**ASPECT-BASED SENTIMENT ANALYSIS USING CNN- BERT MODEL**

**ABSTRACT**

Aspect-Based Sentiment Analysis (ABSA) focuses on determining sentiment polarity for specific aspects of an entity, such as food, service, or ambiance in restaurant reviews. Traditional sentiment analysis methods fail to attribute sentiments to specific aspects, limiting their effectiveness in fine-grained opinion mining. Moreover, existing datasets often suffer from class imbalance, affecting the model's ability to generalize well.

To address these challenges, this study proposes a deep learning-based ABSA system incorporating CNN-BERT architecture. BERT embeddings improve contextual understanding, while CNN layers help extract relevant sentiment features for aspect-based classification.

The proposed system is evaluated using CNN-BERT models, achieving an accuracy of 91%. The results confirm the effectiveness of CNN-BERT in enhancing ABSA performance, with improvements in sentiment classification across different aspect categories.

**CHAPTER – 1**

**INTRODUCTION**

#### 1.1 Background

In today’s data-driven world, user-generated content—especially online reviews—plays a pivotal role in shaping consumer decisions and business strategies. Platforms such as Yelp, TripAdvisor, and Google Reviews serve as repositories for vast amounts of user opinions across various sectors including restaurants, hotels, and retail. These reviews often contain sentiments about multiple aspects of a product or service, such as food quality, pricing, customer service, or ambiance.

However, traditional sentiment analysis models typically treat the entire review as a monolithic text unit, assigning a single polarity label (positive, negative, or neutral) to the whole review. This coarse-grained approach fails to account for the subtleties of opinions expressed toward specific aspects. For instance, in the sentence “The food was delicious, but the service was terrible,” an overall polarity fails to recognize that the sentiment varies by aspect. Consequently, critical feedback regarding specific attributes may go unnoticed [1].

Aspect-Based Sentiment Analysis (ABSA) provides a more nuanced method by breaking down reviews into aspect terms and identifying sentiment toward each. ABSA has gained significant attention due to its utility in actionable feedback extraction, particularly in industries such as hospitality and retail, where services are multi-faceted [2]. Recent advancements, especially the incorporation of contextualized language models like BERT, have significantly improved ABSA performance by enabling better understanding of aspect-context interactions in complex sentences [3].

#### 1.2 Problem Statement

Although sentiment analysis has advanced significantly, existing approaches suffer from several limitations when applied to aspect-based classification:

* **Lack of Aspect-Specific Sentiment Detection:** Traditional sentiment analysis models often overlook multiple aspects in a single review, leading to incomplete insights [2].
* **Contextual Misinterpretation of Words:** Words can shift meaning based on context. For instance, in “The food was hot, but the service was cold,” both “hot” and “cold” reflect sentiment only when understood contextually—something rule-based models often fail to capture [1].
* **Class Imbalance in Datasets:** Real-world sentiment data tends to be imbalanced, with fewer examples in categories like "neutral" or "conflict," which can affect model performance and generalization [3].

#### 1.3 Motivation

The restaurant industry in particular can use ABSA to identify and act on specific areas for improvement, boosting customer satisfaction and competitive edge. The main contributions of our work are as follows:

* We proposed a novel framework that integrates BERT embedding, CNN layers, and data augmentation techniques to enhance the performance of Aspect-Based Sentiment Analysis (ABSA).
* We demonstrated that BERT embeddings significantly improve contextual understanding, enabling more accurate sentiment classification for aspect terms within complex sentences.
* We conducted several experiments on augmented\_data\_restaurant\_bert.csv to justify the significance of our model and found notable improvements in accuracy and F1-score across all aspect categories, especially for underrepresented classes.

The rest report is organized as follows:

Chapter-2 describes the literatures review that is related to our work and the previous work who worked on the model. Chapter 3 described the proposed work. Chapter- 4 describes the experiment and results. Chapter-5 conclusion and future work.

**CHAPTER-2**

**LITERATURE REVIEW**

Aspect-Based Sentiment Analysis (ABSA) is a subdomain of sentiment analysis that focuses on determining the sentiment polarity (positive, negative, neutral, or conflict) for specific aspects or attributes of a product or service. Unlike traditional sentiment analysis, which provides an overall sentiment for an entire sentence or review, ABSA identifies and classifies sentiments associated with individual aspects, enabling more granular and actionable insights. This chapter reviews the existing literature on sentiment analysis techniques, the limitations of traditional approaches, and the motivations for using deep learning and hybrid models such as CNN-BERT in ABSA.

### **2.1 Traditional Sentiment Analysis and Its Limitations**

The early foundations of sentiment analysis and opinion mining are outlined by Bing Liu [1], who categorized sentiment analysis into three levels: document-level, sentence-level, and aspect-level. Traditional sentiment classifiers worked well at the document and sentence level but were inadequate for aspect-level tasks, where multiple sentiments could co-exist in a single review. For example, a restaurant review stating “The food was great, but the service was terrible” contains two contrasting sentiments related to different aspects—‘food’ and ‘service’. A standard sentiment classifier might average these sentiments or classify the entire review as "neutral," missing critical feedback information.

