

## Analisis Sentimen Review Hotel Menggunakan Metode Deep Learning BERT

Vidya Chandradev<sup>1\*</sup>, I Made Agus Dwi Suarjaya<sup>2</sup>, I Putu Agung Bayupati<sup>3</sup>

Program Studi Teknologi Informasi, Fakultas Teknik, Universitas Udayana,  
Bali 80361, Indonesia

Email: <sup>1</sup>vidyaachan14@gmail.com, <sup>2</sup>agussuarjaya@it.unud.ac.id, <sup>3</sup>bayupati@unud.ac.id

**Abstract.** *Hotel Review Sentiment Analysis using BERT Deep Learning Method.* The COVID-19 pandemic has resulted in declining tourism visits and hotel occupancy. Hoteliers must monitor visitor lifestyles to sustain their businesses. One way to achieve this is by understanding the sentiment of hotel visitors through review analysis, enabling better decision-making regarding service and business aspects in the hotel industry. This research applies the natural language processing deep learning model BERT to analyze positive and negative sentiments from hotel visitor reviews in Indonesia. The BERT model undergoes a pre-trained and fine-tuned process to produce accurate sentiment analysis. Evaluation results demonstrate that the fine-tuned SmallBERT model performs well, trained on a dataset of 515k hotel reviews for five epochs. The SmallBERT model achieves an accuracy of 91.40%, precision of 90.51%, recall of 90.51%, and an F1 score of 90.51% when evaluated with manually labelled datasets. Visualizations of the predominantly positive sentiment comparisons are conducted using Tableau.

**Keywords:** *sentiment analysis, SmallBERT, deep learning, natural language processing, fine-tuning*

**Abstrak.** *Pandemi COVID-19 telah menyebabkan penurunan kunjungan pariwisata dan okupansi hotel. Penting bagi pengusaha hotel untuk memantau gaya hidup pengunjung guna menjaga kelangsungan bisnis. Salah satu cara untuk melakukannya adalah dengan memahami sentimen pengunjung hotel melalui analisis review agar mendapatkan pemahaman yang lebih baik dalam pengambilan keputusan terkait layanan dan aspek bisnis di sektor perhotelan. Penelitian ini menerapkan model deep learning natural language processing BERT untuk menganalisis sentimen positif dan negatif dari review pengunjung hotel di Indonesia. Model BERT yang digunakan telah menjalani proses pre-trained dan diterapkan metode fine-tuning untuk menghasilkan analisis sentimen yang akurat. Hasil evaluasi menunjukkan bahwa model fine-tuning SmallBERT yang dilatih menggunakan dataset 515k review hotel selama 5 epoch memberikan performa yang baik. Model SmallBERT mencapai akurasi sebesar 91,40%, presisi 90,51%, recall 90,51%, dan skor f1 90,51% saat dievaluasi dengan dataset yang diberi label secara manual. Visualisasi hasil perbandingan sentimen yang didominasi oleh sentimen positif, dilakukan menggunakan Tableau.*

**Kata kunci:** *analisis Sentimen, SmallBERT, deep learning, natural language processing, fine-tuning*

### 1. Pendahuluan

Sektor pariwisata telah diketahui memiliki peranan penting dalam mengembangkan perekonomian sehingga sektor ini perlu dikembangkan secara maksimal untuk meningkatkan pendapatan masyarakat di kawasan wisata sekitar dengan terciptanya lapangan pekerjaan [1]. Sektor perhotelan memiliki hubungan erat dengan pariwisata dan salah satu indikator penting dalam sektor ini adalah TPK hotel, yaitu perbandingan antara malam kamar terpakai dengan malam kamar yang tersedia untuk memberikan gambaran persentase ketersediaan kamar yang terisi oleh tamu dalam suatu periode waktu tertentu [2]. TPK hotel berbintang di Indonesia mencapai 51,57% pada Desember 2021 dan kunjungan wisman ke Indonesia pada Januari 2022 meningkat 13,62% dibandingkan bulan yang sama tahun sebelumnya, mencapai 143.744

kunjungan [3]. Pemantauan *lifestyle visitor* penting bagi pengusaha atau pemilik hotel untuk menentukan layanan atau fasilitas yang diminati demi kelancaran industri perhotelan dan pariwisata.

Salah satu penelitian terdahulu yang berjudul “*Saving the hotel industry: Strategic response to the COVID-19 pandemic, hotel selection analysis, and customer retention*” menyebutkan bahwa para praktisi di industri perhotelan harus melakukan kegiatan pemasaran yang efektif untuk menciptakan citra hotel yang lebih baik demi mempertahankan pelanggan [4]. Penelitian terdahulu berikutnya berjudul “*Analisis Pengaruh Harga, Produk, Dan Kualitas Layanan Terhadap Kepuasan Wisatawan Di Taman Wisata Air Wendit Kabupaten Malang*” menyimpulkan bahwa suatu daerah tujuan wisata harus memiliki empat komponen utama dalam pariwisata, yaitu *attraction*, *accessibility*, *amenity*, dan *ancillary services*. Semakin baik objek wisata, semakin tinggi kepuasan wisatawan dan akan semakin dikenal oleh wisatawan lain jika mendapat rekomendasi positif [5]. Penelitian ini bertujuan menganalisis sentimen pengunjung hotel berdasarkan *review* untuk memahami perbandingan sentimen masyarakat terhadap layanan perhotelan di Indonesia menggunakan model *deep learning pre-trained* BERT dengan metode *fine-tuning*. Hasil penelitian diharapkan dapat memberikan pemahaman yang lebih baik kepada praktisi hotel dalam pengambilan keputusan terkait layanan dan aspek bisnis di sektor perhotelan yang berhubungan dengan pariwisata.

