Journal of Artificial Intelligence and Data Mining (JAIDM), Vol. 14, No. 1, 2026, 51-60.

Shahrood University of Technology

Journal of Artificial Intelligence and Data Mining (JAIDM)
Journal homepage: http://jad.shahroodut.ac.ir

Research paper

DeGF Network -ABSA: Hybrid Approach with DeBERTa and Gated Fusion

Rojiar Pir mohammadiani[1*], Faezeh Hashemnia[1], Elham Moradizadeh[1] and Soma Solaiman Zadeh[2]

[1] Computer Engineering Department, University of Sanandaj, Sanandaj, Iran.
[2]Cybersecurity Department, Tishk International University, Erbil, Iraq.

Article Info

Article History:
Received 12 January 2025
Revised 01 February 2025
Accepted 15 August 2025

DOI:10.22044/jadm.2025.15593.2673

Keywords:
Aspect-Based Sentiment Analysis (ABSA), Sentiment Classification, Gated Fusion Network.

**Corresponding author: R.pirmohamadiani@uok.ac.ir (R. Pir mohammadian).*

Abstract

Aspect-Based Sentiment Analysis (ABSA) is a vital tool for extracting fine-grained insights from user opinions; however, existing methods often struggle to effectively balance contextual and aspect-specific information or generalize across diverse domains. To address these limitations, we propose DeGF-ABSA (DeBERTa-Gated Fusion for Aspect-Based Sentiment Analysis), a novel architecture that leverages the DeBERTa model's disentangled attention mechanism, which excels at capturing nuanced relationships in text, combined with a gated fusion layer. This layer dynamically weights the contributions of global context features, derived from the DeBERTa [CLS] token (a special token representing the entire input sequence), and aspect-specific features, computed as the average representation of tokens associated with the target aspect. This approach enables precise sentiment classification in complex sentences. Experiments on the SemEval 2014 datasets demonstrate state-of-the-art performance, achieving 86.68% accuracy (84.50% F1-score) for the Laptop domain and 91.43% accuracy (86.83% F1-score) for the Restaurant domain. Given the variability of sentiment expressions across domains—such as "delicious" in food reviews versus "fast performance" in electronics reviews—evaluating DeGF-ABSA on additional domains would further validate its generalization capabilities and broaden its applicability for real-world ABSA tasks.

1. Introduction

Aspect-Based Sentiment Analysis (ABSA) has emerged as a pivotal technique within Natural Language Processing (NLP) for extracting granular opinions from textual data. Unlike traditional sentiment analysis, which typically assigns a single polarity to an entire text, ABSA focuses on identifying the sentiment expressed towards specific aspects or features of an entity, providing a more nuanced understanding of customer feedback and reviews 0. This fine-grained analysis is crucial in scenarios where a single piece of text can contain varying sentiments towards different aspects, such as a restaurant review praising the food but criticizing the service. The field of ABSA encompasses several subtasks, most notably Aspect Term Extraction (ATE), which aims to identify the mentions of aspects in a text, and Aspect Sentiment Classification (ASC), which focuses on determining the sentiment polarity (e.g., positive, negative, or neutral) associated with these identified aspects0. While ABSA represents a broad area of research, this paper specifically concentrates on Aspect-Based Sentiment Classification (ABSC), addressing the task of predicting the sentiment polarity for given aspect terms within a sentence or document. The increasing demand for understanding detailed consumer feedback has driven the evolution of ABSA methodologies from early lexicon-based approaches to sophisticated machine learning and

deep learning techniques. This progression underscores a growing recognition of the need for models capable of capturing sentiment towards specific attributes, as a singular sentiment score for an entire text often lacks the granularity required for actionable insights.
To address these challenges, we introduce the DeFG model which is designed to select creative, high-level ideas from both the overall text and its specific aspects, and then combine them effectively. In our model, global context and aspect features are extracted from the entire text using DeBERTa embeddings. These two sources of information are fused through a Gated Fusion mechanism that dynamically adjusts the contribution of each component.

2. Related Works

ABSA has evolved dramatically since the introduction of deep learning techniques. In this work, we classify the literature concerning five predominant themes describing the development of ABSA methods: attention mechanisms, neural networks, transformers, multi-task learning, and knowledge integration [28].

2.1 Attention Mechanisms in ABSA

Attention mechanisms have been foundational in enhancing sentiment analysis by allowing models to focus selectively on words relevant to the aspect under consideration. Initial works, such as those by Tang et al. 0, introduced memory networks to emphasize key words, outperforming traditional LSTM-based models by making aspect-specific predictions. Expanding on this, Chen et al. 0 developed a Recurrent Attention Network on Memory, which employs a multi-attention mechanism to capture sentiment features even over long distances, thus reducing the need for manual feature engineering. Song et al. 0 proposed the Attentional Encoder Network (AEN), which replaces recurrent layers with attention-based encoders and integrates BERT embeddings, improving sentiment classification accuracy.
Advancing these methods, Li et al. 0 introduced a hierarchical position-aware network (HAPN) that incorporates position embeddings to model the position-sentiment relationship within sentences. Fan et al. 0 proposed a multi-grained attention network (MGAN), capturing both fine-grained and coarse-grained attention patterns, while Tang et al. 0 developed progressive self-supervised attention learning to automatically extract attention supervision signals from the training data, enabling more accurate sentiment prediction for aspects with minimal labeled data.