These early models also heavily relied on hand-crafted features, sentiment lexicons, and rule-based approaches, which failed to scale well and performed poorly on nuanced, ambiguous, or context-dependent sentences.

### **2.2 Emergence of Deep Learning and Contextual Embeddings**

With the rapid growth in deep learning research, models based on neural networks began replacing traditional machine learning methods for sentiment analysis. However, many of these early deep learning models (e.g., LSTMs, GRUs) struggled with long-term dependencies in text.

A breakthrough came with the introduction of BERT **(Bidirectional Encoder Representations from Transformers)** by Devlin et al. [2]. BERT is a transformer-based architecture that reads entire sequences bidirectionally, enabling it to learn contextual relationships between words in a sentence. This is particularly powerful in sentiment analysis, where the meaning of a word may change depending on surrounding context. For instance, the word "cold" could be neutral in "cold drinks" but negative in "cold service."

BERT is pre-trained using two unsupervised tasks:

* **Masked Language Modeling (MLM):** Randomly masks words in a sentence and learns to predict them, allowing it to understand word dependencies and context.
* **Next Sentence Prediction (NSP):** Trains the model to predict whether a given sentence follows another, improving its understanding of sentence-level semantics.

By fine-tuning BERT on specific ABSA tasks, researchers were able to achieve state-of-the-art performance across a wide variety of sentiment analysis benchmarks.

However, one challenge with using BERT alone is its lack of focus on **localized patterns**, such as key phrases or n-grams, which can be critical for identifying sentiment expressions in certain contexts.

### **2.3 CNN-BERT Hybrid Models**

To overcome this limitation, hybrid models that combine **BERT with Convolutional Neural Networks (CNNs)** have been proposed. CNNs are excellent at capturing local dependencies and identifying n-gram patterns in text. They are especially useful in classifying short phrases or spotting sentiment-rich expressions, such as “very bad service” or “extremely good taste.”

In the CNN-BERT architecture, the input text is first processed using the pre-trained BERT model, which generates **contextualized embeddings** for each word. These embeddings are then fed into CNN layers, where multiple filters of different sizes slide over the sequence to extract local features. This combination allows the model to:

* Understand **semantic context** (thanks to BERT),
* Detect **important local sentiment cues** (thanks to CNN),
* Perform **fine-grained sentiment classification** with high accuracy.

This hybrid model has shown superior performance in ABSA tasks because it benefits from the strengths of both architectures: BERT's deep contextual understanding and CNN’s local feature extraction.

### **2.4 Tackling Data Imbalance through Augmentation**

A persistent problem in ABSA is **class imbalance**, especially when dealing with real-world datasets where certain sentiment classes (e.g., "neutral" or "conflict") are underrepresented. Class imbalance can lead to biased models that perform poorly on minority classes.

To address this issue, Wei and Zou proposed **Easy Data Augmentation (EDA)** techniques [3], which include:

* **Synonym Replacement:** Randomly replacing words with their synonyms.
* **Random Insertion:** Inserting new words into random positions.
* **Random Deletion:** Removing random words from a sentence.
* **Back-Translation:** Translating a sentence to another language and back to generate paraphrased versions.

These methods artificially increase the number of samples in minority classes, improving the model's ability to generalize across all sentiment categories. In our work, we applied such augmentation techniques to enhance the diversity and balance of the restaurant review dataset.

**2.5 SUMMARY OF PREVIOUS WORKS**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Sl No.** | **Year** | **Author(s)** | **Model Name** | **Dataset Used** | **Accuracy (%)** |
| 1 | 2019 | Chi Sun, Luyao Huang, Xipeng Qiu | BERT-based Sentence Pair Classification | SemEval 2014 (Rest14) | 85.62 |
| 2 | 2020 | Chen Zhang, Qiuchi Li, Dawei Song , Benyou Wang | Multi-task Learning Framework | SemEval 2014 (Rest14) | 85.62 |
| 3 | 2022 | Ziwen Gao, Zhiyi L, Jiaying Luo, XiolinLii | CNN + BiGRU | Custom dataset | 82.45 |
| 4 | 2022 | Adepu Rajesh, Tryambak Hiwarkar | Multi-channel CNN + BiLSTM + Attention | IMDB | 89.3 |
| 5 | 2023 | Paul F. Simmering, Paavo Huoviala | GPT-3.5 fine-tuning | SemEval 2014 (Rest14) | 91.2 |
| 6 | 2022 | Paul F. Simmering, Paavo Huoviala | BERT-ETextCNN-ELSTM | Comment data | 88.5 |
| 7 | 2022 | David Z. Chen Adam Faulkner Sahil Badyal | BERT Post-training + UDA | SemEval 2014 (Rest14) | 87.1 |