## 2. Tinjauan Pustaka

### 2.1. Sentiment Analysis

*Sentiment analysis* adalah teknik untuk menganalisis pikiran, perasaan, dan penilaian pengguna berdasarkan informasi seperti opini tertulis. Hal ini dilakukan untuk menentukan apa yang diyakini pengguna terkait suatu topik atau subjek tertentu [7]. *Sentiment analysis* memiliki keterkaitan yang erat dengan *Natural Language Processing* (NLP), yang merupakan kemampuan mesin untuk memahami bahasa manusia dengan menggunakan algoritma komputer, matematika, dan linguistik komputasi. Contoh produk NLP adalah Google Assistant dan Siri. Python cocok untuk digunakan pada NLP karena tersedianya *library open source*, mudah digunakan, tidak terlalu rumit, dan optimalisasi sistem berbasis NLP tidak se-kompleks paradigma pemrograman lainnya [6]. NLP adalah teknik untuk menganalisis data tekstual dengan pendekatan *rule-based* dan *machine learning*. Polaritas dan subjektivitas adalah metrik penting untuk menganalisis sentimen. Pelatihan model dapat dipantau melalui diagram garis *training loss*, dan evaluasi akurasi menunjukkan keefektifan model dalam menganalisis sentimen [8]. Model yang digunakan dalam penelitian ini adalah SmallBERT, varian *Bidirectional Encoder Representation from Transformer* (BERT) yang ukurannya lebih kecil tetapi tetap mempertahankan kemampuan pemrosesan bahasa dengan jumlah parameter yang lebih sedikit

### 2.2. Bidirectional Encoder Representation from Transformer (BERT) dan SmallBERT

*Bidirectional Encoder Representation from Transformer* (BERT) adalah representasi *encoder* dari model Transformer, sebuah arsitektur NLP yang menggunakan mekanisme perhatian untuk menggantikan jaringan berulang dan mampu menangkap hubungan antara kata-kata yang jauh secara kontekstual. BERT menghasilkan representasi untuk setiap kata dalam kalimat sebagai output dan dapat meningkatkan kinerja model pada tugas-tugas berurutan kompleks dalam NLP. *Transformer* diperkenalkan dalam paper “*Attention Is All You Need*” [9]. BERT adalah salah satu model penyematan (*embedding*) teks canggih yang diterbitkan oleh Google [10]. BERT adalah model representasi berbasis *fine-tuning* pertama yang mencapai kinerja canggih pada rangkaian besar tugas *sentence-level* dan *token-level*, dan mengungguli banyak arsitektur *task-specific* [11]. Beberapa paper penelitian seperti “*Sentiment Analysis on The Impact of Coronavirus in Social Life using The BERT Model*” [8], “*An intelligent early warning system of analyzing Twitter data using machine learning on COVID-19 surveillance in the US*” [12], “*Transformer-based deep learning models for the sentiment*” [13], “*Predictive Intelligence in Harmful News Identification by BERT-Based Ensemble Learning Model with Text Sentiment Analysis*” [14], “*BERT Fine-Tuning for Sentiment Analysis on Indonesian Mobile Apps Reviews*”

[15], “*Sentimental analysis on user’s reviews using BERT*” [16], dan “*An ensemble-based hotel recommender system using sentiment analysis and aspect categorization of hotel reviews*” [17] menyebutkan bahwa BERT menempati hasil penelitian yang tertinggi dibandingkan model dengan algoritma lainnya. BERT sukses dalam banyak tugas NLP seperti menjawab pertanyaan, pembuatan teks, dan klasifikasi kalimat. BERT menggunakan model *embedding* berbasis konteks yang memahami konteks dan *encoder* dari Transformer untuk membaca kalimat secara *bidirectional*. BERT memberikan representasi setiap kata dalam kalimat sebagai *output*.

SmallBERT adalah arsitektur yang sama dengan BERT, namun dengan jumlah atau ukuran blok *transformer* yang lebih kecil, sehingga memungkinkan untuk mengeksplorasi *tradeoff* antara kecepatan, ukuran, dan kualitas untuk meningkatkan aspek tertentu sambil mengorbankan yang lain. Model ini dipublikasikan oleh Turc *et al.* (2019) dan dibahas dalam paper yang berjudul “*Well-Read Students Learn Better: On the Importance of Pre-training Compact Models*” [18]. Paper ini membahas pentingnya *pre-training* pada model kecerdasan buatan yang lebih kecil untuk meningkatkan kemampuan memahami teks, dan hasil penelitian menunjukkan bahwa *pre-training* tetap penting bahkan pada arsitektur model yang lebih kecil untuk bersaing dengan metode yang lebih kompleks. Penelitian ini juga menyebutkan bahwa Turc *et al.* (2019) telah menyediakan 24 model BERT miniatur terlatih (*pre-trained*) yang dapat digunakan secara umum [18]. Salah satu model yang disediakan oleh Turc *et al.* (2019) dan digunakan pada penelitian ini adalah model *pre-trained* SmallBERT bert\_en\_uncased\_L-4\_H-512\_A-8/1 pada TensorFlow Hub. Model ini memiliki 4 lapisan dengan masing-masing 512 unit pemrosesan dan 8 kepala perhatian. SmallBERT menggunakan arsitektur Transformer dengan 4 lapisan *encoder* yang terdiri dari *multi-head self-attention* dan *feedforward neural network*. Adanya 512 unit pemrosesan pada setiap lapisan dan 8 *attention heads* per lapisan, SmallBERT dapat mempelajari representasi bahasa yang lebih kaya dan efisien. Perbedaan sekaligus hubungan antara Transformer, BERT, dan SmallBERT diringkas dalam Tabel 1 di bawah ini.

