DOI: http://dx.doi.org/10.26483/ijarcs.v16i2.7220

Volume 16, No. 2, March-April 2025

ISSN No. 0976-5697

nternational Journal of Advanced Research in Computer Science

RESEARCH PAPER

Available Online at www.ijarcs.info

ENHANCING ASPECT-BASED SENTIMENT ANALYSIS USING LARGE LANGUAGE MODELS AND DEPENDENCY PARSING

Yougaaraj R, Harris Christopher A, Nandacumaar S
UG Students, Department of Computer Science and Engineering,
Puducherry Technological University, East Coast Road, Pillaichavady, Puducherry, 605014

Dr. Sheeba J I
Associate Professor, Department of Computer Science and Engineering,
Puducherry Technological University, East Coast Road, Pillaichavady, Puducherry, 605014

Abstract: Sentiment analysis is a critical task in Natural Language Processing (NLP) that determines sentiment polarity within textual data. Traditional sentiment analysis primarily focuses on binary classification. However, real-world reviews and social media content often exhibit multiple sentiments within a single sentence. This complexity necessitates Aspect-Based Sentiment Analysis (ABSA), which identifies aspect terms and their corresponding sentiments. Despite advancements, existing ABSA models struggle to capture interdependencies between aspect-opinion pairs, leading to misclassifications in multi-aspect scenarios. To address this, our study proposes enhanced ABSA model which integrates dependency parsing with Large Language Model (LLM)-based learning to incorporate structured semantic knowledge for effective aspect-opinion relationship extraction. The integration of structured feature engineering and domain-specific vocabulary filtering in the proposed work ensures more precise sentiment classification. Experimental evaluations, based on average metrics computed from 5-fold cross validation, demonstrate that the proposed model outperforms existing methods. The model achieves a 3.4% improvement in precision, a 4.9% increase in recall, and a 3.8% boost in F1-score. Additionally, it yields a 5.6% increase in Matthews Correlation Coefficient (MCC), reduces the False Discovery Rate by 3.3%, and lowers the Hamming Loss by 1.7%, ensuring enhanced consistency in multi-aspect sentiment classification. These findings underscore the value of integrating structured semantic knowledge into ABSA, which can significantly enhance the accuracy of sentiment analysis in practical applications.

Keywords: Sentiment Analysis, Aspect-Opinion pair Extraction, Large Language Models, Dependency Parsing, Multi-task Learning.

I. INTRODUCTION

Sentiment Analysis is a fundamental task in Natural Language Processing (NLP) that aims to determine sentiment polarity within textual data. Traditionally, sentiment analysis focused on binary classification, categorizing statements as either positive or negative. However, real-world user-generated content, such as product reviews, social media posts, and customer feedback, often contains multiple sentiments within a single sentence. This complexity necessitates more granular sentiment analysis approaches that can accurately capture varying opinions expressed within the same text.

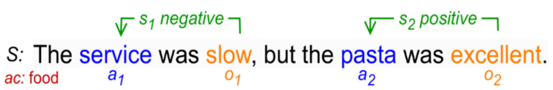

Figure 1. Sample Multi-Aspect, Multi-polarity Scenario

To address this complexity, Aspect-Based Sentiment Analysis (ABSA) has emerged as an advanced framework that identifies specific aspect terms and their corresponding sentiments. The extraction of aspect terms is critical, as reviews and feedback often contain multi-aspect and multi-polarity characteristics (Figure 1). For example, in the restaurant domain, a customer review such as *"The service was slow, but the pasta was excellent"* presents conflicting sentiments: *'service'* conveys a negative sentiment, while *'pasta'* carries a positive one. Standard sentiment analysis models, which do not distinguish between different aspects, often misclassify such statements, leading to inaccurate sentiment predictions [2].

While addressing the necessity of correctly identifying aspect-opinion pairs, recent research in ABSA has focused on structured sentiment extraction. Advances in this area have led to the development of models such as Linking words-guided multidimensional emotional data augmentation and adversarial contrastive training (LWEDA-ACT), which enhances ABSA by generating richer, domain-specific training samples to improve sentiment classification accuracy [1]. Additionally, open-source frameworks such as SETFIT have been widely adopted for aspect and sentiment term extraction. However, these models primarily rely on individually extracting aspects and sentiment terms rather than employing a multi-task learning approach, which limits their ability to capture interdependencies between aspect-opinion pairs. This remains a significant challenge when handling multiple aspects and their corresponding sentiments within a single piece of user feedback.

