Multi Stream Attention Networks with Adaptive Syntactic-Emotional Fusion for Sentiment Analysis

Nidhi Mishra[*], Aakanshi Gupta[*], Shuchi Mala[*], Srishti Das[*],
Naman Tyagi[*], Sanjam Bhardwaj[*], Narayan C Debnath[†], Amit Mishra[‡]

[*]Department of Computer Science and Engineering, Amity School of Engineering Technology, AUUP, Noida, India.

[†]School of Computing and Information Technology Department of Software Engineering, Binh Duong Province, Vietnam.

[‡]Department of Computer Science and Engineering, Jaypee Institute of Information Technology, Noida.

Abstract

Sentiment analysis is a challenging task in natural language processing due to the complexities between the syntactic and emotional structure of human content. This paper presents a dual-stream attention architecture that models both the syntactic representation and emotional intensity through separate attention mechanisms fused together via an adaptive mechanism. We evaluate our architecture on IMDb Movie Review Dataset, demonstrating its effectiveness compared to various lexicon-based solutions such as TextBlob, SentiWordNet and Vader. Additionally, we also examine Machine Learning based approaches with TF-IDF and Bag-of-Words for feature extraction, as well as sequential models like LSTM and BiLSTM combined with embeddings (GloVe). Furthermore, we examine methods to improve classification performance, including stacking and voting techniques. The performance of each solution is examined across various metrics, including accuracy, precision, recall, F1 Score and AUC-ROC. Our results highlight the importance of explicitly modeling both syntactic and emotional representations for sentiment analysis.

Key Words: sentiment, emotion mining, natural language processing, attention.

1 Introduction

Sentiment Analysis refers to the task of determining the emotional tone and opinions expressed in a text and is fundamental to various applications including social media monitoring, customer feedback analysis, and market research. In recent years, sentiment analysis has emerged as one of the important subdomains of NLP (Natural Language Processing) driven by the proliferation of online reviews, social media, and other user-generated content. Movie reviews are a valuable source of public opinion as they provide insights into viewer preferences, emotions, and overall satisfaction with films. Despite significant advances in natural language processing, accurately capturing sentiments remains challenging due to the ways in which humans express emotions through both words and grammatical structures.

Current approaches to sentiment analysis majorly rely on machine learning techniques and deep learning models, which, while successful, face challenges integrating various aspects such as syntactical structure and emotional nuances. Early models like lexicon-based sentiment analysis [13, 30] and machine learning methods focused on keyword matching or feature extraction. While these models provide valuable insights, they often failed to capture the deeper, context-dependent meaning inside a text. More recent approaches, such as Recurrent Neural Networks (RNNs) [17], and long short-term memory (LSTM, BiLSTM) [26, 25] have made strides in handling sequential data. Transformer [32] based models like Bidirectional Encoded Representations from Transformers (BERT) [7] have further pushed boundaries, demonstrating state-of-the-art performance across many NLP tasks. However, these models still struggle to integrate the syntactic and emotional dimensions that are crucial to sentiment in complex scenarios.

To address these challenges, we propose a hybrid architecture, the Multi Stream Attention with Adaptive Syntactic-Emotional Fusion, which combined two different attention streams: one focused on syntactic information and other on emotional representation. By utilizing both syntactic and emotional information, our model aims to enhance the representation of sentiment and overall increase the overall accuracy of sentiment analysis. The dual stream attention mechanism uses dependency guided attention to align syntactic structures, while an emotion enhanced representation integrates external lexicon and emotion embedding to capture the emotional tone of the text.

In this research, we developed a Hybrid Sentiment Analysis model that integrates syntactic structure with emotional semantics through a novel dual-attention mechanism. Our approach combines dependency parsing with emotional lexicon information to create structure-aware and emotion-enhanced attention pathways. The model employs spaCy for dependency parsing and aligns these parse trees with BERT's tokenization scheme to generate syntactic attention masks. Simultaneously, we incorporate sentiment lexicon scores to create emotion intensity embeddings that guide a parallel attention mechanism. These dual pathways are adaptively combined using a trainable alpha parameter that automatically determines the optimal balance between syntactic and emotional information for each input. Implementation in PyTorch with BERT-base as the encoder (8 attention heads, 768 hidden dimensions) demonstrated significant improvements over baseline models, with particularly strong performance on texts where sentiment is expressed through complex syntactic structures or emotionally charged language.

The primary contributions of this paper include development of an adaptive fusion mechanism that combines the two attention streams based on their importance, an approach to token alignment through dependency parsing and the integration of emotion enhanced representations. This paper is organized as follows: Section 2 provides an overview of related work in sentiment analysis, focusing on traditional methods, deep learning-based techniques. Section 3 introduces the architecture in detail, including the dual stream attention mechanism and adaptive fusion method. The experiment setup and evaluation are discussed in Section 4 followed, followed by a comprehensive analysis of results in Section 5. Finally, Section 6 concludes the paper and outlines future work.

2 Related Work

A lexicon-based method is one of the earlier procedures where sentiment analysis is performed by using pre-assigned sentiment lexicons that articulate polarity scores for words in the text. For example, Kiritchenko et al [13] introduced lexicons specifically designed for the analysis of social media texts such as tweets. This approach offers a simple yet effective mechanism, requiring no training data, and it relies on the explicit sentiment values contained in the lexicons. The main advantage of this approach is its ease of implementation and its applicability in situations where annotated training data is scarce. However, a notable limitation is that these lexicons often fail to capture contextual nuances and irony, resulting in misinterpretation when the text is taken out of context. This drawback is especially apparent in texts that incorporate sarcasm or subtle emotional cues that are not represented in the static lexicon. Moreover, while lexicon-based models work well for straightforward cases, they struggle with the polysemy and ambiguity inherent in natural language.