**Tabel 1. Perbedaan Transformer, BERT, dan SmallBERT**

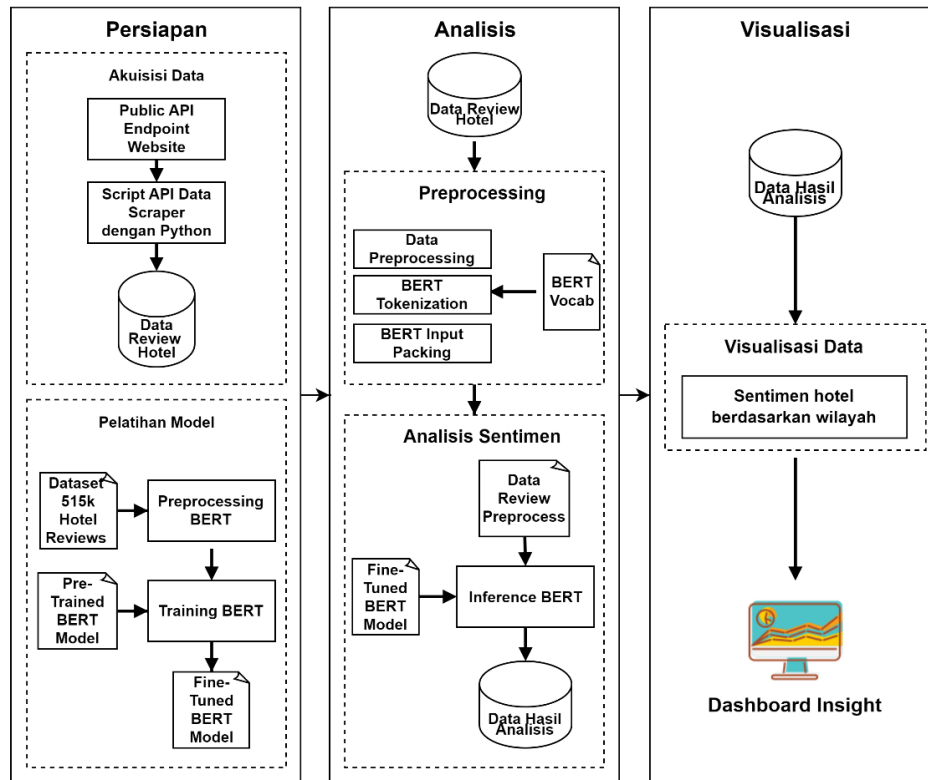
No.	Model	Deskripsi
1.	Transformer	Arsitektur model yang menggabungkan <i>self-attention</i> dan <i>layer-layer neural network</i> untuk pemrosesan sekuensial [9].
2.	BERT	Model <i>pre-trained</i> menggunakan arsitektur Transformer yang dilatih pada tugas pemodelan bahasa berbasis konteks [11].
3.	SmallBERT	Versi tereduksi dari BERT yang memiliki jumlah parameter dan ukuran model yang lebih kecil untuk keperluan efisiensi [18].

### 2.3. Fine Tune

Beberapa *paper* penelitian telah menekankan pentingnya *fine-tuning* BERT pada *dataset* tugas spesifik untuk mencapai kinerja terbaik pada berbagai tugas pemrosesan bahasa alami. Penelitian oleh Devlin *et al.* (2019) menegaskan bahwa *fine-tuning* diperlukan setelah BERT dilatih pada data besar untuk mencapai kinerja terbaik pada tugas-tugas NLP. *Fine-tuning* melibatkan penambahan lapisan *output* khusus dan pembelajaran terawasi, dan secara konsisten meningkatkan kinerja model pada berbagai tugas NLP, bahkan sering kali menghasilkan kinerja terbaik di kelasnya [11]. Penelitian selanjutnya yang dilakukan Colin Raffel *et al.* (2019) menunjukkan bahwa untuk mencapai kinerja terbaik dalam pemrosesan bahasa alami, model Transformer memerlukan proses *fine-tuning* melalui penyesuaian parameter *pre-trained*, pembelajaran terawasi pada data tugas tertentu, dan penambahan *output* layer khusus. *Fine-tuning* secara konsisten meningkatkan kinerja model pada berbagai tugas NLP dan dapat diterapkan pada tugas NLP lainnya [19]. BERT memerlukan penyesuaian untuk mencapai kinerja terbaik pada tugas spesifik seperti *intent classification* dan *slot filling* dalam sistem dialog atau *chatbot*. Penyesuaian tersebut meliputi menambahkan lapisan *output*, melakukan pembelajaran terawasi pada data spesifik, dan penyesuaian parameter *pre-trained* [20]. Penelitian oleh Zhao dan Sun (2022) memberikan hasil model BERT hasil *fine-tuning* efektif dalam memprediksi skor ulasan makanan dengan akurasi yang tinggi [21].

### 3. Metodologi Penelitian

Gambaran umum penelitian terdiri dari tiga bagian yaitu persiapan, analisis, dan visualisasi. Bagian persiapan mencakup akuisisi data dan pelatihan model, sedangkan analisis terdiri dari *preprocessing* dan analisis sentimen menggunakan model *deep learning* BERT. Bagian visualisasi fokus pada memvisualisasikan data sentimen. Gambaran umum penelitian dapat dilihat pada Gambar 1.



