iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 Issue: III Month of publication: March 2026

DOI: https://doi.org/10.22214/ijraset.2026.78484

A Context-Aware Model for Sentiment Analysis of Restaurant Reviews

Shaveta Koundal[1], Sandeep Ranjan[2]
[1]Scholar, [2]Professor, Department of Computer Science and Engineering, CT Institute of Engineering, Management and Technology (CTIEMT), Jalandhar, India

Abstract: Online restaurant platforms feature many customer reviews that discuss various aspects of the dining experience, such as food quality, service, ambiance, and price. It can be hard to draw clear conclusions from these reviews because a single review might express mixed opinions on several aspects. Traditional sentiment analysis often misses these details. This study proposes an Aspect-Based Sentiment Analysis (ABSA) framework that uses transformer-based models to identify sentiment for each aspect in restaurant reviews. A custom dataset was created by collecting restaurant reviews and manually labelling aspect-sentiment pairs as positive, neutral, or negative. For training, aspect terms were combined with the review text in the format "Aspect: Review" to help the model focus on the link between the aspect and its sentiment. This approach was tested using transformer models, specifically BERT and DeBERTa. The results show that DeBERTa outperforms BERT in detecting context and recognizing sentiment in some cases. These results highlight the effectiveness of contextual transformer models for detailed sentiment analysis of restaurant reviews.
Keywords: Aspect-Based Sentiment Analysis (ABSA), Context-aware sentiment analysis, Restaurant reviews, DeBERTa, BERT, Transformer fine-tuning

I. INTRODUCTION

More user-generated content, particularly online product and service reviews, has been produced as digital platforms have expanded. Customers can share their dining experiences and satisfaction after visiting restaurants on platforms such as Google Maps. These reviews have a significant impact on patron decisions as well as the reputation and trust of a restaurant, both of which are critical for commercial success. Food quality, staff conduct, cleanliness, cost, atmosphere, and wait times are all frequently mentioned in restaurant reviews. Both restaurant managers and patrons can gain valuable insights from analyzing this feedback. However, it is not feasible to manually analyze them due to the large number of unstructured reviews. To extract valuable information from this feedback, automated techniques are required [1].

Sentiment analysis seeks to classify textual emotions as either neutral, negative, or positive. Conventional approaches frequently treat the entire sentence or document as a single unit and give it a single label. This provides a broad picture of customer satisfaction but ignores the nuanced aspects of restaurant reviews. For instance, a review may compliment the food but criticize the service. It can be deceptive to use a single general sentiment label that doesn't adequately capture opinions about every aspect of the experience. This emphasizes the need for more thorough sentiment analysis techniques that can identify sentiment for particular review elements [2].

This issue is resolved by Aspect-Based Sentiment Analysis (ABSA), which identifies particular aspects in a review and assigns sentiment to each. In the service sector, where feedback frequently covers a wide range of aspects, this is especially helpful. For instance, ABSA can demonstrate that a review is favourable regarding the quality of the food but unfavourable regarding the service. Restaurant owners can enhance their offerings and boost patron satisfaction with the aid of these comprehensive insights. Aspect-level sentiment analysis provides more valuable information for business decisions than merely examining overall sentiment, according to research [1]. Because customer feedback frequently contains negation, comparisons, sarcasm, and implicit opinions, ABSA is still a difficult task despite its benefits. Additionally, the context can alter a word's meaning. Since restaurant reviews are typically informal and may cover multiple topics in a single sentence, it can be challenging to accurately determine sentiment. Previous research has demonstrated that managing ambiguity, hidden emotions, and context is a significant challenge in ABSA research [3]. Perceptions of the usefulness and reliability of online reviews are heavily influenced by sentiments. As a result, determining sentiment for every component not only improves technical analysis but also makes it easier to comprehend customer feedback on a deeper level. More precise ABSA systems can assist companies in assessing their services and assist consumers in making well-informed decisions [4].

Recent advances in natural language processing have mostly been fuelled by transformer-based architectures. Transformers employ self-attention, which aids models in comprehending the relationships between words in lengthy texts. This is very helpful for ABSA because words that are far away in the sentence may affect how someone feels about a particular aspect. Compared to conventional machine learning techniques that employ manually created features, transformer models have performed better. Additionally, explainable transformer systems have been investigated in recent research to enhance the precision and dependability of sentiment predictions [5].

Because BERT can comprehend context in a text from both directions, it is used as a baseline model. Because it can be optimized on particular datasets to efficiently classify sentiment, BERT is frequently used as a baseline for ABSA tasks. Fine-tuned BERT models can handle the distinctive vocabulary, informal writing, and hidden sentiments frequently found in online reviews, according to studies on restaurant reviews [6].

To improve models' comprehension of context, new transformer models, such as DeBERTa, have recently been developed. DeBERTa represents language more accurately by separating position information from word content. Because of this, DeBERTa is particularly well-suited for sentiment classification tasks where meaning can be altered by minute contextual cues. DeBERTa is therefore a promising tool for evaluating context-dependent opinions in restaurant reviews [7].

Effective sentiment analysis research requires the development of trustworthy datasets and assessment frameworks. Noise, class imbalance, and varying sentiment expressions are common problems with real-world review data. Therefore, developing successful sentiment analysis models requires careful dataset preparation, consistent labelling, and suitable evaluation metrics. These issues have been successfully addressed in restaurant review datasets by fine-tuning transformer models for particular domains [8].

