AWA-GCN: Enhancing Chinese sentiment analysis with a novel GCN model for Triplet and Quadruplet Extraction at SIGHAN 2024 dimABSA Task

Haocheng Xi[1,4], Yutong Wang[1,5], Xinyu Wang[2,6], Zhitao Li[3,7], Bochun Liu[2,8], Yihan Wang[2,9], Chen Ni[1,10]

[1] School of Information Science and Engineering, Yunnan University, Yunnan 650500, China

[2] National Pilot School of software, Yunnan University, Yunnan 650500, China

[3] School of Computer Science and Engineering, Tianjin University of Technology, Tianjin 300384, China

[4] shaunspike813jelly@gmail.com

[5] eggw81616@gmail.com

[6] wangxywy@gmail.com

[7] lnwind598@gmail.com

[8] furfurliang@gmail.com

[9] thomasungaro79@gmail.com

[10] nichen069@gmail.com

Abstract. This study presents the AWA-GCN (Attention-Weighted Affine with Graph Convolutional Networks) model, a new method designed exclusively for the SIGHAN 2024 Chinese Dimensional Aspect-Based Sentiment Analysis (dimABSA) Task. This test requires participants to identify sentiment triples and sentiment quadruples from textual data. The data comprises aspects, sentiment views, categories, and intensities across valence-arousal dimensions. Our model differs from existing models by utilizing a complex multi-layer attention mechanism within a GCN architecture, instead of relying on pipeline techniques or Grid Tagging Schemes (GTS). This design successfully captures the intricate interconnections among many sentiment components. The AWA-GCN model is very innovative in its ability to effectively handle quadruple extraction, which is a complex problem that has not been handled by traditional methods. Our approach enhances the understanding of aspect-sentiment interactions by including sophisticated techniques like word-level attention and semantic graph representations. The AWA-GCN model outperforms previous baselines in precision, recall, and F1 score on the dimABSA dataset, as demonstrated by empirical assessments. Additionally, the model exhibits significant enhancements in capturing the dimensional features of sentiment expressions. The results validate the model's exceptional ability to handle the complex nuances necessary for successful sentiment analysis in simple Chinese.

Keywords: Graph Convolutional Network(GCN), Chinese Sentiment Analysis, Sentiment Triple Extraction, Sentiment Quadruple Extraction, Valence-Arousal Dimensions.

1. Introduction

In the field of Natural Language Processing (NLP), Aspect-Based Sentiment Analysis (ABSA) is essential for dissecting the nuances of sentiment by pinpointing specific aspects of entities, assessing the emotions they evoke, and measuring the intensity of these emotions (Pontiki et al.[1]). In aspect

sentiment analysis, aspect sentiment triple extraction (ASTE) and aspect sentiment quadruple extraction (ASQE) are new variants of aspect-based sentiment analysis. Most current aspect sentiment extraction tasks only include simple positive, negative or neutral sentiment predictions. The ACL SIGHAN 2024 dimABSA task we will solve represents a major advancement in this field. Unlike the traditional simple sentiment polarity prediction, this task includes a more continuous valence-arousal numerical prediction. It also includes predicting aspect and opinion items as the fourth dimension of the aspect class. What is more worth mentioning is that this task uses the Chinese language, which is more difficult to segment, so we need to have a more detailed explanation of the Chinese emotional expression. In the figure1, we will use an example to illustrate the definition of ASTE and ASQE based on simple Chinese.
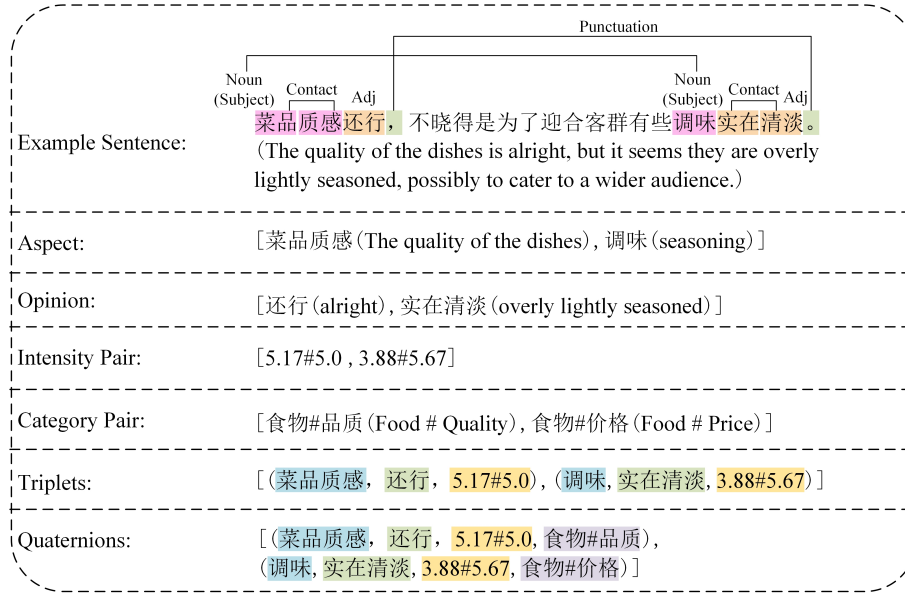

Figure 1. A Chinese example sentence from the ASTE and ASQE tasks. In the triple and quadruple tasks, aspect terms, opinion terms, and valence-arousal values are highlighted in blue, green, and orange, respectively, and aspect classes unique to the quadruple task are highlighted in purple. Multiple colors are also used to mark Chinese nouns, adjectives, and punctuation marks in the example sentence.

Pipeline models have been the main tool used in traditional ABSA approaches[2]. These models handle aspect extraction, sentiment classification, and intensity prediction. Although these models might be useful, they frequently have error propagation problems and cannot give a complete picture of sentiment exchanges. On the other hand, Grid Tagging Schemes (GTS)[3] provide a comprehensive method, but they might not fully capture the intricate dynamics of emotion intensity in two dimensions.

A powerful technique for encapsulating relational data in a variety of NLP applications is the graph convolutional network (GCN). However, their ability to tackle the particular difficulties of dimABSA has not been fully exploited, particularly in continuous intensity prediction and sentiment quadruple extraction. Our work intends to cover a significant gap in using naïve GCNs to do these complex tasks, as revealed by a thorough literature analysis.

With a unique GCN architecture created to intricately describe the interactions among aspect words, sentiment expressions, and their intensities, our suggested AWA-GCN model innovates. In order to concurrently learn and extract aspect phrases and sentiment expressions and map their respective intensities within the valence-arousal framework, this model uses the sophisticated capabilities of GCNs. In addition, our method closes a gap created by other approaches by effectively tackling the challenge of

extracting sentiment triplets and quadruplets from Chinese texts. Using these developments, our model offers a thorough and reliable resolution to the problems presented by the dimABSA assignment.

2. Related Work

Natural language processing (NLP) has made Aspect-Based Sentiment Analysis (ABSA) a major subject, emphasising removing subtle sentiment components from the text. This section examines the development of ABSA from its inception to the most advanced versions available now.

Scholars such as Pang et al.[4], who were among the first to break out the components of opinion mining, set the groundwork for the field of ABSA. They established a precise framework for comprehending sentiment at a fine level by identifying aspect terms and opinion terms as the fundamental components of sentiment analysis. This groundbreaking study established the conceptual basis for ABSA and emphasised the significance of differentiating between the mood and its objective, or the aspect.

After the first research, the pipeline models used by classic ABSA techniques mostly addressed the problem as a series of discrete sub-tasks: sentiment classification (SC), opinion term extraction (OTE), and aspect term extraction (ATE). Hu et al.[5] presented an early rule-based system that, although useful, brought to light the drawbacks of such models, most notably the problem of error propagation since subtasks are not interdependent. The necessity for a more comprehensive method of ABSA that might more effectively capture the intricate interactions between elements, viewpoints, and attitudes was highlighted by this work.

Lu Xu et al.[6] created a position-aware tagging model for Aspect Sentiment Triplet Extraction (ASTE), building on the shortcomings of conventional approaches. This approach uses a unique tagging system that captures the positional interrelations between the terms inside a phrase, integrating the extraction of aspect terms, opinion terms, and sentiment polarities into a cohesive framework. The enhanced accuracy in triplet extraction over conventional approaches proves that Xu et al.'s novel strategy produced more accurate and contextually aware sentiment analysis. However, the absence of a thorough investigation of the model's scalability across many languages or domains restricts the study's application; this is a limitation that future research might address to improve the model's adaptability in various contexts.

In recent years, the ABSA research community has emphasised sentiment element prediction in a more organised style, like the triplet format. The Aspect Sentiment Triplet Extraction (ASTE) task was established by Peng et al.[7] and has had a major impact on future research. The approach for extracting aspect-sentiment pairs and their corresponding opinion terms has been refined by Huang et al.[8], Jing et al.[9], and Chen et al.[10]. Moreover, Wan et al.[11] introduced the Target-Aspect Sentiment Detection (TASD) task, which is capable of concurrently predicting the sentiment polarity, aspect word, and aspect category—even in cases when the aspect term is conveyed implicitly. An important step has been taken in comprehending the implicit manifestations of sentiment in text with the work of Wan et al.