Gambar 1. Gambaran Umum Penelitian

#### 3.1. Tahap Persiapan

Proses persiapan dimulai dengan mengumpulkan 2.886.047 data *review* hotel dari seluruh Indonesia. Metode pengumpulan data yang digunakan dalam penelitian ini adalah dengan cara *text mining* menggunakan *endpoint API public* di *website* penyedia layanan *booking* hotel terbesar di Indonesia. Rentang waktu data yang dikumpulkan dimulai dari 8 September 2014 hingga 14 Januari 2023. Hasil akuisisi data *review* dapat dilihat pada Gambar 2.

hotelid	overallScore	cleanlinessScore	comfortScore	serviceScore	locationScore	language	timestamp	translated	reviewId	profileId	reactionSummaries	reviewText
012273	9.7	10.0	10.0	10.0	10.0	UNKNOWN	2023-01-06 11:48:22.118	False	127141192	19497096.0	{'reactionSummaryMap': {'LIKE': {}}}	NaN
012273	10.0	10.0	10.0	10.0	10.0	ENGLISH	2022-11-02 03:16:08.495	False	126993390	255317127.0	{'reactionSummaryMap': {'LIKE': {}}}	Comfortable to stay, service good, room clean.
012273	8.5	8.0	8.0	8.0	8.0	UNKNOWN	2022-10-31 09:53:16.157	False	126990135	63330782.0	{'reactionSummaryMap': {'LIKE': {}}}	NaN
012273	9.4	10.0	8.0	10.0	10.0	UNKNOWN	2022-10-26 08:08:54.922	False	126973772	24162217.0	{'reactionSummaryMap': {'LIKE': {}}}	NaN
012273	9.7	10.0	10.0	10.0	10.0	ENGLISH	2022-10-17 08:52:07.411	False	126961781	3636013.0	{'reactionSummaryMap': {'LIKE': {}}}	The room is clean. The staffs are friendly and...
...	...	...	...	...	...	...	...	...	...	...	...	...
217381	6.0	6.0	6.0	6.0	6.0	ENGLISH	2015-09-07 03:40:29.435	True	107662628	5222074.0	{'reactionSummaryMap': {'LIKE': {}}}	The hot water in the room was broken and was n...

Gambar 2. Hasil Akuisisi Data Review Hotel

Selanjutnya, dilakukan pelatihan model klasifikasi sentimen menggunakan *dataset* 515k *hotel reviews* dari Kaggle [22]. *Dataset* ini terdiri dari *review* dengan sentimen positif dan negatif. *Dataset* yang digunakan untuk proses pelatihan tersebut dapat dilihat di Tabel 3.

**Tabel 3. Dataset 515k Hotel Reviews Beserta Labelnya**

text	label
perfect	1
Beautiful roof views excellent clean rooms	1
The breakfast was luke warm	0
ve made top marks if there was more variety at the breakfast buffet as it became very repetitive at day 3	0
Location is very good	1
areas where my room was located The lighting in the room was far too low definitely needs to be brigh...	0
oncurrent users Some daytime noise due to what I assume was nearby construction starting up at 7am...	0
e who are hungry missing variety in the menu Overall atmosphere in the restaurant and hall very stylish ...	0
vas great Very small but they have designed it in such a way that it s not important The rooms are superb	1
Na	0
way The rooms are comfortable too Most importantly the staff is very helpful Especially those guys at...	1

*Dataset* untuk pelatihan dibagi menjadi data *train*, data validasi, dan data uji dengan persentase 70:15:15 dari total 867.640 data. 70% untuk data *train*, yaitu 607.348 data, 15% untuk data validasi yaitu 130.146 data, dan 15% untuk data tes yaitu 130.146 data. Subset yang digunakan pada penelitian ini adalah 70% data *train* dan 15% data validasi karena data uji akan menggunakan *dataset* lain. Pelatihan model menggunakan *pre-trained* SmallBERT bert\_en\_uncased\_L-4\_H-512\_A-8/1 dari TensorFlow *model garden*. Dalam proses *fine-tuning*, *encoder* layer dari model BERT diatur sebagai *trainable* untuk menyesuaikan bobot dan bias dengan *dataset* spesifik. Penggunaan *optimizer* Adam *Weighted* dengan *learning rate scheduling* 3e-5 digunakan untuk memperbaiki model dengan efisiensi perhitungan yang baik. Selama pelatihan, model dilatih dengan menggunakan *callback* ModelCheckpoint untuk menyimpan *checkpoint* model terbaik. *Hyperparameter* yang digunakan adalah jumlah *epoch* lima, fungsi *loss* BinaryCrossentropy, dan *metrics* BinaryAccuracy untuk meningkatkan performa analisis sentimen *review* hotel. Proses *fine-tuning* ini bertujuan untuk meningkatkan kemampuan model dalam mengklasifikasikan sentimen *review* hotel dengan lebih akurat. Model hasil *fine-tuning* ini selanjutnya disimpan agar dapat dimuat untuk proses *testing* dan analisis sentimen. Data uji yang digunakan adalah *dataset review* hotel hasil akuisisi berjumlah 4.032 data yang telah dilabeli secara manual sebagai positif atau negatif dengan label 1 dan 0. *Dataset* ini dipilih secara selektif. Penelitian ini menggunakan 15% dari 4.032 data uji, yaitu 605 baris data uji yang dilabeli secara manual. *Dataset* yang digunakan untuk data uji ini dapat dilihat pada Tabel 4.

**Tabel 4. Dataset Manually Labeled**

	text	label
0	We stayed in the family suite. There room are s...	1
1	After staying for two days, there were a numbe...	0
2	The hotel is dirty, the room is dirty.	0
3	Customer service provides personal information...	0
4	Comfortable, clean and friendly.	1
...	...	...
2817	The hotel is near the Bali Galeria Mall, only ...	0
2818	The security guard is dirty, there is no smile...	0
2819	Love the room, the staffs are friendly and the...	1

### 3.2. Tahap Analisis Sentimen

Analisis dilakukan dengan *preprocessing* data *review* terlebih dahulu. *Preprocessing* yang dilakukan adalah (1) menghapus baris *review* yang *null*, (2) menerjemahkan ke dalam bahasa Inggris, (3) melakukan HTML *unescape* untuk mengonversi karakter ASCII dengan karakter khusus HTML, (4) menghapus tanda baca yang berulang seperti tanda titik yang muncul

sebanyak tiga kali, dan (5) menghapus *emoticon* pada teks yang akan dianalisis untuk memberikan hasil analisis yang optimal. Proses analisis sentimen dimulai dengan memuat model hasil *fine-tuning* yang telah dilakukan pada tahap persiapan sebelumnya. Selanjutnya, dilakukan tahap tokenisasi menggunakan BERT Tokenizer untuk mengubah teks menjadi representasi angka yang sesuai dengan *vocabulary* BERT.