There are a number of research gaps in restaurant review analysis, despite the fact that Aspect-Based Sentiment Analysis has advanced significantly. Rather than restaurant-level sentiment modelling, the majority of the literature in the field concentrates on product reviews or general sentiment classification. Furthermore, despite the widespread use of transformer-based architectures (like BERT) in sentiment analysis, few studies have carried out systematic comparisons between traditional baseline models and high-quality transformer models (like DeBERTa) in the context of restaurant reviews. To ascertain whether improved contextual representations could enhance sentiment classification in restaurant-oriented ABSA tasks, it is important to close this gap [9].

A. Objective

This study's primary objective is to develop and evaluate a context-aware sentiment analysis framework for restaurant reviews using sophisticated transformer models. DeBERTa's performance is compared with a baseline BERT model under the same conditions after it has been optimized for aspect-level sentiment classification. To ensure a fair assessment across all sentiment categories, the model's performance is evaluated using Accuracy, Macro-Precision, Macro-Recall, and Macro-F1 score.

B. Contributions

This study offers three primary contributions. First, a structured dataset of restaurant reviews is made for sentiment analysis at the aspect level. To better predict sentiment, the aspect term and review text are put together into an aspect-aware input format. Second, we create a transformer-based ABSA framework that uses DeBERTa as the main model and BERT as a baseline for comparison. Finally, Accuracy, macro precision, macro recall, and macro F1-score are used to test how well DeBERTa can capture aspect-level sentiments in restaurant reviews.

II. RELATED WORK

Aspect-Based Sentiment Analysis (ABSA) is an important part of opinion mining, especially in the service and hospitality industries. In this case, customers often talk about more than one thing in a single comment. Research on ABSA in hospitality underscores the necessity to adjust to domain-specific contexts, model contextual dependencies, and employ structured evaluation methods to attain precise sentiment classification in restaurant reviews. These studies underscore the necessity of models that transcend general sentiment polarity to capture the interrelations among various aspects and sentiments present in actual service data [10].

Aspect-level analysis is helpful for restaurant management because it uses online reviews to give useful information. ABSA frameworks help improve service and support informed decision-making by finding feelings connected to things like food quality, service behaviour, atmosphere, and prices. Studies indicate that aspect-level modelling markedly improves comprehension and operational decisions in restaurants [11]. The creation of annotated restaurant datasets has been essential for the advancement of ABSA research. These structured datasets are meant to find aspects, opinion targets, and sentiment labels all at once. They can be used as benchmarks for judging detailed extraction tasks in real reviews. They back up detailed annotation plans for testing contextual dependency modelling in restaurant reviews [12].

Joint detection methods improve ABSA evaluation by combining opinion target extraction and sentiment classification in unified modelling frameworks. This creates benchmark settings that allow for systematic comparisons of architectures based on how well they capture the contextual relationships between aspects and sentiments [13].

In addition to resources in one language, researchers have created multilingual datasets to study cross-lingual generalization in aspect-based modelling. These structured multilingual datasets help researchers understand how contextual representations change across linguistic and cultural differences in expressing sentiment [14].

Recent advances in multilingual tourism analytics have introduced hybrid data augmentation strategies to tackle challenges from dataset imbalance and linguistic variety in ABSA. These strategies combine generative models that create synthetic reviews with semantic filtering, which employs SBERT embeddings and masked language models to maintain contextual consistency. By balancing generative creativity with semantic accuracy, this approach reduces misleading samples and boosts F1-scores for less represented aspects like sanitation and amenities, showing robustness in multilingual tourism reviews [15].

Research has also explored counterfactual data augmentation to support sentiment learning in ABSA. This method finds sentiment-carrying tokens using integrated gradients, masks them, and generates reversed-polarity versions with a pre-trained language model. By selectively altering key opinion words while keeping aspect terms intact, this framework enhances sample diversity without compromising contextual meaning. Experiments with Restaurant and Laptop datasets demonstrate consistent improvements in accuracy and Macro-F1 scores, highlighting the effectiveness of careful polarity changes for sentiment modelling [16].

Large-scale domain datasets have been utilized to assess the performance of advanced language models in business and marketing contexts. Integrating large language models into aspect-level sentiment analysis shows promising flexibility across various review datasets and industrial applications [17].

Recent progress in transformer architectures has significantly improved the learning of contextual representations. Innovations like disentangled attention allow models to process positional and semantic information separately, enhancing their capacity to manage complex sentences involving contrasts and long-distance dependencies. These architectural upgrades establish a stronger foundation for structured aspect-level sentiment classification [18].

Comparative studies indicate that transformer-based ABSA systems outperform traditional machine learning and earlier neural methods in capturing long-distance contextual dependencies. Self-attention mechanisms enable transformers to achieve better performance and generalization on standard sentiment datasets [7].

Explainability has also become an essential element of transformer-based ABSA systems. Techniques like attention visualization and attribution analysis help pinpoint important tokens that influence sentiment predictions, improving understanding and trust in sentiment analytics for service-oriented applications [5].

Multimodal ABSA approaches broaden contextual modelling by fusing visual and relational information with text. Multimodal fusion and conditional relation modelling enhance sentiment detection in complex reviews while maintaining manageable architectural complexity [19].

Knowledge-enhanced ABSA models seek to improve contextual reasoning by integrating external knowledge and multi-perspective filtering. While these methods strengthen representation robustness, they often require extra preprocessing and resource integration [20].

