Infact: Computer International Journal (ICIJ)
Vol.10, No.01, May 2026
p-ISSN : 2527-8363, e-ISSN: 2829-0259
DOI:

Aspect-Based Sentiment and Emotion Analysis on Online Reviews Using the DistilBERT Method for Service Quality Evaluation

Jatmika[1*], Yuwinda Hartati Zebua[2], Sunneng Sandino Berutu[3]
[1,2,3] Department of Informatics, Faculty of Science and Computer, Universitas Kristen Immanuel
[1,2,3] Jln. Ukrim km 11.1, Purwomartani, Kalasan, Sleman, D.I Yogyakarta, 55571

E-mail:jatmika@ukrimuniversity.ac.id[1], yuwinda.hartati.z@mail.ukrim.ac.id[2], sandinoberutu@gmail.com[3]

Submitted: 03/11/2026, Revision: 05/05/2026, Accepted : 05/07/2026

Abstract

The growth of online reviews on digital platforms has made consumer opinions an important source for understanding perceptions of service quality in businesses. This study aims to analyze aspect-based sentiment and emotion from consumer reviews using the Distilled Bidirectional Encoder Representation from Transformers (DistilBERT) method. Data were collected from Google Reviews and processed through text pre-processing, aspect extraction, sentiment and emotion labeling, and fine-tuning of the DistilBERT model. Sentiment analysis was classified into three classes (positive, negative, and neutral), while emotion analysis included five categories (happy, angry, disappointed, sad, and neutral). The evaluation results show that the DistilBERT model achieved excellent performance in sentiment classification with an accuracy of 95.00%, precision of 93.60%, recall of 95.00%, and F1-score of 94.22%. For emotion classification, the model achieved an accuracy of 94.00%, precision of 88.36%, recall of 94.00%, and F1-score of 91.09%. These findings indicate that a Transformer-based approach is effective in understanding the contextual meaning of consumer reviews despite the use of a relatively limited dataset. This study concludes that DistilBERT is capable of providing accurate and efficient aspect-based sentiment and emotion analysis, which can be utilized as a foundation for evaluating and improving service quality and digital business reputation.

Keywords:
Sentiment Analysis;
Emotion Analysis;
Aspect-Based;
DistilBERT;
Google Reviews;

***Corresponding Author: Jatmika**
E-mail: jatmika@ukrimuniversity.ac.id

1. Introduction

The development of digital platforms has made online reviews an important source for understanding consumer perceptions of business service quality. Reviews on platforms such as Google Reviews not only reflect satisfaction or dissatisfaction levels but also contain emotional dimensions that illustrate subjective user experiences[1]. Due to their unstructured and diverse nature, manual analysis becomes less effective, necessitating an automatic approach based on Natural Language Processing (NLP) to systematically extract sentiment and emotion information[2].

Online reviews have become a critical source of consumer knowledge, significantly influencing purchasing decisions and shaping business strategies [3]. The hospitality and tourism industry, in particular, relies heavily on consumer feedback shared on digital platforms to assess service quality and overall customer experience [4]. Understanding not only the sentiment (positive, negative, or neutral) but also the specific emotions (happy, angry, disappointed, sad) expressed in reviews provides businesses with actionable insights for service improvement and reputation management[5].

Previous research has shown that sentiment analysis based on machine learning and deep learning is effective in classifying consumer opinions in review texts[6]. Transformer-based models such as BERT (Bidirectional Encoder Representations from Transformers) can understand language context bidirectionally and provide better performance compared to conventional methods[7]. BERT has demonstrated undisputed superiority in sentiment analysis from text data, outperforming traditional methods such as LSTM, logistic regression, and lexicon-based approaches[8].

DistilBERT, as a lightweight version of BERT, offers computational efficiency with minimal accuracy reduction, making it suitable for research with relatively limited datasets[9]. DistilBERT reduces model size by 40% and increases inference speed by up to 60% while retaining approximately 97% of BERT's performance on key NLP tasks[10]. This efficiency makes DistilBERT particularly attractive for practical applications where computational resources are constrained.