Traditional machine learning models such as Naïve-Bayes [11], Support Vector Machines [9], and Random Forest [3] have been widely employed for sentiment classification tasks. These models typically work by extracting a range of features from textual data such as n-grams, part-of-speech tags, and syntactic structures to predict sentiment labels. One of the key strengths of these approaches is their ability to achieve high accuracy, particularly in controlled environments where feature engineering is optimized. However, a major disadvantage is that their performance is highly sensitive to the choice of feature selection and pre-processing techniques. The reliance on manual feature engineering makes these methods less adaptable to the subtle nuances of language and context found in diverse datasets. Despite this, machine learning techniques remain popular due to their efficiency and the relative simplicity of their implementation in comparison to more complex deep learning methods.

The introduction of word embeddings has revolutionized the way textual data is represented for sentiment analysis. Techniques such as word2vec [19] and GloVe [23] transform words into high-dimensional vectors that encapsulate both semantic and syntactic information. These embeddings enable models to capture relationships between words and understand the context in which they appear, thereby providing a richer and more nuanced representation of the text. Despite their success, word embeddings are generally trained on large, general-purpose corpora and may not perform optimally when applied directly to domain-specific sentiment analysis tasks without fine-tuning. This limitation highlights the importance of adapting these embeddings to the particularities of the task at hand, ensuring that they capture the unique characteristics and subtleties present in specialized datasets.

Deep learning has introduced a new paradigm in sentiment analysis through models that can automatically learn hierarchical representations of text. Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks have proven effective in handling sequential data, as they are capable of modeling dependencies over long sequences. For example, Rao et al [26] demonstrated that LSTM networks outperform traditional classifiers by capturing long-range dependencies in text data. More recently, transformer-based models, such as BERT, have further advanced the field by employing self-attention mechanisms that more accurately capture contextual relationships. Although these models represent a significant leap forward in performance, they also require substantial computational resources and large volumes of training data. As such, while deep learning approaches have set new benchmarks in sentiment analysis, their complexity and resource demands can be a barrier to their widespread adoption in all contexts.

Hybrid models aim to capitalize on the strengths of both traditional machine learning techniques and deep learning approaches. For instance, dos Santos and Gatti [8] introduced a hybrid model that integrates Convolutional Neural Networks (CNNs) with character-to-word embeddings for sentiment classification. This approach leverages the robust feature extraction capabilities of traditional methods and the powerful representation learning of deep learning models. By combining these methodologies, hybrid approaches are able to create a more robust framework for sentiment analysis that mitigates the weaknesses of each individual method. Such models are

Table 1: Comparison of Related Studies in Sentiment Analysis

Study	**Approach**	**Key Technique**	**Dataset**	**Performance**
Pang et al. (2002)	Classical ML	SVM, Naive Bayes (bag-of-words)	Movie reviews	82.9% (Acc.)
Taboada et al. (2011)	Lexicon-based	SO-CAL lexicon (with intensifiers)	Various domains	78.7% (Acc.)
Socher et al. (2013)	DL (Recursive)	Recursive Neural Tensor Network (RNTN)	Stanford Sentiment Treebank	85.4% (Acc.)
Mohammad et al. (2013)	Lexicon-based	NRC Emotion Lexicon (features for SVM)	SemEval-2013 Task 2 (Twitter)	69.0% (F1)
Tang et al. (2016)	DL (LSTM)	Target-dependent LSTM	Twitter (target-dependent)	71.5% (Acc.)
Howard & Ruder (2018)	DL (Transfer)	ULMFiT (fine-tuned AWD-LSTM)	IMDb movie reviews	95.4% (Acc.)
Peters et al. (2018)	DL (Contextual)	ELMo embeddings	SST-5 (Stanford Sentiment Treebank)	54.7% (Acc.)
Devlin et al. (2019)	DL (Transfer)	BERT pre-training	GLUE benchmark (e.g., SST-2)	94.9% (Acc.)
Yang et al. (2019)	DL (Transfer)	XLNet (Transformer)	IMDb movie reviews	96.2% (Acc.)
Liu et al. (2019)	DL (Transfer)	Multi-task Deep NN (MT-DNN)	GLUE benchmark (e.g., SST-2)	95.6% (Acc.)

particularly effective in capturing both the overt sentiment cues and the underlying contextual subtleties present in complex textual data.

3 Methodology

3.1 Input Representation

Our architecture enhances traditional transformer-based models by introducing a dual-stream attention mechanism that processes both syntactic and emotional features simultaneously. The model takes as input a sequence of tokens. The input sequence consists of tokens arranged in order, where each token is an element of the sequence, and the total number of tokens corresponds to the sequence length.

These tokens are then processed by a pre-trained BERT encoder to obtain contextual representations. The BERT encoder transforms the input token sequence into a matrix of contextual embeddings that capture the semantic and syntactic context of each token.

$$\mathbf{H} = \mathrm{BERT}(\mathbf{X}) \tag{1}$$

where $\mathbf{X}$ is the tokenized input sequence (in embedding form) and $\mathbf{H}$ is the resulting matrix of contextual embeddings capturing the semantic and syntactic context of each token.

3.2 Syntactic Attention Stream

The syntactic stream leverages dependency parsing to capture grammatical relationships between tokens. For each input sequence, a dependency tree is constructed using a syntactic parser. This tree is then converted into an adjacency matrix where entries indicate whether tokens are syntactically related or not. This matrix guides the syntactic attention mechanism, which extends the standard multi-head attention by incorporating dependency information. Attention scores are computed taking into account the syntactic relationships between tokens, allowing the model to focus on grammatically relevant connections.