In their investigation of the subtleties of ASQP, Hu et al.[12] questioned the standard method of converting phrases into structured target sequences using a set template order. Using the pre-trained language model to determine the best template orders that reduce entropy, they suggested a unique technique for template-order data augmentation that greatly improved ASQP's performance, especially in low-resource circumstances. Hu et al.'s creative method demonstrates the adaptability provided by ASQP's order-free structure and how it may be used to enhance model performance.

A new Grid Tagging Scheme (GTS) for Aspect-oriented Fine-grained Opinion Extraction (AFOE) was presented by Wu et al.[3] in response to the shortcomings of early pipeline models, which frequently led to error propagation and inferior performance. The novel method uses a unified grid tagging job that simultaneously extracts all opinion elements, eschewing conventional pipeline methods. Through extensive experimentation across many model implementations, including CNN, BiLSTM, and BERT, GTS consistently outperforms robust baselines and achieves state-of-the-art performance, demonstrating its efficacy. The study's shortcomings, however, may restrict its application in more varied linguistic situations. These include its reliance on large amounts of training data and its possible decreased

adaptability to languages with fewer structural similarities to those examined.

Later studies have expanded on the groundwork established by GTS and other integrated models, delving more into the subtleties of ABSA. For example, Zhang et al.[13] presented a multitask learning system that simultaneously extracts sentiment polarities, opinion words, and aspect terms. This method makes use of the relationships that exist between the various parts of an emotion expression, which results in a more cohesive and contextually aware extraction procedure. Similarly, Zhou et al.[14] presented a bidirectional machine reading comprehension framework that considers both the aspect-to-opinion and opinion-to-aspect directions, thereby improving the comprehension of the sentiment structure within the text. This framework collects data that helps extract the aspect terms, opinion terms, and sentiment polarities.

With models such as the Syntax-Aware Transformer (SA-Transformer) presented by Yuan et al.[15], improving aspect-based sentiment triplet extraction (ASTE) by fully exploiting syntactic information, the inclusion of syntactic information has been an important topic of emphasis. By considering the different kinds of syntactic relationships and giving distinct weights to these edges, this technique outperforms conventional graph convolutional networks (GCNs). It avoids the improper pairing of unrelated words. The novel aspect of this approach is the adjacent edge attention (AEA) mechanism, which integrates edge representations into word-pair representations to encode syntactic information. It dynamically learns edge representations based on the dependence types of nearby edges. Based on many benchmark datasets, experimental results show that the SA-Transformer performs better than current methods. The research does not discuss the increased model size and possible computational complexity brought about by the AEA method, which might limit the approach's scalability for longer texts or bigger datasets.

Frameworks like Zhang et al.'s OTE-MTL [16] simultaneously extract aspect terms, opinion terms, and their sentiment dependencies using a Bi-Affine scorer, therefore addressing the shortcomings of previous aspect-based sentiment analysis techniques. The method stands out because it separates aspect and opinion extraction from sentiment prediction, enabling more accurate sentiment dependency parsing. This is especially helpful for managing the overlapping and intricate structures that sentiment analysis jobs sometimes involve. Their tests on a series of SemEval benchmarks show that OTE-MTL performs noticeably better than many strong baselines, demonstrating its efficacy. The research does point out that more advanced interaction modelling between extracted pieces might be beneficial in situations involving complicated phrase structures and overlapping sentiment contexts.

The Multi-Branch Graph Convolutional Network (MBGCN) was presented by Shi et al.[17]. It is a unique design that combines structure-biased BERT to generate enhanced semantic features and Bi-Affine attention to refining these features according to different language connections. This method embeds syntactic dependencies, part-of-speech combinations, and other linguistic linkages as neighbouring matrices in a GCN, significantly enhancing the interactions between various textual feature types. This technique outperformed state-of-the-art algorithms on many benchmarks in extracting sentiment-related triplets from text. The article indicates a possible area for improvements in future model robustness and accuracy. However, it also argues that despite these achievements, greater optimisation in handling many types of textual characteristics simultaneously might potentially boost performance.

The introduction of pretrained transformer-based models such as BERT has transformed ABSA. Devlin et al. [18] have shown that BERT is effective in various NLP tasks, and its application to ABSA has been revolutionary. BERT was creatively used for the aspect sentiment classification challenge by Sun et al.[19], turning it into a language inference issue. Applying BERT and sequence-to-sequence techniques have greatly improved the granularity and accuracy of sentiment element prediction.

Zhang et al.[20] proposed the Aspect Sentiment Quad Prediction (ASQP) challenge, which marked a paradigm change in the field of ABSA. They argued for a unified approach to paraphrase generation. They converted the job into a sequence-to-sequence learning issue, facilitating end-to-end training and more comprehensive use of semantic information. The suggested paraphrase modelling framework shows good cross-task transfer capabilities and performs better on benchmark datasets. The robustness

and generalizability of the model may be impacted by possible issues with modelling complexity and generation mistakes brought on by the free nature of text production, as noted in the study.

Hu et al.[12] concentrated on improving aspect sentiment quad prediction (ASQP) performance by tackling the strict template order commonly employed in this assignment. They provide a novel approach that makes use of several template ordering and a pre-trained language model to identify the best sequences for lowering entropy in predictions. Their data augmentation method through several template orders greatly enhances ASQP performance, especially in low-resource environments. The main contribution is to demonstrate how rearranging ASQP's parts may yield fresh perspectives and enhance model performance, demonstrating the method's efficacy compared to cutting-edge models. However, the study's dependence on certain pre-trained models and the best choice of template orders may restrict its generalizability and point to areas needing modification and more research.

Moreover, the Aspect Category Opinion Sentiment (ACOS) challenge was presented by Cai et al. [21], which greatly motivated me to complete the aspect class prediction job in quadruple extraction. Their main goal was to extract the overall opinion feeling of the aspect category from text reviews. It used cutting-edge models, including Extract-Classify-ACOS, TAS-BERT-ACOS, and JET-ACOS. By utilising BERT-based architecture and innovative tagging strategies, these models combined aspect opinion joint extraction with category sentiment classification, greatly improving the extraction process. The inclusion of a conditional extraction approach in TAS-BERT-ACOS is a noteworthy breakthrough. This method improves the detection of implicit aspects and views using the context supplied by category sentiment pairings. Furthermore, two new datasets, Restaurant-ACOS and Laptop-ACOS, are produced, which include annotations of explicit and implicit aspects. These datasets are a significant resource for future study and benchmarking. The article does admit several difficulties, though, such as the difficulty of modelling and the scarcity of fully annotated datasets. Building on this, my study intends to enhance the scalability and accuracy of the category extraction component, particularly in handling nuanced emotional expressions.

Motivated by the paradigm-shifting effects of these investigations, we provide a unique method for sentiment triplet and quadruplet extraction specifically designed for the SIGHAN competition assignment. Our approach seeks to provide a more thorough and reliable sentiment analysis in textual data by utilising the advantages of pre-trained models via a paraphrase-generating framework. Unlike conventional methods, our model can accurately and comprehensively analyse the sentiment expressed in reviews by capturing the continuous sentiment intensity using a valence-arousal representation.

3. Methodology

Our work presents the AWA-GCN (Attention, Word-Level Analysis, and Graph Convolutional Network) model, a sophisticated architecture created especially for the SIGHAN 2024 Chinese Dimensional Aspect-Based Sentiment Analysis (dimABSA) Task, building on the advances described in earlier studies. This competition requires the detection of sentiment triplets and quadruplets and their rating on a continuous valence-arousal scale beyond the limits of typical aspect-based sentiment analysis.

In order to ensure accurate detection of Chinese word edges, the suggested model particularly executes distinct BERT processing for Chinese characters (comparable to English letters) and words (comparable to English words). Multi-layer attention fusion is carried out, especially for character processing after BERT embedding, which addresses the issue of information loss by focusing on more than one layer of attention. Together, the graph convolutional network (GCN) and parent attention may successfully extract sentiment triplets and quadruplets from Chinese text. A thorough sentiment analysis requires the consideration of these three factors: aspect, emotion, and intensity. The parent attention method allows accurate modelling of the interdependencies between text components, and incorporating BERT embedding fills the model's input with deep contextual information. With the help of the GCN layer, sentiment analysis may be accomplished more accurately by improving the capacity to recognize and utilize intricate links inside words.

With its lack of clear word boundaries and implicit sentiment expressions, the Chinese language

presents unique challenges that the AWA-GCN model is particularly well-suited to tackle. Often overlooked by simpler models, its multi-layer attention mechanism enhances the identification and contextual interpretation of subtle attitude cues.