Context-aware transformer models have been developed to include user intent signals alongside textual sentiment data. By incorporating structured contextual features, these models recognize subtle sentiment signals and minority classes better in aspect-level sentiment classification tasks [21].

Studies on emotional expression in restaurant reviews support the notion that contextual and emotional signals play a crucial role in shaping sentiment interpretation and the usefulness of reviews. These findings underline the significance of modelling contextual interactions instead of relying only on lexical polarity indicators in ABSA systems for hospitality [4].

Beyond text reviews, systems for voice-enabled multilingual sentiment analysis have been proposed to enhance usability and enable real-time feedback processing in restaurants. These systems integrate speech-to-text technology with rule-based sentiment classification, accommodating both voice and text inputs. While they mainly target overall sentiment rather than detailed aspect-level analysis, they highlight the increasing importance of multimodal feedback systems in restaurant sentiment analytics [22].

Despite these advancements in multilingual datasets, augmentation methods, and transformer architectures, gaps remain in restaurant-focused ABSA research. Many augmentation techniques aim to diversify datasets but are rarely tested specifically on imbalanced aspect-level restaurant data. Moreover, extensive comparisons between baseline transformer models like BERT and more advanced architectures, such as DeBERTa, are still limited in structured restaurant review datasets. Existing applications often prioritize general sentiment classification rather than precise aspect-level modelling.

Therefore, a comprehensive empirical evaluation of transformer models, using balanced assessment metrics in restaurant-specific ABSA settings, is vital. This study addresses this gap by comparing BERT and DeBERTa for context-aware aspect-level sentiment classification in restaurant reviews.

III. PROPOSED METHODOLOGY

This section presents a context-sensitive framework designed for Aspect-Based Sentiment Analysis (ABSA) in restaurant reviews. Unlike traditional sentiment analysis, which assigns a single sentiment label to an entire review, this framework classifies sentiment at the aspect level. This allows for separate sentiment predictions for various elements such as food quality, service, ambiance, pricing, and hygiene.

The framework employs transformer-based deep learning models, which are effective at capturing long-range contextual dependencies. This capability is essential for restaurant reviews, where sentiment may depend on negation, contrasting statements, or contextual modifiers.

Two transformer architectures are evaluated:

1) BERT, the baseline model, and
2) DeBERTa, the context-aware model.

BERT serves as the baseline due to its strong performance in NLP tasks, while DeBERTa is chosen for its disentangled attention mechanism. This mechanism separates semantic and positional information to improve the context used in analysis [18].

A. Dataset Preparation

Restaurant reviews were collected from publicly available Google Maps feedback using automated extraction methods through the Outscraper Google Maps Review Scraper tool [23]. The raw dataset included review text and several metadata attributes. Since the goal is aspect-level sentiment classification, irrelevant metadata fields were removed to reduce noise and improve consistency. After preprocessing, each dataset entry includes: Restaurant Name, Location, Rating, Review Text, Aspect Category (Food, Service, Ambience, Price, etc.), and Sentiment Label (Positive, Neutral, Negative).

The task is structured as a supervised multi-class classification problem, where the review-aspect pair serves as the input, and the sentiment label is the prediction target. Sentiment labels are numerically encoded as:

Negative = 0, Neutral = 1, Positive = 2.

This encoding enables cross-entropy loss-based optimization during transformer fine-tuning.

B. Aspect-Aware Input Representation

To enable aspect-level sentiment prediction, the framework uses an aspect-aware input format. The aspect category is explicitly included with the review text:

X=Aspect: Review Text.

This representation allows the model to focus on a specific aspect of the review, enabling independent sentiment predictions for each aspect.

Table I: Illustration Of Aspect-Aware Input Formatting For Multi-Aspect Sentiment Classification (Self–Generated)

Review Text	Aspect-Aware Input
The pizza was delicious, but the service was slow. The pizza was delicious, but the service was slow.	Food: The pizza was delicious, but the service was slow. Service: The pizza was delicious, but the service was slow.
The ambiance was beautiful, although the food was overpriced. The ambiance was beautiful, although the food was overpriced.	Ambiance: The ambiance was beautiful, although the food was overpriced. Price: The ambiance was beautiful, although the food was overpriced.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue III Mar 2026- Available at www.ijraset.com

As shown in Table I, a single review can generate multiple aspect-specific inputs, allowing the model to learn contextual sentiment for each aspect. Figure 1 illustrates this process.