The syntactic attention output is obtained by multiplying the attention matrix with the value matrix for the syntactic stream (Derived from equation 2). The query, key, and value matrices for the syntactic stream are computed by applying learnable weight matrices to the contextual embeddings produced by the BERT encoder. Equation 3 describes the mathematical formulation for the operation.

$$M_{\text{syntactic}}[i,j] = \begin{cases} 0, & \text{if tokens } i \text{ and } j \text{ are related,} \\ -\infty, & \text{otherwise.} \end{cases} \tag{2}$$

$$A_s = \mathrm{softmax}\left(\frac{Q_s K_s^\top}{\sqrt{d_{\text{head}}}} + M_{\text{syntactic}}\right) \tag{3}$$

3.3 Emotional Attention Stream

The overall sentiment analysis process can be summarized in four phases. First, the raw text sequence is tokenized using the BERT tokenizer with subword alignment, and contextual embeddings are generated using the BERT encoder. Second, the dual attention streams operate: the syntactic stream builds the dependency adjacency matrix based on parsing, computes attention scores considering syntactic relationships, while the emotional stream retrieves emotional intensity scores from a sentiment lexicon and adjusts attention weights accordingly. Third, the outputs of both syntactic and emotional streams are combined adaptively using a weighted fusion mechanism. Fourth, the fused features are pooled using mean aggregation and passed through a classifier layer to predict the sentiment class.

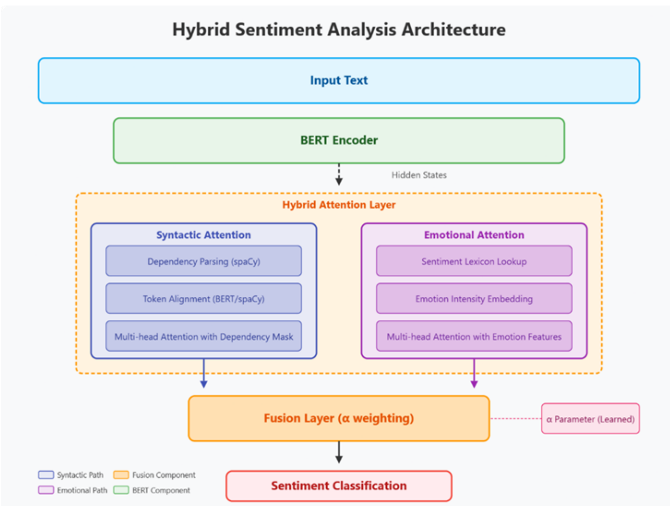

Figure 1: Proposed architecture

The adaptive fusion mechanism described in equation 4 combines the outputs of the syntactic and emotional attention streams using a weighted sum controlled by a learnable parameter. This parameter balances the contribution of each stream by applying a sigmoid function, resulting in a final integrated output that effectively captures both grammatical structure and emotional context for improved sentiment analysis performance.

$$O_{\text{fused}} = \sigma(\alpha) \cdot O_s + \big(1 - \sigma(\alpha)\big) \cdot O_e \tag{4}$$

where O_{fused} is the final attention output, O_s is the output from the syntactic attention stream, and O_e is the output from the emotional attention stream.

An overview of our pipeline is described by Algorithm 1.

Algorithm 1 Sentiment Analysis with Syntactic-Emotional Fusion

---

Require: Raw text sequence $A = \{a_1, a_2, \dots, a_n\}$
Ensure: Predicted sentiment class y

Phase 1: Input Processing
Tokenize A using BERT tokenizer and align subwords
Generate contextual embeddings H using BERT (see Fig. 1)
Phase 2: Dual Attention Streams
Dependency-Guided Syntactic Stream:
Parse dependencies using spaCy; build adjacency matrix $M_{\text{syntactic}}$
Compute attention scores for syntactically related tokens
Emotion-Aware Attention Stream:
Retrieve emotional intensity scores from AFINN lexicon
Adjust attention weights for sentiment-bearing tokens
Phase 3: Feature Fusion
Combine syntactic and emotional outputs adaptively
Phase 4: Classification
Pool features using mean aggregation
Predict sentiment class via classifier layer

---

4 Experimental Setup

This study utilized the IMDb movie review dataset [19] consisting of 50,000 reviews evenly distributed between positive and negative sentiments. With an 80:20 ratio, the dataset was divided into training and testing sets, ensuring there were an equal number of positive and negative reviews in both sets. Figure 2 shows the distribution of length and frequency of words.

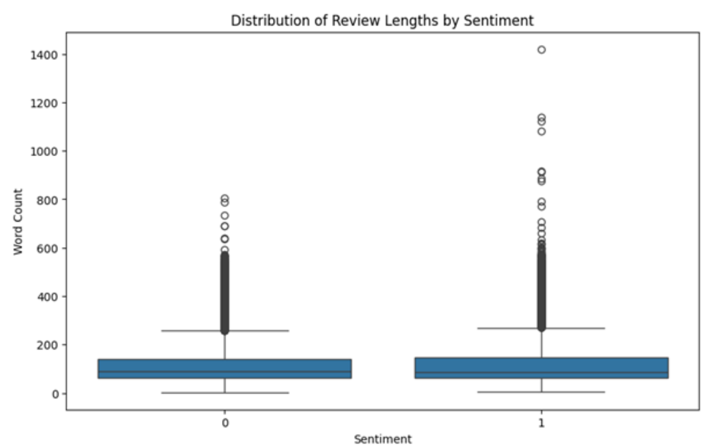

Figure 2: Distribution of sentence length

Given the unstructured nature of text data in movie reviews, preprocessing plays a crucial role in preparing the dataset for analysis. The key preprocessing steps applied include tokenization, which involves splitting the text into individual words while preserving sentence boundaries; lowercasing, which converts all text to lowercase for consistency; and stopword removal, which eliminates common words like “the,” "is," and "in" that do not carry sentiment. Additionally, lemmatization is used to convert words into their base forms, and cleaning procedures are employed to remove special characters, punctuation, and normalize repeated characters, thereby refining the text for more accurate analysis.