Figure 2 below shows the AWA-GCN model architecture, demonstrating how these sophisticated approaches are integrated into a coherent and efficient system for complex ABSA tasks, especially well-suited to Chinese texts.
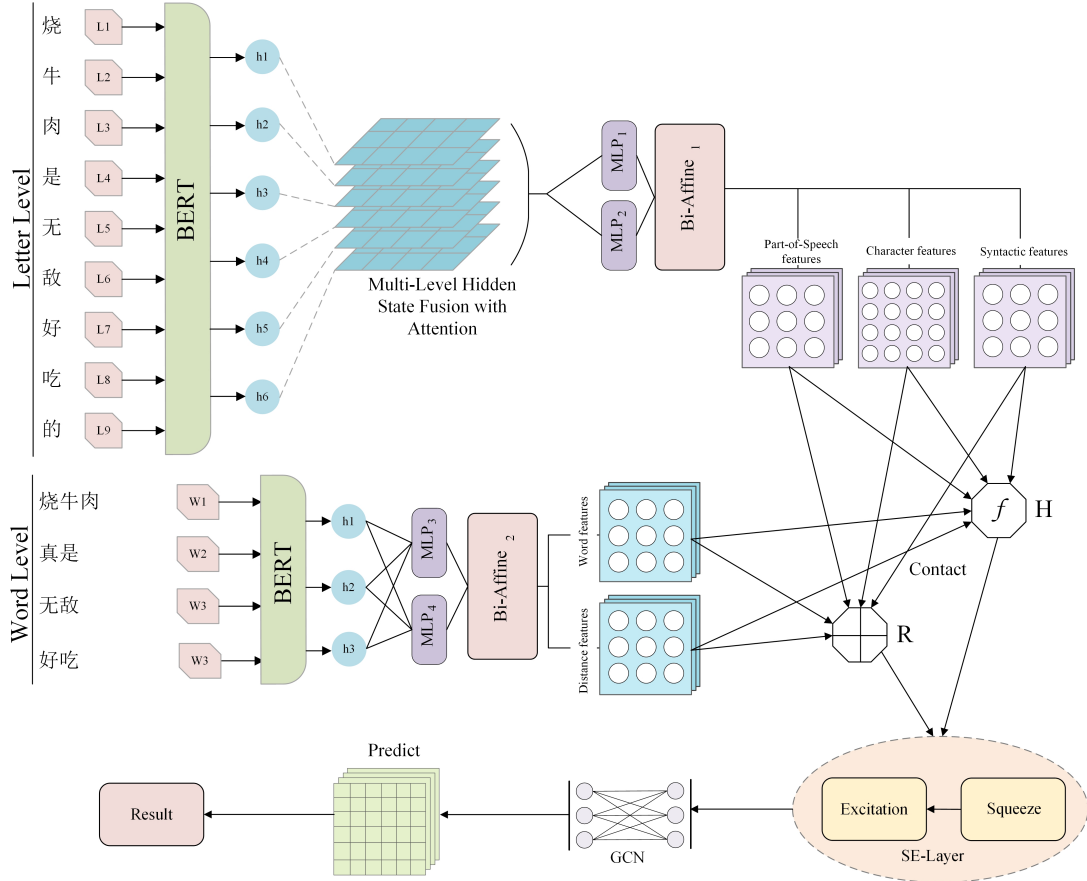

Figure 2. Our Whole Model

3.1. BERT-fine-tuning

In our model, BERT[18] functions as the foundational mechanism for converting raw Chinese text into a structured format that can be effectively utilized for extracting emotional triplets and quartets. The following subsections detail the mathematical transformations and the role of BERT in the sentiment analysis model.

3.1.1. Text Tokenization and Encoding

The initial step in processing the text with BERT involves tokenizing the input into sequences of both characters and words, enhanced by adding special tokens. This dual tokenization approach ensures comprehensive coverage of Chinese linguistic features.

For character-level tokenization:

$$\text{CharTokenizer}(text) = \{[\text{CLS}], c_1, c_2, \ldots, c_m, [\text{SEP}]\} \tag{1}$$

For word-level tokenization:

$$\text{WordTokenizer}(text) = \{[\text{CLS}], w_1, w_2, \ldots, w_n, [\text{SEP}]\} \tag{2}$$

Using a pre-trained embedding model, each token c_i or w_i is then mapped to a high-dimensional vector space. This embedding captures the semantic meaning of each token and includes subword information, which is crucial for understanding the morphological nuances of the Chinese language.

3.1.2. Incorporating Context and Positional Information   To encode the order of tokens and distinguish between different segments in tasks that involve multiple sentences, segment and positional embeddings are added to the token embeddings. This stage is crucial for maintaining the contextual integrity of the text and helping the model recognize sentence structures and links.

For character-level processing:

$$\text{Char Segment Embeddings} = \{s_1, s_2, \ldots, s_m\} \tag{3}$$

where s_i is generally 0 for tokens in the first sentence and 1 for the second sentence, assisting the model in context differentiation.

$$\text{Char Positional Embeddings} = \{\text{PosEmbed}(1), \text{PosEmbed}(2), \ldots, \text{PosEmbed}(m)\} \tag{4}$$

These embeddings offer the model with relative or absolute token placement information.

For word-level processing:

$$\text{Word Segment Embeddings} = \{s_1, s_2, \ldots, s_n\} \tag{5}$$

$$\text{Word Positional Embeddings} = \{\text{PosEmbed}(1), \text{PosEmbed}(2), \ldots, \text{PosEmbed}(n)\} \tag{6}$$

The whole sentence representation for each level is created by summing the associated embeddings:

$$\text{Char Sentence Embedding} = \text{Char Embeddings} + \text{Char Segment Embeddings} + \text{Char Positional Embeddings} \tag{7}$$

$$\text{Word Sentence Embedding} = \text{Word Embeddings} + \text{Word Segment Embeddings} + \text{Word Positional Embeddings} \tag{8}$$

3.1.3. Fine-tuning for Specific Sentiment Analysis Tasks   During the fine-tuning step, BERT is customized to the unique job of extracting sentiment-related information from the text. This entails changing the pre-trained weights via back-propagation based on the gradients obtained using task-specific loss functions. Fine-tuning allows the model to specialize in identifying and interpreting emotional cues in Chinese text.

The fine-tuning procedure changes the weights by minimizing the loss function, which assesses the gap between the predicted and real sentiment labels.

$$\theta_{\text{new}} = \theta_{\text{old}} - \eta \cdot \nabla_\theta L(\theta; \text{data}) \tag{9}$$

Where θ represents the parameters of the model, η is the learning rate, $\nabla_\theta L$ is the gradient of the loss function L with respect to the parameters, and "data" refers to the training data comprising both the inputs and the predicted outputs.

This stage is crucial as it enables BERT to shift from a broad comprehension of language to a subtle comprehension of emotional expressions in text[22], thereby improving the model's capacity to forecast sentiment patterns in unfamiliar data accurately.

By following these steps, BERT effectively becomes tailored to the sentiment analysis task within the Chinese linguistic context, enabling it to capture and interpret complex emotional cues essential for accurate sentiment extraction and classification.

3.2. Integration of Attention Mechanisms in Multi-Layer Hidden State Fusion

After preprocessing the Chinese text for input into BERT, the next step involves further refinement of the character-level embeddings using advanced attention mechanisms. This process enhances the model's ability to capture deeper contextual relationships within the text, which is crucial for nuanced sentiment analysis tasks. Here is a detailed exposition of how multi-layer hidden state fusion using attention mechanisms is employed.

The fusion of hidden states from multiple layers of a BERT model allows the system to integrate diverse aspects of language understanding captured at different levels of the network. This multi-layer fusion employs attention mechanisms to weigh the significance of features from each layer, enabling a more robust representation of text semantics and syntax.

From the BERT model, we extract the hidden states $\mathbf{H}^{(l)}$ for each layer l, where $l = 1, 2, \ldots, L$, and L is the total number of transformer layers in the model.

$$\mathbf{H}^{(l)} = \text{BERT}^{(l)}(\text{Input Embeddings}) \tag{10}$$

The fusion process applies an attention mechanism to combine these hidden states. We calculate the attention weights α_l for each layer's output, reflecting the importance of that layer's features in the context of the specific sentiment analysis task.

$$\alpha_l = \frac{\exp(\mathbf{w}^T \mathbf{h}_l)}{\sum_{k=1}^{L} \exp(\mathbf{w}^T \mathbf{h}_k)} \tag{11}$$

Here, $\mathbf{w}$ is a trainable parameter vector, and $\mathbf{h}_l$ is a summary vector for layer l, typically obtained by applying a non-linear transformation (e.g., tanh) to the pooled output of the layer:

$$\mathbf{h}_l = \tanh(\mathbf{W}_h \mathbf{H}^{(l)} + \mathbf{b}_h) \tag{12}$$

where $\mathbf{W}_h$ and $\mathbf{b}_h$ are trainable parameters.

The final fused hidden state $\mathbf{H}_{\text{fused}}$ is computed as a weighted sum of all layer outputs, where the attention scores give the weights:

$$\mathbf{H}_{\text{fused}} = \sum_{l=1}^{L} \alpha_l \mathbf{H}^{(l)} \tag{13}$$

Depending on the model architecture, this fused representation is then used as the input for subsequent layers or directly for classification.

Different from the previous approach of only focusing on the last layer of the attention mechanism[23], this approach allows the model to dynamically focus on the most relevant linguistic features for sentiment analysis, combining low-level lexical details captured in earlier layers with more abstract textual interpretations from higher layers. By integrating these diverse representations, the model can better understand complex linguistic constructs such as irony, negation, and emotional subtext, which are critical for accurate sentiment prediction.

The attention-based fusion of multi-layer hidden states thus significantly enhances the model's ability to discern subtle sentiment cues embedded in the text, leading to more robust and accurate sentiment analysis outcomes. This method also adds a layer of interpretability to the model, as the attention weights provide insights into which layers of the BERT model are most influential for specific sentiment determinations.

3.3. Bi-Affine Layers

After processing through BERT and embedding layers, the character-level and word-level features of the text are further analyzed using Bi-Affine attention mechanisms[24]. These mechanisms are crucial

for capturing complex dependencies between textual elements, which are particularly significant in tasks like sentiment analysis, where understanding the relationships between different text parts is essential.