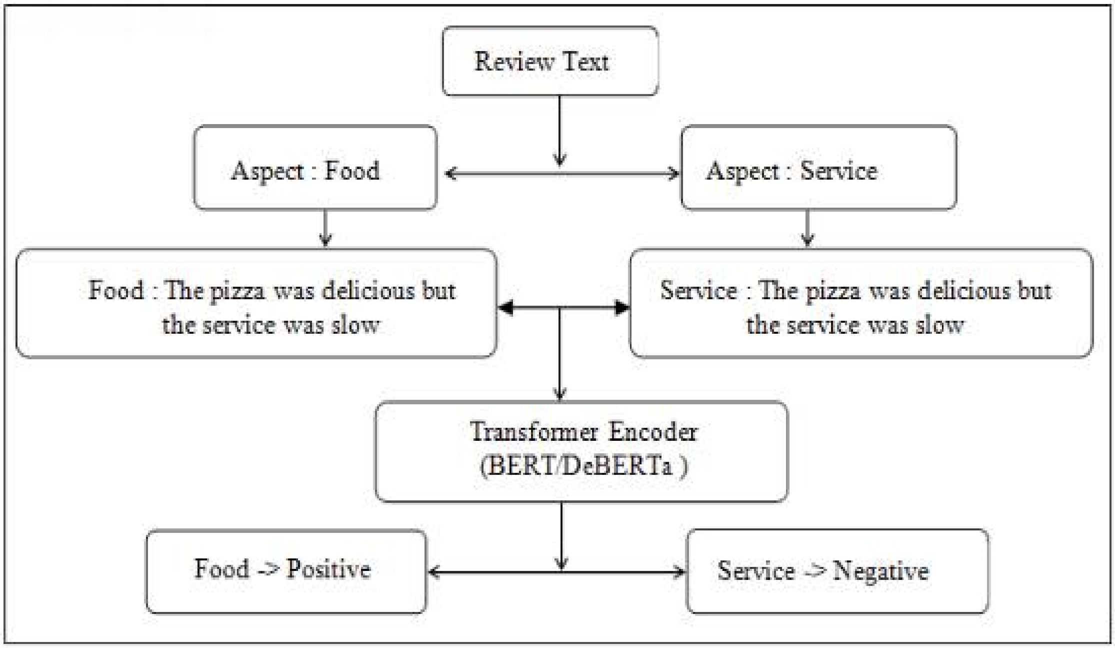

Figure 1. Illustration of aspect-aware input construction and aspect-level sentiment prediction. (self-generated)

In this process, as shown in Figure 1, the review text is paired with different aspects to generate aspect-specific inputs. A transformer encoder, such as BERT or DeBERTa, processes these inputs to generate contextual representations. Sentiment is then predicted for each aspect independently.

C. Model Architecture

The proposed framework uses transformer encoders to process aspect-aware inputs. The architecture includes four stages:
Data preprocessing
Aspect-aware input construction
Transformer-based encoding
Sentiment classification
The review-aspect pair is first tokenized into numerical representations, including input_ids and attention_mask. Padding and truncation ensure consistent sequence length throughout the dataset. The transformer encoder processes these inputs to capture contextual connections across the entire text sequence.
Two transformer models are evaluated: BERT uses multi-head self-attention to learn contextual relationships between tokens. DeBERTa improves representation learning with disentangled attention and relative positional encoding.
This design helps DeBERTa better capture long-range dependencies and contextual details, which improves aspect-level sentiment classification. Figure 2 presents the overall architecture of the proposed framework.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue III Mar 2026- Available at www.ijraset.com

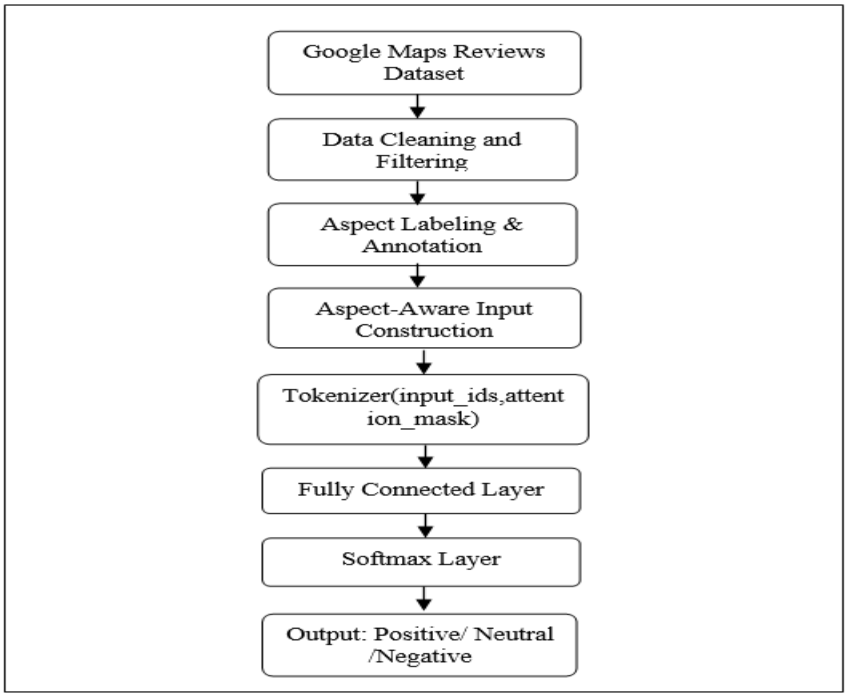

Figure 2. Overall architecture of the proposed context-aware ABSA framework. (self-generated)

Figure 2 illustrates that the proposed framework comprises four stages: data preprocessing, aspect-aware input construction, transformer-based encoding, and sentiment classification. Initially, the review aspect pair is transformed into an aspect-aware format and tokenized into numerical tensors. A transformer encoder then processes these tensors to capture contextual dependencies, followed by a fully connected layer and a softmax classifier for the final sentiment prediction. Figure 3 presents the architectural comparison between BERT and DeBERTa encoders.