2.2 Neural Network-Based Models

Neural network architectures have played a central role in ABSA. Early contributions included LSTM-based models tailored for sentiment analysis. For instance, Do 0 introduced bitmask bidirectional LSTM networks to maintain focus on specific aspects, while Li et al. 0 designed transformation networks combining CNN and RNN layers to address the limitations of simple attention mechanisms. Target-dependent LSTM models, as proposed by Tang et al. 0, integrated aspect information directly, resulting in significant performance improvements over traditional LSTM-based approaches.
Beyond LSTM models, memory networks have been leveraged for ABSA. Majumder et al. 0 presented Inter-Aspect Relation Modeling with Memory Networks (IARM), which enhances sentiment prediction by capturing neighboring aspects' influences. Graph-based approaches, including Zhao et al.'s 0 graph convolutional network, introduced relational sentiment modeling across multiple aspects. Bai et al. 0 used graph attention neural networks to incorporate typed syntactic dependencies, improving performance for aspect-targeted sentiment analysis.

2.3 Transformer-Based Approaches

The architecture of transformers such as BERT has recently led to important advances in ABSA[25]. Early work by Xu et al. 0 extended BERT for review comprehension to post-training BERT, which could capture domain-specific sentiment patterns. Dai et al. 0 also pointed out the syntactic power of RoBERTa and showed that syntax-aware embeddings can improve ABSA, although syntactic features are not included. Rietzler et al. 0 also investigated domain adaptation in fine-tuning BERT on aspect-target sentiment classification, showing the contribution of domain knowledge in improving performance.

Advanced Transformer-based models continue to refine ABSA. Cheng et al. 0 proposed a variational semi-supervised Transformer for aspect-term sentiment classification using semi-supervised learning to use a small number of annotated data. Karimi et al. 0 proposed a model further augmenting BERT by parallel-and-hierarchical aggregation that can make more fine-tuned aspect-level sentiment distinctions.

2.4 Multi-Task and Joint Learning

Multi-task learning has made progress in ABSA possible by tackling several related tasks simultaneously, which, in turn, helps to improve the model's generalization [26]. He et al. 0

proposed an interactive multi-task learning network, with extraction and sentiment classification used jointly in an end-to-end model. This solution significantly minimizes error leakage among tasks. Yang et al. 0 formulated a model based on the Chinese ABSA task, which employs multi-task learning to achieve concurrent context term extraction and polarity classification.

Transfer learning has also been used for ABSA, particularly in situations of limited resources[24]. Li et al. 0 introduced a coarse-to-fine task transfer method in which models are first trained on more general sentiment classification problems. Then, they are fine-tuned on more specific aspect classification problems. This method is adapted to solve the data scarcity problem and improve the model's performance.

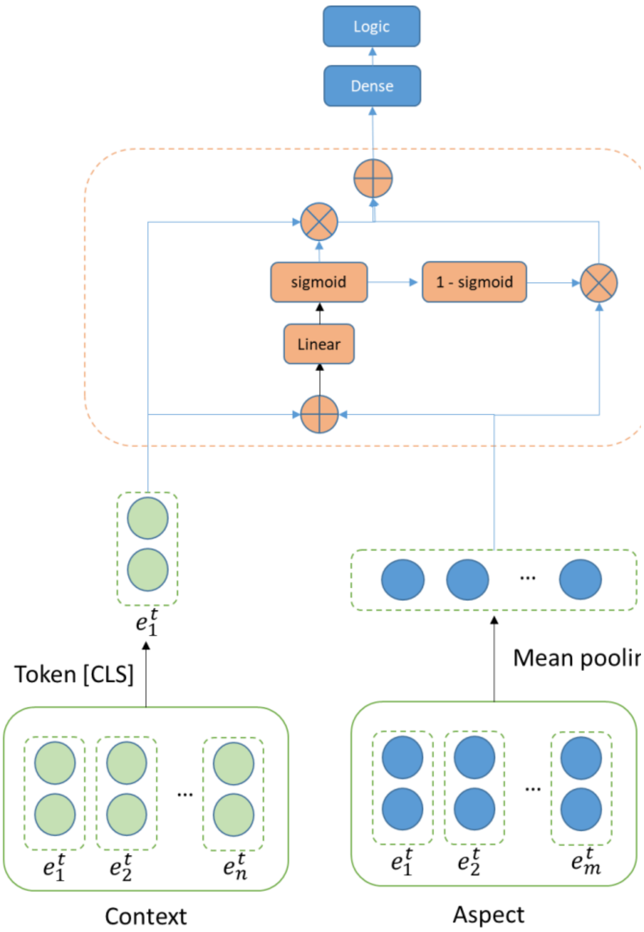

Figure 1. Architecture of the proposed model.

2.5 Incorporation of External Knowledge

More recently, research has also introduced external knowledge, including aspect-specific and syntactic information, into neural models to enhance sentiment prediction further. Xing and Tsang 0 proposed a knowledge-aware gated recurrent memory network that encodes aspect knowledge and syntactic information and that, as a result, provides more contextually appropriate sentiment predictions. In this method, knowledge-based components are directly embedded into the model structure, which leads to a richer interpretation of sentiment on the sentiment-rich fine-grained sentence level Gated Fusion unit to allow for adaptive,aspect-sensitive sentiment representation.