3.3.1. Bi-Affine Attention for Character-Level Features (Bi-Affine1)   After processing with BERT, each character in the text is represented by a high-dimensional vector that captures the character's inherent semantic properties and its contextual relationship with its neighbors. The next step involves using a Bi-Affine[25] attention mechanism to analyze and capture interactions between pairs of characters.

For Character Feature Vectors: Let $\mathbf{X}$ denote the matrix of character feature vectors output by BERT, where each row $\mathbf{x}_i$ corresponds to a character in the text.

$$\mathbf{X} = \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \vdots \\ \mathbf{x}_n \end{bmatrix} \tag{14}$$

For Bi-Affine Transformation: The Bi-Affine attention layer computes a score for each pair of characters using a Bi-linear form which includes a trainable parameter tensor $\mathbf{W}$ and bias vectors $\mathbf{b}_x$ and $\mathbf{b}_y$.

$$S_{ij} = \mathbf{x}_i^T \mathbf{W} \mathbf{x}_j + \mathbf{b}_x^T \mathbf{x}_i + \mathbf{b}_y^T \mathbf{x}_j + b \tag{15}$$

Here, S_{ij} is the score that models the relationship between the i-th and j-th characters, potentially capturing dependencies relevant to sentiment expressions like negations or intensifiers.

3.3.2. Bi-Affine Attention for Word-Level Features (Bi-Affine2)   Simultaneously, word-level embeddings are derived through another set of BERT transformations or specialized embedding techniques tailored for whole words or phrases. These embeddings are crucial for capturing semantics at a higher granularity.

For Word Feature Vectors: Let $\mathbf{Y}$ denote the matrix of word-level feature vectors, where each row $\mathbf{y}_k$ corresponds to a word or phrase.

$$\mathbf{Y} = \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_m \end{bmatrix} \tag{16}$$

For Bi-Affine Transformation: Similarly, an affine transformation is applied to the word-level features to compute interactions between words:

$$T_{kl} = \mathbf{y}_k^T \mathbf{W}' \mathbf{y}_l + \mathbf{b}'^T_x \mathbf{y}_k + \mathbf{b}'^T_y \mathbf{y}_l + b' \tag{17}$$

This setup allows the model to infer higher-level semantic relationships, such as the sentiment associated with specific product or service aspects mentioned in the text.

Both character-level and word-level Bi-Affine scores are then integrated to provide a comprehensive view of the text structure and semantics. This integration is crucial as it allows the model to leverage the fine-grained syntactic cues available at the character level and the broader semantic cues at the word level. The outputs of these Bi-Affine layers serve as inputs to further layers, such as Graph Convolutional Networks (GCN), which utilize these relationships to predict sentiment labels effectively.

This sophisticated use of Bi-Affine attention layers, adapted from neural dependency parsing, underscores the model's capability to understand and analyze complex linguistic patterns, making it highly effective for the nuanced task of sentiment analysis[26].

Following the detailed analysis through Bi-Affine attention mechanisms for both character and word-level features, the model proceeds to synthesize and consolidate these insights into a more structured representation, which is essential for subsequent processing layers.

The output from the character-level Bi-Affine attention layer includes enhanced character vectors that now encapsulate richer semantic and POS features. Each character vector $\mathbf{x}_i$ is augmented to $\mathbf{x}_i'$ which includes POS tagging information:

$$\mathbf{x}_i' = \text{Concat}(\mathbf{x}_i, \text{POS}_i) \tag{18}$$

Similarly, the output from the word-level Bi-Affine attention captures syntactic roles and relations. Each word vector $\mathbf{y}_k$ is augmented to $\mathbf{y}_k'$ incorporating syntactic dependency information:

$$\mathbf{y}_k' = \text{Concat}(\mathbf{y}_k, \text{Syntax}_k) \tag{19}$$

Additionally, distance features $\mathbf{D}$ are calculated to represent the relative or absolute distances between significant textual elements, providing spatial context that can influence the interpretation of relationships and sentiments:

$$\mathbf{D}_{ij} = \text{Distance}(\mathbf{x}_i, \mathbf{x}_j) \tag{20}$$

Where $\mathbf{D}_{ij}$ measures the positional offset between characters or words, enhancing the model's ability to understand ordering and proximity, which are crucial for tasks like scope detection in sentiment analysis.

The augmented feature vectors from both the character and word levels are concatenated with the distance features to form a comprehensive feature set ready for the subsequent layers:

$$\mathbf{F} = \text{Concat}(\mathbf{x}_i', \mathbf{y}_k', \mathbf{D}_{ij}) \tag{21}$$

This concatenated feature matrix $\mathbf{F}$ now serves as the input to further processing stages, such as a Graph Convolutional Network (GCN), which can effectively utilize these structured representations to identify and classify sentiment expressions based on the complex interplay of semantic, syntactic, and spatial features.

This sophisticated integration of multi-level linguistic features with attention-derived character and word embeddings ensures that the model is well-equipped to handle the intricacies of sentiment analysis in the text. The model achieves a deeper understanding of textual structure and dynamics by leveraging Bi-Affine attention to enrich the feature space with POS and syntactic information and integrating these with contextual distance metrics.

3.4. SE-layer

The integration of the Squeeze-and-Excitation (SE) layer[27] plays a pivotal role in refining the feature representation for the nuanced task of emotion triplet and quadruplet extraction, encompassing elements such as aspect, sentiment, intensity, category, and opinion. The SE layer fundamentally transforms the feature processing approach by implementing a dynamic recalibration mechanism that adjusts feature activations across the model, thereby significantly enhancing the extraction process.

The primary function of the SE layer is to recalibrate the feature maps generated by prior network layers, which are enriched with linguistic and semantic cues through the Bi-Affine attention mechanisms. This recalibration is critical as it directly influences the model's sensitivity to relevant textual features that denote emotional content and contextual relationships. By employing a global information embedding followed by a channel-wise excitation process, the SE layer adaptively enhances features crucial for identifying specific emotional expressions and suppresses less informative features that could potentially introduce noise or ambiguity into the sentiment analysis.

This dynamic adjustment is executed through a computationally efficient yet powerful gating mechanism embedded within the SE layer. Each channel of the feature map is independently adjusted using learned recalibration weights, which are derived from the global contextual information aggregated across all features. These weights are applied to the original features, effectively allowing the model to focus more on semantically rich features such as intense adjectives, nuanced aspect terms, and specific opinion expressions essential for accurately capturing the complex structure of emotion-related information within the text.

The recalibration process can be expressed as:

$$\mathbf{s} = \sigma(\mathbf{W}_2 \delta(\mathbf{W}_1 \mathbf{z})), \quad \hat{\mathbf{X}} = \mathbf{X} \otimes \mathbf{s} \tag{22}$$

Where σ is the sigmoid activation, δ represents the ReLU function, and $\mathbf{W}_1$, $\mathbf{W}_2$ are trainable parameters of the fully connected layers within the SE block. This formulation ensures that the recalibration is deeply embedded within the network's architecture, enabling the model to adaptively enhance or suppress features based on their contextual relevance to the task.

Furthermore, the ability of the SE layer to dynamically adapt to the specific content of sentences enhances the model's versatility across different textual contexts[28], making it robust to variations in language use and expression styles. This is especially beneficial in tasks involving fine-grained emotional assessments where the precise understanding of context, sentiment polarity, and intensity levels is critical.

In summary, the SE layer's capability to modulate feature importance adaptively ensures that the sentiment analysis model not only becomes more efficient in distinguishing between subtle nuances of language that convey emotion but also becomes more effective in generalizing across diverse linguistic contexts. This leads to improved accuracy in the automatic extraction of emotion triplets and quadruplets, which are crucial for advanced sentiment analysis applications, thereby enhancing both the reliability and depth of the analytical insights derived from the model.

The figure3 illustrates the architecture of our SE Layer.
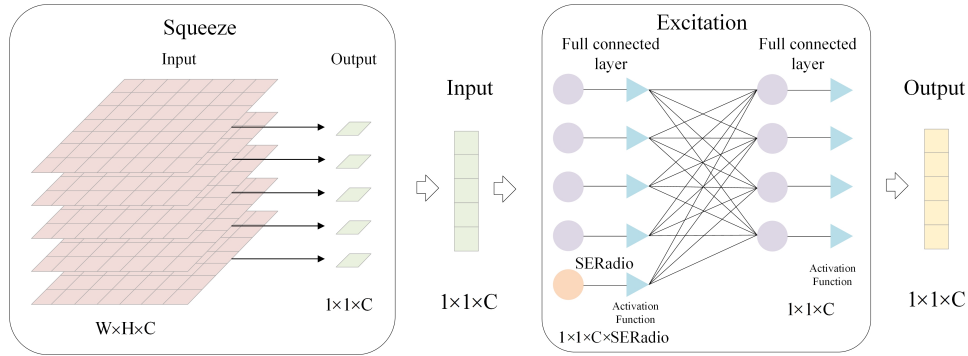
Figure 3. Our SE Layer

3.5. GCN model

(GCNs)[29], which use the improved feature representations generated from the Squeeze-and-Excitation (SE) layer, is essential for the extraction of emotion triplets and quadruplets from the text. GCNs efficiently use graph-based representations to extract and comprehend intricate connections from text.