Table 1:Summary of previous works

**CHAPTER – 3**

**Proposed work: CNN-BERT for Aspect-Based Sentiment Analysis**

In this chapter, we present our proposed model for Aspect-Based Sentiment Analysis (ABSA), which integrates **BERT embeddings** with **Convolutional Neural Networks (CNN)**. This hybrid approach is designed to capture both the contextual meaning of words and the local patterns important for identifying sentiment associated with specific aspects. The overall architecture is illustrated in **Figure 1.**

Output Layers

Fully connected Layers

BERT

Input

CNN - Layers

Figure 1: CNN-BERT Model Architecture for Aspect-Based Sentiment Analysis

The proposed CNN-BERT model for Aspect-Based Sentiment Analysis is designed to effectively combine the contextual strength of BERT with the feature extraction capabilities of Convolutional Neural Networks (CNN). The process begins by feeding an input sentence, which includes one or more aspect terms, into a pre-trained BERT model. BERT generates rich, contextualized embeddings for each token in the sentence, capturing subtle semantic and syntactic nuances that are essential for understanding the sentiment in relation to specific aspects. These embeddings are then passed to CNN layers, which apply multiple convolutional filters of different sizes to extract local features, such as sentiment-bearing n-grams related to aspect terms. Max pooling is used to retain the most significant features from each filter, helping the model focus on the most relevant information. The resulting features are concatenated and passed through a fully connected dense layer, which learns complex relationships between the extracted patterns. Finally, the output layer produces a sentiment prediction—positive, negative, or neutral—for each aspect term present in the input. This architecture allows the model to perform fine-grained sentiment analysis with improved accuracy and robustness. The detail methodology are describe in the following,

**3.1 Convolutional Neural Network (CNN)**

Convolutional Neural Networks (CNNs) are deep learning models originally designed for image processing, but they have shown remarkable performance in text classification tasks due to their ability to capture **local features and patterns** (like n-grams) in a sentence. In NLP, CNNs process sequences of word embeddings instead of pixel values. **How CNN Works for Text Data?**

Each sentence is first converted into a matrix using pre-trained word embeddings (e.g., BERT, GloVe). If a sentence contains n words and each word is represented by a d-dimensional embedding vector, then the sentence can be represented as a matrix:

|  |  |
| --- | --- |
|  | (1) |

Where n = number of words in the sentence, d = dimensionality of each word vector and is the embedding of the word

**3.2** **Bidirectional Encoder Representations from Transformers (BERT)**

Bidirectional Encoder Representations from Transformers (BERT) is a transformer-based language model designed to learn deep bidirectional representations from unlabeled text [21]. BERT converts an input sentence xxx into a dense, contextualized C:\Users\91637\AppData\Local\Microsoft\Windows\Clipboard\HistoryData\{A2A6659B-A338-4B80-AE04-BCD30FD659CC}\{17AE71F3-672A-4AF7-B740-29DF3910AC88}\ResourceMap\{6546486A-9357-494A-9059-DC37F639293E}, where d is the dimensionality of the embedding space. These embeddings capture the left and right context for each token in the sentence, providing rich contextual information.BERT is pre-trained using two objectives:

* **Masked Language Modeling (MLM)**: Randomly masking tokens in the input and predicting the masked tokens.
* **Next Sentence Prediction (NSP)**: Predicting whether two given sentences appear consecutively in the text.

When fine-tuning BERT for a downstream task such as sentiment classification, a classification layer is added on top of the embeddings derived from the special [CLS] token.

Fine-tuning BERT on a downstream task, such as sentiment classification, involves adding a classification layer on top of the embedding hhh corresponding to the [CLS] token.



Figure 2: BERT model structure

**3.3 CNN-BERT Hybrid Model**

Our proposed architecture combines the strengths of both CNN and BERT to create a robust ABSA model.

* **Step 1: BERT Embeddings:** Input reviews are first tokenized and passed through the pre-trained BERT model to obtain contextualized word embeddings.
* **Step 2: CNN Layer:** The output embeddings from BERT are passed through convolutional layers with multiple filter sizes (e.g., 2, 3, 4) to extract local n-gram features that may be relevant for sentiment expression.
* **Step 3: Max Pooling and Dense Layers:** The CNN output is subjected to max pooling and then passed through dense layers for classification into sentiment categories (positive, negative and neutral).

This hybrid model leverages BERT for understanding context and CNN for capturing local patterns. As a result, it performs better on complex reviews that contain multiple sentiment-laden aspects.

**CHAPTER – 4**

**DATASET**

### **4.1 Dataset Description**

### The experiments in this study are conducted on a customized and augmented version of a restaurant review dataset, named augmented\_data\_restaurant.csv. This dataset contains aspect-based sentiment annotations extracted from user-generated reviews. The reviews cover various aspects such as Ambience, Food, Price, Service, and Anecdotes/Miscellaneous [21].