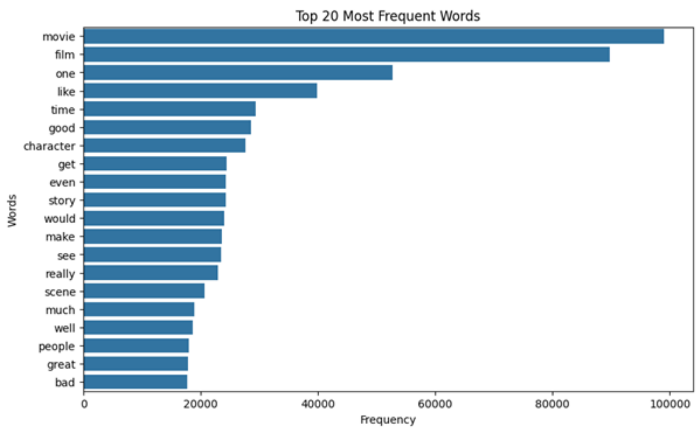

Figure 3: Distributions of frequency of words

The English model from spaCy is used to perform dependency parsing and emotional scores are derived from the AFINN sentiment lexicon (normalized to the range [-1,1]). During training, all parameters are optimized using the AdamW optimizer with a learning rate of 2e-5 and linear warmup. The hybrid architecture combines structural and affective information in text which enables more comprehensive sentiment analysis.

Three lexicon-based approaches were evaluated, each utilizing different dictionaries and techniques for sentiment scoring:

- **TextBlob**: Provides sentiment analysis through a simple lexicon-based approach by assigning polarity scores from negative one to positive one. It considers word polarity while accounting for basic grammatical structures.
- **VADER**: Optimized for social media and informal text, it incorporates rules for punctuation, capitalization, and negation. It uses a comprehensive sentiment lexicon with empirically evaluated sentiment intensity scores.
- **SentiWordNet**: Extends WordNet by assigning three sentiment scores—positive, negative, and neutral—to each sentence.

Traditional machine learning models were implemented using both Bag-of-Words and TF-IDF features for text representation:

- **Naïve Bayes**: This probabilistic classifier assumes feature independence and calculates class probability using Bayes' theorem.
- **Logistic Regression**: It models the probability of positive sentiment using a sigmoid function applied to a linear combination of features.
- **SVM**: Finds an optimal hyperplane to separate sentiment classes in the feature space, maximizing the margin between classes.
- **Random Forest**: This ensemble method combines multiple decision trees to reduce overfitting and improve generalization.

Advanced neural network architectures were implemented to capture complex linguistic patterns:

- **LSTM**: Handles sequential data by maintaining a cell state that captures long-term dependencies.
- **BERT**: Revolutionizes text processing by using bidirectional context and self-attention mechanisms.
- **BiLSTM**: A specialized version of LSTM that processes input sentences in both directions simultaneously.

To assess the sentiment analysis models' performance on movie reviews, we employ six key metrics:

- Accuracy
- Precision
- Recall
- F1 Score
- AUC-ROC
- Confusion Matrix

5 Results

The models are compared based on key metrics discussed earlier. The confusion matrices for each model are demonstrated in the figures. The comparison of accuracy, precision, and recall across different models is summarized below.

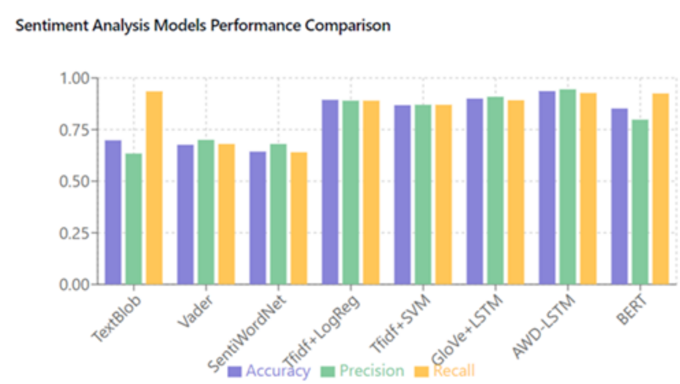

Figure 4: Comparison of Accuracy, Precision, and Recall for Various Sentiment Analysis Models: TextBlob, Vader, SentiWordNet, TfIdf+LogReg, TfIdf+SVM, GloVe+LSTM, AWD-LSTM, and BERT

Table 2 provides a summary of the evaluation metrics for all the models tested, illustrating the quantitative performance of each approach.

Table 2: Evaluation metrics for models tested

Model	Accuracy	Precision	Recall	F1 Score	AUC ROC
TextBlob	0.698	0.634	0.935	0.756	0.698
Vader	0.676	0.700	0.680	0.727	0.676
SentiWordNet	0.643	0.680	0.640	0.620	0.638
Tfid + Logistic Regression	0.894	0.890	0.890	0.890	0.961
Tfid + SVM	0.868	0.870	0.870	0.870	0.947
GloVe + LSTM	0.900	0.909	0.892	0.900	0.967
BERT	0.852	0.798	0.925	0.856	0.855
Proposed model	0.913	0.912	0.916	0.912	0.9632

Figure 5: Confusion Matrices for Sentiment Analysis Models: TextBlob, VADER, Stanford NLP, and BERT

Lexicon-based approaches like TextBlob, Vader, and SentiWordNet are characterized by their simplicity and ease of implementation. They rely on predefined dictionaries to evaluate sentiment, making them computationally efficient and interpretable. However, their reliance on fixed lexicons presents significant limitations:

- **Accuracy:** Lexicon-based models often yield moderate accuracy. For example, TextBlob achieved an accuracy of 69.8%, while Vader reached 67.57%.
- **Strengths:** They are straightforward to implement, requiring minimal computational resources and providing quick insights into sentiment.
- **Weaknesses:** These models struggle with the nuances of language, particularly in detecting sarcasm, idiomatic expressions, and mixed sentiments. For instance, in a review stating, "The acting was excellent, but the plot was terrible," lexicon models may misclassify the sentiment as positive, failing to recognize the contrasting sentiments expressed.