3. Model Architecture

The model architecture of DeGF is specifically handcrafted to cope with issues of ABSA. It combines global and aspect-wise embeddings by the powerful pre-trained DeBERTa model which can deliver deep contextualized word representations. Using a Gated Fusion mechanism, DeGF achieves an intelligently balanced relevance between the different embeddings, enabling it to produce truly relevant and sensitive sentiment predictions to specific content within the texts. This structure enables the model to capture nuanced sentiment patterns and deliver robust performance on complex ABSA tasks.

3.1 DeBERTa for Contextualized Embedding Extraction

At the core of DeGF, DeBERTa serves as a pre-trained feature extractor to capture both global and aspect-specific information:

Global Context Features: Extracted from the whole text input.

Aspect Features: Extracted from the text's specific aspect term or phrase.

For an input text $T=[t_1.t_2......t_n]$ and aspect $A=[a_1.a_2......a_m]$, DeBERTa generates contextual embeddings:

$$\mathbf{H}_{global} = \mathrm{DeBerta}(T).\mathbf{H}_{aspect} = \mathrm{DeBerta}(A) \quad (1)$$

where $\mathbf{H}_{global} \in \mathbb{R}^{n\times d}$ and $\mathbf{H}_{aspect} \in \mathbb{R}^{m\times d}$, with d being the embedding dimension.

The symbol ℝ refers to the set of real numbers.

3.2 Extract CLS Token From Global Context Features

In DeBERTa's architecture, the [CLS] token (Classification Token) is specifically designed to summarize the global meaning of a sentence and extract its general features. Positioned at the start of the input sequence, this token learns during pre-training to aggregate information from all tokens via multi-head self-attention layers, creating a unified representation of the text. The choice of the

[CLS] token for capturing general textual features (e.g., overall sentiment or key contextual cues) is optimal for the following reasons:

- Unlike standard BERT, which computes coupled attention between token content and positional embeddings, DeBERTa decouples these components (token content and relative positions). This allows the [CLS] token to capture global semantic interactions while filtering out positional noise.
- In aspect-based sentiment analysis, the [CLS] token's global features act as a reference framework, providing the context needed to interpret individual aspect sentiments.

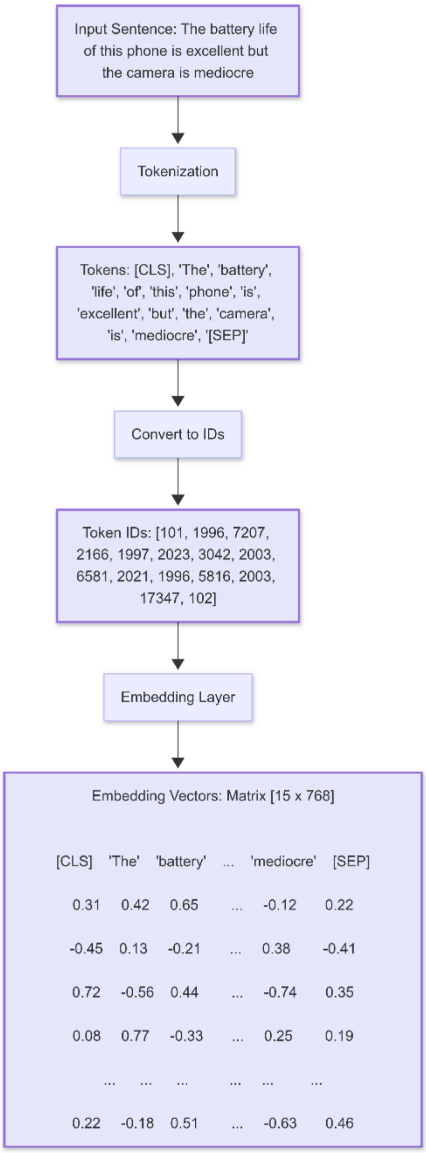

Figure 2. Embedding Extraction from DeBerta.

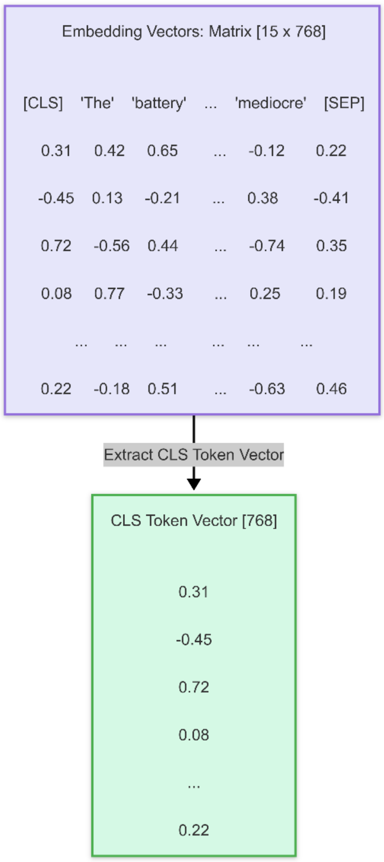

Figure 3. Extract CLS Token From Global Context Features.