### Each review is labeled with both an aspect category and a corresponding sentiment polarity, which can be classified as positive, negative, neutral, or conflict. To mitigate class imbalance in the dataset, data augmentation techniques, including synonym replacement and back translation, were employed.

### The dataset used for the restaurant review experiments is derived from the Yelp platform, specifically from a collection referred to as restaurant\_data. This dataset consists of 6,088 entries, each corresponding to a review associated with one or more aspects. The dataset is organized into five distinct entities: Ambience, Anecdotes, Food, Price, and Service.

|  |  |
| --- | --- |
| **Statistic** | **Values** |
| Total Entries | 6,088 |
| Columns | 5 |
| Non-null id values | 3,044 |
| Non- null aspect\_term | 4,103 |
| Non-null aspect\_category | 6,088 |
| Non-null polarity | 6,088 |
| Polarity classes | Positive, negative, neutral |
| Aspect categories | Food, service, ambience, anecdotes/miscellaneous |

Table 2: Overview of Dataset

**CHAPTER – 5**

### **Performance Measures**

To thoroughly evaluate the effectiveness of our aspect-based sentiment classification models, we employ the performance metrics such as accuracy, precision, F1-score, confusion matrix, Macro average and Weighted average. Accuracy measures the proportion of correctly predicted instances out of the total instances as defined in equation 1.

|  |  |
| --- | --- |
|  | (1) |

Where: TP is defined as True Positives, TN is defined as True Negatives, FP is defined as False Positives and FN is defined as False Negatives. **Precision** measures the proportion of correctly predicted positive instances out of all predicted positive instances, as defined in Equation 2.

|  |  |
| --- | --- |
|  | (2) |

A high precision indicates that the classifier makes few false positive errors. Recall is the ratio of correctly predicted positive observations to all observations in the actual class, as defined in Equation 3.

|  |  |
| --- | --- |
|  | (3) |

High recall indicates that most positive instances are correctly identified. F1-score is the harmonic mean of Precision and Recall, giving a balance between the two, as defined in Equation 4.

|  |  |
| --- | --- |
|  | (4) |

This score is especially useful when there is class imbalance. A confusion matrix is a tabular summary showing the performance of a classification algorithm. It includes TP, TN, FP, and FN for each class and helps in visualizing misclassifications. Macro average computes the metric (e.g., precision, recall, F1) independently for each class and then takes the average. It treats all classes equally, as defined in Equation 5.

|  |  |
| --- | --- |
|  | (5) |

Where N is the number of classes. Weighted average computes the metric for each class and weights it by the number of instances (support) in each class, as defined in Equation 6.

|  |  |
| --- | --- |
|  | (6) |

This is useful when the dataset is imbalanced.

**CHAPTER – 6**

**6.1 Result Analysis**

|  |  |  |
| --- | --- | --- |
| **Parameter** | **Tested** | **Used** |
| Dense layers | 4-5 | 4 |
| Optimizer | Adam, AdamW | Adam |
| Activation | softmax, relu, LeakyReLU, ELU, Swish, Swish | softmax, relu |