Graph Convolutional Networks function on graph structures, which may be thought of as nodes (words or characters) and edges (syntactic or semantic associations) in the context of text [30]. In our sentiment analysis model, GCNs' main job is to spread information around these nodes to uncover dependencies essential for figuring out sentiment structures.

Every word or character that comes from the SE layer's output has a high-dimensional feature vector assigned to it. Furthermore, the associations found via syntactic parsing or learned embeddings may be used to initialize edge features:

$$\mathbf{H}^{(0)} = \text{SE Output}, \quad \mathbf{E}^{(0)} = \text{Initial Edge Features} \tag{23}$$

GCNs get data from their neighbors to update the node features:

$$\mathbf{H}^{(l+1)} = \sigma \left(\mathbf{D}^{-\frac{1}{2}} \mathbf{A} \mathbf{D}^{-\frac{1}{2}} \mathbf{H}^{(l)} \mathbf{W}^{(l)} \right) \tag{24}$$

Where the node features at layer l are represented by $\mathbf{H}^{(l)}$; the adjacency matrix of the graph, potentially enhanced by self-connections, is represented by $\mathbf{A}$; the degree matrix of $\mathbf{A}$ is represented by $\mathbf{D}$; the weight matrix for layer l is represented by $\mathbf{W}^{(l)}$, and a non-linear activation function, like ReLU, is represented by σ.

As new data is analyzed, the model may dynamically modify the connections by updating edge characteristics depending on node properties as well:

$$\mathbf{E}^{(l+1)} = \text{EdgeUpdate}(\mathbf{E}^{(l)}, \mathbf{H}^{(l)}) \tag{25}$$

The precise role of EdgeUpdate may differ, including tasks like using neural networks to merge attributes of interconnected nodes.

Through convolution, each node (word or letter) has its feature representation improved by its neighbors' weighted contributions, reflecting local and extended contextual dependencies. As the network digs deeper into the textual structure, edge features may be updated simultaneously to represent changing relationships dynamically. A sophisticated comprehension of language patterns is made possible by this repeated refining.

Thus, GCNs have two functions in our model: one is to identify and highlight entities that significantly impact sentiment expressions in the text; the other is to clarify the complex web of relationships between these entities, such as verb-object interactions or adjective-noun modifications. This thorough analysis helps identify the main entities and their characteristics and guarantees that these entities are situated within their larger textual context.

GCNs work well in this tiered technique because they smoothly include the recalibrated features from the Squeeze-and-Excitation layer, which improves the model's capacity to reduce irrelevant noise and concentrate on important information. With the help of this integrated method, text may be deeply semantically explored, allowing the model to extract structured emotion descriptors (quadruples adding a category dimension or triplets combining aspect, sentiment, and intensity) that are essential for in-depth sentiment analysis.

The figure 4 illustrates the architecture of our Graph Convolutional Network (GCN) model.
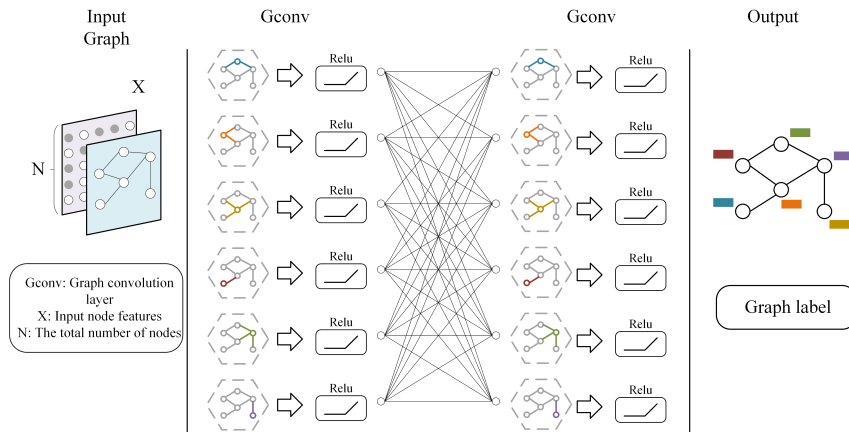

Figure 4. Our GCN Model

3.6. the Dual Prediction matrix

The design and implementation of the predictive matrices for intensity and category in sentiment analysis involve carefully considering the relationships and features extracted by the Graph Convolutional Network (GCN). Here, we explore in depth how these matrices are generated and the rationale behind their specific constructions for effective emotion triplet and quadruplet extraction.

3.6.1. Intensity Prediction Matrix Construction

The intensity prediction matrix is designed to quantify the emotional strength or depth associated with each pair of nodes (representing words or phrases) identified as relevant in the text. This matrix is crucial for tasks that require detecting sentiments and measuring their intensity, which can vary significantly even within the same sentence.

- The intensity matrix $\mathbf{Y}_{intensity}$ is formulated using a bilinear model, which is particularly suited for capturing interactions between node features that the GCN outputs:

$$\mathbf{Y}_{intensity} = \sigma(\mathbf{H}^T \mathbf{W}_{intensity} \mathbf{H}) \tag{26}$$

where: - $\mathbf{H}$ represents the node features output from the GCN, - $\mathbf{W}_{intensity}$ is a trainable weight matrix that models the pairwise interactions between the features, - σ is typically a sigmoid function that scales the output to a [0,1] range, representing the intensity level.

Each element of $\mathbf{Y}_{intensity}$, say y_{ij}, represents the predicted intensity of sentiment between node i and node j. The sentiment intensity is reflected using continuous real-valued scores in the valence-arousal dimensions. The valence represents the degree of pleasant and unpleasant (i.e., positive and negative) feelings, while the arousal represents the degree of excitement and calm. Both the valence and arousal dimensions use a nine-degree scale. Value 1 on the valence and arousal dimensions denotes extremely high-negative and low-arousal sentiment, respectively. In contrast, 9 denotes extremely high-positive and high-arousal sentiment, and 5 denotes a neutral and medium-arousal sentiment. Valence-arousal values are separated by a hashtag (symbol "#") for a mark.

3.6.2. Category Prediction Matrix Construction

The category prediction matrix extends the analysis by categorizing the aspects and sentiments into predefined classes. The predefined categories can be split into entities and attributes using the symbol "#." These categories are tailored for the restaurant domain and include entities and attributes such as "restaurant#prices," "food#prices," and so forth.

- For categorical predictions, a similar approach is used but with a focus on multiclass classification for each node pair:

$$\mathbf{Y}_{category} = \text{softmax}(\mathbf{H}^T \mathbf{W}_{category} \mathbf{H}) \tag{27}$$

Where: - $\mathbf{W}_{category}$ is structured to capture the multi-dimensional aspect of sentiment categorization, - softmax ensures that the output for each pair sums to one, effectively turning the output into a probability distribution over possible categories.

This matrix helps assign each sentiment expression to a specific category, facilitating nuanced understanding and response mechanisms for automated systems such as customer service bots or content moderation tools.

3.6.3. Integration in Sentiment Analysis

The generation of these matrices from the GCN outputs provides a structured and quantitative representation of the sentiment information extracted from the text. By using bilinear and softmax transformations[31], these matrices not only capture the strength and category of sentiments but also ensure that the interactions between different entities in the text are accurately represented and utilized for downstream applications.

The dual predictive framework allows the model to not only predict the presence of specific sentiments and their intensities but also categorize these sentiments in ways that align with predefined categories relevant to the domain of application.

This predictive approach benefits from the rich, relationally infused feature sets provided by the GCN, ensuring that the predictions are contextually informed and nuanced. The intensity matrix helps understand the depth of sentiment, which is crucial for applications requiring detailed emotional insights. In contrast, the category matrix adds a layer of specificity, which is crucial for tailoring responses or actions based on sentiment analysis.

3.6.4. Two prediction matrices   We define nine relations between words in a sentence for this dimABSA task. Compared with Yuan et al.'s ESIT[15] and Chen et al.'s EM-GCN[32] models, the relations we define introduce more precise strength and clearer Chinese word boundary information to our model. The specific relations are shown in Table 1.

Table 1. Labels for the Nine Relations Defined

#	Relation	Meaning
1	S-A	Start of aspect term.
2	M-A	Middle of aspect term.
3	A	Vocabulary pair (v_i, v_j) belongs to the same aspect term.
4	S-O	Start of opinion term.
5	M-O	Middle of opinion term.
6	O	Vocabulary pair (v_i, v_j) belongs to the same opinion term.
7	I-A # I-O	(valence of aspect # arousal of opinion)
8	C-A # C-O	(category of aspect # category of opinion)
9	$\perp$	No above relations between word pair (v_i, v_j).

(The relations and meanings are clearly defined for reference in this table.)

In row 7 of the table, the Meaning section has the following additions: v_i and v_j of the word pair (v_i, v_j) belong to an aspect term and an opinion term; the specific score values (valence and arousal rating) will be given at the corresponding positions. In row 8 of the table, the Meaning section has the following additions: w_i and w_j of the word pair (w_i, w_j) respectively belong to an aspect term and an opinion term; the specific classification results will be given at the corresponding positions.