3.3 Mean Pooling for Aspect Feature Aggregation

To represent aspect features as a single vector, we apply mean pooling over the tokens in $\mathbf{H}_{aspect}$:

$$\mathbf{h}_{aspect} = \frac{1}{m} \sum_{i=1}^{m} \mathbf{H}_{aspect}[i] \tag{2}$$

This operation yields a fixed-size vector $\mathbf{h}_{aspect} \in \mathbb{R}^{d}$, summarizing the information in the aspect phrase.

3.4 Gated Fusion Mechanism

Fusion layers in neural networks serve the fundamental purpose of integrating information from multiple sources or different representational spaces to create a unified representation.In the DeGF-ABSA model, the gated fusion layer plays a crucial role in integrating the general contextual information derived from the entire input text with the specific information related to the target aspect.The model extracts the CLS token vector from the embedded representation of the overall text using the pre-trained DeBERTa, which serves as the general contextual feature. Simultaneously, for each aspect, the model extracts the embedded vectors corresponding to the tokens of that aspect and calculates their average to obtain a specific aspect feature vector.

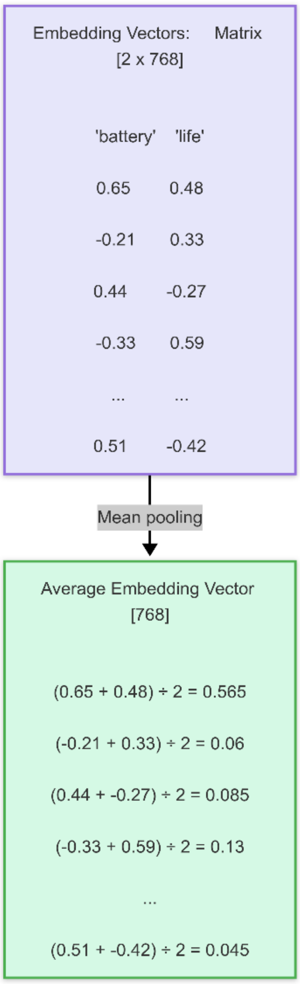

Figure 4. Embedding Extraction from DeBerta.

The gated fusion layer then takes these two feature vectors (the CLS token vector and the averaged aspect vector) as input. Below, we formalize these steps mathematically:

1. Concatenation of the global contextual representation $\mathbf{h}_{global}$ and the pooled aspect representation $\mathbf{h}_{aspect}$:

$$\mathbf{f}_{concat}=[\mathbf{h}_{global};\mathbf{h}_{aspect}] \quad (3)$$

2. Gating Vector calculation using a sigmoid function, which learns the importance of global and aspect features:

$$\mathbf{g}=\sigma(\mathbf{w}_g.\mathbf{f}_{concat}+\mathbf{b}_g) \quad (4)$$

where $\mathbf{w}_g \in \mathbb{R}^{d\times 2d}$ and $\mathbf{b}_g \in \mathbb{R}^{d}$ are learnable parameters.

3. Fused Representation combines $\mathbf{h}_{global}$ and $\mathbf{h}_{aspect}$ weighted by the gate :

$$\mathbf{h}_{fused}=\mathbf{g}\odot\mathbf{h}_{global}+(1-\mathbf{g})\odot\mathbf{h}_{aspect} \quad (5)$$

Here, $\odot$ denotes element-wise multiplication. This operation ensures that the fused representation $\mathbf{h}_{fused}$ dynamically adapts based on the relevance of the aspect and global context information.This approach offers several potential benefits for aspect-based sentiment analysis. Firstly, by allowing the model to weight the importance of both the overall context and the specific aspect, the gated fusion layer can potentially enhance the accuracy of sentiment prediction. The model can learn to prioritize the general sentiment expressed in the sentence when the aspect sentiment aligns with it, or to focus more on the aspect-specific information when the sentiment towards the aspect differs from the overall sentiment. Secondly, the fusion mechanism facilitates a more nuanced understanding of the sentiment expressed towards the particular aspect by explicitly considering it within the broader context of the sentence. This can be particularly useful in handling cases where the sentiment towards an aspect is subtle or context-dependent. For instance, in the sentence "The food was great, but the service was slow," the gated fusion mechanism can help the model correctly identify the positive sentiment towards "food" and the negative sentiment towards "service" by considering both the overall sentence structure and the specific mentions of these aspects..The Gated Fusion mechanism is a critical component of DeGF, allowing the model to adapt to global and aspect-specific information.

3.5 Final Classification Layer

The fused representation $\mathbf{h}_{fuused}$ is passed through a dropout layer to prevent overfitting:

$$\mathbf{h}_{drop}=\text{Dropout}(\mathbf{h}_{fused}) \quad (6)$$

A linear layer then maps $\mathbf{h}_{drop}$ to the output logits, representing sentiment classes:

$$\mathbf{y}_{logits}=\mathbf{w}_{out}\,.\,\mathbf{h}_{drop}+\mathbf{b}_{out} \quad (7)$$

Where $\mathbf{w}_{out} \in \mathbb{R}^{output\_dim \times d}$ and $\mathbf{b}_{out} \in \mathbb{R}^{output\_dim}$ are parameters of the classification layer.