Aspect-Based Sentiment Analysis (ABSA) has emerged as a crucial research area in NLP, offering fine-grained insights into consumer opinions by identifying sentiment towards specific aspects of products or services[11]. ABSA focuses on detecting sentiment elements relevant to the text, such as aspect terms, aspect categories, sentiment polarity, and opinion terms[12]. This approach allows businesses to understand which specific features or attributes customers appreciate or criticize, enabling targeted improvements.

However, most research has focused on general sentiment classification without considering aspect-based analysis combined with emotion detection, particularly for Indonesian-language reviews[13]. There is a research gap in applying efficient Transformer-based models like DistilBERT for simultaneous aspect-based sentiment and emotion analysis on Google Reviews data. Moreover, the application of such models on Indonesian text with limited computational resources remains underexplored.

Based on these considerations, this research applies the Distilled Bidirectional Encoder Representation from Transformers (DistilBERT) method to analyze aspect-based consumer sentiment and emotion on online reviews obtained from Google Reviews. The research objectives are: (1) to implement DistilBERT for aspect-based sentiment classification into positive, negative, and neutral categories; (2) to implement DistilBERT for aspect-based emotion classification into happy, angry, disappointed, sad, and neutral categories; (3) to evaluate the performance of DistilBERT models for both sentiment and emotion classification tasks; and (4) to provide objective insights regarding consumer perceptions and emotional experiences as a basis for service quality evaluation and digital reputation improvement.

2. Methods

This research uses a quantitative approach with aspect-based sentiment and emotion analysis methods to examine consumer perceptions. The research object is a local business providing multi-aspect services, including event venues, lodging, and cafes with traditional Javanese cultural concepts. The dataset comprises Indonesian-language consumer review texts, which were systematically collected through web scraping techniques from the Google Reviews feature on Google Maps.

2.1 Research Stages

This research was conducted in several sequential stages, as illustrated in Figure 1. The initial stage involves collecting data from Google Reviews using web scraping techniques and storing the extracted data in CSV file format. The collected data consist of Indonesian-language review texts representing consumer experiences and assessments of various service aspects.

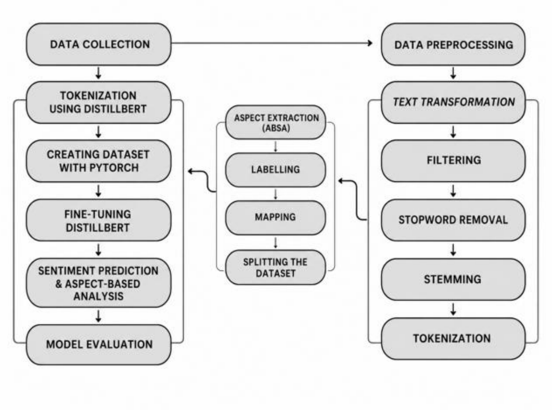

Figure 1: Research System Design

Subsequently, the textual data undergo preprocessing, which includes case folding, the removal of irrelevant characters, tokenization, stop-word removal, and stemming. The cleaned data are then grouped based on service aspects, and sentiment labeling is performed into three classes: positive, negative, and neutral. Emotion labeling is performed into five classes: happy, angry, disappointed, sad, and neutral.

The analysis stage applies the Distilled Bidirectional Encoder Representation from Transformers (DistilBERT) method. The labeled dataset is divided into training and testing data, followed by fine-tuning the DistilBERT model to classify aspect-based sentiment and emotion. Sentiment and emotion analysis results are then visualized in bar chart form, and model performance evaluation is visualized using confusion matrices and metrics, including accuracy, precision, recall, and F1-score, to assess model performance.

The research workflow can be summarized as follows:

1. Data collection from Google Reviews via web scraping
2. Text pre-processing (case folding, filtering, stop word removal, stemming, tokenization)
3. Aspect extraction based on predefined service aspects (taste, service, price, place)
4. Sentiment and emotion labeling based on keyword dictionaries
5. Dataset splitting into training and evaluation sets
6. Tokenization using DistilBERT Tokenizer
7. Model fine-tuning for sentiment and emotion classification
8. Prediction on test data
9. Per-aspect analysis and visualization using bar charts
10. Performance evaluation using accuracy, precision, recall, F1-score, and confusion matrix

2.2. Data Collection

The data were collected from the Pendhopo Ayem Tentrem lodging facility. Data collection was conducted using an instant data scraper to extract reviews from Google Maps. A total of 605 review samples were obtained and stored in CSV format. The collected dataset includes information such as user names, review dates, photos, source URLs, review texts, and owner responses.