Machine learning models utilize TF-IDF and Bag-Of-Words for feature extraction. These techniques demonstrate significant improvements in sentiment classification:

- **Accuracy:** Models like Tfidf + Logistic Regression achieved an accuracy of 89.4%, significantly outperforming lexicon-based models.
- **Strengths:** These models can learn patterns from the data, allowing them to capture subtle distinctions between positive and negative sentiments. The effective weighting of words through techniques like TF-IDF contributes to better sentiment discrimination.
- **Weaknesses:** While machine learning models manage simpler linguistic patterns well, they may still struggle with the broader context or dependencies within longer sentences, limiting their effectiveness in nuanced sentiment analysis.

Deep learning approaches, particularly those utilizing LSTM and BiLSTM architectures, significantly outperformed both lexicon-based and traditional machine learning models:

- **Accuracy:** The GloVe + LSTM model achieved 90.02%, while AWD-LSTM reached 93.64%.
- **Strengths:** These models excel in capturing long-range dependencies, allowing them to understand the context surrounding individual words and phrases. Their ability to leverage pre-trained word embeddings enhances semantic understanding, making them adept at handling complex sentiment structures.
- **Weaknesses:** While deep learning models are powerful, they require significant computational resources and extensive datasets for training, which may limit their accessibility for some applications.

Our proposed model demonstrates robust performance across all evaluation metrics, achieving an accuracy of 91.3% and an AUC-ROC score of 0.9632. The model's confusion matrix reveals strong classification capabilities, with 4,834 correct negative predictions and 4,296 correct positive predictions out of 10,000 samples.

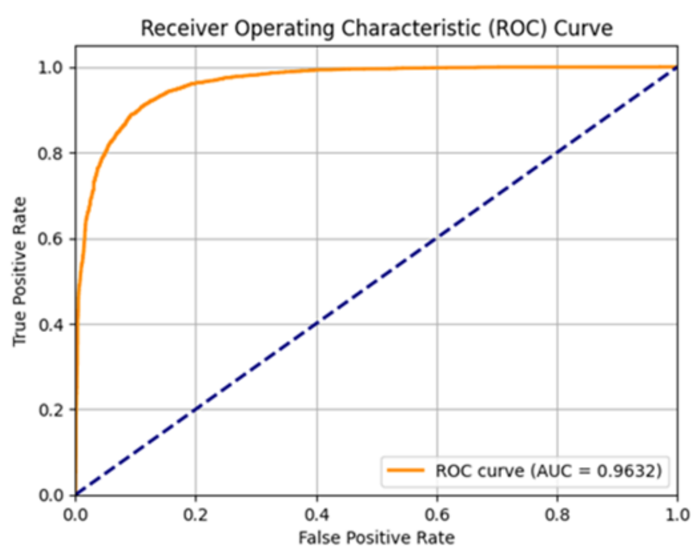

Figure 6: AUC ROC Curve for our model

The model exhibits balanced performance across sentiment classes, with class-specific metrics highlighting its effectiveness. For negative sentiment, it achieves a precision of 0.9593 and recall of 0.8791, indicating high reliability in negative sentiment identification. For positive sentiment, the model shows strong recall (0.9545) with good precision (0.8660), demonstrating effective positive sentiment detection.

Validation Set			
TARGET / OUTPUT	Positive	Negative	SUM
Positive	4296 42.96%	665 6.65%	4961 86.60% 13.40%
Negative	205 2.05%	4834 48.34%	5039 95.93% 4.07%
SUM	4501 95.45% 4.55%	5499 87.91% 12.09%	9130 / 10000 91.30% 8.70%

Figure 7: Confusion matrix for our model

The balanced F1-scores (0.9174 for negative and 0.9081 for positive) indicate consistent performance across classes, suggesting that our hybrid architecture effectively combines syntactic and emotional features. This balanced performance, coupled with the high AUC-ROC score, validates our approach of integrating dependency-guided attention with emotion-aware processing for sentiment analysis.

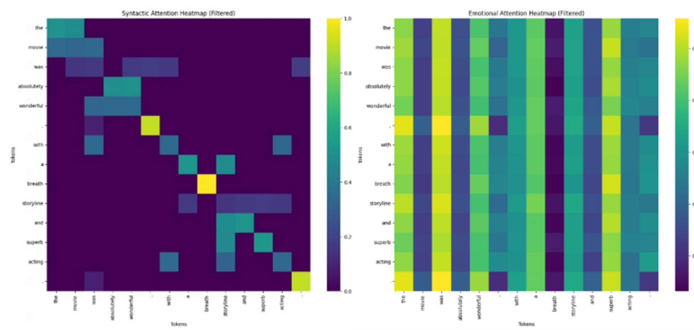

Figure 8: Syntactic (left) and emotional (right) attention heatmaps for the input sentence

The model's error patterns, particularly the lower false negative rate (205 cases) compared to false positives (665 cases), suggest that while the model is more conservative in assigning negative sentiment, it maintains high overall classification accuracy. These results demonstrate that our hybrid approach successfully captures both structural and affective aspects of sentiment, leading to robust and reliable classification performance.