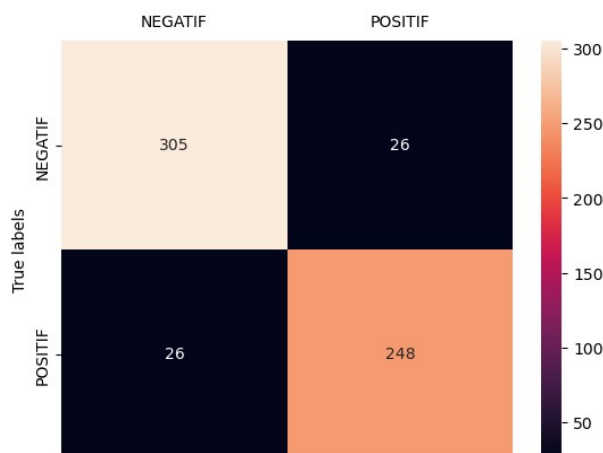
BERT *Tokenizer* menggantikan kata dengan indeks kata yang terdapat dalam *vocabulary*. Selanjutnya, dilakukan *input packing* untuk mengubah struktur data menjadi *input* yang sesuai dengan model BERT. *Input packing* melibatkan beberapa hal seperti representasikan kata-kata dalam bentuk angka setelah tokenisasi, representasi biner yang menunjukkan keberadaan kata dalam indeks *array* tokenisasi, dan indeks kalimat dari kata-kata tersebut. Setelah itu, model melakukan klasifikasi sentimen dengan *binary classification* menjadi label negatif dan positif. Prediksi dilakukan dengan menghasilkan nilai antara 0 dan 1 yang diinterpretasikan sebagai probabilitas. Nilai prediksi dibandingkan dengan ambang batas yang telah ditentukan, yaitu 0,5. Jika probabilitas lebih besar dari ambang batas, sentimen ulasan dikategorikan sebagai "Positif". Jika probabilitas lebih kecil atau sama dengan ambang batas, sentimen ulasan dikategorikan sebagai "Negatif". Jika nilai prediksi lebih kecil dari ambang batas, variabel *confidence* dihitung sebagai selisih antara 1 dan nilai prediksi. Hal ini dilakukan karena nilai prediksi dianggap sebagai probabilitas kelas negatif. Nilai *confidence* pada prediksi kelas negatif dibalik menggunakan formula  $1 - \text{confidence}$ . Sebagai contoh, jika nilai prediksi adalah 0,3, maka nilai *confidence* dibalik menjadi 0,7 karena model 70% yakin bahwa itu adalah kelas negatif.

### 3.3. Tahap Visualisasi

Setelah proses analisis data selesai, langkah selanjutnya adalah melakukan visualisasi data menggunakan perangkat lunak Tableau. Data yang telah dianalisis diekspor ke format CSV, kemudian dilakukan *filtering* untuk memilih fitur atau kolom yang akan digunakan untuk visualisasi. Selanjutnya, dilakukan pemilihan jenis diagram yang paling sesuai untuk memvisualisasikan data perbandingan sentimen, yaitu diagram batang dan *pie chart*.

### 4. Hasil dan Diskusi

Model SmallBERT dilatih menggunakan proses *fine tuning pre-trained* model dan *dataset 515k hotel reviews*. Model mencapai nilai *loss* 0.0879 dan *binary\_accuracy* 97,25% pada data *training*, serta *val\_loss* 0.1429 dan *val\_binary\_accuracy* 96.18% pada data validasi pada *epoch* terakhir. Model tidak *overfitting* atau *underfitting* dan evaluasi silang dilakukan pada tahap selanjutnya.



Gambar 5. *Confusion Matrix* Model Hasil *Fine-Tuning*

Evaluasi silang dilakukan pada 15% data evaluasi yaitu 605 data yang telah diberi label secara manual. Hasilnya direpresentasikan dalam bentuk *confusion matrix* pada model yang

dilatih dengan *dataset 515k hotel reviews* pada data uji yang dapat dilihat pada Gambar 5. Terdapat 305 ulasan yang diklasifikasikan dengan benar sebagai negatif (*True Negative/TN*), 26 ulasan yang salah diklasifikasikan sebagai positif (*False Positive/FP*), 26 ulasan yang salah diklasifikasikan sebagai negatif (*False Negative/FN*), dan 248 ulasan yang diklasifikasikan dengan benar sebagai positif (*True Positive/TP*). Performa model diukur melalui beberapa metrik evaluasi seperti akurasi, *precision*, *recall*, dan *F1-score* yang dihitung pada Persamaan 1-4 .