To mitigate this limitation, extracting aspect-opinion pairs forms a critical basis for managing multi-aspect and multi-polarity sentiment analysis, thereby ensuring that sentiment classification remains both accurate and context-aware. In recent years, Large Language Models (LLMs) such as LLAMA have been explored for multi-task learning in ABSA. While these models demonstrate strong generalization capabilities, they often fail to capture the intricate semantic relationships between aspects and opinions effectively. This challenge arises because current methods struggle to model the dependencies between aspect-opinion pairs in a structured manner, leading to suboptimal sentiment classification.

Resolving this challenge necessitates an enhanced method that embeds structured semantic knowledge into the process. NLP-driven data manipulation techniques, such as structured feature engineering and domain-specific vocabulary filtering, play a crucial role in refining aspect-opinion linking and improving sentiment classification accuracy. These methods help models better understand the relationships between different aspects and their associated sentiments, thereby enhancing the overall sentiment analysis process.

This study proposes an improved ABSA model that integrates dependency parsing with LLM-based learning to incorporate structured semantic knowledge for aspect-opinion relationship extraction. By utilizing dependency parsing, the model can more effectively identify the semantic connections between aspects and their associated opinions, which leads to enhanced accuracy and robustness in sentiment classification, especially in scenarios involving multiple aspects. Integrating structured semantic knowledge into LLM-based learning helps to bridge the gap between aspect extraction and opinion classification, thereby enabling sentiment analysis models to handle real-world text data more efficiently.

The rest of the paper is organized as follows. Section 2 reviews the literature by discussing current approaches and identifying their shortcomings. In Section 3, the proposed methodology is detailed, followed by Section 4, which examines the experimental results and analysis. Finally, Section 5 concludes the paper with key findings.

II. LITERATURE REVIEW

Aspect-Based Sentiment Analysis (ABSA) has gained significant attention due to its ability to extract fine-grained sentiments related to specific aspects within textual data. This granular approach to sentiment analysis has widespread applications in domains such as e-commerce, social media monitoring, and customer feedback analysis. To enhance the effectiveness of ABSA, various techniques have been proposed, ranging from deep learning-based models and generative approaches to syntax-aware adaptations and external knowledge integration.

Huang et al. (2025) introduced a linking words-guided emotional augmentation approach to improve ABSA by generating in-domain samples and employing clustering-based semantic boundary estimation for adversarial contrastive training. Their approach enhances performance under data scarcity conditions and improves handling of complex emotional dimension interactions[1]. Similarly, Zou and Wang (2025) proposed a syntax-aware domain adaptation method integrated with large language models and soft prompt learning, which successfully improves ABSA performance across domains. However, these studies are bounded to only two domain topics, limiting the breadth of its applicability[2].

Meanwhile, Xu et al. (2025) introduced a syntax-aware graph attention network combined with BiLSTM, improving ABSA performance, particularly in cases involving conflicting sentiments across multiple aspects[5]. He et al. (2025) addressed information loss in ABSA models by proposing a feature-driven layer freezing technique with multi-head self-attention. These methods preserve implicit feature information across layers and outperforms traditional models, but they struggles with dataset imbalance and requires significant computational resources[6].

On the other side, Lam et al. (2025) tackled this by combining LIME and LORE with similarity-based sampling to interpret the inner workings of state-of-the-art ABSA models. Their findings provide valuable insights into model behaviour, yet their approach is currently confined to the very few linguistic structures in text, limiting its generalizability[3]. In a related study, Chen et al. (2024) explored aspect-oriented semantic and syntactic knowledge integration using Graph Convolutional Networks (GCN) with BERT. Their approach leverages self-attention mechanisms to enhance performance, but it still faces challenges in reducing false positives, which can negatively impact sentiment classification accuracy[4].