Several potential threats to validity must be considered:

- **Dataset Bias:** The IMDb dataset primarily consists of movie reviews, which may not generalize well to other domains such as financial news or social media sentiment analysis.
- **Preprocessing Impact:** The choice of tokenization, stopword removal, and lemmatization could influence model performance. Alternative preprocessing techniques might yield different results.
- **Hyperparameter Sensitivity:** Deep learning models, including our proposed architecture, rely on numerous hyperparameters. Small variations in learning rates, batch sizes, or attention mechanisms may impact final performance.
- **Model Interpretability:** Despite improved interpretability via attention mechanisms, neural networks remain complex and somewhat opaque. Further analysis is needed to improve explainability.
- **Real-world Deployment:** The evaluation was conducted in a controlled environment. Performance in real-time applications with noisy or adversarial data remains an open challenge.

6 Conclusions

This paper provides a review of sentiment analysis techniques applied to IMDb reviews, with a focus on our proposed architecture combining syntactic and emotional components. The comparative analysis across different methodologies reveals clear differences in the performance and capabilities of methods. Our experimental results show that lexicon-based approaches, while interpretable and computationally efficient, face big challenges in handling the nuanced complexity of movie reviews. Machine learning models using TF-IDF and BoW for feature extraction showed improvements. They achieved accuracies above 85% but still struggled with maintaining contextual understanding in larger sentences.

Deep learning architectures, particularly our architecture, demonstrated better performance in maintaining long-range dependencies and relationships. By using syntactic dependencies with emotional features via a dual stream attention mechanism, our model achieved accuracy of 89.73% and an AUC ROC score of 0.9632. These results validate our design choices and show the effectiveness of combining syntactic and emotional features.

The current investigation is constrained by its consideration of only one domain, i.e., IMDb movie reviews, and such narrowing might limit the applicability of the findings to other text types or use cases. The use of English-language corpora also forecloses direct application to multilingual or code-switching scenarios without further modification. The suggested architecture relies on the correctness of external tools like dependency parsers and sentiment lexicons and is thus vulnerable to parser errors and lexicon coverage issues. Additionally, the use of BERT-based encoding, though effective, brings with it significant computational and memory demands that can prevent deployment in environments with less resources. Lastly, while attention visualizations provide some interpretability, the decision process itself remains opaque, hindering full transparency of model reasoning.

Future work includes expanding this effort to utilize the suggested dual-stream attention architecture in varied domains like financial news, healthcare stories, and social media posts to test domain generalization. Code-switched and multilingual sentiment analysis will be investigated by integrating with models like mBERT or XLM-R and language-specific syntactic parsers and emotion lexicons. The approach can be further tested with newer transformer variants (e.g., RoBERTa, DeBERTa) and augmented emotion representations obtained from context-sensitive embeddings or knowledge graphs. Real-time sentiment analysis applications such as streaming data will be explored with low latency optimized inference. Apart from that, future work will target better model interpretability through attention visualization and explanation techniques, robustness against noisy and adversarial inputs, and multimodal sentiment analysis extension to text, visual, and audio modalities.

References

[1] Wajdi Aljedaani, Furqan Rustam, Mohamed Wiem Mkaouer, Abdullatif Ghallab, Vaibhav Rupapara, Patrick Bernard Washington, Ernesto Lee, and Imran Ashraf. Sentiment analysis on twitter data integrating textblob and deep learning models: The case of us airline industry. *Knowledge-Based Systems*, 255:109780, 2022.

[2] Stefano Baccianella, Andrea Esuli, and Fabrizio Sebastiani. SentiWordNet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining. In Nicoletta Calzolari, Khalid Choukri, Bente Maegaard, Joseph Mariani, Jan Odijk, Stelios Piperidis, Mike Rosner, and Daniel Tapias, editors, *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10)*, Valletta, Malta, May 2010. European Language Resources Association (ELRA).

[3] Bahrawi Bahrawi. Sentiment analysis using random forest algorithm-online social media based. *Journal of Information Technology and Its Utilization*, 2:29, 12 2019.

[4] Jeremy Barnes, Roman Klinger, and Sabine Schulte im Walde. Projecting embeddings for domain adaptation: Joint modeling of sentiment analysis in diverse domains, 2018.

[5] Piotr Bojanowski, Edouard Grave, Armand Joulin, and Tomas Mikolov. Enriching word vectors with subword information. *Transactions of the Association for Computational Linguistics*, 5:135–146, 2017.

[6] Peng Chen, Zhongqian Sun, Lidong Bing, and Wei Yang. Recurrent attention network on memory for aspect sentiment analysis. In Martha Palmer, Rebecca Hwa, and Sebastian Riedel, editors, *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pages 452–461, Copenhagen, Denmark, September 2017. Association for Computational Linguistics.

[7] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. BERT: Pre-training of deep bidirectional transformers for language understanding. In Jill Burstein, Christy Doran, and Thamar Solorio, editors, *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 4171–4186, Minneapolis, Minnesota, June 2019. Association for Computational Linguistics.

[8] Cícero dos Santos and Maíra Gatti. Deep convolutional neural networks for sentiment analysis of short texts. In Junichi Tsujii and Jan Hajic, editors, *Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers*, pages 69–78, Dublin, Ireland, August 2014. Dublin City University and Association for Computational Linguistics.

[9] Teng-Kai Fan and Chia-Hui Chang. Blogger-centric contextual advertising. In *Proceedings of the 18th ACM Conference on Information and Knowledge Management*, CIKM '09, page 1803–1806, New York, NY, USA, 2009. Association for Computing Machinery.