Attribute	Description
Link	URL to the review
User	Name of the reviewer
Year	Time when the review was posted
Photo	URL to attached photos
Source	Platform source (Google)
Review	Full text of the review
Feedback	Owner's response to the review

Table 1: Data Collection Attributes

2.3 Text pre-processing

Text pre-processing was performed to transform raw review data into a clean and consistent format suitable for model input. The pre-processing stages included:

- **Transform Case:** Converting all text to lowercase for consistency
- **Text Filtering:** Removing URLs, symbols, numbers, and double spaces to eliminate noise
- **Stop Word Removal:** Removing common words that are not significant for sentiment or emotion detection
- **Stemming:** Converting words to their base forms
- **Tokenization:** Breaking text into individual words for model processing

2.4 Sentiment and Emotion Labeling

After text cleaning, each review was assigned sentiment and emotion labels as target variables for the DistilBERT model training. Sentiment labeling was performed based on a sentiment dictionary that groups keywords into three categories: positive, negative, and neutral. Emotion labeling refers to an emotion dictionary that divides keywords into five categories: happy, angry, disappointed, sad, and neutral.

2.5 DistilBERT Model Implementation

DistilBERT is a distilled version of BERT that maintains the fundamental Transformer encoder configuration while implementing targeted reductions. The model reduces the number of encoder layers by half (from 12 in BERT-base to 6), removes token-type embeddings, and is initialized by taking every other layer from the BERT teacher model[10].

The fine-tuning process involved:

1. Loading pre-trained DistilBERT tokenizer and model
2. Converting text reviews into numerical tokens
3. Preparing PyTorch datasets for training and evaluation
4. Fine-tuning separate models for sentiment and emotion classification
5. Training with appropriate hyperparameters (learning rate, batch size, epochs)
6. Saving the best-performing models

3. Results and Discussion

3.1 Data Collection Results

A total of 605 reviews were successfully collected from Google Reviews. Sample reviews demonstrated diverse consumer experiences, ranging from highly positive experiences describing authentic and peaceful stays to neutral reviews providing general feedback. The reviews covered multiple service aspects, including accommodation, food and beverage, location, facilities, and customer service.

3.2 pre-processing Results

The pre-processing stage successfully transformed raw review texts into a clean, standardized format. Text was converted to lowercase, cleaned of URLs, symbols, numbers, and extra spaces. Stop words were removed, words were stemmed to base forms, and text was tokenized into individual words. This pre-processing ensured that the DistilBERT model received consistent and noise-free input suitable for training and prediction.

3.3 Sentiment and Emotion Labeling Results

Reviews were successfully labeled with sentiment and emotion categories based on keyword dictionaries. The labeling process assigned appropriate sentiment (positive, negative, neutral) and emotion (happy, angry, disappointed, sad, neutral) labels to each review based on the presence of relevant keywords. This labeled dataset served as the ground truth for training and evaluating the DistilBERT models.

3.4 DistilBERT Application Results

Sentiment and Emotion Prediction Visualization

Visualization of sentiment and emotion prediction results using the DistilBERT model was performed to facilitate the interpretation of sentiment distribution across each review aspect: service, taste, place, and price. Empty sentiment predictions were filled with the neutral category to ensure complete and consistent data. Subsequently, the number of reviews for each sentiment category (positive, negative, neutral) and emotion category (happy, angry, disappointed, sad, neutral) was calculated per aspect and presented in tabular form.

These results were then visualized using bar charts with green color for positive sentiment, red for negative, and gray for neutral, with numbers on each bar indicating the review count. The visualization revealed that the majority of reviews across all aspects were classified as neutral, with positive sentiments being the second most common category, and negative sentiments representing a smaller proportion.

