CHAPTER 1

INTRODUCTION

1.1 Background

Sentiment analysis is the process of extracting insights from subjective text, such as opinions on individuals, events, and products [1, 2]. Despite deep learning advancements, models still struggle with complex linguistic structures, long-range dependencies, and attention utilization for better contextual understanding [3]. A key challenge is text length variation, where short texts (e.g., tweets) require effective sentiment extraction from minimal context, while long texts (e.g., movie reviews) demand strong dependency retention. CNN and RNN-based models often fail to address these differences, highlighting the need to explore attention mechanisms for improved performance across varying text lengths.

Deep learning utilises deep artificial neural networks, such as Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), Long Short-Term Memory (LSTM), and Gated Recurrent Units (GRU), which have shown outstanding performance in various domains, including speech recognition, signal and EEG analysis, computer vision, emotion recognition, and disease detection, including cancer diagnosis. In addition, in text classification and sentiment analysis, artificial neural networks have been shown to achieve outstanding accuracy and effectiveness [4]. Compared to transform-based models that are computationally intensive, deep learning is more lightweight for computers with small computations and can be used flexibly.

In natural language processing (NLP), CNNs and RNNs are widely used for feature extraction and sequential data processing. CNNs apply convolution and pooling layers over sequences to capture local features and produce fixed-length outputs [5]. However, they struggle with long-range dependencies and lose contextual and word position information due to pooling operations. In contrast, RNNs handle variable-length sequences more effectively but face challenges such as vanishing gradients and limited parallel processing capability, especially in long texts [6, 7].

To address these limitations, LSTMs were introduced with gated mechanisms to retain long-term dependencies and mitigate the vanishing gradient problem [8]. BiLSTMs further enhance text representation by capturing both past and future contexts, improving performance in sentiment classification and other text-based tasks [4].

However, even with BiLSTMs, capturing the most relevant words in a sentence remains a challenge. To enhance contextual understanding, attention mechanisms have been widely adopted in NLP tasks such as text summarization, review analysis, and question-answering systems [9, 10]. Attention selectively focuses on important words, making it particularly effective for handling text length variations. In short texts, it highlights key sentiment words that might be overshadowed, in long texts, it preserves crucial contextual relationships that could otherwise be lost in sequential processing. By addressing these issues, attention mechanisms significantly improve sentiment classification, especially in longer textual data where traditional CNN and BiLSTM models struggle to maintain context [11]. Positioning attention strategically within a deep learning model can optimize its ability to focus on relevant features [12].

Research in this field has employed various hybrid deep learning approaches to enhance sentiment classification performance. Rehman et al. (2019) utilized a CNN-LSTM architecture on the IMDB Movie Reviews dataset, achieving an accuracy of 91% [13]. Wang and Li (2023) enhanced a CNN-BiLSTM model using attention mechanisms, achieving improved contextual understanding on the IMDB dataset with accuracy of 90.8% [11]. On shorter texts, Kamyab et al. (2021) applied a CNN-BiLSTM model integrated with attention on the Twitter US Airline Sentiment dataset and obtained an accuracy of 81.2% [5]. Similarly, Jang et al. (2020) validated the effectiveness of the CNN-BiLSTM architecture enhanced by attention, achieving approximately 91.4% accuracy on IMDB reviews [14]. These studies show the significant potential of the attention mechanism, especially its placement in the CNN-BiLSTM and CNN-LSTM hybrid models, to effectively address the challenges of sentiment analysis across various text lengths. Therefore, this study further explores how different attention placements in these architectures affect sentiment analysis performance on short text (US Airline Twitter) and long text (IMDB Movie Reviews).

1.2 Theoretical Framework

Several studies have examined sentiment classification using machine learning methods such as Logistic Regression, Support Vector Machines (SVM), Naïve Bayes, Random Forest, and Gradient Boosting. While these models achieve competitive accuracy, with SVM and Logistic Regression reaching up to 86.22% in sentiment classification [15], they often struggle with handling negative sentiment and complex semantic relationships.