3.4 Advantages of the DeGF Design

The architecture combines the following strengths:

- Context-Specific Sentiment Adaptability: The Gated Fusion mechanism enables a nuanced balance between global context and aspect information, essential for accurate sentiment classification at an aspect level.
- Efficient Feature Representation: Mean pooling and gating mechanisms minimize computational costs, making it practical for large-scale ABSA tasks.
- Enhanced Generalization: Dropout regularization and the gating mechanism improve the model's robustness by mitigating overfitting and dynamically prioritizing relevant information.

4. Experiments and Results

4.1 Data Collection

The SemEval-2014 dataset was created as part of the **International Workshop on Semantic Evaluation** (SemEval), specifically for **Task 4: Aspect-Based Sentiment Analysis**. The dataset focuses on reviews with sentiment annotations tied to specific aspects within each domain. Here is a breakdown of its main features:

Table 1. Overview of Datasets for ABSA.

Domain	Total Reviews	Total Aspects	Sentiment Labels
Restaurant	3845	4473	Positive, Negative, Neutral
Laptops	3045	3830	Positive, Negative, Neutral

4.1.1 Structure of the Dataset

Each review in the Dataset contains the following information:

- **Sentence ID**: Unique identifier for each sentence.
- **Aspect Term**: A specific term or phrase representing the aspect (e.g., "service" in restaurant reviews).
- **Aspect Category**: High-level categories representing the aspect type, like *Food*, *Service*, or *Price* for restaurants.
- **Sentiment Label**: The sentiment label assigned to each aspect. Labels include:
  - **Positive**: Indicates positive sentiment toward the aspect.
  - **Negative**: Indicates negative sentiment toward the aspect.
  - **Neutral**: A neutral sentiment with no strong opinion expressed.

Table 2. Example of Restaurant dataset.

Sentence	Aspect Term	Aspect Category	Sentiment
The pizza was excellent, but the service was slow.	Pizza	Food	Ppositive
The pizza was excellent, but the service was slow.	Service	Service	Negative
Great value for money, but the ambiance was lacking.	Value	Price	Positive
Great value for money, but the ambiance was lacking.	Ambiance	Ambiene	Negative

4.2 Results

To ensure rigorous evaluation and equitable comparison with existing approaches, we implemented our experiments using PyABSA (Python for Aspect-Based Sentiment Analysis), a widely recognized toolkit for ABSA research. This framework provides standardized benchmarks, preconfigured state-of-the-art baselines (e.g., LSTM, BERT, and RoBERTa variants), and unified evaluation protocols—critical for ensuring reproducibility and reducing implementation bias in cross-method comparisons. We run our model on 3060TI GPU. By aligning with this community-adopted platform, our results directly contextualize DeGF-ABSA's advancements against contemporary methods while maintaining consistency with established ABSA evaluation practices.

Our proposed model achieved the following results:

- Restaurant Domain Accuracy: 91.43%
- Laptop Domain Accuracy: 86.68%

- Mean Accuracy: 89.55% (calculated as the average of both domains)

Table 3. Accuracy Comparison for various ABSA.

Model	Mean Accuracy	Restaurant Accuracy	Laptop Accuracy	Year
LSA+DeBERTa-V3-Large	88.27	90.33	86.21	2021
LCF-PC	86.24	90.18	82.29	2019
ABSA-DeBERTa	86.11	89.46	82.76	2021
DPL-BERT	85.75	89.54	81.96	2022
RoBERTa+MLP	85.58	87.37	83.78	2021
KaGRMN-DSG	84.61	87.35	81.87	2021
YORO	84.48	87.14	81.82	2024
BERT-ADA	84.06	87.89	80.23	2019
RGAT+	83.92	86.59	81.25	2020
PH-SUM	82.96	86.37	79.55	2020
MaskedABSA	86.95	87.65	86.24	2024
BAT	82.69	86.03	79.35	2020
InstructABSA	81.5	82.44	80.56	2023
DeGF	**89.55**	**91.43**	**86.68**	2024

Based on mean accuracy across both domains, the performance hierarchy is as follows:

1. Proposed Model (89.55%)
2. InstructABSA (89.065%)
3. LSA+DeBERTa-V3-Large (88.27%)
4. LCF-ATEPC (86.24%)
5. Other models (< 86%)

Table 4. Classification report of Laptop.

Sentiment	Precision	Recall	F1 Score	Support
Negative	0.67	0.91	0.77	128
Nature	0.81	0.53	0.64	169
Positive	0.9	0.93	0.92	341

Table 5. Classification report of Resturant.

Sentiment	Precision	Recall	F1 Score	Support
Negative	0.85	0.86	0.86	196
Nature	0.78	0.65	0.71	196
Positive	0.92	0.96	0.94	728

The results from The two datasets highlight the effectiveness of the sentiment analysis model while revealing some domain-specific challenges. Both datasets demonstrate strong classification performance, with ROC-AUC values consistently above 0.89 for all classes. However, subtle differences exist in the model's behavior across the datasets.

