_cleaned*) secara bertahap (*batch processing*) untuk efisiensi memori. Model bertugas mengklasifikasikan setiap ulasan ke dalam tiga label sentimen *positive, negative,* atau *neutral*, serta memberikan skor probabilitas (*sentiment score*) untuk setiap prediksi. Output dari tahapan ini adalah *dataset* yang telah diperkaya dengan kolom label dan skor sentimen, yang kemudian menjadi dasar bagi analisis distribusi sentimen dan evaluasi kinerja sistem selanjutnya.

Tabel 7. Output Klasifikasi Model BERT

Teks *Review* (*Cleaned*)	*Sentiment Label*	*Sentiment Score*
mantapp	*positive*	0.422
makan diresto harga tidak sesuai yg tera daftar harga	*neutral*	0.654
lumayan harga	*positive*	0.449
with the price thats okay tolong perhati bersih sampah...	*neutral*	0.406
tempat yang strategis akses belanja dekat layan hotel...	*positive*	0.561

Berdasarkan hasil klasifikasi model terhadap keseluruhan *dataset* ulasan hotel, diperoleh distribusi sentimen sebagai berikut.



Gambar 2. Distribusi Sentimen Hasil Prediksi Model BERT

Dari total 394 ulasan yang dianalisis, sentimen positif mendominasi sebanyak 202 ulasan (51,3%). Hal ini mengindikasikan bahwa secara umum tamu memiliki persepsi yang baik terhadap pengalaman menginap mereka, sentimen negatif teridentifikasi sebanyak 99 ulasan (25,1%). Meskipun persentasenya lebih kecil, ulasan negatif ini mengandung informasi krusial

mengenai keluhan operasional yang perlu ditindaklanjuti, sentimen netral sebanyak 93 ulasan (23,6%).

Analisis Sentimen Berbasis Aspek

Untuk memberikan wawasan yang lebih spesifik dibandingkan sekadar sentimen umum, penelitian ini melakukan pemetaan sentimen ke dalam aspek-aspek layanan hotel.

Tabel 8. Persentase Distribusi Sentimen Per Aspek

Aspek	Netral (%)	Negatif (%)	Positif (%)
Fasilitas	3,23	19,35	77,42
Harga	3,70	11,11	85,19
Kamar	9,52	23,81	66,67
Kamar Mandi	11,11	66,67	22,22
Lokasi	0,00	14,29	85,71
Pelayanan	3,39	5,08	91,53
Sarapan	6,25	18,75	75,00

Berdasarkan perhitungan persentase sentimen per aspek yang disajikan dalam Tabel 8, terlihat adanya polarisasi kinerja yang signifikan antar departemen. Aspek 'Pelayanan' mencatatkan kinerja terbaik dengan persentase sentimen positif tertinggi mencapai 91,53%, diikuti oleh aspek 'Lokasi' (85,71%) dan 'Harga' (85,19%). Hal ini menunjukkan bahwa keramahan staf dan posisi strategis hotel merupakan kekuatan utama (*key strengths*) yang diapresiasi tamu. Sebaliknya, aspek 'Kamar Mandi' teridentifikasi sebagai titik lemah utama dengan persentase sentimen negatif tertinggi mencapai 66,67%, jauh melampaui sentimen positifnya yang hanya 22,22%. Tingginya keluhan pada aspek ini menjadi sinyal bagi manajemen untuk memprioritaskan evaluasi pada fasilitas tersebut.

Analisis Akar Masalah

Guna mendiagnosis penyebab tingginya sentimen negatif pada aspek kritis, sistem melakukan ekstraksi kata kunci dominan pada ulasan negatif.