Another promising research direction is the incorporation of external knowledge to improve ABSA performance. Lia et al. (2024) integrated knowledge from Wiktionary and SenticNet into a fusion model, enhancing sentiment polarity accuracy by leveraging entity definitions and sentiment lexicons. However, this dependency on external databases makes the approach less adaptable to real-time applications where external knowledge sources may not always be available[8]. Shen et al. (2024) introduced the Aspect Term Information Enhancement (ATIE) model, combining word-level BERT with dual adjacency matrices in GCN. This approach effectively addresses word vector misalignment and improves sentiment classification. However, it requires substantial computational resources, making it less feasible for deployment in real-world applications with limited infrastructure[9].

Generative models have also emerged as promising solutions for ABSA improvement. Zhu et al. (2025) proposed a generative aspect-based sentiment analysis approach using grid-based tag matching, which enhances sentiment relationship estimation while minimizing error propagation. However, the method faces difficulties in classifying similar aspect categories, particularly in longer sentences[7].

Despite the remarkable progress made in ABSA research, several challenges remain unaddressed. The ability to generalize across varying linguistic structures of the selected domains continues to be a pressing issue, as many models perform well in linguistically-constrained settings but fail when applied to diverse datasets. Most importantly, estimating aspect-opinion relationship pairs in multi-aspect contexts remains a persistent challenge. As ABSA models become more sophisticated, future research must focus on developing robust and scalable solutions that accurately capture complex aspect-opinion relationships while addressing computational and generalization challenges.

III. METHODOLOGY

The proposed system (Figure 2) follows a modular pipeline that integrates structured semantic knowledge with advanced language modelling techniques. Each module is designed to progressively refine the input data, from raw text to high-quality aspect-opinion-sentiment triples before fine-tuning a state-of-the-art LLaMA model using LoRA. The overall architecture comprises the following key components:

1. Data Preprocessing
2. Augmentation & Adversarial Sample Generation
3. Dependency Parsing
4. Filtering of Aspect-Opinion Pairs Using Zero-Shot Prompting
5. Fine-Tuning LLaMA-LoRA based model
6. Evaluation using 5-Fold cross validation

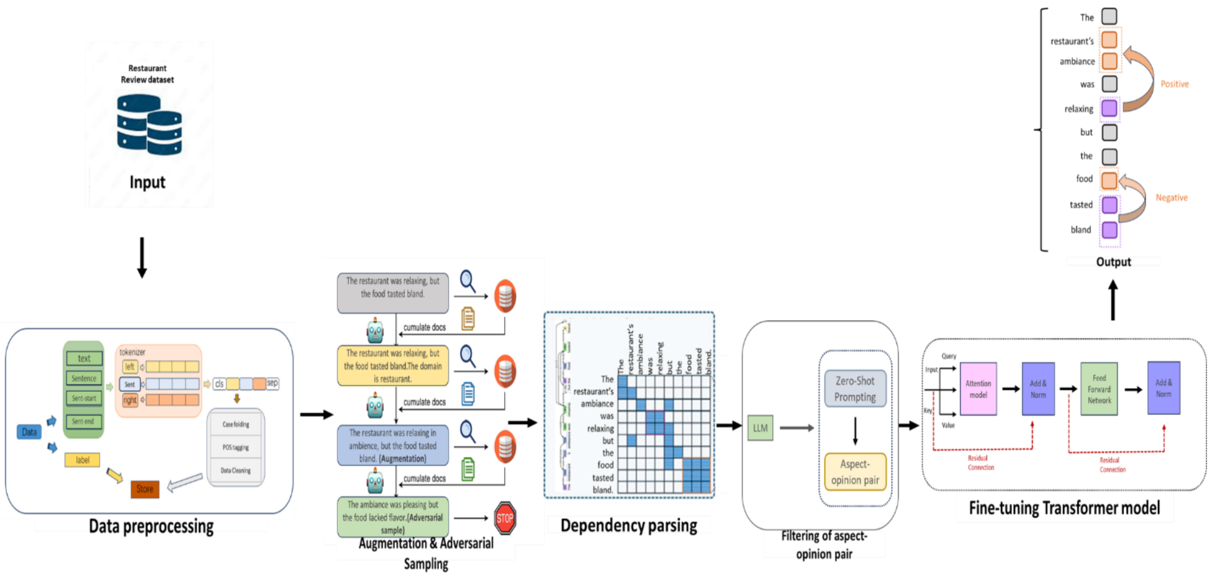

Figure 2. Proposed LLM and Dependency Parsing Integrated Architecture

A. Data Preprocessing

Effective data preprocessing is essential to ensure that the subsequent sentiment analysis steps can operate on clean and uniform text.