The **Laptop dataset** achieved AUC values of 0.95 (negative), 0.89 (neutral), and 0.96 (positive). In contrast, the **Restaurant dataset** showed slightly higher performance, with AUC values of 0.98 (negative), 0.91 (neutral), and 0.97 (positive). This suggests that the model performs slightly better on restaurant reviews, possibly due to the richer and more explicit sentiment expressions in the restaurant domain than laptop reviews' technical and potentially more nuanced nature.

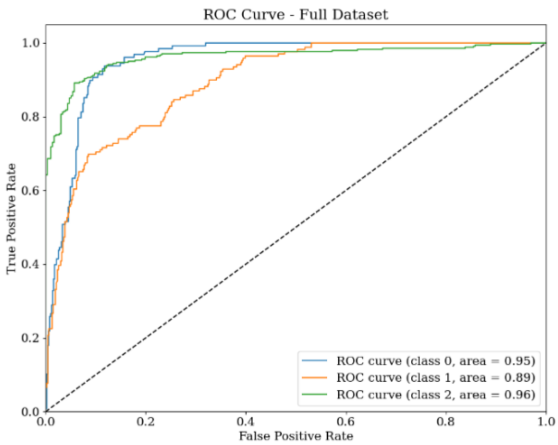

Figure 5. ROC Curve for Laptop dataset.

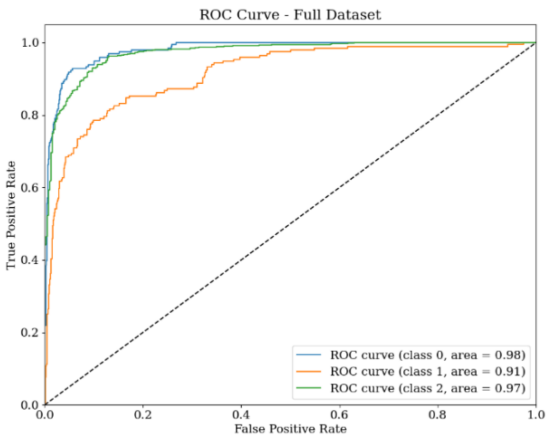

Figure 6. ROC Curve for Restaurant dataset.

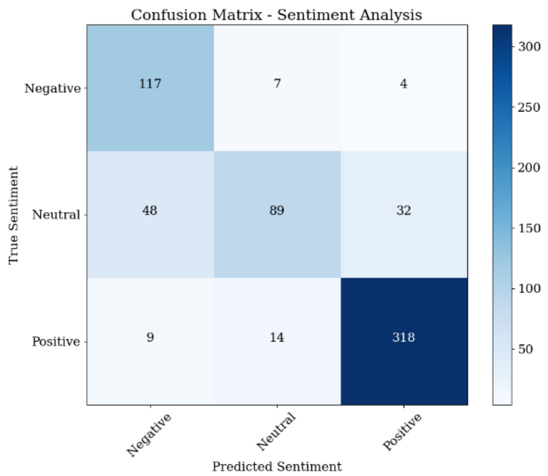

Figure 7. Confusion Matrix for Laptop dataset.

A comparison of the confusion matrices further highlights these differences. Both datasets demonstrate strong performance in classifying positive sentiments, with the restaurant dataset correctly identifying 728 positive samples (98.2%) compared to 701 (97.1%) for the laptop dataset. However, the neutral sentiment remains the most challenging class in both domains. For laptops, 53 neutral samples were misclassified as positive, while the restaurant dataset showed the same pattern, with 53 misclassified neutral samples. This indicates that neutral reviews often overlap semantically with positive ones in both domains, but the impact is slightly less pronounced in the restaurant dataset.

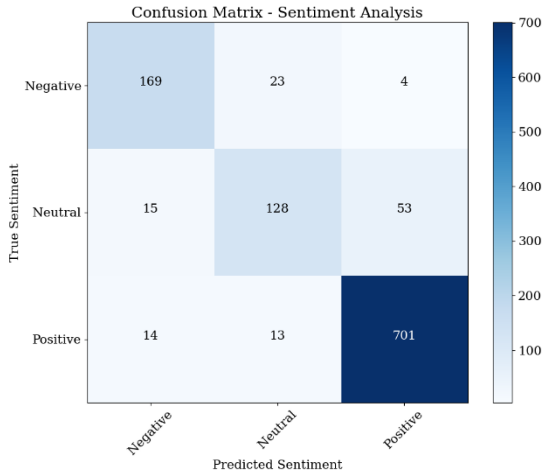

Figure 8. Confusion Matrix for Restaurant dataset.

5. Conclusion

The proposed model also has multiple advantages in terms of domains, demonstrating its negotiability and ability to grasp sentiment nuances, as is illustrated by the following:

High Domain-Specific Accuracy: DeGF produced the highest domain accuracy in the restaurant dataset (91.43%), underscoring its remarkable ability to detect delicate sentiment in this domain.

Consistent Cross-Domain Performance: The model obtained an acceptable accuracy of 86.68% on the laptop domain, reflecting its capability and effectiveness in different domains.

Overall Mean Accuracy Advantage: With an average mean accuracy of 89.06%, DeGF is slightly better than InstructABSA's average accuracy, confirming its promise as a next-generation ABSA model.