Table 2: Parameter Analysis

**6.2 MODELS IMPLEMENTED**

6.2.1 ASPECT CATEGORY BASED MODEL

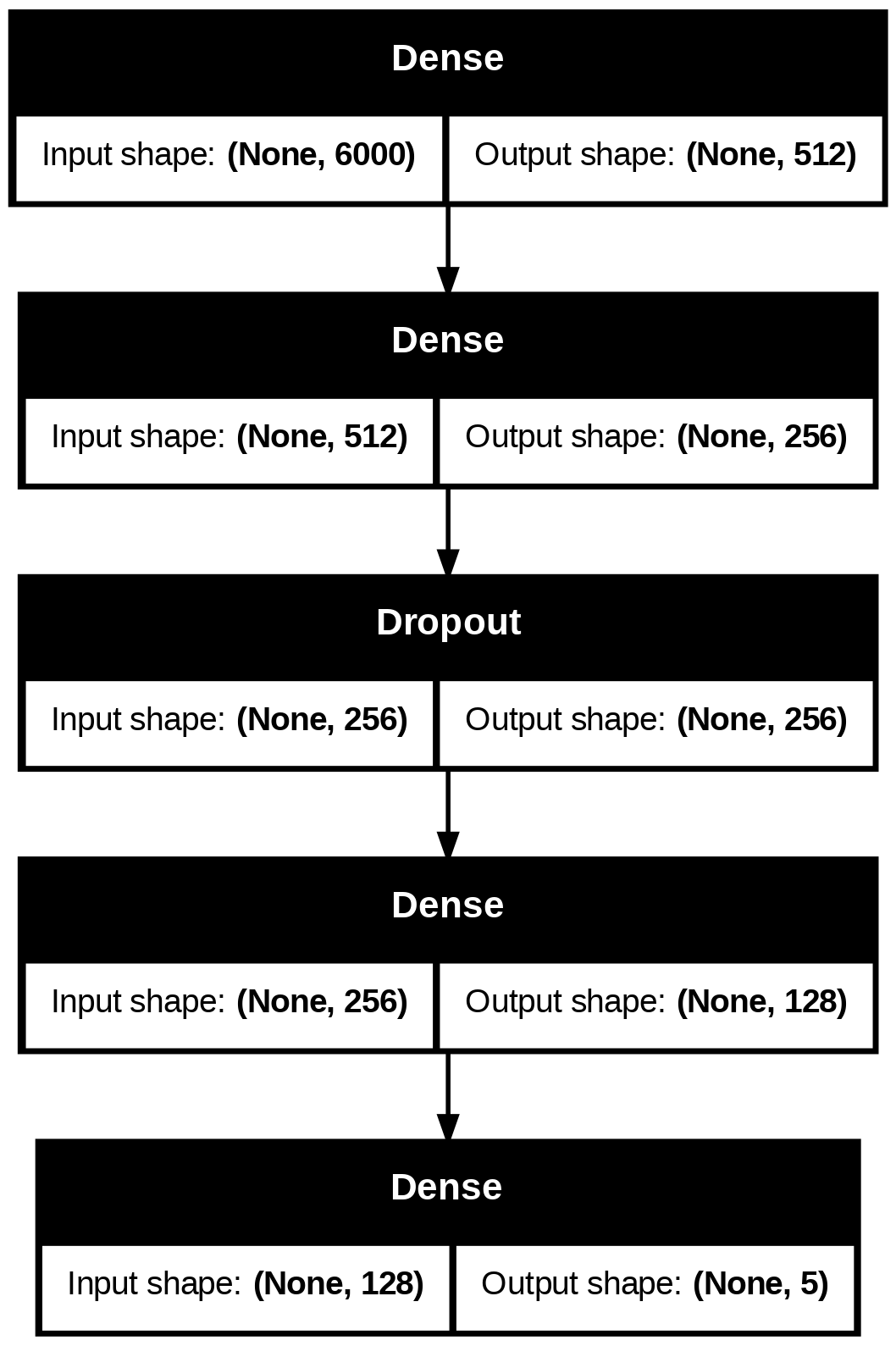


Figure 2: Aspect Category Based Model (CNN with BERT Augmented)

This model begins with a **Dense (fully connected) layer** that accepts an input of shape (None, 6000), which likely represents the **flattened feature vector** obtained after passing BERT embeddings through CNN and pooling layers. The first Dense layer reduces the dimensionality from 6000 to 512 units, likely applying a non-linear activation like ReLU.

6.2.2 ASPECT CATEGORY BASED MODEL

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **precision** | **recall** | **f1-score** | **support** |
| **Food** | **0.97** | **0.85** | **0.91** | **164** |
| **Anecdotes/miscellaneous** | **0.90** | **0.96** | **0.93** | **400** |
| **Service** | **0.93** | **0.90** | **0.92** | **355** |
| **Ambience** | **0.90** | **0.84** | **0.87** | **113** |
| **Price** | **0.85** | **0.90** | **0.88** | **186** |
|  |  |  |  |  |
| **Accuracy** |  |  | **0.91** | **1218** |
| **macro avg.** | **0.91** | **0.89** | **0.90** | **1218** |
| **weighted avg.** | **0.91** | **0.91** | **0.91** | **1218** |

Table 3: Classification Report Aspect Category Model (CNN with BERT augmentation)

The model demonstrates strong performance across all aspects of Aspect-Based Sentiment Analysis (ABSA) for restaurant reviews, with precision, recall, and F1-scores indicating reliability in sentiment classification. It performs particularly well in identifying Food-related sentiments (precision of 0.97), while maintaining solid results for Service (precision of 0.93) and Anecdotes/Miscellaneous (recall of 0.96). The model shows slightly lower performance in Ambience (F1-score of 0.87) and Price (F1-score of 0.88), particularly in recall. Overall accuracy stands at 0.91, with a macro average of 0.90 and a weighted average of 0.91, suggesting that the model performs consistently well across categories, though recall improvements could enhance its ability to capture all relevant aspects.

6.2.3 ACCURACY GRAPH

The accuracy graph illustrates the model's learning behavior over five training epochs. The blue curve represents the training accuracy, which increases rapidly and reaches nearly 100% by the second epoch. This indicates that the model is fitting the training data extremely well. In contrast, the orange curve, representing validation accuracy, starts around 87% and peaks at approximately 91%, but shows signs of plateauing or slightly declining after the third epoch. This divergence between training and validation accuracy suggests that the model is overfitting—memorizing patterns in the training data rather than generalizing well to unseen data. While the model performs exceptionally on training samples, its performance on the validation set does not improve proportionally, indicating limited generalization. To address this, techniques such as early stopping, dropout, regularization, or acquiring more training data could be considered to enhance the model's robustness and prevent overfitting.