1) Data Cleaning: Data cleaning eliminates unwanted noise and fixes formatting issues by removing extraneous symbols (such as emojis), consolidating multiple punctuation marks into a single instance, and trimming unnecessary whitespace.

Example:

Input: "The restaurant is amazing!!! 🍕 🍝 Loved the ambiance!!!"

Output: "The restaurant is amazing! Loved the ambiance!"

2) Case Folding: Case folding maintains uniformity across the dataset. This is simple and yet a critical step that prevents the model from treating the same words as different words. It helps normalize the text, especially when user-generated content often mixes uppercase and lowercase letters arbitrarily.

3) Part-of-Speech (POS) Tagging: Assigning parts-of-speech (POS) tags (e.g., NOUN, VERB, ADJ) to each token enables the system to determine its grammatical role, which in turn facilitates the accurate linking of relevant aspect terms to their corresponding opinion terms.

Example:

The: **DET** restaurant: **NOUN** is: **AUX** amazing: **ADJ** !: **PUNCT** Loved: **VERB** the: **DET** ambiance: **NOUN** !: **PUNCT**

B. Augmentation and Adversarial Sample Generation

1) Augmentation: To address data scarcity and improve linguistic consistency, we employ a structured augmentation procedure that transforms each original sentence into a refined variant. Specifically, sentences are split at commas to isolate sub-phrases, each of which is reconstructed to include a clear subject and opinion. This reconstruction enforces a consistent format, commonly [Subject][connecting verb][rest of the sentence], thereby standardizing sentence structure. In addition, **informal abbreviations and slang are replaced with their full, formal equivalents** to minimize irregularities.

Example:

Input: I love the drinks, esp lychee martini and the food is also very good.

Output: The drinks are loved by me especially lychee martini, and the food is also excellent.

2) Adversarial Sample Generation: Adversarial sample generation rephrases the input sentence while preserving the critical aspect-sentiment pairs intact. The process focuses on modifying only the surrounding descriptive language, generating adversarial examples for all the reviews in the dataset. This targeted rephrasing introduces subtle linguistic variations, challenging the model to maintain its focus on key information while enhancing robustness.

Example:

Input: The restaurant was relaxing in ambience, but the food tasted bland.

Aspect-Sentiment pairs: (“ambience", “positive") , (“food", “negative")

Output**:** The ambiance was pleasing ,but the food lacked flavor.

C. Dependency Parsing

Dependency Parsing module leverages a transformer-based spacy model to extract aspect-opinion pairs from sentences in a dataset. It employs multiple syntactic patterns to reliably capture and associate aspects with their corresponding opinions.

The following patterns are being considered:

Pattern A:Direct adjectival modifier (amod) (Figure 3) - Captures adjectives that are directly modifying a noun.

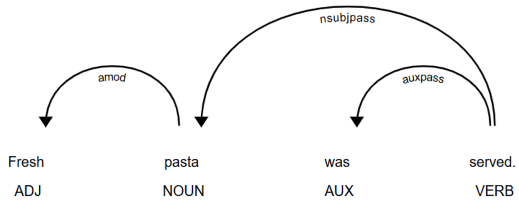

Figure 3. Illustrating Direct adjectival modifier (amod)

Extraction: ("pasta", "Fresh")

Pattern B: Copular constructions (acomp, noun) (Figure 4) - Identifies adjectives linked to a noun via a copular verb.