The model exhibits performance disparities across sentiment classes, particularly in the Restaurant domain, where the Neutral class achieves significantly lower recall (0.53) compared to Positive (0.93). This stems from two key challenges:

- Class Imbalance: Dominant Positive-class samples (e.g., 341 vs. 128 Negative in Restaurants) bias the model toward majority classes, reducing sensitivity to under-represented sentiments.

- Domain-Specific Nuances: Neutral sentiment in Restaurant reviews (e.g., “The pasta was acceptable”) often relies on subtle contextual cues, contrasting with explicit sentiment markers in Laptop domains (e.g., “The battery is mediocre”).

In conclusion, our DeGF model competes with and surpasses many leading ABSA models across diverse domains. These results indicate that DeGF is well suited to generalize across domains such as complex sentiment morphology. Future research could see the optimization of this model to other ABSA tasks or expanding its generalizability to categories comprising more specific aspects.

References

[1] D. Jayakody, A. V. A. Malkith, K. Isuranda, V. Thenuwara, N. de Silva, S. R. Ponnamperuma, G. G. N. Sandamali, and K. L. K. Sudheera, "Instruct-DeBERTa: A Hybrid Approach for Aspect-Based Sentiment Analysis on Textual Reviews," arXiv preprint arXiv:2408.13202, 2024.

[2] G. M. Shafiq, T. Hamza, M. F. Alrahmawy and R. El-Deeb, "Enhancing Arabic Aspect-Based Sentiment Analysis Using End-to-End Model," in IEEE Access, vol. 11, pp. 142062-142076, 2023.

[3] M. Pontiki, D. Galanis, J. Pavlopoulos, H. Papageorgiou, I. Androutsopoulos, and S. Manandhar, "SemEval-2014 Task 4: Aspect-Based Sentiment Analysis," in Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval, 2014), 2014.

[4] P. Chen, Z. Sun, L. Bing, and W. Yang, "Recurrent Attention Network on Memory for Aspect Sentiment Analysis," in Proc. EMNLP, 2017.

[5] Y. Song, J. Wang, T. Jiang, Z. Liu, and Y. Rao, "Attentional Encoder Network for Targeted Sentiment Classification," arXiv preprint arXiv:1902.09314, 2019.

[6] L. Li, Y. Liu, and A. Zhou, "Hierarchical Attention Based Position-Aware Network for Aspect-Level Sentiment Analysis," in Proc. CONLL, 2018.

[7] F. Fan, Y. Feng, and D. Zhao, "Multi-grained Attention Network for Aspect-Level Sentiment Classification," in Proc. EMNLP, 2018.

[8] J. Tang, Z. Lu, J. Su, Y. Ge, L. Song, L. Sun, and J. Luo, "Progressive Self-Supervised Attention Learning for Aspect-Level Sentiment Analysis," in Proc. ACL, 2019.

[9] B. T. Do, "Aspect-Based Sentiment Analysis Using Bitmask Bidirectional Long Short Term Memory Networks," in Proc. FLAIRS-31, 2018.

[10] X. Li, L. Bing, W. Lam, and B. Shi, "Transformation Networks for Target-Oriented Sentiment Classification," in Proc. ACL, 2018.

[11] D. Tang, B. Qin, X. Feng, and T. Liu, "Effective LSTMs for Target-Dependent Sentiment Classification," in Proc. COLING, 2016.

[12] N. Majumder, S. Poria, A. Gelbukh, et al., "IARM: Inter-Aspect Relation Modeling with Memory Networks in Aspect-Based Sentiment Analysis," in Proc. EMNLP, 2018.

[13] P. Zhao, L. Hou, and O. Wu, "Modeling Sentiment Dependencies with Graph Convolutional Networks for Aspect-level Sentiment Classification," arXiv preprint, 2019.

[14] X. Bai, P. Liu, and Y. Zhang, "Investigating Typed Syntactic Dependencies for Targeted Sentiment Classification Using Graph Attention Neural Network," arXiv preprint, 2020.

[15] H. Xu, B. Liu, L. Shu, and P. S. Yu, "BERT Post-Training for Review Reading Comprehension and Aspect-based Sentiment Analysis," in Proc. NAACL, 2019.

[16] J. Dai, H. Yan, T. Sun, P. Liu, and X. Qiu, "Does Syntax Matter? A Strong Baseline for Aspect-based Sentiment Analysis with RoBERTa," in Proc. NAACL, 2021.

[17] A. Rietzler, S. Stabinger, P. Opitz, and S. Engl, "Adapt or Get Left Behind: Domain Adaptation through BERT Language Model Finetuning for Aspect-Target Sentiment Classification," in Proc. LREC, 2020.

[18] X. Cheng, W. Xu, T. Wang, and W. Chu, "Variational Semi-supervised Aspect-term Sentiment Analysis via Transformer," in Proc. CONLL, 2019.

[19] A. Karimi, L. Rossi, and A. Prati, "Improving BERT Performance for Aspect-Based Sentiment Analysis," in Proc. ICNLSP, 2021.

[20] R. He, W. S. Lee, H. T. Ng, and D. Dahlmeier, "An Interactive Multi-Task Learning Network for End-to-End Aspect-Based Sentiment Analysis," in Proc. ACL, 2019.