S and M denote the start of and middle of the term, respectively, while -A and -O subtags aim to determine the role of the term, i.e., an aspect or an opinion. The A and O relations in Table 1 are used to detect whether the word pair formed by two different words belongs to the same aspect or opinion term, respectively. We study the matrix tagging method of Miwa and Sasaki[33], using a table-filling approach to build a relation table for each tagged sentence. The specific example is shown in the following figure5, including the intensity prediction matrix (left) and the category prediction matrix (right), taking two Chinese sentences in the dataset as examples. Using the tags in the relationship table, the color block filling table accurately marks the triple and quadruple extraction results of our AWA-GCN model for Chinese sentences.
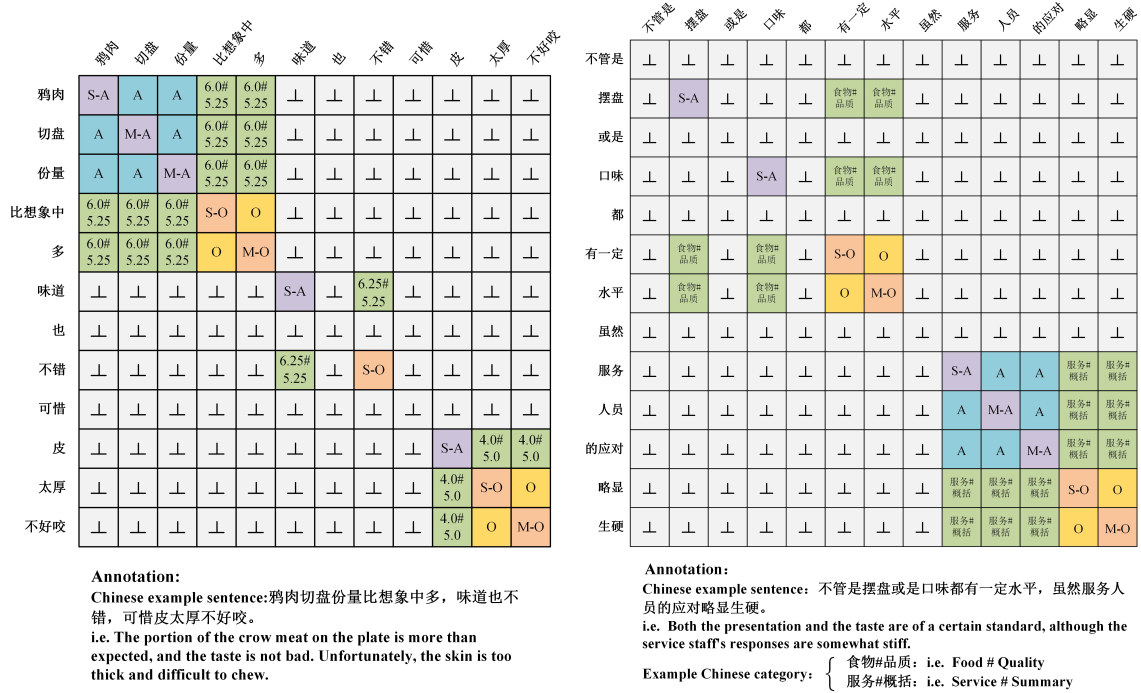
Intensity Prediction

	鸦肉	切盘	份量	比想象中	多	味道	也	不错	可惜	皮	太厚	不好咬
鸦肉	S-A	A	A	6.0#5.25	6.0#5.25	⊥	⊥	⊥	⊥	⊥	⊥	⊥
切盘	A	M-A	A	6.0#5.25	6.0#5.25	⊥	⊥	⊥	⊥	⊥	⊥	⊥
份量	A	A	M-A	6.0#5.25	6.0#5.25	⊥	⊥	⊥	⊥	⊥	⊥	⊥
比想象中	6.0#5.25	6.0#5.25	6.0#5.25	S-O	O	⊥	⊥	⊥	⊥	⊥	⊥	⊥
多	6.0#5.25	6.0#5.25	6.0#5.25	O	M-O	⊥	⊥	⊥	⊥	⊥	⊥	⊥
味道	⊥	⊥	⊥	⊥	⊥	S-A	⊥	6.25#5.25	⊥	⊥	⊥	⊥
也	⊥	⊥	⊥	⊥	⊥	⊥	⊥	⊥	⊥	⊥	⊥	⊥
不错	⊥	⊥	⊥	⊥	⊥	6.25#5.25	⊥	S-O	⊥	⊥	⊥	⊥
可惜	⊥	⊥	⊥	⊥	⊥	⊥	⊥	⊥	⊥	⊥	⊥	⊥
皮	⊥	⊥	⊥	⊥	⊥	⊥	⊥	⊥	S-A	4.0#5.0	4.0#5.0	
太厚	⊥	⊥	⊥	⊥	⊥	⊥	⊥	⊥	⊥	4.0#5.0	S-O	O
不好咬	⊥	⊥	⊥	⊥	⊥	⊥	⊥	⊥	⊥	4.0#5.0	O	M-O

Annotation:
Chinese example sentence:鸦肉切盘份量比想象中多，味道也不错，可惜皮太厚不好咬。
i.e. The portion of the crow meat on the plate is more than expected, and the taste is not bad. Unfortunately, the skin is too thick and difficult to chew.

Category Prediction

	不管是	摆盘	或是	口味	都	有一定	水平	虽然	服务	人员	的应对	略显	生硬
不管是	⊥	⊥	⊥	⊥	⊥	⊥	⊥	⊥	⊥	⊥	⊥	⊥	⊥
摆盘	⊥	S-A	⊥	⊥	⊥	食物#品质	食物#品质	⊥	⊥	⊥	⊥	⊥	⊥
或是	⊥	⊥	⊥	⊥	⊥	⊥	⊥	⊥	⊥	⊥	⊥	⊥	⊥
口味	⊥	⊥	⊥	S-A	⊥	食物#品质	食物#品质	⊥	⊥	⊥	⊥	⊥	⊥
都	⊥	⊥	⊥	⊥	⊥	⊥	⊥	⊥	⊥	⊥	⊥	⊥	⊥
有一定	⊥	食物#品质	⊥	食物#品质	⊥	S-O	O	⊥	⊥	⊥	⊥	⊥	⊥
水平	⊥	食物#品质	⊥	食物#品质	⊥	O	M-O	⊥	⊥	⊥	⊥	⊥	⊥
虽然	⊥	⊥	⊥	⊥	⊥	⊥	⊥	⊥	⊥	⊥	⊥	⊥	⊥
服务	⊥	⊥	⊥	⊥	⊥	⊥	⊥	⊥	S-A	A	A	服务#概括	服务#概括
人员	⊥	⊥	⊥	⊥	⊥	⊥	⊥	⊥	A	M-A	A	服务#概括	服务#概括
的应对	⊥	⊥	⊥	⊥	⊥	⊥	⊥	⊥	A	A	M-A	服务#概括	服务#概括
略显	⊥	⊥	⊥	⊥	⊥	⊥	⊥	⊥	服务#概括	服务#概括	服务#概括	S-O	O
生硬	⊥	⊥	⊥	⊥	⊥	⊥	⊥	⊥	服务#概括	服务#概括	服务#概括	O	M-O

Annotation：
Chinese example sentence：不管是摆盘或是口味都有一定水平，虽然服务人员的应对略显生硬。
i.e. Both the presentation and the taste are of a certain standard, although the service staff's responses are somewhat stiff.

Example Chinese category： { 食物#品质：i.e. Food # Quality
服务#概括：i.e. Service # Summary }

Figure 5. Prediction Matrices

4. Experiments and Results

4.1. Dataset

Our study utilized the Simplified Chinese dataset from the dimABSA dataset, which is tailored for the Chinese Dimension ABSA shared task organized during the SIGHAN 2024 workshop. This dataset is meticulously structured to facilitate the fine-grained sentiment intensity prediction associated with various aspects of restaurant reviews. The dataset not only distinguishes between different aspects of the dining experience but also measures the intensity of sentiments expressed towards these aspects on a detailed valence-arousal scale.

The dimABSA dataset is designed to enable deep analyses of restaurant reviews through the lens of dimensional sentiment analysis. It comprises annotations for aspect terms, aspect categories, opinion terms, and sentiment intensities, providing a comprehensive framework for evaluating the nuances of sentiment in the context of dining experiences.

Aspect Term (A): Refers to specific entities within the restaurant domain that are the targets of opinions. Unmentioned aspects are labeled as "NULL". Aspect Category (C): Categorizes each aspect into one of twelve predefined categories, further detailed by combining entity and attribute descriptions. Opinion Term (O): Describes the sentiment expressions directed at the aspects. Sentiment Intensity (I): Measured on a continuous scale, reflecting sentiment valence (pleasantness or unpleasantness) and arousal (excitement or calmness) on a scale from 1 to 9.

The dataset's primary tasks are divided into extracting sentiment triplets and quadruplets from the text, which challenges the model to accurately identify and quantify sentiment structures.