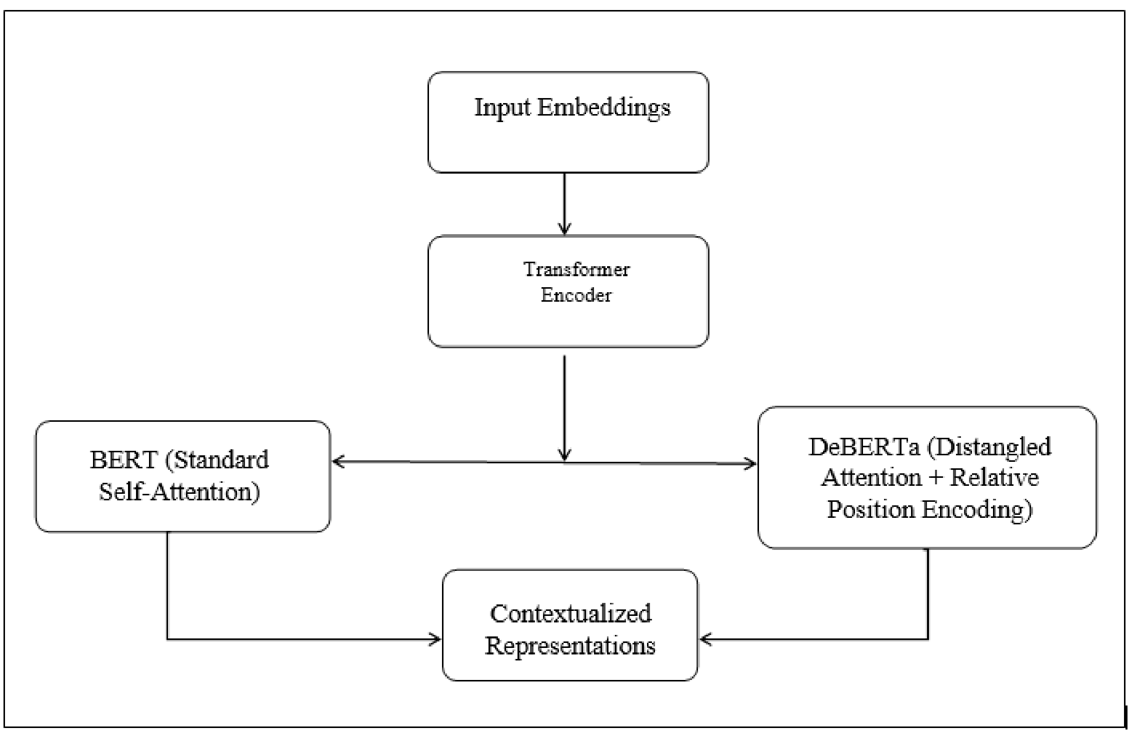

Figure 3. Architectural comparison between BERT and DeBERTa transformer encoders. (self-generated)

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue III Mar 2026- Available at www.ijraset.com

As shown in Figure 3, both models use input embeddings processed through a transformer encoder to create contextual representations. Bert applies standard multi-head self-attention to find contextual relationships among tokens. In contrast, Deberta uses disentangled attention and relative positional encoding, which separately manage semantic and positional data. This method improves the model's ability to recognize long-range dependencies and subtle contextual differences. As a result, it enhances performance in aspect-based sentiment classification.

D. Training Procedure

Both models are fine-tuned on the restaurant review dataset using supervised learning. Fine-tuning adjusts the pretrained transformer parameters to the specific task of aspect-level sentiment classification.

BERT is trained for 3 epochs, while DeBERTa is trained for 5 epochs to achieve better contextual adaptation. The classification head consists of a fully connected layer followed by a softmax function, which predicts one of three sentiment categories: Positive, Neutral, or Negative. The experimental workflow is shown in Figure 4.

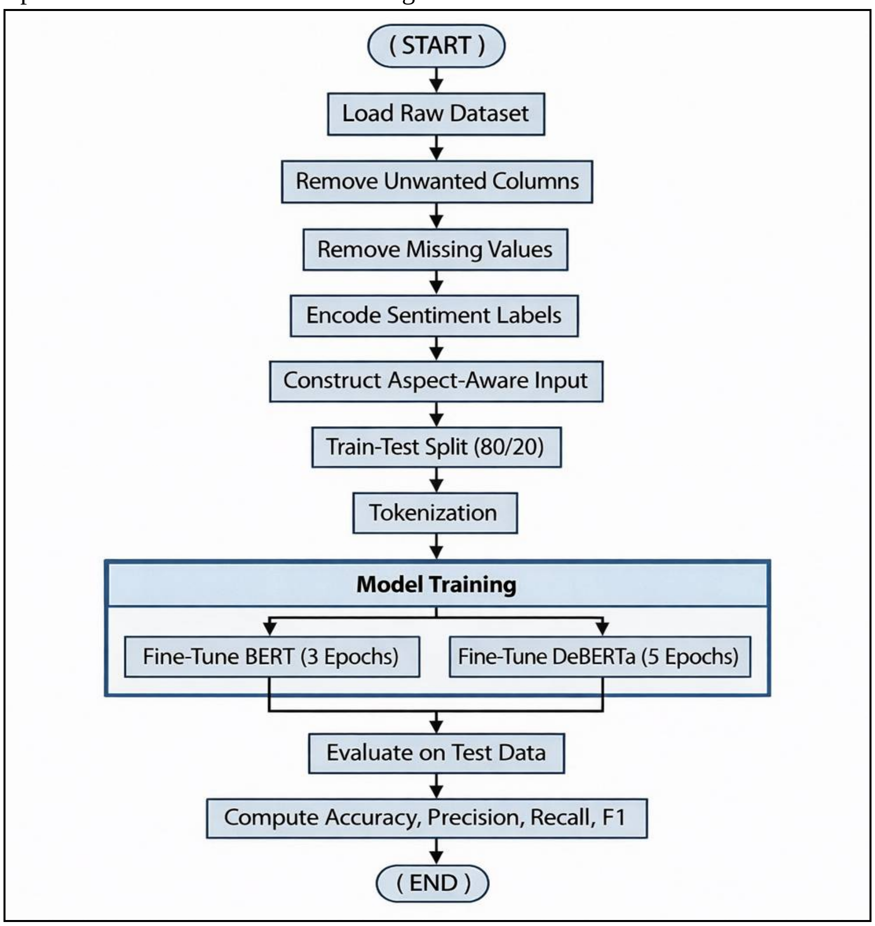

Figure 4. Algorithmic workflow of the proposed context-aware ABSA framework. (self-generated)