[21] H. Yang, B. Zeng, J. Yang, Y. Song, and R. Xu, "A Multi-task Learning Model for Chinese-oriented Aspect Polarity Classification and Aspect Term Extraction," arXiv preprint, 2019.

[22] Z. Li, Y. Wei, Y. Zhang, X. Zhang, X. Li, and Q. Yang, "Exploiting Coarse-to-Fine Task Transfer for Aspect-level Sentiment Classification," in Proc. AAAI, 2019.

[23] B. Xing and I. W. Tsang, "Understand me if you refer to Aspect Knowledge: Knowledge-aware Gated Recurrent Memory Network," arXiv preprint arXiv:2108.02105, 2021.

[24] S. Dehghan, R. Pir Mohammadiani, and S. Mohammadi,"The credibility assessment of Twitter/X users based on organization objectives by heterogeneous resources in big data life cycle", Computers in Human Behavior, Jan. 2025.

[25] S. Bastami, R. Pirmohamadiani, M. B. Dowlatshahi, and A. Abdollahpouri,"Enhanced high-dimensional data classification by combining fuzzy learning integration and graph transformers," Iranian Journal of Fuzzy Systems, vol. 22, no. 2, pp. 129-146, Apr. 2025.

[26] R. P. Mohammadiani and S. Mohammadi,"Criteria for evaluating influence value of social media users: A framework based on social media mining,"Iranian Communication and Information Technology, vol. 11, no. 3940, pp. 109-125, Jan. 2019.

[27] S. Dehghan, S. Mohammadi, and R. P. Mohammadiani,"The main components of evaluating the credibility of users according to organizational goals in the life cycle of big data,"Journal of Information and Communication Technology, vol. 55, no. 15, pp. 141-164, 2023.

[28] R. Pir Mohammadiani and Z. Malik,"The effects of interactivity of electronic word of mouth systems on value creation practices on social media,"in Proceedings of the AMA Winter Academic Conference, American Marketing Association, 2019.

شبکه DeGF - ABSA : رویکرد ترکیبی با مدل DeBERTa و ادغام دروازه‌ای

روجیار پیر محمدیانی[*،۱]، فائزه هاشم نیا[۱]، الهام مرادی زاده[۱] و سوما سلیمان زاده[۲]

[۱] دانشکده مهندسی کامپیوتر، دانشگاه سنندج، سنندج، سنندج، ایران.

[۲] دانشکده امنیت سایبری، دانشگاه بین‌المللی تیشک، اربیل، عراق.

ارسال ۲۰۲۵/۰۱/۱۲؛ بازنگری ۲۰۲۵/۰۲/۰۱؛ پذیرش ۲۰۲۵/۰۸/۱۵

چکیده:

تحلیل احساس مبتنی بر جنبه (Aspect-Based Sentiment Analysis یا ABSA) ابزار حیاتی برای استخراج بینش‌های دقیق از نظرات کاربران است؛ اما روش‌های موجود غالباً در متعادل‌سازی مؤثر اطلاعات متنی کلی و اطلاعات خاصِ جنبه‌ای یا در تعمیم‌پذیری در حوزه‌های گوناگون با مشکل مواجه هستند. برای رفع این محدودیت‌ها، ما معماری جدیدی به نام DeGF-ABSA (تحلیل احساس مبتنی بر جنبه با استفاده از DeBERTa و ادغام دروازه‌ای) پیشنهاد می‌دهیم. این معماری از مکانیزم توجه گسسته (Disentangled Attention) مدل DeBERTa بهره می‌برد که در درک روابط ظریف در متن بسیار مؤثر است و آن را با یک لایه ادغام دروازه‌ای ترکیب می‌کند. این لایه به‌صورت پویا میزان تأثیر ویژگی‌های زمینه‌ای کلی (که از توکن ویژه [CLS] مدل DeBERTa به‌دست می‌آیند و نماینده کل جمله هستند) و ویژگی‌های خاصِ جنبه (که به‌صورت میانگین برداری از توکن‌های مرتبط با جنبه هدف محاسبه می‌شوند) را تعیین می‌کند. این رویکرد امکان طبقه‌بندی دقیق احساسات در جملات پیچیده را فراهم می‌سازد. آزمایش‌ها روی داده‌های مجموعه SemEval 2014 نشان‌دهنده عملکردی در سطح بهترین روش‌های موجود است، به‌طوری‌که برای حوزه لپ‌تاپ دقت ۸۶٫۶۸٪ (امتیاز F1 برابر ۸۴٫۵۰٪) و برای حوزه رستوران دقت ۹۱/۴۳ % (امتیاز F1 برابر ۸۶/۸۳ %) به‌دست آمده است. با توجه به تفاوت بیان احساسات در حوزه‌های مختلف — مانند خوشمزه" در نقدهای غذایی در مقابل "عملکرد سریع" در نقدهای الکترونیکی — ارزیابی مدل DeGF-ABSA در حوزه‌های بیشتر می‌تواند توانایی تعمیم آن را بهتر اثبات کرده و کاربردپذیری آن را در وظایف واقعی ABSA گسترش دهد.

کلمات کلیدی: تحلیل احساس مبتنی بر جنبه (ABSA)، طبقه‌بندی احساس، شبکه ادغام دروازه‌ای (Gated Fusion).