The specific content of the ABSA dataset (Simplified Chinese) is shown in Figure6 :

ID	Sentence	Aspect	Opinion	Intensity
R1166	清爽的昆布汤。	昆布汤	清爽	6.0 - 5.12
R1166	石头汤底不错喝。	石头汤底	不错喝	5.88 - 5.5
R1166	汤底实在太油。	汤底	实在太油	3.67 - 6.0
R1168	肉不错吃。	肉	不错	5.62 - 5.12
R1169	蕃薯椪外皮酥软。	蕃薯椪外皮	酥软	6.0 - 5.0
R1170	着肉口感爽滑细嫩。	着肉口感	爽滑细嫩	6.25 - 5.88
R1173	店内每张桌子都不大。	桌子	不大	4.75 - 5.0
R1174	牛奶巧克力超有层次感。	牛奶巧克力	超有层次感	6.88 - 7.0
R1174	冰摩卡味道很特别。	冰摩卡味道	很特别	5.67 - 6.0
R1174	冰摩卡很有层次。	冰摩卡	很有层次	6.0 - 6.0

Figure 6. some content of the ABSA dataset

4.2. Experimental Setting

Our project primarily aims to properly extract sentiment-based triplets and quadruplets from individual phrases using our recently created AWA-GCN model. This endeavor is particularly difficult because of the inherent intricacy of Chinese writing, where subtle differences in meaning are of utmost importance. A model with the ability to comprehend and analyze subtle linguistic cues is required to handle structured outputs, such as aspects, views, intensities, and categories.

The first phase in our experimental setup requires careful data preparation, where each input phrase goes through a number of preprocessing processes to standardize and optimize the text for future processing. This involves the process of segmentation, which aims to accurately detect and separate phrases and individual words in Chinese text despite the lack of unambiguous delimiters. After the segmentation process, every phrase is divided into individual tokens. The tokens are inputted into a transformer-based BERT to produce dense vector representations that capture both the contextual and semantic subtleties contained in the text.

The AWA-GCN model combines graph convolutional networks with attention methods to improve the model's capacity to concentrate on relevant text parts dynamically. This configuration enables a subtle comprehension and extraction of sentiment-related elements. The model undergoes fine-tuning using a tailored training regimen incorporating a precisely regulated learning rate to enhance both the convergence speed and the model's accuracy. In order to address the disparity in training data often seen in natural language problems, we use several strategies, including weighted loss functions, to guarantee that less frequent but important categories are accurately identified.

The training phase consists of numerous epochs during which the model acquires the ability to recognize and categorize the intricate connections among various components of the phrases. We use a batch size that strikes a compromise between computational efficiency and the model's capacity to generalize from the training data. After each epoch, there is a validation phase when the model's performance is evaluated on a distinct subset of the data that was not used for training. This process aids in optimizing the hyperparameters and modifying the model architecture to enhance performance.

We perform our experiments using a hardware configuration that includes GPUs to accelerate the training process due to the high computing demands of the tasks and the deep learning models used. To guarantee the repeatability of our findings, we use fixed seeds for the random number generators utilized

in the data shuffling and model initialization procedures. To provide consistent performance and avoid differences in execution settings, the software environment is standardized throughout all studies.

This experimental setting aims to thoroughly assess our AWA-GCN model's skills in extracting organized sentiment information from textual input.

4.3. Evaluation Metrics

The classifier's performance is evaluated using precision, recall, and F1 score statistics. These metrics provide a comprehensive assessment of the model's ability to identify sentiment components within the text correctly.

First, the valence and arousal values are rounded to an integer. A triplet or quadruplet is regarded as correct if and only if all three or four elements and their combination match those in the gold triplet or quadruplet. On this basis, we calculate the Precision, Recall, and F1 score as the evaluation metrics.

Each metric for the valence and arousal dimensions is calculated and ranked either independently or in combination. Precision is defined as the percentage of triplets or quadruplets extracted by the system that are correct. Recall is the percentage of triplets or quadruplets present in the test set found by the system. The F1 score is the harmonic mean of precision and recall. All metrics range from 0 to 1, with higher Precision, Recall, and F1 scores indicating more accurate performance.

The formulas for these metrics are as follows:

$$Precision = \frac{TP}{TP + FP} \tag{28}$$

$$Recall = \frac{TP}{TP + FN} \tag{29}$$

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{30}$$

Here, TP (True Positives) represents the number of correctly identified triplets or quadruplets, FP (False Positives) denotes the number of incorrect triplets or quadruplets identified, and FN (False Negatives) represents the number of correct triplets or quadruplets that the model did not identify.

4.4. Baseline

Our work involves a comparison of the performance of our AWA-GCN model with other existing models, including EMC-GCN [32] and a set of GTS-based approaches: GTS-BERT, GTS-Transformer, GTS-CNN, and GTS-BiLSTM [3]. The selection of these models is based on their suitability for effectively addressing intricate sentiment analysis problems, namely in extracting organized sentiment information from text. We duplicated the techniques used by these organizations and customized them to suit the specific demands of aspect-based sentiment analysis in Chinese literature.

The main objective of EMC-GCN, which stands for Entity-Mention Construction with Graph Convolutional Networks, is to establish explicit connections between entity mentions and create graph-based representations. These representations are then processed using a GCN. This architecture is specifically designed to improve the process of capturing and analyzing relational data inside texts, hence enabling more precise extraction of emotion and entity relationships.

GTS Models use a Grid Tagging Scheme to tackle aspect-based sentiment analysis. The GTS framework is specifically built to predict aspect and opinion words, their relations, and sentiment expressions straight from the text. This method avoids the need to break down the job into smaller subtasks, which helps to prevent the usual problem of error propagation in pipeline techniques.

GTS-BERT incorporates the BERT architecture into the GTS framework, using its deep contextual embeddings to improve the precision of sentiment triplet and quadruplet extraction. This model performs very well in situations when a detailed comprehension of sentence structure and meaning is of utmost importance.

GTS-Transformer is a variant of GTS-BERT that employs a conventional Transformer design, specifically emphasizing self-attention mechanisms for handling textual input. This model excels at capturing distant relationships within text, which is advantageous for thorough sentiment analysis.

GTS-CNN utilizes convolutional neural networks (CNNs) in the GTS framework to accurately detect certain characteristics in text data, such as spatial and temporal elements. This is particularly beneficial for extracting sentiment-related information from organized and semi-structured text.

The GTS-BiLSTM model integrates bidirectional Long Short-Term Memory networks to effectively process sequences by examining data points concerning both preceding and subsequent components. This is especially valuable in sentiment analysis since the context may greatly impact the understanding of the text.

These models serve as a strong foundation for assessing the effectiveness of our proposed AWA-GCN approach. Through several techniques, each possessing distinct strengths and approaches to sentiment analysis, our comparative research seeks to emphasize the practical and theoretical gains that our model provides compared to current alternatives.

4.5. Ablation Study

Our work included conducting an ablation study to analyze the individual impacts of certain components in the ABFF_Word_ATT GCN (AWA-GCN) model and determine whether the combined model is best. The model is specifically developed to extract sentiment triplets and quadruplets from Chinese text precisely. The model integrates a multi-layer attention fusion mechanism (ATT), parallel bi-affine processing, and a Chinese word segmentation (Word) technique that focuses on lexical aspects instead of individual word attributes, improving border recognition. This research aims to validate the need and efficacy of these components by assessing their influence on the model's performance when they are eliminated.

Our work explored the effectiveness of the AWA-GCN model (Aspect-Based Fine-grained Feature extraction with Word-level Attention) in overcoming the inherent difficulties of sentiment analysis, specifically in the setting of Chinese text. This model incorporates many sophisticated strategies aimed at enhancing the level of contextual comprehension and refining the accuracy of sentiment analysis.

The AWA-GCN model introduces a multi-layer attention fusion mechanism as one of its core innovations. This mechanism addresses the limits of layer-specific attention often present in typical designs. This method allows the model to capture various contextual interactions across the whole network effectively. Furthermore, the model utilizes parallel Bi-Affine attention to enhance the identification and mapping of correlations among sentiment components, hence enhancing the precision of sentiment analysis. Importantly, the model also includes improved segmentation approaches specifically designed for Chinese text, focusing on extracting features at the word level rather than analyzing individual characters. This differentiation maintains the meaning of the text and improves the model's capacity to identify and delineate the boundaries of linguistic components precisely.

In order to assess the individual contributions of these components, we performed a series of ablation experiments. We systematically eliminated each component in these experiments and observed the subsequent effects on important performance metrics, including accuracy, recall, and F1-score. These metrics were evaluated for both triplet and quadruplet extraction tasks.

The findings indicated that the whole AWA-GCN model consistently obtained the top scores for all criteria, highlighting the usefulness of its integrated components. The exclusion of the multi-layer attention fusion resulted in a significant decrease in accuracy, recall, and F1 score, emphasizing the importance of multi-layer attention in capturing intricate textual linkages. In addition, when the word-level processing component was removed, there was a significant decline in performance measures. This notably impacted the model's capacity to effectively partition and parse complex Chinese text, taking into account both semantic and syntactic clues.

The ablation study's assessment measures, which demonstrate the quantifiable influence on performance, are presented in Table 2 and Table 3. It is evident that the decrease in performance seen

after deleting these components emphasizes their collective role in enhancing the resilience and precision of the model.

The ablation experiment conclusively demonstrated the significance of each component in the AWA-GCN model, showcasing their combined contribution to its exceptional performance in sentiment analysis tasks. This confirms that our model-building combination is the most optimal.