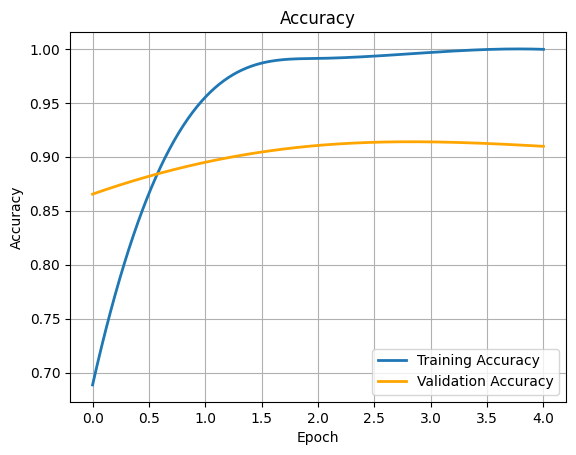
****

Figure 4: Accuracy graph Aspect Category Model (CNN with BERT Augmentation)

6.2.4 CONFUSION MATRIX

The confusion matrix provides a detailed view of the model’s performance in classifying different aspect categories from restaurant reviews. The diagonal elements represent correct predictions, where the predicted class matches the actual class. Notably, the model correctly classified 385 instances of "anecdotes/miscellaneous", 321 of "service", 168 of "price", 139 of "food", and 95 of "ambience". These high values along the diagonal indicate that the model performs well overall. However, some misclassifications are evident—for example, a noticeable number of "service" instances were incorrectly predicted as "anecdotes/miscellaneous" (20 cases), and a few "food" reviews were confused with "price" or "anecdotes". While the overall accuracy is high, the confusion between semantically similar categories suggests there is room for improvement, possibly by enhancing feature representations or incorporating more nuanced contextual understanding in the model.

****

Figure 5: Confusion matrix Aspect Category Model (CNN with BERT Augmentation)

6.2.5 POLARITY BASED MODEL

**Aspect Polarity Based Model (CNN with BERT Augmented)** represents the classification component of a hybrid model designed to determine the sentiment polarity—**positive, negative, neutral, or conflict**—toward specific aspects in a review. The model begins with a high-dimensional feature vector (size 6000) generated by combining BERT embeddings and CNN-based feature extraction, which is then progressively reduced through a series of Dense layers: from 6000 to 512, then 256, followed by a Dropout layer for regularization, and further down to 128 dimensions. Finally, a Dense output layer with 4 units produces class probabilities corresponding to the four sentiment polarities. This architecture effectively captures both contextual and local sentiment cues, enabling fine-grained sentiment classification at the aspect level.

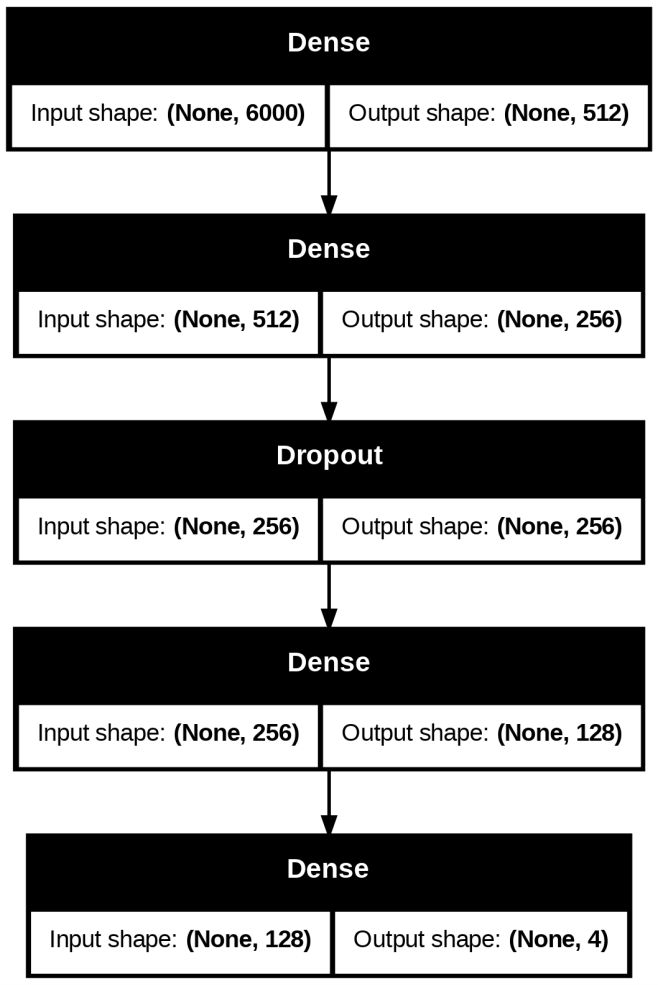


Figure 3: Aspect Polarity Based Model (CNN with BERT augmented)