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} = \frac{248 + 305}{248 + 305 + 26 + 26} = \frac{553}{605} = 0,914049 \text{ atau } 91,41\% \tag{1}$$

$$Precision = \frac{TP}{TP + FP} = \frac{248}{248 + 26} = \frac{248}{274} = 0,905109 \text{ atau } 90,51\% \tag{2}$$

$$Recall = \frac{TP}{TP + FN} = \frac{248}{248 + 26} = \frac{248}{274} = 0,905109 \text{ atau } 90,51\% \tag{3}$$

$$F1 - score = \frac{2TP}{2TP + FP + FN} = \frac{2 * 248}{(2 * 248) + 26 + 26} = \frac{496}{548} = \frac{248}{274} = 0,905109 \text{ atau } 90,51\% \tag{4}$$

Tabel 2 menunjukkan hasil evaluasi kinerja model analisis sentimen SmallBERT yang dilatih dengan *dataset 515k hotel reviews* terhadap data uji yang telah dilabeli secara manual. Model tersebut mencapai akurasi 91,40%, *precision* 90,51%, *recall* 90,51%, dan *F1-score* 90,51%, menunjukkan kinerja yang baik dalam klasifikasi sentimen pada 605 *review* hotel. Model SmallBERT tanpa *fine-tuning* memiliki performa evaluasi rendah dengan akurasi 45,45% dan *F1 score* 54,17%. *Fine-tuning* penting untuk mempelajari karakteristik data, memperbaiki kesalahan, dan meningkatkan performa model dengan menyesuaikan bobot pada layer tertentu saat proses analisis. Perbandingan hasil evaluasi ini menegaskan pentingnya *fine-tuning* untuk meningkatkan performa model.

**Tabel 2. Perbandingan Matriks Performa Model**

No.	Dataset	Data Train	Data Validation	Epoch	Accuracy	Precision	Recall	F1 score
1	Model Hasil <i>Fine-Tuning</i>	607.348 data	130.146 data	5	91,41%	90,51%	90,51%	90,51%
2	Model Tanpa <i>Fine-Tune</i>	-	-	-	45.45%	43.72%	71.17%	54.17%

Kumpulan *review* hotel yang telah diakuisisi harus melalui tahap *preprocessing* sebelum dianalisis. Tahap pertama adalah penghapusan data yang *null*. Selanjutnya, kalimat-kalimat dalam *review* yang bukan dalam bahasa Inggris perlu diterjemahkan karena data *training* menggunakan bahasa Inggris. Langkah penting berikutnya adalah melakukan *HTML unescape* untuk mengembalikan karakter-karakter khusus yang dienkripsi dalam format HTML. Hasil dari *HTML unescape* terlihat pada Gambar 6 yang menunjukkan tanda “%#39;” sudah berhasil diubah menjadi tanda “'” pada proses *HTML unescaped*. Hasil penghapusan tanda baca berulang ditunjukkan oleh Gambar 7 yang memperlihatkan tanda baca berulang seperti titik telah berhasil dihapus, sehingga hanya titik pertama yang tersisa. *Emoticon* yang sebelumnya ada sudah terhapus pada proses penghapusan *emoticon*, seperti yang tampak pada Gambar 8.

The Glory really good, it's so cool  
 The Glory really good, it's so cool

**Gambar 6. Hasil HTML Unescape**

```
remove_redundant_punc("test kata. ini kalimat. ini kalimat nyambung. ini kalimat...apaya... yah... jij!0  

'test kata. ini kalimat. ini kalimat nyambung. ini kalimat. apaya. yah. jij!0Jo 0@J 0!2@'
```

**Gambar 7. Hasil Hapus Tanda Baca Berulang**

This dog 🤪  
This dog

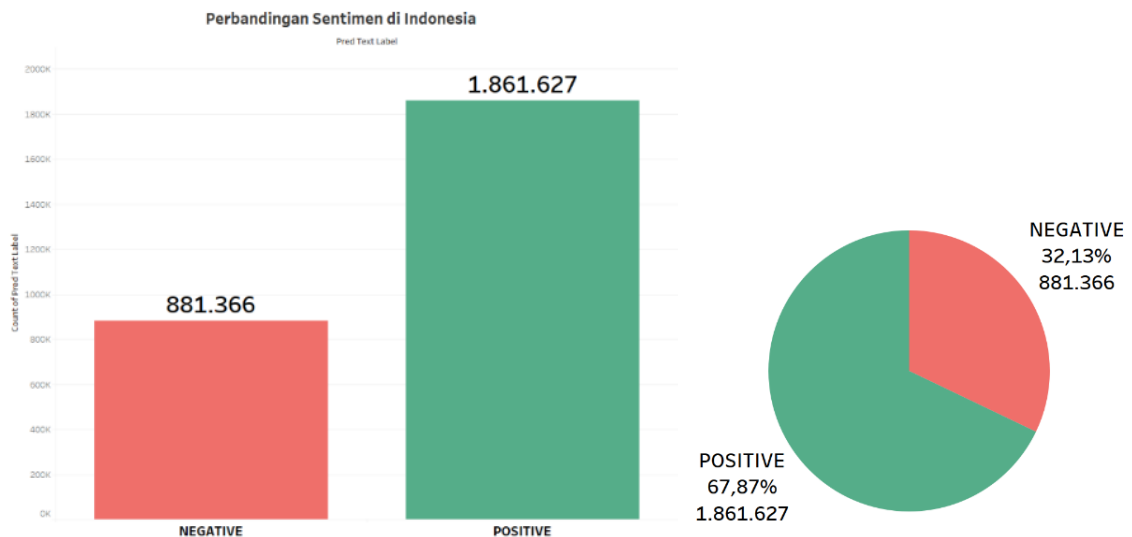
Gambar 8. Hasil Hapus Emoticon

Sebanyak 2.743.029 data *review* hotel yang merupakan hasil dari proses *preprocessing* dianalisis dengan model *smallBERT* hasil *fine-tuning*. Hasil analisis ditampilkan dalam bentuk tabel yang menunjukkan penambahan kolom baru, yaitu *pred\_probability*, *pred\_confidence*, *pred\_rounded\_label*, dan *pred\_text\_label*. Skor probabilitas mendekati 1 menunjukkan sentimen positif, sedangkan mendekati 0 menunjukkan sentimen negatif. Gambar 9 adalah hasil dari analisis sentimen.