Tabel 9. Hasil Analisis Akar Masalah

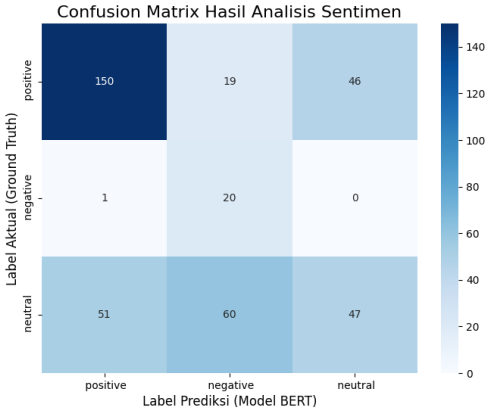
Aspek	**Kata_Kunci_Negatif**	**Jumlah_Kemunculan**
kamar mandi	bau	4
kamar mandi	apek	2
kamar mandi	kotor	2
kamar mandi	panas	2
kamar mandi	berisik	1
kamar mandi	kecewa	1
kamar mandi	noda	1

Pada aspek 'Kamar Mandi' yang memiliki sentimen negatif tertinggi (66,67%), ditemukan bahwa kata kunci "bau" memiliki jumlah kemunculan terbanyak (4 kali), diikuti oleh "kotor", "panas", dan "apek". Temuan ini mengindikasikan bahwa masalah utama bukan sekadar pada kerusakan fisik, melainkan pada kebersihan dan aroma yang mengganggu kenyamanan. Berdasarkan data tersebut, rekomendasi paling mendesak bagi manajemen Hotel XYZ adalah melakukan perbaikan pada sistem saluran pembuangan air dan meningkatkan standar prosedur pembersihan harian untuk menghilangkan bau tidak sedap.

Evaluation

Hasil perbandingan antara prediksi model dengan label manual divisualisasikan menggunakan *Confusion Matrix* untuk melihat detail kesalahan klasifikasi. *Confusion Matrix*

adalah matrik untuk pengukuran kinerja untuk model *supervised learning* khususnya klasifikasi. *Outcome* dari klasifikasi terdapat 4 perbandingan antara nilai prediksi dan aktual, yaitu *True Positive* (TP), *True Negative* (TN), *False Positive* (FP), dan *False Negative* (FN) (Rahman Salam dkk., 2023)



Gambar 3. *Confusion Matrix*

Berdasarkan Gambar 3, kinerja model paling optimal terlihat pada kelas sentimen positif dengan 147 prediksi tepat dari total 203 data. Secara krusial, model menunjukkan sensitivitas yang tinggi pada kelas negatif dengan berhasil mengenali 28 dari 37 ulasan keluhan aktual, sehingga meminimalkan risiko keluhan tamu yang terlewat (*False Negative*). Sebaliknya, kelas netral mencatatkan tingkat kesalahan tertinggi (hanya 50 prediksi benar) akibat adanya ambiguitas bahasa atau sentimen campuran yang seringkali sulit dipisahkan secara tegas oleh model.

Berdasarkan *Confusion Matrix* di atas, kinerja model diukur menggunakan empat parameter standar *Accuracy, Precision, Recall*, dan *F1-Score* (Singgalen, 2025).

Tabel 9. Hasil Analisis Akar Masalah

	precision	***recall***	***f1-score***	***support***
negative	0.28	0.76	0.41	37
neutral	0.54	0.32	0.40	154
positive	0.73	0.72	0.73	203
accuracy			0.57	394
macro avg	0.52	0.60	0.51	394
weighted avg	0.61	0.57	0.57	394

Hasil evaluasi menunjukkan model mencapai akurasi global sebesar 57%. Meskipun akurasi ini tergolong moderat akibat ketidakseimbangan data (*imbalanced data*) di mana data positif jauh lebih banyak, model menunjukkan performa yang sangat baik dalam metrik *Recall* untuk kelas negatif yang mencapai 76%. Tingginya nilai sensitivitas (*Recall*) ini berarti sistem sangat andal dalam mendeteksi dan "menangkap" ulasan negatif yang masuk, meminimalkan risiko adanya keluhan tamu yang terlewat oleh manajemen. Sebaliknya, kelas netral memiliki performa terendah dengan *Recall* 32%, yang wajar terjadi karena ambiguitas bahasa pada ulasan netral yang seringkali tercampur dengan sentimen lain. Secara keseluruhan, dengan *Recall* negatif 76% dan *F1-Score* positif 73%, model ini dinilai efektif untuk tujuan bisnis dalam menjaga reputasi hotel.