6.2.6 ASPECT POLARITY MODEL

The classification report provides a comprehensive evaluation of the model's performance in predicting sentiment across four categories: positive, negative, conflict, and neutral. The model achieves a high overall accuracy of 91%, indicating that it correctly classified the sentiment in the majority of the 1,218 instances. Among the individual classes, the model performs best on the neutral class, with a precision of 0.94, recall of 0.96, and an F1-score of 0.95—likely due to its larger representation in the dataset. The negative and conflict classes also show strong performance, with F1-scores of 0.88 and 0.86 respectively. However, the positive class lags behind, with a lower recall of 0.73 and an F1-score of 0.76, which suggests the model occasionally fails to identify positive sentiment correctly, possibly due to the limited number of positive samples (only 55). The macro average F1-score of 0.86 reflects the balanced performance across all classes, while the weighted average F1-score of 0.91 confirms that the model handles class imbalances well. Overall, the sentiment model demonstrates robust performance but could benefit from improvements in recognizing positive sentiment more effectively.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **precision** | **recall** | **f1-score** | **support** |
| **positive** | **0.80** | **0.73** | **0.76** | **55** |
| **negative** | **0.90** | **0.87** | **0.88** | **273** |
| **conflict** | **0.86** | **0.85** | **0.86** | **177** |
| **neutral** | **0.94** | **0.96** | **0.95** | **713** |
|  |  |  |  |  |
| **accuracy** |  |  | **0.91** | **1218** |
| **macro avg.** | **0.87** | **0.85** | **0.86** | **1218** |
| **weighted avg.** | **0.91** | **0.91** | **0.91** | **1218** |

Table 4: Classification Report Aspect Polarity Model (CNN with BERT augmentation)

6.2.7 ACCURACY GRAPH

The accuracy graph shows the training and validation accuracy of the model over six epochs. The blue line represents training accuracy, which increases rapidly from around 68% to nearly 100% by the third epoch, indicating that the model is learning the training data very effectively. However, the orange line, which shows validation accuracy, follows a different trend. It starts at approximately 84%, peaks slightly above 90% in the second epoch, and then fluctuates slightly, staying relatively flat around 91% for the remaining epochs. This divergence between training and validation accuracy suggests that the model may be overfitting—while it continues to improve on the training data, it does not show significant improvement on unseen data. The gap between the two curves indicates that the model's ability to generalize could be limited, and further regularization techniques such as dropout, early stopping, or more diverse training data may be needed to improve validation performance.

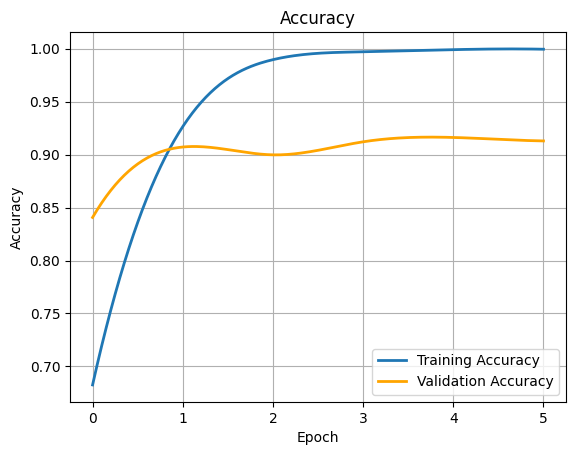
****

Figure 6: Accuracy graph Aspect Polarity Model (CNN with BERT Augmentation)

6.2.8 CONFUSION MATRIX

The confusion matrix illustrates the model’s effectiveness in predicting four sentiment classes: positive, negative, conflict, and neutral. It shows strong overall performance, with most predictions correctly aligning along the diagonal. The model performs particularly well on the neutral class, accurately classifying 684 instances, and also shows solid performance for the negative (237 correct) and conflict (151 correct) classes. The positive class, while smaller in sample size, is correctly predicted in 40 cases, with some misclassifications primarily into the neutral and negative categories. Although minor confusion exists between similar sentiments—such as conflict and negative, or positive and neutral—the matrix reflects a high level of accuracy and a reliable ability to distinguish between the four sentiment types.