reviewTextRegex	overallScore	language	pred_probability	pred_confidence	pred_rounded_label	pred_text_label
The bad spring which is on the edge is crumbli...	6.0	en	0.002665	0.997335	0.0	NEGATIVE
First time staying here, only 1 here is good, ...	6.3	en	0.008925	0.991075	0.0	NEGATIVE
The rooms are not clean, the air conditioner i...	5.8	en	0.006296	0.993704	0.0	NEGATIVE

Gambar 9. Hasil Analisis Sentimen

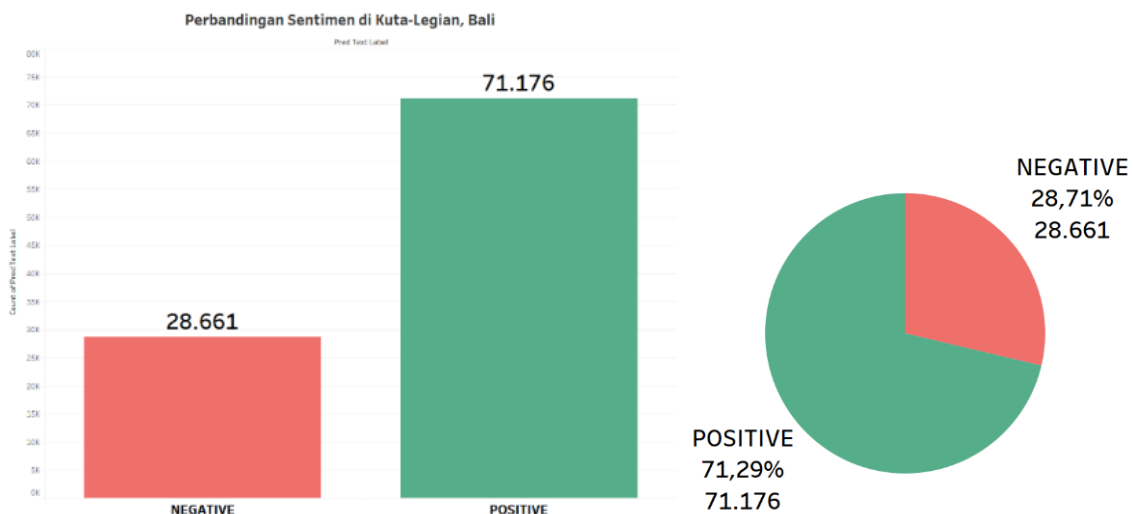
Visualisasi perbandingan antara sentimen positif dan negatif dibagi menjadi dua bagian, yaitu untuk seluruh wilayah Indonesia dan untuk wilayah tertentu, yaitu Kuta dan Legian di Bali. Perbandingan sentimen positif dan negatif di seluruh Indonesia memberikan hasil sebanyak 1.861.627 ulasan atau 67,9% dianggap positif dan 881.366 ulasan atau 32,1% dianggap negatif dari total 2.743.993 ulasan. Meskipun mayoritas tamu memberikan ulasan positif, hasil analisis tersebut tidak mewakili seluruh preferensi tamu sehingga membutuhkan penelitian lebih lanjut untuk mendapatkan gambaran yang lebih akurat. Visualisasi disajikan dalam bentuk grafik batang dan grafik lingkaran pada Gambar 10.



Gambar 10. Bar dan Pie Chart Perbandingan Sentimen Positif dan Negatif di Seluruh Indonesia

Visualisasi pada Gambar 11 menunjukkan hasil analisis sentimen pada wilayah Kuta dan Legian di Bali bahwa mayoritas ulasan hotel memiliki sentimen positif sebesar 71,29% dan sentimen negatif hanya 28,71%. Hal ini dapat mempengaruhi peningkatan jumlah kunjungan wisata ke wilayah tersebut, sehingga wilayah Kuta dan Legian di Bali dianggap sebagai tujuan wisata yang menarik bagi wisatawan yang ingin merasakan pengalaman positif.





Gambar 11. Bar dan Pie Chart Perbandingan Sentimen Positif dan Negatif di Kuta-Legian, Bali

## 5. Kesimpulan dan Saran

Kesimpulan penelitian ini adalah metode *fine-tuning* pada model *deep learning* BERT, khususnya SmallBERT dapat digunakan untuk klasifikasi sentimen dengan hasil yang baik tanpa banyak *preprocessing*. Model SmallBERT dilatih dengan *dataset* 515k *hotel reviews* selama lima *epochs* dan berhasil mencapai hasil yang memuaskan. Evaluasi silang pada data uji menunjukkan akurasi, *precision*, *recall*, dan *f1 score* yang baik. Mayoritas ulasan hotel di Indonesia dan wilayah Kuta-Legian menunjukkan sentimen positif selama delapan tahun empat bulan terakhir, yang dapat mempengaruhi peningkatan jumlah kunjungan wisata ke wilayah tersebut. Penelitian selanjutnya perlu mencoba variasi konfigurasi SmallBERT, metode *fine-tuning*, dan kombinasi *hyperparameter* dengan data yang lebih beragam dan pembagian data yang berbeda.