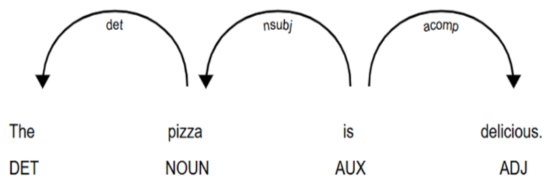

Figure 4. Illustrating Copular constructions (acomp, noun)

Extraction: (“pizza", “delicious")

Pattern C: Verbal constructions with negation (Figure 5) - Extracts pairs where negation modifies a descriptive verb

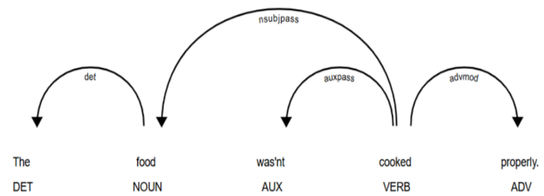

Figure 5. Illustrating Verbal construction with negation

Extraction: ("food", "not cooked")

Pattern D: Loose adjectives in descriptive clauses (Figure 6) - Captures adjectives from additional descriptive clauses

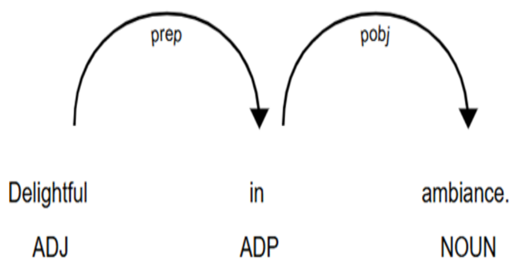

Figure 6. Illustrating Loose adjectives in descriptive clauses

Extraction:(“ambience", “delightful")

Example:

Input: It wasn't an amazing evening, but the food I ate was good

Output: [["food", "good"], ["evening", "amazing"]]

D. Filtering of Aspect-Opinion Pairs Using Zero-Shot Prompting

Dependency parsing can yield syntactically valid but contextually irrelevant aspect-opinion pairs . For example, “It was an amazing evening and the food I ate was good” may produce ["food", "good"] and ["evening", "amazing"].While ["food", "good"] is valid, ["evening", "amazing"] is not contextually relevant in a restaurant review . Our filtering module checks if the opinion genuinely applies to the aspect restaurant reviews context . Pairs that fail this contextual validity check are discarded . This ensures only meaningful aspect-opinion pairs remain for accurate sentiment analysis**.**

E. Fine-Tuning LLaMA-LoRA based model

Fine-tuning the LLaMA-LoRA model on a restaurant dataset involves adapting a pre-trained LLaMA model using Low-Rank Adaptation (LoRA) to capture restaurant review sentiment nuances. The dataset, which contains restaurant reviews annotated with aspects, opinions, and sentiments, is preprocessed into prompts that embed these annotations . For example, ### Human: The food is fresh at a good price, and the place is clean and hygienic. ###Assistant: Aspect detected: food, price, place ## Opinion detected: fresh, good, clean and hygienic ## Sentiment detected: positive, positive, positive

1) Training : In the query projection layer, W has shape d×k . LoRA leaves this matrix unchanged during training . Instead of modifying W directly, LoRA introduces two smaller matrices A and B:

ΔW=A×B

Where,

A – matrix of shape d×r
B – matrix of shape r×k
r - is the low-rank factor, chosen to be much smaller than d and k, thereby keeping additional parameters low.

The effective weight used in the forward pass becomes W+ΔW. The loss is computed based on the output using these combined weights, and the gradients update only A and B during the backpropagation of the loss.

2) Merging LoRA and Base Weights: After training, ΔW=A×B is computed and added to the original weight matrix W to form a new, fully adapted weight matrix .

W new=W+ΔW

F. Evaluation using 5-Fold cross validation

Our approach utilizes 5-fold cross validation via scikit-learn’s KFold. The dataset is split into five equal segments, with each segment serving as the validation set at least once while the others are used for training. This method enables us to assess the model's performance across various unseen data splits, even when certain classes have fewer samples.Performance is summarized by aggregating metrics like precision, recall, F1 score (F1), Matthews Correlation Coefficient (MCC), hamming loss, and False Discovery Rate (FDR) using the average.

IV. EXPERIMENTAL RESULTS AND DISCUSSION

A. Dataset

To evaluate the effectiveness and robustness of our proposed framework, we conducted experiments on a well-established benchmark restaurant review dataset from SemEval-2014 Task 4: Aspect-Based Sentiment Analysis[12]. This dataset comprises over 3,000 English restaurant reviews, reflecting customer experiences across various dining scenarios. It includes annotations for aspect terms and their corresponding polarities, making it a standard resource for aspect-based sentiment analysis tasks.