As indicated in Figure 4, the process starts with dataset preprocessing, followed by the construction of aspect-aware inputs. The dataset is split into training (80%) and test (20%) sets, and tokenized inputs are used to fine-tune the transformer models. The trained models produce sentiment predictions for the test set.

E. Evaluation Metrics

Model performance is assessed using standard multi-class classification metrics:
Accuracy
Macro-Precision
Macro-Recall
Macro-F1 Score.

Accuracy shows the share of correctly classified samples. Given the dataset's three sentiment classes and significant imbalance, macro-averaged metrics are used. These metrics calculate precision and recall for each class separately and then average the results.
Macro-Precision: average precision across all classes
Macro-Recall: average recall across all classes
Macro-F1 Score: average F1-score across all classes
Macro-averaging gives equal importance to each sentiment category. In this study, the Macro-F1 score is the main evaluation metric because it balances precision and recall.

IV. EXPERIMENTAL SETUP

A. Description of the Dataset

The dataset includes five categories: Food, Service, Ambience, Price, and Cleanliness. Each entry is given one of three sentiment labels: Positive, Neutral, or Negative. For supervised learning, these labels were converted to numbers: Negative = 0, Neutral = 1, and Positive = 2.
The dataset was divided into training and test sets at an 80:20 ratio, using a fixed random seed (42) to ensure reproducibility. The training set was used for fine-tuning the model, while the test set was used to evaluate performance. Table III shows the class distribution of the dataset.

TABLE II: CLASS DISTRIBUTION OF THE DATASET (SELF-GENERATED)

Sentiment Class	Number of Instances	Percentage (%)
Positive	7590	84.71
Neutral	336	3.75
Negative	1034	11.54
Total	8960	100

As shown in Table II, the dataset is significantly imbalanced, with positive samples predominating. Consequently, the Macro-F1 score is used as the primary evaluation metric to ensure balanced assessment across sentiment categories.

B. Implementation and Training Configuration

The proposed framework was built with the Hugging Face Transformers library. Two pre-trained transformer models were adjusted for aspect-level sentiment classification:
bert-base-uncased (baseline model)
microsoft/deberta-v3-base (proposed model)
Tokenization was done with the specific pre-trained tokenizers. Input sequences were padded and truncated to a maximum of 128 tokens, forming the input_ids and attention_mask tensors required by the models. Training was performed using the Hugging Face Trainer API, with the AdamW optimizer and the Cross-Entropy loss function.

TABLE III: HYPERPARAMETER CONFIGURATION (SELF-GENERATED)

Parameter	BERT (Baseline)	DeBERTa (Proposed)
Pre-trained Model	bert-base-uncased	microsoft/deberta-v3-base
Epochs	3	5
Batch Size	8	8
Learning Rate	5×10^{-5}	1×10^{-5}
Maximum Sequence Length	128	128
Optimizer	AdamW	AdamW
Loss Function	Cross-Entropy	Cross-Entropy

As shown in Table III, both models were trained under the same experimental conditions, including batch size, sequence length, optimizer, and loss function. Variations in epochs and learning rates follow common fine-tuning guidelines for each model. Training began with publicly available pre-trained checkpoints, with no additional data augmentation applied.

C. Evaluation Metrics

Model performance was measured on the reserved test set using standard multi-class classification metrics:

- Accuracy
- Macro-Precision
- Macro-Recall
- Macro-F1 Score

Because the dataset is significantly imbalanced, Macro-F1 was chosen as the main evaluation metric. Macro-averaging calculates performance for each class separately and then averages the results. This approach ensures that minority classes have equal influence on the final evaluation. Both BERT (baseline) and DeBERTa (proposed) were evaluated using the same dataset split, preprocessing steps, and evaluation protocol. This setup allows for a fair comparison between the two transformer-based models.

V. RESULTS AND DISCUSSION

A. Assessment of Quantitative Performance.

The proposed context-aware ABSA model was evaluated with the reserved test data. Accuracy, Macro-Precision, Macro-Recall, and Macro-F1 score were used for this assessment. BERT, the baseline model, and DeBERTa, the proposed model, were both trained and evaluated under the same experimental conditions. This ensured a fair comparison. Table III presents the quantitative findings.

Table IV: Performance Comparison on Test Dataset(self-generated)

Model	Accuracy	Macro-Precision	Macro-Recall	Macro-F1
BERT	0.9319	0.7074	0.6173	0.6300
DeBERTa	0.9497	0.7321	0.6806	0.6989

Table IV demonstrates that DeBERTa is consistently superior to BERT, especially for Macro-Recall and Macro-F1, which are higher in class-specific sentiment prediction. To get a clearer picture of how the models perform, the evaluation metrics for BERT and DeBERTa are illustrated in Figure 5.