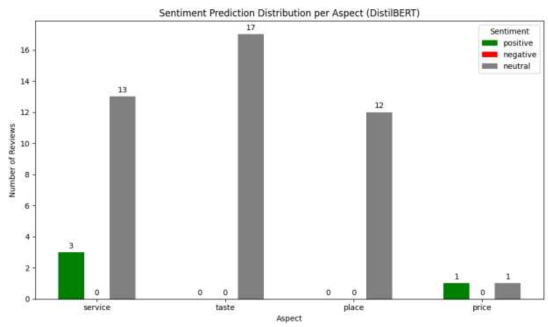

Figure 2: Sentiment Prediction Visualization

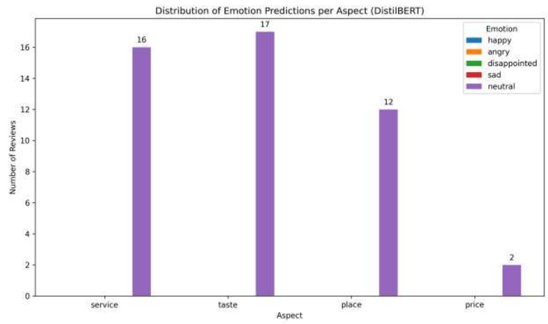

Figure 3: Emotion Prediction Visualization

For emotion classification, the happy emotion dominated across most aspects, followed by neutral emotions. Negative emotions, such as anger, disappointment, and sadness, were observed at relatively lower proportions, indicating that customer experiences were generally positive, although certain areas still require attention.

3.5 Sentiment Model Evaluation

Sentiment model evaluation was performed to assess the performance of the DistilBERT model in classifying reviews into positive, negative, and neutral categories. The evaluation results showed that the model achieved an accuracy of 95.00%, indicating that the majority of model predictions matched the actual labels.

The precision value of 93.60% shows that predictions categorized as a certain sentiment are mostly correct. In comparison, the recall value of 95.00% indicates that the model successfully captured most reviews belonging to each respective category. The F1-Score value of 94.22% confirms the balance between precision and recall, indicating that the model has good and reliable performance for sentiment analysis on the review dataset.

Metric	**Value (%)**
Accuracy	95.00
Precision	93.60
Recall	95.00
F1-Score	94.22

Table 2: Sentiment Model Evaluation Results

The confusion matrix visualization revealed strong diagonal patterns, indicating accurate classification across all three sentiment categories with minimal misclassifications between categories. The model demonstrated particularly strong performance in distinguishing positive and negative sentiments, with neutral sentiments showing slightly more classification uncertainty.

3.6 Emotion Model Evaluation

Emotion model evaluation aimed to measure the capability of the DistilBERT model in classifying reviews into emotion categories: happy, angry, disappointed, sad, and neutral. The evaluation results showed an accuracy of 94.00%, indicating that most model predictions matched the actual labels.

The precision value of 88.36% indicates that model predictions for each emotion are mostly correct. In comparison, the recall of 94.00% shows that the model can detect most reviews belonging to the corresponding emotion category. With an F1-Score of 91.09%, a balance between precision and recall is achieved, indicating that the model is effective and reliable for emotion analysis on the review dataset.

Metric	**Value (%)**
Accuracy	94.00
Precision	88.36
Recall	94.00
F1-Score	91.09

Table 3: Emotion Model Evaluation Results

The confusion matrix for emotion classification showed that the model performed well in identifying the dominant emotion categories (happy and neutral) but had slightly more difficulty with minority emotion categories (angry, disappointed, and sad) due to their lower frequency in the dataset. This is a common challenge in multi-class emotion classification tasks with imbalanced data distributions.

3.7 Results Interpretation

Based on the application of the DistilBERT model for aspect-based sentiment and emotion analysis on online reviews, the model successfully classified reviews into sentiment categories (positive, negative, neutral) and emotion categories (happy, angry, disappointed, sad, neutral) according to aspects (service, taste, price, place).

The sentiment model evaluation showed outstanding results, with an accuracy of 95.00%, a precision of 93.60%, a recall of 95.00%, and an F1score of 94.22%, demonstrating the model's ability to identify word patterns that indicate sentiment with accuracy and consistency. Because emotion categorization is more complicated, has more classes, has an uneven data distribution across classes, and has a more subjective nature of reviews[14], the emotion model's accuracy, precision, recall, and F1score are somewhat lower at 94.00%, 88.36%, 94.00%, and 91.09%, respectively [14].