B. Metrics

To conduct a comprehensive evaluation of our framework, we use six primary metrics: Precision, Recall, F1 Score, Matthews Correlation Coefficient (MCC), Hamming Loss, and False Discovery Rate (FDR).

Precision

Measures the proportion of predicted positive labels that are actually correct.

Precision = TP / (TP + FP)

Recall
Assesses the ability of the model to capture all actual positive labels.

Recall = TP / (TP + FN)

F1 Score

The harmonic mean of Precision and Recall, balancing the trade-off between false positives and false negatives.

F1 Score = 2 × (Precision × Recall) / (Precision + Recall)

Matthews Correlation Coefficient (MCC)

Provides a balanced evaluation of the model's prediction quality, particularly useful for imbalanced datasets.

MCC = (TP × TN – FP × FN) / sqrt[(TP + FP) × (TP + FN) × (TN + FP) × (TN + FN)]

Hamming Loss

Quantifies the fraction of incorrect predictions in multi-label classification by comparing the total number of false predictions (false positives and false negatives) against the total number of predictions.

Hamming Loss = (FP + FN) / (TP + TN + FP + FN)

False Discover Rate(FDR)

Indicates the proportion of false positive predictions among all positive predictions.

FDR = FP / (TP + FP)

C. Table and Graph

In this section, we compare the performance of our proposed framework against existing approach by analysing all six metrics across different folds. The evaluation highlights improvements and trade-offs between methods, with results presented using tables and graphs for clarity.

The above graph and table (Table 1, Figure 7) clearly shows that our proposed work outperforms the existing approach across all five folds, as evidenced by the improvements in precision, recall, and F1 scores .

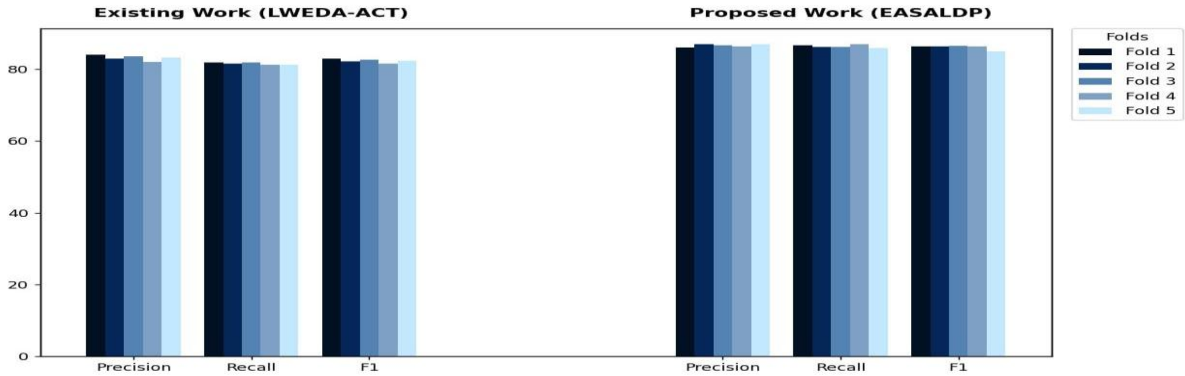

Figure 7. Comparison of Cross-Validation Performance Metrics (Precision,Recall,F1)

Table 1. Comparison of Cross-Validation Performance Metrics (Precision,Recall,F1)

Fold	***Existing Work(LWEDA-ACT)***			***Proposed Work(EASALDP)***		
	Precision	***Recall***	***F1***	***Precision***	***Recall***	***F1***
1	84.07	81.91	83.03	86.09	86.80	86.39
2	82.99	81.63	82.30	86.99	86.29	86.45
3	83.60	81.91	82.70	86.79	86.30	86.53
4	82.05	81.30	81.65	86.39	87.01	86.46
5	83.33	81.29	82.32	87.01	86.02	85.03

Table 2. Comparison of Cross-Validation Performance Metrics (MCC, Hamming Loss ,FDR)