[10] Jeremy Howard and Sebastian Ruder. Universal language model fine-tuning for text classification, 2018.

[11] Hanhoon Kang, {Seong Joon} Yoo, and Dongil Han. Senti-lexicon and improved naïve bayes algorithms for sentiment analysis of restaurant reviews. *Expert Systems with Applications*, 39(5):6000–6010, April 2012. Funding Information: This work was supported by the Korea Research Foundation (KRF) grant funded by the Korea government (MEST) (No. 2010-0015842).

[12] Pei Ke, Haozhe Ji, Siyang Liu, Xiaoyan Zhu, and Minlie Huang. SentiLARE: Sentiment-aware language representation learning with linguistic knowledge. In Bonnie Webber, Trevor Cohn, Yulan He, and Yang Liu, editors, *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 6975–6988, Online, November 2020. Association for Computational Linguistics.

[13] S. Kiritchenko, X. Zhu, and S. M. Mohammad. Sentiment analysis of short informal texts. *Journal of Artificial Intelligence Research*, 50:723–762, Aug 2014.

[14] Qizhi Li, Xianyong Li, Yajun Du, Yongquan Fan, and Xiaoliang Chen. A new sentiment-enhanced word embedding method for sentiment analysis. *Applied Sciences*, 12(20), 2022.

[15] Xin Li, Lidong Bing, Wenxuan Zhang, and Wai Lam. Exploiting bert for end-to-end aspect-based sentiment analysis. pages 34–41, 01 2019.

[16] Ning Liu, Bo Shen, Zhenjiang Zhang, Zhiyuan Zhang, and Kun Mi. Attention-based sentiment reasoner for aspect-based sentiment analysis. *Human-centric Computing and Information Sciences*, 9(1):35, 2019.

[17] Pengfei Liu, Xipeng Qiu, and Xuanjing Huang. Recurrent neural network for text classification with multi-task learning, 2016.

[18] Xiaodong Liu, Pengcheng He, Weizhu Chen, and Jianfeng Gao. Multi-task deep neural networks for natural language understanding. In Anna Korhonen, David Traum, and Lluís Màrquez, editors, *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 4487–4496, Florence, Italy, July 2019. Association for Computational Linguistics.

[19] Andrew L. Maas, Raymond E. Daly, Peter T. Pham, Dan Huang, Andrew Y. Ng, and Christopher Potts. Learning word vectors for sentiment analysis. In Dekang Lin, Yuji Matsumoto, and Rada Mihalcea, editors, *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, pages 142–150, Portland, Oregon, USA, June 2011. Association for Computational Linguistics.

[20] Tomas Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean. Efficient estimation of word representations in vector space, 2013.

[21] Saif Mohammad, Svetlana Kiritchenko, and Xiaodan Zhu. NRC-Canada: Building the state-of-the-art in sentiment analysis of tweets. In Suresh Manandhar and Deniz Yuret, editors, *Second Joint Conference on Lexical and Computational Semantics (*SEM), Volume 2: Proceedings of the Seventh International Workshop on Semantic Evaluation (SemEval 2013)*, pages 321–327, Atlanta, Georgia, USA, June 2013. Association for Computational Linguistics.

[22] Bo Pang, Lillian Lee, and Shivakumar Vaithyanathan. Thumbs up? sentiment classification using machine learning techniques. In *Proceedings of the 2002 Conference on Empirical Methods in Natural Language Processing (EMNLP 2002)*, pages 79–86. Association for Computational Linguistics, July 2002.

[23] Jeffrey Pennington, Richard Socher, and Christopher Manning. GloVe: Global vectors for word representation. In Alessandro Moschitti, Bo Pang, and Walter Daelemans, editors, *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 1532–1543, Doha, Qatar, October 2014. Association for Computational Linguistics.

[24] Matthew E. Peters, Mark Neumann, Mohit Iyyer, Matt Gardner, Christopher Clark, Kenton Lee, and Luke Zettlemoyer. Deep contextualized word representations. In Marilyn Walker, Heng Ji, and Amanda Stent, editors, *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers)*, pages 2227–2237, New Orleans, Louisiana, June 2018. Association for Computational Linguistics.

[25] Barbara Plank, Anders Søgaard, and Yoav Goldberg. Multilingual part-of-speech tagging with bidirectional long short-term memory models and auxiliary loss. In Katrin Erk and Noah A. Smith, editors, *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 412–418, Berlin, Germany, August 2016. Association for Computational Linguistics.

[26] Guozheng Rao, Weihang Huang, Zhiyong Feng, and Qiong Cong. Lstm with sentence representations for

document-level sentiment classification. *Neurocomputing*, 308:49–57, 2018.

[27] Elena Rudkowsky, Martin Haselmayer, Matthias Wastian, Marcelo Jenny, Štefan Emrich, and Michael Sedlmair. More than bags of words: Sentiment analysis with word embeddings. *Communication Methods and Measures*, 12(2-3):140–157, 2018.

[28] Richard Socher, Alex Perelygin, Jean Wu, Jason Chuang, Christopher D. Manning, Andrew Ng, and Christopher Potts. Recursive deep models for semantic compositionality over a sentiment treebank. In David Yarowsky, Timothy Baldwin, Anna Korhonen, Karen Livescu, and Steven Bethard, editors, *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*, pages 1631–1642, Seattle, Washington, USA, October 2013. Association for Computational Linguistics.

[29] Chi Sun, Luyao Huang, and Xipeng Qiu. Utilizing BERT for aspect-based sentiment analysis via constructing auxiliary sentence. In Jill Burstein, Christy Doran, and Thamar Solorio, editors, *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 380–385, Minneapolis, Minnesota, June 2019. Association for Computational Linguistics.