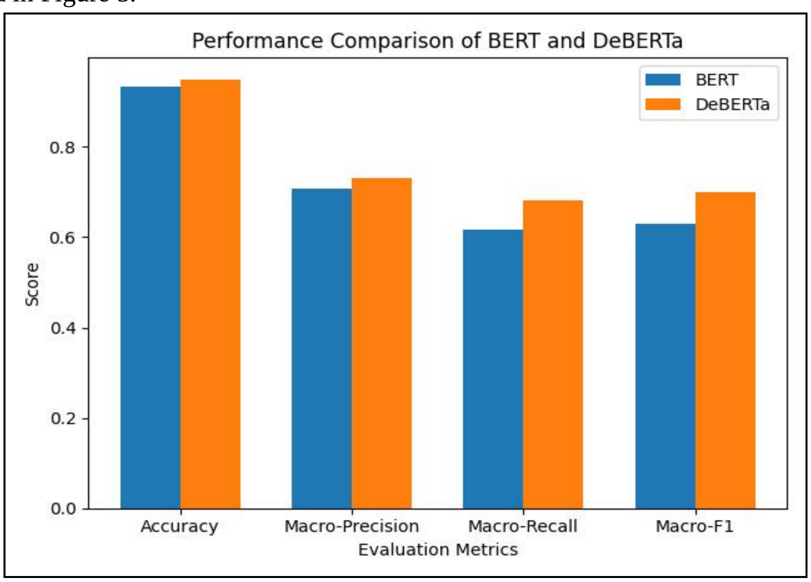

Figure 5. Performance comparison of BERT and DeBERTa across evaluation metrics. (self-generated)

As shown in Figure 5, DeBERTa consistently achieved higher scores in Accuracy, Macro-Precision, Macro-Recall, and Macro-F1. This further confirms its superior contextual modelling skills.

B. Discussion

The results show that the DeBERTa-based model performs better than the original BERT model in most evaluation metrics. DeBERTa reached 94.97% accuracy, while BERT reached 93.19% accuracy, showing better classification performance. DeBERTa also achieved a higher Macro-F1 score of 0.6989 compared to BERT's 0.6300. This highlights DeBERTa's stronger performance across all sentiment categories.

Given that the dataset is highly imbalanced, Macro-F1 is a more trustworthy measure than accuracy alone. Even with its high overall accuracy, BERT has a lower Macro-Recall of 0.6173, showing less effectiveness in identifying minority sentiment categories. In contrast, DeBERTa has a higher Macro-Recall of 0.6806, meaning it is more aware of underrepresented classes like Neutral and Negative sentiments.

This improvement comes from DeBERTa's use of disentangled attention, which separates content and position data. This design change allows for better understanding of context, particularly in contrasting sentences like "The food was great, but the service was slow," which frequently appear in restaurant reviews.

In summary, the experimental results confirm that the context-aware transformer model achieves stronger and fairer aspect-level sentiment classification than the baseline BERT model.

VI. CONCLUSIONS

This study proposed a context-sensitive Aspect-Based Sentiment Analysis (ABSA) system to analyze restaurant reviews. It used transformer-based models for this purpose. The structure included generating aspect-aware input and applying contextual modelling for multi-class sentiment classification with Positive, Neutral, and Negative classes.

The research compared two transformer models: BERT, which served as the baseline model, and DeBERTa, the proposed model. Both models underwent fine-tuning under the same test conditions. They were evaluated based on Accuracy, Macro-Precision, Macro-Recall, and Macro-F1 score to ensure a fair comparison.

The results indicated that DeBERTa outperformed BERT in most evaluation metrics. It achieved an accuracy of 94.97 and a higher Macro-F1 score of 0.6989. The improved Macro-Recall and Macro-F1 scores suggest that DeBERTa provides a more balanced classification across sentiment categories. This is especially true when dealing with minor classes in the imbalanced dataset.

The better performance of DeBERTa likely stems from its disentangled attention mechanism. This mechanism enables more effective contextual representation by modelling content and positional information independently. This feature is particularly beneficial in restaurant reviews, where sentiment often depends on context, such as negation, contrasting conjunctions, and intensity modifiers.

A. Future Work

Future work may focus on improving contextual modelling by including clear methods for aspect-opinion interactions. Better strategies for managing class imbalance could improve the performance of minority classes. Additionally, using explainability methods would increase the model's clarity and understanding. Expanding the framework to include larger or multilingual datasets would also help assess its generalizability.

REFERENCES

[1] Maroof, S. Wasi, S. I. Jami, and M. S. Siddiqui, "Aspect-Based Sentiment Analysis for Service Industry," IEEE Access, vol. 12, pp. 109702–109713, 2024, doi: 10.1109/ACCESS.2024.3440357.

[2] V. Gupta and P. Rattan, "Enhancing Sentiment Analysis in Restaurant Reviews: A Hybrid Approach Integrating Lexicon-Based Features and LSTM Networks," International Journal of Intelligent Systems and Applications in Engineering, vol. 12, pp. 185–198, 2024.

[3] L. Davoodi, J. Mezei, and M. Heikkilä, "Aspect-based sentiment classification of user reviews to understand customer satisfaction of e-commerce platforms," Electronic Commerce Research, Springer US, 2025, pp. 1-43, doi: 10.1007/s10660-025-09948-4.