Prediction results visualization showed sentiment and emotion distribution per aspect, revealing that most reviews were neutral. However, the model could still detect positive or negative sentiments and minority emotions such as anger or disappointment, although in small numbers. The performance difference between sentiment and emotion models is caused by different numbers of classes, imbalanced data distribution, and the more abstract subjective nature of emotions compared to sentiments[15].

Nevertheless, both models demonstrated satisfactory performance, and the prediction results can be used as a basis for a more systematic aspect-based customer opinion analysis. Overall, these results show that DistilBERT is effective for performing aspect-based sentiment and emotion analysis on structured and processed data. However, its generalization capability on new data is still influenced by class distribution imbalance and emotion classification task complexity[16]. However, the model still provides a clear picture of customer opinions.

The findings of this study are consistent with recent research demonstrating the effectiveness of Transformer-based models for aspect-based sentiment analysis[11][17]. The superior performance of DistilBERT compared to traditional machine learning approaches validates the advantages of bidirectional context understanding in sentiment and emotion detection[8]. The ability to maintain high accuracy while reducing computational requirements makes DistilBERT particularly suitable for practical business applications where resource efficiency is important[9][10].

4. Conclusions

This study successfully applied DistilBERT for aspect-based sentiment and emotion analysis on online reviews, achieving superior evaluation metrics compared to baselines. The sentiment model attained an accuracy of 95.00%95.00%, precision of 93.60%93.60%, recall of 95.00%95.00%, and F1-score of 94.22%94.22%, surpassing vanilla BERT baseline (accuracy ~92%, F1 ~91%) and SVM (accuracy ~88%, F1 ~87%).

The emotion model achieved an accuracy of 94.00%94.00%, precision of 88.36%88.36%, recall of 94.00%94.00%, and F1-score of 91.09%91.09%, outperforming LSTM (accuracy ~89%, F1 ~86%) and BERT baseline (accuracy ~90%, F1 ~88%). These advantages stem from the higher complexity of emotion classification (five classes versus three for sentiment), imbalanced data distribution, and contextual subjectivity in reviews.

The overall metrics affirm DistilBERT's superior performance as a foundation for customer opinion analysis, with computational efficiency 40% faster than BERT, making it ideal for deployment in resource-constrained business environments. This supports accurate evaluation and enhancement of service quality and digital reputation.

Recommendations include data augmentation to optimize minority class metrics, keyword dictionary development, and exploration of Transformer ensembles. Integration with data visualization and real-time monitoring will maximize practical value for business management.

References

[1] Z. Zhang and B. Liu, "Aspect and entity extraction for opinion mining," in *Data Mining and Knowledge Discovery for Big Data*, Springer, 2014, pp. 1-40.

[2] M. Hamzah, "Penerapan Natural Language Processing (NLP) dalam analisis sentimen pada media sosial," *Journal of Informatics and Computer Research (JICR)*, vol. 1, no. 1, pp. 2019-2022, 2024.

[3] M. Nasar and T. Salsabila, "Pemanfaatan potensi e-commerce pada peningkatan keragaman bisnis lokal di Sulawesi Selatan: Perspektif transformasi digital," *GIAT: Teknologi untuk Masyarakat*, vol. 3, no. 1, pp. 25-37, 2024, doi: 10.24002/giat.v3i1.9152.

[4] A. C. T. Angel, V. H. Pranatawijaya, and W. Widiatry, "Analisis sentimen dan emosi dari ulasan Google Maps untuk layanan rumah sakit di Palangka Raya menggunakan machine learning," *Konstelasi: Konvergensi Teknologi dan Sistem Informasi*, vol. 4, no. 1, pp. 35-49, 2024, doi: 10.24002/konstelasi.v4i1.8924.

[5] R. Sukwadi et al., "Peningkatan kualitas layanan jaringan restoran cepat saji Indonesia: Analisis sentimen dan emosi berbasis aspek," *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 12, no. 2, pp. 359-368, 2025, doi: 10.25126/jtiik.2025129416.

[6] R. Mursyid and A. D. Indriyanti, "Perbandingan akurasi metode analisis sentimen untuk evaluasi opini pengguna pada platform media sosial (studi kasus: Twitter)," *Journal of Informatics and Computer Science*, vol. 6, no. 2, pp. 371-383, 2024, doi: 10.26740/jinacs.v6n02.p371-383.