## Referensi

- [1] S. Wibowo, O. Rusmana, S. Muslim, and S. Tinggi Pariwisata Bandung, "Pengembangan Ekonomi Melalui Sektor Pariwisata Tourism," *Jurnal Kepariwisata: Destinasi, Hospitalitas dan Perjalanan*, vol. 1, no. 2, pp. 93–99, Dec. 2017, doi: 10.34013/JK.V1I2.13.
- [2] Badan Pusat Statistik, "Tingkat Penghunian Kamar (TPK) Hotel," *Sistem Informasi Rujukan Statistik*. <https://sirusa.bps.go.id/sirusa/index.php/indikator/60> (accessed Apr. 18, 2022).
- [3] Kemenparekraf/Baparekraf, "Statistik Kunjungan Wisatawan Mancanegara Bulan Januari 2022," Mar. 11, 2022. <https://kemenparekraf.go.id/statistik-wisatawan-mancanegara/statistik-kunjungan-wisatawan-mancanegara-bulan-januari-2022> (accessed Apr. 08, 2022).
- [4] J. J. Kim and H. Han, "Saving the hotel industry: Strategic response to the COVID-19 pandemic, hotel selection analysis, and customer retention," *Int J Hosp Manag*, vol. 102, p. 103163, Apr. 2022, doi: 10.1016/J.IJHM.2022.103163.
- [5] S. Alvianna, "Analisis Pengaruh Harga, Produk, Dan Kualitas Layanan Terhadap Kepuasan Wisatawan Di Taman Wisata Air Wendit Kabupaten Malang" *Jurnal Pariwisata Pesona*, vol. 2, no. 1, Jun. 2017, doi: 10.26905/JPP.V2I1.1263.
- [6] J. Thanaki, *Python natural language processing : explore NLP with machine learning and deep learning techniques*. Birmingham: Packt Publishing Ltd., 2017.
- [7] P. Chinnasamy, V. Suresh, K. Ramprathap, B. J. A. Jebamani, K. Srinivas Rao, and M. Shiva Kranthi, "COVID-19 vaccine sentiment analysis using public opinions on Twitter," *Mater Today Proc*, Apr. 2022, doi: 10.1016/J.MATPR.2022.04.809.

- [8] M. Singh, A. K. Jakhar, and S. Pandey, "Sentiment analysis on the impact of coronavirus in social life using the BERT model," *Soc Netw Anal Min*, vol. 11, no. 1, pp. 1–11, Dec. 2021, doi: 10.1007/s13278-021-00737-z.
- [9] A. Vaswani et al., "Attention Is All You Need," *Adv Neural Inf Process Syst*, vol. 2017-December, pp. 5999–6009, Jun. 2017, Accessed: Apr. 24, 2023. [Online]. Available: <https://arxiv.org/abs/1706.03762v5>
- [10] S. Ravichandiran, *Getting started with Google BERT : build and train state-of-the-art natural language processing models using BERT*. Birmingham: Packt Publishing Ltd., 2021.
- [11] J. Devlin, M. W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," *NAACL HLT 2019 - 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies - Proceedings of the Conference*, vol. 1, pp. 4171–4186, Oct. 2018, doi: 10.48550/arxiv.1810.04805.
- [12] Y. Zhang, K. Chen, Y. Weng, Z. Chen, J. Zhang, and R. Hubbard, "An intelligent early warning system of analyzing Twitter data using machine learning on COVID-19 surveillance in the US," *Expert Syst Appl*, vol. 198, p. 116882, Jul. 2022, doi: 10.1016/J.ESWA.2022.116882.
- [13] S. Tabinda Kokab, S. Asghar, and S. Naz, "Transformer-based deep learning models for the sentiment analysis of social media data," *Array*, vol. 14, p. 100157, Jul. 2022, doi: 10.1016/J.ARRAY.2022.100157.
- [14] S. Y. Lin, Y. C. Kung, and F. Y. Leu, "Predictive intelligence in harmful news identification by BERT-based ensemble learning model with text sentiment analysis," *Inf Process Manag*, vol. 59, no. 2, pp. 1–18, Mar. 2022, doi: 10.1016/J.IPM.2022.102872.
- [15] K. S. Nugroho, A. Y. Sukmadewa, H. W. DW, F. A. Bachtiar, and N. Yudistira, "BERT Fine-Tuning for Sentiment Analysis on Indonesian Mobile Apps Reviews," *ACM International Conference Proceeding Series*, pp. 258–264, Jul. 2021, doi: 10.1145/3479645.3479679.
- [16] B. Selvakumar and B. Lakshmanan, "Sentimental analysis on user's reviews using BERT," *Mater Today Proc*, Apr. 2022, doi: 10.1016/J.MATPR.2022.03.678.
- [17] B. Ray, A. Garain, and R. Sarkar, "An ensemble-based hotel recommender system using sentiment analysis and aspect categorization of hotel reviews," *Appl Soft Comput*, vol. 98, p. 106935, Jan. 2021, doi: 10.1016/J.ASOC.2020.106935.
- [18] I. Turc, M.-W. Chang, K. Lee, and K. Toutanova, "Well-Read Students Learn Better: On the Importance of Pre-training Compact Models," *Google Research*, pp. 1–13, Aug. 2019, Accessed: Apr. 05, 2023. [Online]. Available: <https://arxiv.org/abs/1908.08962v2>
- [19] C. Raffel et al., "Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer," *Journal of Machine Learning Research*, vol. 21, pp. 1–67, Oct. 2019, Accessed: Apr. 10, 2023. [Online]. Available: <https://arxiv.org/abs/1910.10683v3>
- [20] Q. Chen, Z. Zhuo, and W. Wang, "BERT for Joint Intent Classification and Slot Filling," Feb. 2019, Accessed: Apr. 10, 2023. [Online]. Available: <https://arxiv.org/abs/1902.10909v1>
- [21] X. Zhao and Y. Sun, "Amazon Fine Food Reviews with BERT Model," *Procedia Comput Sci*, vol. 208, pp. 401–406, Jan. 2022, doi: 10.1016/J.PROCS.2022.10.056.
- [22] J. Liu, "515K Hotel Reviews Data in Europe | Kaggle," 2017. <https://www.kaggle.com/datasets/jiashenliu/515k-hotel-reviews-data-in-europe> (accessed Apr. 21, 2023).