Fold	***Existing Work(LWEDA-ACT)***			***Proposed Work(EASALDP)***		
	MCC	***Hamming Loss***	***FDR***	***MCC***	***Hamming Loss***	***FDR***
1	78.70	0.0712	0.1593	83.36	0.0579	0.1396
2	76.92	0.0822	0.1701	83.20	0.0549	0.1299
3	78.40	0.0704	0.1634	83.18	0.0563	0.1322
4	78.34	0.0701	0.1796	83.96	0.0476	0.1361
5	76.21	0.0662	0.1667	82.74	0.0590	0.1356

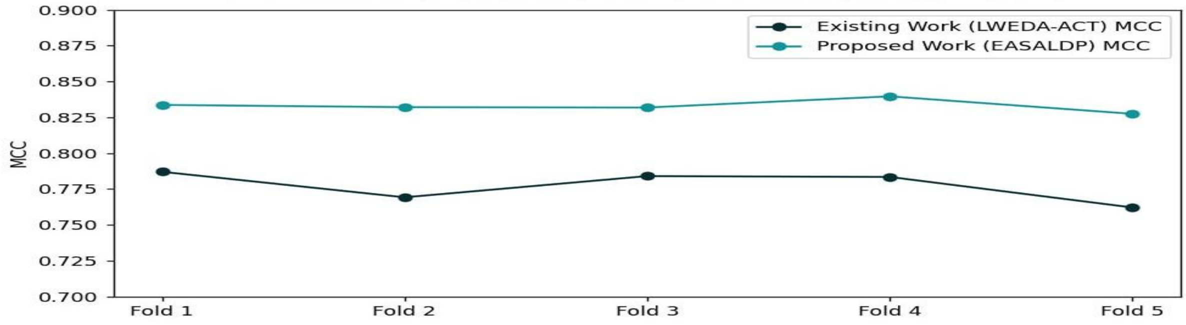

Figure 8. Comparison of Cross-Validation Performance Metrics (Matthews Correlation Coefficient)

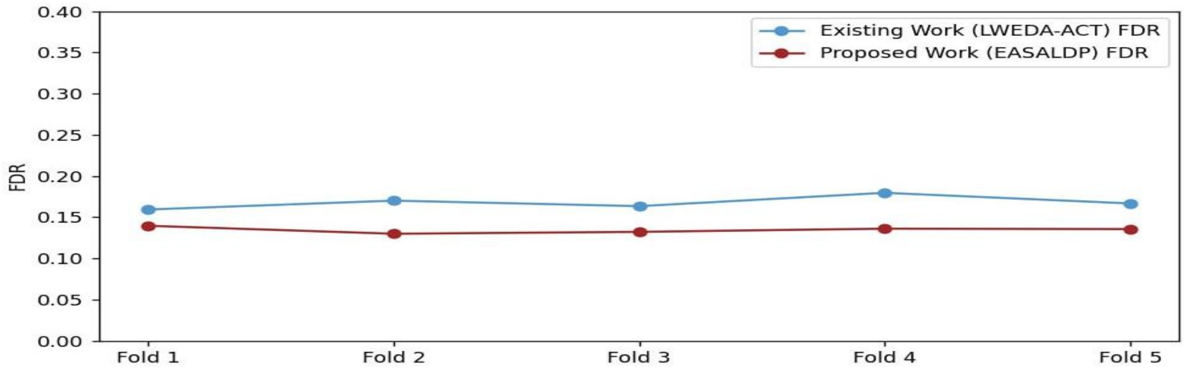

Figure 9. Comparison of Cross-Validation Performance Metrics (False Discovery Rate)

The above graphs and table (Table 2, Figure 8, Figure 9) indicate that our proposed method achieves higher MCC, lower FDR, and reduced Hamming Loss values. These results emphasize the effectiveness of the new technique and its capacity to deliver more accurate predictions.

V. CONCLUSION

We propose an approach that integrates dependency parsing and zero-shot prompting to effectively include opinion as a target label in aspect-based sentiment analysis. By leveraging dependency parsing, our method captures meaningful aspect–opinion relationships, while zero-shot prompting filters out irrelevant or misleading pairs, ensuring each aspect is accurately linked to its corresponding sentiment. Moreover, the adoption of a Llama-LoRA based model enables resource-efficient training, enhancing the scalability of our framework.