KESIMPULAN

Penelitian ini menyimpulkan bahwa implementasi metode *Deep Learning* menggunakan algoritma *Bidirectional Encoder Representations from Transformers* (BERT) efektif untuk

menganalisis ulasan tamu guna mendukung strategi peningkatan layanan. Sistem yang dibangun berhasil menjawab tujuan penelitian dengan menyediakan klasifikasi sentimen otomatis yang memiliki sensitivitas tinggi dalam mendeteksi keluhan pelanggan, sebagaimana dibuktikan oleh capaian nilai *Recall* sentimen negatif sebesar 76%. Meskipun akurasi global model terukur sebesar 57% akibat tantangan ketidakseimbangan data, kemampuan model dalam mengidentifikasi ulasan negatif secara presisi menegaskan bahwa pendekatan berbasis konteks ini andal untuk digunakan sebagai alat evaluasi operasional yang objektif, menggantikan proses analisis manual yang tidak efisien.

Temuan penelitian ini memberikan rekomendasi konkret bagi manajemen hotel, di mana analisis sentimen berbasis aspek mengungkapkan bahwa prioritas perbaikan harus difokuskan pada fasilitas fisik, khususnya aspek kamar mandi dan pendingin ruangan. Hal ini didukung oleh hasil analisis akar masalah yang secara konsisten mengidentifikasi kata kunci "bau", "kotor", dan "panas" sebagai pemicu utama ketidakpuasan tamu, sementara aspek pelayanan dan lokasi teridentifikasi sebagai kekuatan utama yang perlu dipertahankan. Integrasi seluruh hasil analisis ini ke dalam sistem informasi berbasis web memungkinkan manajemen untuk beralih ke pola pengambilan keputusan berbasis data (*data-driven decision making*) yang lebih cepat dan tepat sasaran dalam merespons dinamika kepuasan pelanggan.

Hasil penelitian ini dapat menambah penerapan *Natural Language Processing* (NLP) untuk pemecahan masalah industri nyata. Penggunaan *Aspect-Based Sentiment Analysis* (ABSA) yang dikombinasikan dengan deteksi akar masalah (*Root Cause Analysis*) pada data ulasan yang tidak terstruktur dan multibahasa menunjukkan bahwa algoritma kompleks seperti BERT dapat dikemas menjadi solusi praktis. Sistem ini tidak hanya berfungsi sebagai alat monitoring reputasi, tetapi juga sebagai instrumen strategis yang menjembatani kesenjangan antara data ulasan digital yang masif dengan kebutuhan wawasan operasional yang spesifik untuk keberlanjutan bisnis

REFERENSI

Aji Riyantoko, P., Maulana Fahrudin, T., Arman Prasetya, D., & Dari Timur, T. (2022). *Analisis Sentimen Sederhana Menggunakan Algoritma LSTM dan BERT untuk Klasifikasi Data Spam dan Non-Spam*.

Alfarobby, A. N., & Irawan, H. (2024). *Analisis Sentimen Kepuasan Konsumen Pengguna Transportasi Online Pada Ulasan Google Playstore Menggunakan Indobert Dan Topic Modeling (Studi kasus: Gojek dan Grab)* (Vol. 11, Nomor 1).

Amna, S, W., Gede Iwan Sudipa, I., Andi Putra, T. E., Jurnaidi Wahidin, A., Alfa Syukrilla, W., Khrisna Wardhani, A., Heryana, N., Indriyani, T., Willyanto Santoso Tutuk Indriyani, L., & Willyanto Santoso, L. (2023). *DATA MINING*. www.globaleksekutifteknologi.co.id

Ariansyah, F. (2024). *PENERAPAN BIDIRECTIONAL ENCODER REPRESENTATIONS FROM TRANSFORMERS (BERT) PADA ANALISIS SENTIMEN VAKSIN BOOSTER COVID-19*.

Astriningsih, W. (2023). *Identifikasi Multi Aspek dan Sentimen Analisis pada Review Hotel Menggunakan Deep learning*.