[30] Maite Taboada, Julian Brooke, Milan Tofiloski, Kimberly Voll, and Manfred Stede. Lexicon-based methods for sentiment analysis. *Computational Linguistics*, 37(2):267–307, June 2011.

[31] Duyu Tang, Bing Qin, and Ting Liu. Document modeling with gated recurrent neural network for sentiment classification. In Lluís Màrquez, Chris Callison-Burch, and Jian Su, editors, *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, pages 1422–1432, Lisbon, Portugal, September 2015. Association for Computational Linguistics.

[32] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. In *Proceedings of the 31st International Conference on Neural Information Processing Systems*, NIPS'17, page 6000–6010, Red Hook, NY, USA, 2017. Curran Associates Inc.

[33] Zhilin Yang, Zihang Dai, Yiming Yang, Jaime Carbonell, Russ R Salakhutdinov, and Quoc V Le. Xlnet: Generalized autoregressive pretraining for language understanding. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 32. Curran Associates, Inc., 2019.

[34] Liang-Chih Yu, Jin Wang, K. Robert Lai, and Xuejie Zhang. Refining word embeddings for sentiment analysis. In Martha Palmer, Rebecca Hwa, and Sebastian Riedel, editors, *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pages 534–539, Copenhagen, Denmark, September 2017. Association for Computational Linguistics.

Authors

Dr. Nidhi Mishra is working as an Assistant Professor in the Department of Computer Science and Engineering at Amity University, Noida, Uttar Pradesh, India. She received her Ph.D. degree from Banasthali University, Jaipur, India in 2016. She has published more than 55 articles in peer-reviewed international journals and conferences. She is a member of IEEE and ACM society. She has guided 3 Ph.D. and 15 M.Tech research scholars in the area of Artificial Intelligence, Machine Learning, Deep Learning, Neural Networks, and NLP. Her research interests are Artificial Intelligence, Information Retrieval, Data Science, and Natural Language Processing.

Dr. Aakanshi Gupta is working is as an Assistant Professor in the department of Computer Science and Engineering at Amity University, Noida, Uttar Pradesh, India. She has an experience of more than 15 years in academics and research She received her Ph.D. degree from GGSIPU, Delhi; under the guidance Prof. Bharti Suri. Her research interests are Artificial Intelligence, Information retrieval, Data Science and Software Reliability.

Dr. Shuchi Mala is working as an Assistant Professor in the department of Computer Science and Engineering at Amity University, Noida, Uttar Pradesh, India. She has an experience of more than 6 years in academics and 4 years in full-time research She received her Ph.D. degree from Malaviya National Institute of Technology Jaipur (MNIT), India in 2018. She has published more than 32 articles in peer reviewed journals, international conferences, and book chapters. She is a member of IET and ACM society. Her research interests are spatial data modelling, health informatics and intelligent systems. She actively serves as a reviewer for esteemed international journals.

Srishti Das is currently pursuing a B.Tech in Computer Science and Engineering from Amity University, Noida, Uttar Pradesh, India. Her interests include deep learning, computer vision, and artificial intelligence. She is enthusiastic about research in emerging AI technologies and aims to contribute to the field through innovative applications.

Naman Tyagi is currently pursuing a B.Tech in Computer Science and Engineering from Amity University, Noida, India. His research interests include deep learning, model optimization, and artificial intelligence. He is passionate about exploring efficient and practical applications of AI techniques.

Sanjam Bhardwaj is pursuing a B.Tech in Computer Science and Engineering at Amity University, Noida. His research interests include Artificial Intelligence, Machine Learning, and Cybersecurity.

Dr. Narayan C Debnath is currently the Founding Dean of the School of Computing and Information Technology at Eastern International University, Vietnam. He is also serving as the Head of the Department of Software Engineering at Eastern International University, Vietnam. Formerly, Dr. Debnath served as a Full Professor of Computer Science at Winona State University, Minnesota, USA for 28 years, and the elected Chairperson of the Computer Science Department at Winona State University for 7 years. Dr. Debnath has been the Director of the International Society for Computers and their Applications (ISCA), USA since 2014. Professor Debnath made significant contributions in teaching, research, and services across the academic and professional communities. He has made original research contributions in Software Engineering, Artificial Intelligence and applications, and Information Science, Technology and Engineering. He is an author or co-author of over 500 research paper publications in numerous refereed journals and conference proceedings in Computer Science, Information Science, Information Technology, System Sciences, Mathematics, and Electrical Engineering. He is also an author of over 15 books published by well-known international publishers including Elsevier, CRC, Wiley, Bentham Science, River Publishing, and Springer. Dr. Debnath has made numerous teaching, research and invited keynote presentations at various international conferences, industries, and teaching and research institutions in Africa, Asia, Australia, Europe, North America, and South America. He has been a visiting professor at universities in Argentina, China, India, Sudan, and Taiwan. He has been maintaining an active research and professional collaborations with many universities, faculty, scholars, professionals and practitioners across the globe. Dr. Debnath is an active member of the IEEE, IEEE Computer Society, and a Senior Member of the International Society for Computers and their Applications (ISCA), USA.

Amit Mishra has eleven years of teaching experience and research experience. He has done PhD from NIT Kurukshetra, Haryana in the area of Natural Language Processing. He has worked on a couple of research projects such as developing a Semantic Question Answering System, and an Intention Mining of potential buyers on E-Commerce Websites. His Area of interest includes Data Science, Natural Language Processing, and Machine Learning. He has published papers in various reputed International conferences and Journals. He has technical expertise in C, Java (J2SE J2EE), Android programming, MySQL, IBM DB2, ORACLE, and MATLAB