Research on sentiment analysis through social media has been found using various algorithms. Research [16] compares classification using several machine learning and deep learning algorithms on sentiment analysis. The deep learning model obtained the highest results in comparing the two classification methods. The following research was conducted by combining two deep learning algorithms, CNN and BiLSTM (CNN-BiLSTM), in research [6]. While the CNN-BiLSTM hybrid model effectively captures both local and sequential dependencies, it still struggles with long-range dependencies and lacks the ability to prioritize crucial sentiment words.

Proper word embedding can improve the performance of deep learning models. Deng and Liu (2025) [17], emphasized RoBERTa as a core component for generating contextually rich word embeddings through an optimized masked language modeling strategy.

This robust embedding forms a strong foundation and enables subsequent components to effectively model long-term dependencies. Meanwhile the attention mechanism selectively highlights textual elements that are important for sentiment detection.

Integrating attention mechanisms into deep learning models can improve contextual understanding, address remote dependencies, and enhance interpretation capabilities. Studies such as Kardakis dkk. [4] dan Tan dkk. [18] demonstrate the effectiveness of attention mechanisms in deep learning models for sentiment analysis, achieving high accuracy and better context understanding [18]. Research [10] confirms that attention mechanisms perform very well when applied to sentiment classification of longer textual data, overcoming the limitations faced by traditional CNN and LSTM models in handling extended context. However, the attention mechanism may perform differently based on text length-short texts require a focus on key words, while longer texts require continuous context retention.

Kardakis et al [4]. (2021) analyzed self-attention, global attention, and hierarchical attention, showing that attention-based models boost accuracy by up to 3.5%. Self-attention outperforms global attention, especially in long-sequence texts, while hierarchical attention enhances sentiment classification by capturing word—and sentence-level context. These findings confirm that integrating attention mechanisms significantly improves sentiment analysis performance.

Several studies have explored the role of attention mechanisms in enhancing deep learning models, particularly in optimizing feature selection and improving classification accuracy. The positioning of attention within a hybrid model architecture plays a crucial role in determining its effectiveness.

Wei et al. (2021) [12] introduced Position-Aware Self-Attention (PSA) in a BiLSTM-CRF sequence labeling model, demonstrating that incorporating self-attention before or after the recurrent layer impacts performance. Their findings indicated that self-attention positioned after BiLSTM better captured long-range dependencies in text, leading to improved accuracy in Named Entity Recognition (NER) and POS tagging tasks. This study confirms that positioning attention strategically within a deep learning model can optimize its ability to focus on relevant features.

In another study by Zeng et al.(2019) [19] proposed a Position-Aware Attention LSTM (PosATT-LSTM) model for aspect-level sentiment classification, incorporating positional context into the attention mechanism. This reserach integrates a Gaussian-based distance vector to emphasize words closer to the target aspect. This approach significantly improved accuracy over baseline models, demonstrating that strategic attention positioning—especially after LSTM with distance-aware weighting—enhances model performance by focusing on contextually relevant features. Their findings reinforce the importance of attention placement within hybrid architectures for better semantic understanding

Sifak et al. [20] (2023) developed a hybrid CNN-BiGRU model for Twitter hate speech detection, incorporating an attention mechanism to prioritize important words. Their findings showed that the BiGRU-CNN model with optimal attention position on both layers achieved the highest accuracy of 88.12%, improving performance compared to the baseline model. This confirms that attention with optimal position effectively enhance deep learning models by focusing on critical features.

This study systematically investigates the impact of attention mechanism with different placement in hybrid CNN-BiLSTM and CNN-LSTM architectures for sentiment classification. While prior work supports the effectiveness of attention, most only explore a single position without comparing placements before, between, or after core components.Using the IMDB dataset (long texts) and Twitter US Airline dataset (short texts), this study aims to fill that gap and provide insights for optimizing hybrid models across varied text types.

1.3 Conceptual Framework/Paradigm

Hybrid deep learning models have been widely applied in sentiment classification due to their ability to integrate feature extraction capabilities with the modeling of long-term dependencies. Contextual embeddings like RoBERTa have also been explored to enhance text representation by providing dynamic and context-sensitive word embeddings, as shown in studies such as Rahman et al. [8], which achieved favorable results on benchmark datasets. Furthermore, the addition of attention mechanisms has assisted models in identifying and focusing on relevant parts of the text, potentially improving interpretability and performance [5, 18].