Bayu Baskoro, B., Susanto, I., Khomsah, S., Informatika, P., Sains Data, P., Teknologi Telkom Purwokerto Jl Panjaitan, I. DI, & Tengah, J. (2021). *Analisis Sentimen Pelanggan Hotel di Purwokerto Menggunakan Metode Random Forest dan TF-IDF (Studi Kasus: Ulasan Pelanggan Pada Situs TRIPADVISOR)*. *3*(2), 21–029. https://doi.org/10.20895/INISTA.V3I2

Danuwangsa, R., Ayu, G., Suwintari, E., & Suadnyana, W. (2025). The Influence of Online Reviews, Prices, and Locations on Guest Interests to Stay. *Jurnal Pariwisata dan Bisnis)*, *04*(1), 1953–1960. https://doi.org/10.22334/paris.v4i1

Dwicahyo, K., & Indah Ratnasari, C. (2023). *Perbandingan Metode Web Scraping Dalam Pengambilan Data: Kajian Literatur*.

Ilham, F., & Maharani, W. (2022). Analyze Detection Depression In Social Media Twitter Using Bidirectional Encoder Representations from Transformers. *Journal of Information System Research (JOSH)*, *3*(4), 476–482. https://doi.org/10.47065/josh.v3i4.1885

Nurfitri, W., & Chowanda, A. (2024). *Analisis Sentimen Pada Kasus Positif Covid-19 Berdasarkan Pemberitaan Media Di Indonesia Menggunakan Indobert*.

Prasetyo, B., Ahmad Yusuf Al-Majid, & Suharjito. (2024). A Comparative Analysis of MultinomialNB, SVM, and BERT on Garuda Indonesia Twitter Sentiment. *PIKSEL: Penelitian Ilmu Komputer Sistem Embedded and Logic*, *12*(2), 445–454. https://doi.org/10.33558/piksel.v12i2.9966

Rachmadana Ismail, A., Bagus, R., Hakim, F., & Artikel, R. (2023). Implementasi Lexicon Based Untuk Analisis Sentimen Dalam Mengetahui Trend Wisata Pantai Di DI Yogyakarta Berdasarkan Data Twitter P-ISSN E-ISSN. Dalam *Emerging Statistics and Data Science Journal* (Vol. 1, Nomor 1).

Rahman Salam, R., Fajri Jamil, M., Ibrahim, Y., Rahmaddeni, Soni, & Herianto. (2023). *MALCOM: Indonesian Journal of Machine Learning and Computer Science Sentiment Analysis of Cash Direct Assistance Distribution for Fuel Oil Using Support Vector Machine Analisis Sentimen Terhadap Bantuan Langsung Tunai (BLT) Bahan Bakar Minyak (BBM) Menggunakan Support Vector Machine*. *3*, 27–35.

Ratu Bilqis, L. D., Taviprawati, E., & Ciu, P. Y. (2024). Pengaruh Kualitas Pelayanan dan Electronic Word of Mouth Terhadap Keputusan Pembelian di Hotel Ra Suite Simatupang. *Jurnal Ilmiah Pariwisata*, *29*(3), 264. https://doi.org/10.30647/jip.v29i3.1828

Ravichandiran, Sudharsan. (2021). *Getting started with Google BERT: build and train state-of-the-art natural language processing models using BERT*. Packt Publishing Ltd.

Rizal Chandra Rivaldi, & T.D. Wismarini. (2024). Analisis Sentimen Pada Ulasan Produk Dengan Metode Natural Language Processing (NLP). *Elkom: Jurnal Elektronika dan Komputer*, *17*(1), 120–128. https://doi.org/10.51903/elkom.v17i1.1680

Singgalen, Y. A. (2025). Performance Analysis of IndoBERT for Sentiment Classification in Indonesian Hotel Review Data. *Journal of Information System Research*, *6*(2), 976–986. https://doi.org/10.47065/josh.v6i2.6505

T Angel, A. C., & Pranatawijaya, V. H. (2024). *Analisis Sentimen dan Emosi dari Ulasan Google Maps untuk Layanan Rumah Sakit di Palangka Raya Menggunakan Machine Learning* (Vol. 4, Nomor 1).

Wafda, A. (2025). *Aspect-Based Sentiment Analysis terhadap Cuitan Platform X tentang Kurikulum Merdeka Menggunakan IndoBERT*.

Yunitasari Yessi. (2023). *Teori Dan Implementasi Analisis Sentimen Menggunakan Python*.