[4] D. Jeong, D. Kim, J. Yan, and J. Li, "Effect of Aspect-Level Emotion Expression of Online Restaurant Reviews on Perceived Helpfulness: An Aspect-Based Sentiment Analysis Perspective," SAGE Open, vol. 15, no. 2, pp. 1–22, 2025, doi: 10.1177/21582440251337195.

[5] I. Perikos and A. Diamantopoulos, "Explainable Aspect-Based Sentiment Analysis Using Transformer Models," Big Data and Cognitive Computing, vol. 8, no. 11, 2024, doi: 10.3390/bdcc8110141.

[6] A. G. Prabhune, V. R. Shri Hari, and N. K. Sethiya, "Adaptation of a BERT Model to the India Restaurant Data Using Rule-based Approach for Aspect-based Sentiment Analysis," In 2023 2nd International Conference on Smart Technologies for Smart Nation, SmartTechCon 2023, pp. 457–461, 2023, doi: 10.1109/SmartTechCon57526.2023.10391510.

[7] D. Jayakody et al., "Aspect-based Sentiment Analysis Techniques: A Comparative Study," in Proceedings Moratuwa Engineering Research Conference (MERCon 2024), pp. 205–210, 2024, doi: 10.1109/MERCon63886.2024.10688631.

[8] L. A. A. Rahim, K. A. F. A. Samah, N. H. I. Teo, and A. A. Shari, "Enhance Fake Review Detection: a Hybrid Approach of Implicit Absa and Imbalanced Dataset Handling," Journal of Theoretical and Applied Information Technology, vol. 103, no. 13, pp. 4631–4649, 2025.

[9] A. Rahman et al., "Multilingual sentiment analysis in restaurant reviews using aspect-focused learning," Scientific Reports, vol. 15, no. 1, pp. 1–22, 2025, doi: 10.1038/s41598-025-12464-y.

[10] I. C. Sahin and C. Eyupoglu, "Aspect-Based Sentiment Analysis for Hospitality Industry Applications: A Systematic Literature Review," Applied Computer Systems, vol. 30, no. 1, pp. 53–67, 2025, doi: 10.2478/acss-2025-0007.

[11] P. Carrasco and S. Dias, "Enhancing Restaurant Management through Aspect-Based Sentiment Analysis and NLP Techniques," Procedia Computer Science, vol. 237, pp. 129–137, 2024, doi: 10.1016/j.procs.2024.05.088

[12] N. C. Hellwig, J. Fehle, M. Bink, and C. Wolff, "GERestaurant: A German Dataset of Annotated Restaurant Reviews for Aspect-Based Sentiment Analysis," in 20th Conference on Natural Language Processing, KONVENS 2024 - Proceedings of the Conference, pp. 123–133, 2024.

[13] S. U. S. Chebolu, F. Dernoncourt, N. Lipka, and T. Solorio, "OATS: A Challenge Dataset for Opinion Aspect Target Sentiment Joint Detection for Aspect-Based Sentiment Analysis," in Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024), pp. 12336–12347, 2024.

[14] J. Šmíd, P. Přibáň, O. Pražák, and P. Král, "Czech Dataset for Complex Aspect-Based Sentiment Analysis Tasks," in 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation, LREC-COLING 2024 - Main Conference Proceedings, pp. 4299–4310, 2024.

[15] N. M. S. Iswari and N. Afriliana, "Enhancing Aspect-Based Sentiment Analysis in Tourism Reviews Through Hybrid Data Augmentation," Journal of Applied Data Science, vol. 6, no. 3, pp. 2192–2206, 2025, doi: 10.47738/jads.v6i3.842.

[16] D. Wu, L. Wen, C. Chen, and Z. Shi, "A Novel Counterfactual Data Augmentation Method for Aspect-Based Sentiment Analysis," Proceedings of Machine Learning Research, pp. 1479–1493, 2024.

[17] O. Silcenco, M. R. Machad, W. C. Ugulino, and D. Braun, "A Retail-Corpus for Aspect-Based Sentiment Analysis with Large Language Models," Proceedings of the 8th International Conference on Natural Language and Speech Processing (ICNLSP-2025), 2025.

[18] P. He, X. Liu, J. Gao, and W. Chen, "DeBERTa: Decoding-enhanced BERT with Disentangled Attention," in International Conference on Learning Representations (ICLR), 2021.

[19] X. Liu, R. Li, S. Ye, G. Zhang, and X. Wang, "Multimodal Aspect-Based Sentiment Analysis under Conditional Relation," Proceedings - International Conference on Computational Linguistics, COLING, pp. 313–323, 2025.

[20] J. Yang, Y. Xiao, and X. Du, "Enhancing aspect-based sentiment analysis with multiple-knowledge promotion and multi-perspective noise filtering," Complex & Intelligent Systems, vol. 11, no. 9, 2025, doi: 10.1007/s40747-025-02034-0.

[21] H. N. Chaudhry, F. Kulsoom, and Z. U. Khan, "TASCI : transformers for aspect-based sentiment analysis with contextual intent integration," *PeerJ Computer Science* pp. 1–27, 2025, doi: 10.7717/peerj-cs.2760.

[22] G. Venkateswari, K. Meghamala, B. Devika, and D. Sujitha, "Voice-Based Sentimental Analysis for Restaurant Review," International Advanced Research Journal in Science, Engineering and Technology, vol. 12, no. 4, pp. 459–463, 2025, doi: 10.17148/iarjset. 2025.12470.

[23] Outscraper, "Google Maps Review Scraper," Outscraper. [Online]. Available: https://outscraper.com/google-maps-reviews-scraper/.