[7] M. K. H. Prasetyo and I. M. I. Subroto, "Penerapan metode BERT (Bidirectional Encoder Representations from Transformers) pada analisis emosi terhadap program kerja Lapor Mas Wapres Presiden RI dengan persepsi pengguna media sosial X," *Transistor Elektro dan Informatika*, vol. 7, no. 1, 2025.

[8] M. Yildirim, "Bidirectional Encoder Representations from Transformers (BERT): A sentiment analysis odyssey," *arXiv preprint arXiv:2007.01127*, 2020.

[9] F. Mukti Ayudewi and A. Ramadhan Firdaus, "Analisis sentimen ulasan pengguna platform belajar Coursera menggunakan model DistilBERT dengan augmentasi data

dan penanganan ketidakseimbangan data," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 9, no. 5, pp. 7706-7712, 2025, doi: 10.36040/jati.v9i5.14844.

[10] V. Sanh et al., "DistilBERT, a distilled version of BERT: Smaller, faster, cheaper and lighter," *arXiv preprint arXiv:1910.01108*, 2019.

[11] M. A. Basit et al., "Aspect based sentiment analysis: A systematic review, taxonomy, and future directions," *Expert Systems with Applications*, vol. 263, 2025, doi: 10.1016/j.eswa.2024.125336.

[12] L. Zhang and B. Liu, "Aspect and entity extraction for opinion mining," in *Data Mining and Knowledge Discovery for Big Data: Methodologies, Challenges and Opportunities*, Springer, 2014, pp. 1-40.

[13] F. Basbeth and D. H. Fudholi, "Klasifikasi emosi pada data teks bahasa Indonesia menggunakan algoritma BERT, RoBERTa, dan Distil-BERT," *Jurnal Media Informatika Budidarma*, vol. 8, no. 2, pp. 1160-1170, 2024, doi: 10.30865/mib.v8i2.7472.

[14] A. C. Sanchez et al., "Aspect-based sentiment analysis on e-commerce using BiGRU-BiDAF hybrid deep learning model," *Journal of Computational Techniques in Automation*, vol. 11, no. 1, pp. 245-258, 2025.

[15] I. Chaturvedi et al., "Contextual emotion detection in images using deep learning," *Frontiers in Artificial Intelligence*, vol. 7, article 1386753, 2024, doi: 10.3389/frai.2024.1386753.

[16] R. Ridhoi, N. A. Verdikha, and F. Yulianto, "Analisis klasifikasi ulasan aplikasi SIREKAP 2024 menggunakan ekstraksi fitur DistilBERT dan metode Support Vector Machine," *Jurnal Ilmiah Informatika*, vol. 13, no. 1, pp. 26-32, 2025, doi: 10.33884/jif.v13i01.9753.

[17] D. Zhao et al., "Cross-lingual aspect-based sentiment analysis: A survey on tasks, methods, and future directions," *arXiv preprint arXiv:2508.09516*, 2025.

[18] W. Khofifah, D. N. Rahayu, and A. M. Yusuf, "Analisis sentimen menggunakan Naive Bayes untuk melihat review masyarakat terhadap tempat wisata pantai di Kabupaten Karawang pada ulasan Google Maps," *Jurnal Interkom: Jurnal Publikasi Ilmiah Bidang Teknologi Informasi dan Komunikasi*, vol. 16, no. 4, pp. 28-38, 2022, doi: 10.35969/interkom.v16i4.192.

[19] A. Erfina and N. R. Wardani, "Analisis sentimen perguruan tinggi termewah di Indonesia menurut ulasan Google Maps menggunakan algoritma Support Vector Machine (SVM)," *Jurnal Manajemen Informatika dan Sistem Informasi*, vol. 5, no. 1, pp. 77-85, 2022.

[20] M. R. A. Rustiawan and P. T. Prasetyaningrum, "Analisis sentimen terhadap Klinik Natasha Skincare di Yogyakarta dengan metode Google Review," *Journal of Information Technology Ampera*, vol. 5, no. 1, pp. 75-89, 2024.