VI. REFERENCES

[1] Deling Huang, Ling Ren, Zanxiong Li,"Enhancing aspect-based sentiment analysis with linking words-guided emotional augmentation and hybrid learning", Neurocomputing,(2025),Volume 612,2025,128705,ISSN 0925-2312, https://doi.org/10.1016/j.neucom.2024.128705

[2] Haochen Zou, Yongli Wang,"Large language model augmented syntax-aware domain adaptation method for aspect-based sentiment analysis", Neurocomputing,(2025),Volume 625,2025,129472,ISSN 0925-2312, https://doi.org/10.1016/j.neucom.2025.129472.

[3] Stefan Lam, Yin Liu, Max Broers, Jasper van der Vos, Flavius Frasincar, David Boekestijn, Finnvan der Knaap,"Local interpretation of deep learning models for Aspect-Based Sentiment Analysis", Engineering Applications of Artificial Intelligence,(2025),Volume 143,2025,109947,ISSN 0952-1976, https://doi.org/10.1016/j.engappai.2024.109947

[4] Y. Chen, K. Zhang, Z. Liu, M. Bao, X. Liu and Q. Liu, "SISK: Integrating Aspect-oriented Semantic Information and Syntactic Knowledge for Aspect-based Sentiment Analysis," 2024 International Conference on Computational Linguistics and Natural Language Processing (CLNLP), Yinchuan, China, (2024), pp. 64-70, doi: 10.1109/CLNLP64123.2024.00020.

[5] Xu, G., Chen, Z. & Zhang, Z. "Aspect category sentiment analysis based on pre-trained BiLSTM and syntax-aware graph attention network". *Sci Rep* 15, 3333 (2025), https://doi.org/10.1038/s41598-025-86009-8

[6] Bo He, Ruoyu Zhao, Dali Tang,CABiLSTM-BERT: "Aspect-based sentiment analysis model based on deep implicit feature extraction", Knowledge-Based Systems,(2025),Volume 309,2025,112782,ISSN0950-7051, https://doi.org/10.1016/j.knosys.2024.112782

[7] Linan Zhu, Xiaolei Guo, Zhechao Zhu, Yifei Xu, Zehai Zhou, Xiangfan Chen, Xiangjie Kong,"Generative aspect-based sentiment analysis with a grid tagging matching auxiliary task", Computer Speech & Language,(2025),Volume 92,2025,101776,ISSN 0885-2308, https://doi.org/10.1016/j.csl.2025.101776.

[8] Y. Li, Z. Su, K. Chen and K. Jiang, "Aspect-based Sentiment Analysis via Knowledge Enhancement," 2024 International Joint Conference on Neural Networks (IJCNN), Yokohama, Japan, (2024), pp. 1-8, doi: 10.1109/IJCNN60899.2024.10650800.

[9] Y. Shen, Z. Chen, J. Di and Y. Meng, "Aspect Term Information Enhancement Network for Aspect-Based Sentiment Analysis," 2024 9th International Conference on Intelligent Computing and Signal Processing (ICSP), Xian, China, (2024), pp. 1198-1204, doi:10.1109/ICSP62122.2024.10743826.

[10] X. Wang, X. Li, Y. Yin and Y. Li, "Implicit aspect-based generative model for sentiment analysis based on prompt learning," 2024 5th International Conference on Big Data & Artificial Intelligence & Software Engineering (ICBASE), Wenzhou, China, (2024), pp. 94-97, doi: 10.1109/ICBASE63199.2024.10762313

[11] Chao Zhu and Qiang Ding. 2024. "Aspect-based sentiment analysis via dual residual networks with sentiment knowledge", J. Supercomput. 81, 1 (2025), https://doi.org/10.1007/s11227-024-06546-3

[12] M. Pontiki, D. Galanis, J. Pavlopoulos, H. Papageorgiou, I. Androutsopoulos,S. Manandhar, SemEval-2014 task 4: Aspect based sentiment analysis, in: P.Nakov, T. Zesch (Eds.), Proceedings of the 8th International Workshop on Semantic Evaluation, SemEVal 2014, Association for Computational Linguistics,Dublin, Ireland, 2014, pp. 27–35, http://dx.doi.org/10.3115/v1/S14-2004.