Although hybrid models are widely used, few studies have examined how attention placement affects sentiment classification, especially across texts of different lengths. This study uses RoBERTa-based CNN-BiLSTM and CNN-LSTM models with attention placed at various points, tested on Twitter (short texts) and IMDB (long texts), to evaluate its impact on sentiment detection across different text structures.

1.4 Statement of the Problem

Hybrid deep learning models, such as CNN-BiLSTM and CNN-LSTM, have been effective in sentiment analysis tasks due to their ability to capture both local textual features and long-term contextual dependencies [21]. However, despite these capabilities, these models still face challenges in accurately capturing context, particularly when handling texts of varying lengths and different levels of linguistic complexity [3]. Previous research indicates that inadequate contextual understanding can negatively affect sentiment classification accuracy, especially when applied to datasets with contrasting characteristics, such as short informal texts (e.g., Twitter posts) and longer formal texts (e.g., IMDB movie reviews).

Although recent studies have suggested attention mechanisms as a potential solution to enhance contextual understanding, the optimal placement of these attention layers within hybrid architectures remains unclear [20]. Specifically, there is limited research examining whether attention mechanisms should be placed before, after, or both before and after the BiLSTM/LSTM layers to maximize sentiment classification performance. Additionally, it is still unknown how the effectiveness of attention placement may vary based on the text length and complexity.

1.5 Objective and Hypotheses

1.5.1 Objective

This research aims to evaluate the impact of attention mechanisms and their optimal placement within CNN-BiLSTM and CNN-LSTM hybrid deep learning models for sentiment classification.

Specifically, the study will:

1. Identify the most effective hyperparameter configurations for hybrid models with attention mechanisms, considering dataset characteristics such as text length and linguistic complexity.
2. Assess the effect of incorporating attention mechanisms on the classification performance of hybrid deep learning models across datasets with varying text lengths (short texts from Twitter US Airline and long texts from IMDB Movie Reviews).
3. Determine the optimal attention and position (before, after, or both before and after BiLSTM/LSTM) and its impact on sentiment classification, particularly in relation to text length variations.

This study does not propose a new model but explores how attention mechanisms influence classification accuracy in hybrid architectures and how their effectiveness varies across different text lengths.

1.5.2 Hypotheses

The hypotheses formulated in this research are as follows:

- **H1:** Integrating attention mechanisms improves classification accuracy in hybrid deep learning models especially in longer textual data [11].
- **H2:** Proper attention selection followed by optimal attention positioning (before, after, or before and after the BiLSTM/LSTM layers) can affect sentiment classification performance [20].
- **H3:** Attention mechanisms are more effective in longer texts (IMDB Movie Reviews) than in shorter texts (Twitter US Airline) due to their ability to capture long-range dependencies [10].

1.6 Assumption

This study examines the role of attention mechanisms in hybrid models for sentiment analysis, assuming that the placement of attention (before, after, or both around BiLSTM/LSTM) affects feature learning and classification accuracy. It also assumes that models with attention outperform those without, especially on long-text datasets.

1.7 Scope and Delimitation

This study evaluates **attention mechanisms** and their placement in **CNN-BiLSTM** and **CNN-LSTM** models for sentiment analysis. It does not propose a new model but examines how attention mechanisms influence classification.

1.7.1 Scope

1. **Models:** Limited to CNN-BiLSTM and CNN-LSTM.
2. **Attention Mechanisms:** Evaluates self-attention and global attention at different placements.
3. **Word Embedding:** Uses RoBERTa without comparison to other embeddings.
4. **Datasets:** Uses two sentiment analysis datasets:
    - US Airline Sentiment (short text)
    - IMDB Movie Review (long text)
5. **Hyperparameters:** Focuses on learning rate and hidden units.

1.7.2 Delimitation

This study does not introduce a new model but optimizes an existing hybrid architecture by evaluating the best attention placement for different text lengths. It is limited to two datasets (Twitter US Airline and IMDB) and uses only RoBERTa embeddings, without comparison to alternatives like Word2Vec or GloVe. The focus is on tuning learning rate and hidden units, with no broader hyperparameter optimization. Other attention types were not tested due to time and computational constraints.