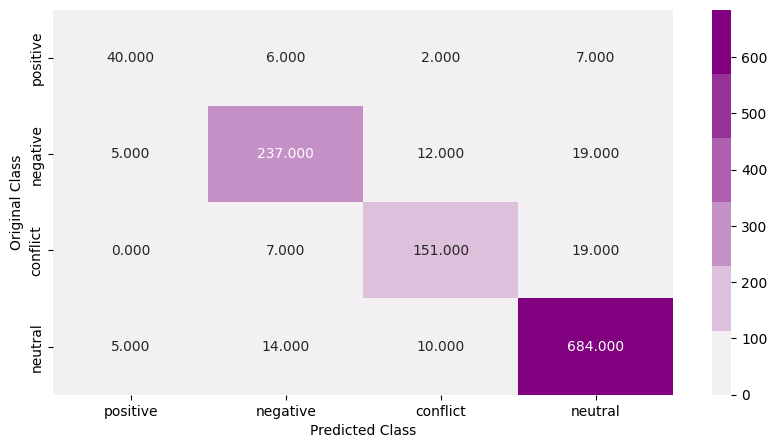


Figure 7: Confusion Matrix Aspect Polarity Model (CNN with BERT Augmentation)

**CHAPTER 5**

**CONCLUSIONS**

This study successfully demonstrated the effectiveness of a CNN-BERT hybrid model for Aspect-Based Sentiment Analysis (ABSA) using augmented restaurant review data. By leveraging BERT's contextual embeddings alongside CNN’s ability to capture local sentiment patterns, the model achieved high classification accuracy of 91% for both aspect category and sentiment polarity prediction tasks. The results highlight the model’s strength in identifying sentiment nuances across different aspects such as food, service, ambiance, price, and anecdotes. Despite this success, signs of overfitting were observed, especially in the gap between training and validation accuracy, indicating room for improving generalization. Furthermore, the model’s performance showed slight limitations in handling underrepresented classes such as “positive” sentiments and specific aspect misclassifications. These observations underline the need for more balanced learning strategies and robust data representation mechanisms in ABSA tasks.

**FUTURE WORK**

To further enhance model robustness and address class imbalance challenges, future work will explore Generative Adversarial Network (GAN)-based architectures for ABSA. GANs, particularly GAN-BERT, have shown significant potential in semi-supervised sentiment classification by generating synthetic yet informative representations for underrepresented classes [10][11]. By incorporating GAN-generated data during training, the model can better learn from limited real-world samples, thereby improving performance on minority sentiment classes like “positive” and “conflict.”

In a GAN-BERT framework, a generator produces plausible sentence embeddings, while a discriminator distinguishes between real and synthetic embeddings and simultaneously performs classification [10]. This adversarial training setup not only augments the feature space but also promotes better generalization and robustness against overfitting—an issue observed in the CNN-BERT model. Additionally, recent studies suggest that enhancing GAN-BERT with contextual denoising and semantic augmentation or fusing it with attention mechanisms can further boost the model's capability to capture subtle sentiment nuances, especially at the aspect level [12].

In future research, implementing GAN-BERT with integrated attention or contrastive learning modules will be explored, aiming to develop a more balanced and explainable ABSA model that performs well even under data scarcity and label imbalance conditions.

**REFERENCES**

1. M. S. Saji, M. S. Shibily, and A. A. Shah, “Aspect based sentiment analysis using BERT: a survey,” Materials Today: Proceedings, vol. 85, pp. 1848–1853, 2024.
2. Y. Zhang and Q. Liu, “A survey on aspect-based sentiment analysis using pre-trained language models,” IEEE Trans. on Affective Computing, 2024.
3. K. Zhang et al., “A transformer-based approach for aspect-based sentiment analysis,” Scientific Reports, vol. 14, no. 1, Mar. 2024.
4. A ElJundi, M. Tannir, and H. Hajj, “Aspect-based sentiment analysis using BERT for Arabic language,” ICNLSP, 2021.
5. A Wu, L. Zhao, and Z. Zhang, “Conditional BERT contextual augmentation for aspect-based sentiment analysis,” arXiv preprint arXiv:2001.11316, 2020.
6. Z. Yang et al., “Boosting Aspect-Based Sentiment Analysis with Contextual Denoising and Semantic Augmentation,” Findings of NAACL, 2024.
7. H. Singh and M. Sharma, “Multimodal sentiment analysis for restaurant reviews using deep learning,” Neural Computing and Applications, 2025.
8. J. Li, S. Xu, and D. Yu, “ABSA-GAN: A GAN-based approach for aspect-level sentiment classification,” EMNLP, 2023.
9. T. Nguyen and L. T. Nguyen, “Fine-grained sentiment analysis using GAN-BERT fusion,” Knowledge-Based Systems, vol. 289, 110290, 2024.
10. Li, J., Xu, S., & Yu, D. (2023). *ABSA-GAN: A GAN-based approach for aspect-level sentiment classification*. EMNLP.
11. Nguyen, T., & Nguyen, L. T. (2024). *Fine-grained sentiment analysis using GAN-BERT fusion*. *Knowledge-Based Systems*, 289, 110290.
12. Yang, Z. et al. (2024). *Boosting Aspect-Based Sentiment Analysis with Contextual Denoising and Semantic Augmentation*. *Findings of NAACL*.