Table 2. Results of the Ablation Study for ASTE(Triplet)

Configuration	Precision (%)	Recall (%)	F1 Score (%)
Full Model	66.39	81.82	73.31
- ATT	65.74	81.01	72.58
- ATT, - Word	65.72	80.32	72.29

'-xxx' means to remove the corresponding component 'xxx'

Table 3. Results of the Ablation Study for ASQE(Quadruplet)

Configuration	Precision (%)	Recall (%)	F1 Score (%)
Full Model	63.11	77.77	69.68
- ATT	62.61	77.15	69.12
- ATT, - Word	62.35	76.21	68.59

'-xxx' means to remove the corresponding component 'xxx'

4.6. Case Study

This case study showcases the efficacy of several sentiment analysis models, including our AWA-GCN, in effectively dealing with the complexities of the Chinese language for both triplet and quadruplet extraction tasks in food reviews. In Figure7, we marked important aspect phrases (such as "lion's head") and opinion terms that describe the quality and quantity of food, as well as the eating experience, in the sample sentences. These terms are highlighted with distinct colored blocks. The model differentiates these phrases by highlighting them in blue and yellow and represents the relationships between them using red lines to signify attitudes and entities. Our AWA-GCN model excels in the comparative sentiment analysis of food reviews because of its precise identification and correlation of aspect phrases and sentiment expressions. This enables the formation of sentiment triplets or quadruplets that include aspect categories and intensity measures.

AWA-GCN performs well in accurately determining the intensity and category of emotions, with two category labels predicted correctly in the given example phrases, where the valence arousal value of "food# portion and style" is 6.5#6.38, which is very close to the given answer of 6.54#6.23. It gets a very good score after rounding according to the task accuracy evaluation requirements. The same is true for the two "food# quality" labels, which the model evaluates correctly. This high accuracy shows that the model can understand Chinese's complex expressions and deep semantic connections. Chinese writing often lacks clear word boundaries and relies on context for interpretation, so it has these characteristics.

The comparison examination also exposes the constraints of other models. For example, GTS-BERT is skilled at collecting overall sentiment but has difficulty recognizing opinion phrases far removed in context, such as "soft and juicy." These terms are essential for a thorough sentiment analysis. While skilled at extracting features, the GTS-CNN model tends to disregard subtle connections between aspect

and opinion words, particularly when they are not located adjacent to each other in the text. The GTS-BLSTM model, which can capture sequential dependencies, may not be able to identify sentiment-bearing triplets that are not positioned in a linearly contiguous manner within the text. Although EMC-GCN has the capability to construct connections based on sentiment relation and linguistic traits, it may not always be able to capture the range of sentiment expressions fully, especially when opinion phrases are indicated rather than clearly mentioned.

On the other hand, AWA-GCN is more proficient in identifying and linking far-reaching aspects and opinion phrases. For example, it can establish a connection between "very amazingly big" and "lion head" even if they are separated in the text. It is worth emphasizing that the model's complex attention process enables it to capture the implicit sentiment in the word "very popular," which other models may ignore. Because in Chinese expressions, there are cases where the subject is ignored, just like the first sentence in the example sentence: "because it is really popular." this sentence is very colloquial and does not indicate what is really popular. Other models can easily fail to recognize the adjective "very popular." In contrast, our model can recognize it and intelligently indicate that this hidden aspect class is NULL when predicting the classification, which aligns with the task requirements. In summary, the architecture of AWA-GCN is very good at handling the complexity of sentiment expression in food reviews, where sentiment is often conveyed through a mixture of explicit and implicit language.
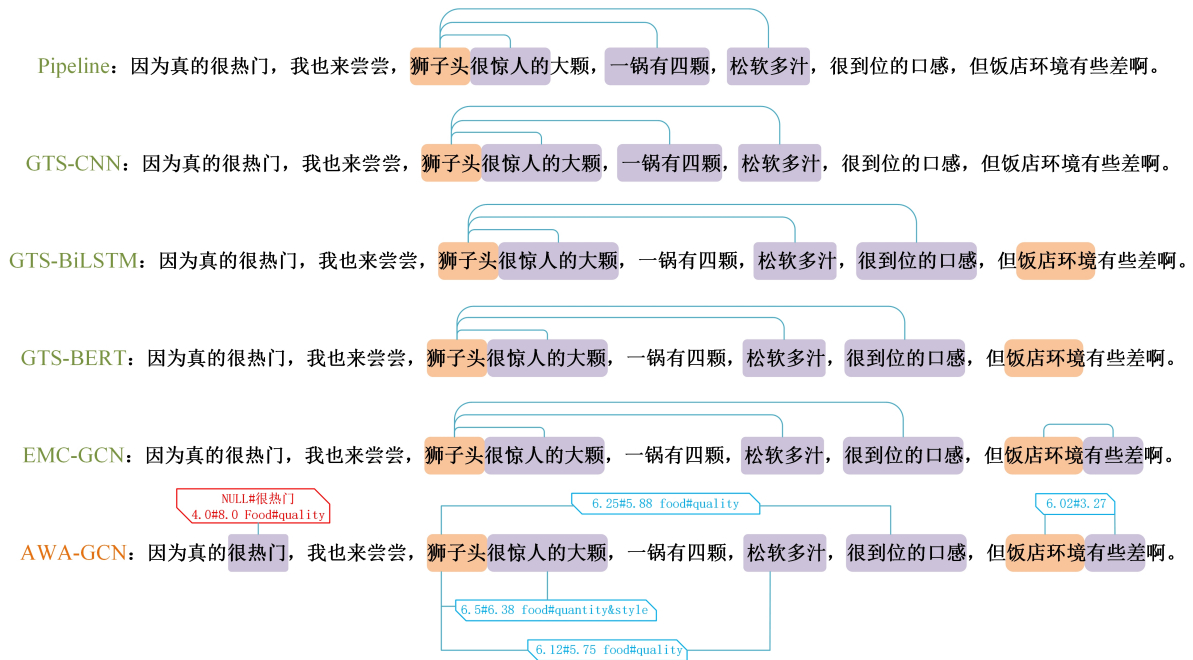

Figure 7. Different models outputs for a given sentence

4.7. Results

The results for each model we compared are presented in Tables 4 and 5, directly comparing all evaluated metrics for both triplet and quadruplet extraction tasks. Our groundbreaking Attention-Weighted Affine GCN model surpasses conventional deep learning techniques and extraction methods in Sentiment recognition and extraction of Chinese text. Through comprehensive evaluation, including metrics such as Precision, Recall, and F1 score statistics, our model exhibits exceptional performance. While the Precision value (71.23%) is slightly lower than that of the GTS-BERT Model (72.02%) in the task of Sentiment triple extraction, when considering all metrics collectively, our model demonstrates outstanding performance with Recall=81.82% and F1=76.16% in the task of Sentiment triple extraction and Precision=69.11%, Recall=77.77% and F1=73.20% in the task of Sentiment quadruple extraction.

Notably, the key performance index F1 of our model in Sentiment quadruple extraction is 14.28% higher than that of the optimal F1 in the control group, and the index Recall of our model in Sentiment triple extraction is 9.32% higher than that of the optimal Recall in the control group. This highlights its innovative capability, offering a promising approach for precisely and effectively identifying depressive tendencies in textual data. The results clearly demonstrate that the AWA-GCN particularly excels in the more complex quadruplet extraction task. This performance underscores the effectiveness of our model's innovative architecture, which integrates advanced attention mechanisms and word-level processing to handle the intricacies of sentiment analysis in Chinese text.

Table 4. Performance of Sentiment Analysis Models on Triplet Extraction Task

Model	Precision (%)	Recall (%)	F1-Score (%)
AWA-GCN	<u>71.23</u>	**81.82**	**76.16**
EMC-GCN	61.39	<u>74.84</u>	67.45
GTS-BERT	**72.02**	68.43	<u>70.18</u>
GTS-Transformer	23.16	43.52	30.24
GTS-CNN	33.45	42.40	37.40
GTS-BiLSTM	50.43	43.77	46.87

(In this table, the best performance is highlighted in bold, and the second-best performance is underlined.)

Table 5. Performance of Sentiment Analysis Models on Quadruplet Extraction Task

Model	Precision (%)	Recall (%)	F1-Score (%)
AWA-GCN	**69.11**	**77.77**	**73.20**
EMC-GCN	54.90	<u>66.94</u>	60.33
GTS-BERT	<u>65.73</u>	62.45	<u>64.05</u>
GTS-Transformer	19.28	36.24	25.17
GTS-CNN	23.18	29.39	25.92
GTS-BiLSTM	40.38	35.06	37.53

(In this table, the best performance is highlighted in bold, and the second-best performance is underlined.)

5. Conclusion

This research has presented the AWA-GCN model as a resilient framework designed particularly to address the intricate difficulties of sentiment analysis in Chinese text. The model's advanced design, which combines multi-layer attention mechanisms with word-level processing, has proven to be highly effective in accurately parsing and analyzing intricate sentiment structures. These structures are intricately woven into the syntactic and semantic layers of the Chinese language.

Our model was evaluated using the dimABSA dataset, which was given as a component of the SIGHAN 2024 Shared Task for Chinese Dimensional Aspect-Based Sentiment Analysis (dimABSA). This dataset is essential because it has been carefully selected and organized to test and assess systems'

capability to do complex and diverse sentiment analysis, particularly in the setting of Chinese language text. The dataset comprises a wide range of utterances obtained from real-life situations, including intricate sentiment structures that need advanced analytical methods to interpret accurately.

Unlike English, which has explicit word delimiters like spaces, Chinese poses distinctive difficulties for sentiment analysis. The syntax of Chinese does not always adhere to the Subject-Verb-Object sequence. Instead, it incorporates other structures such as Subject-Predicate and Topic-Comment, which greatly complicates the task of detecting and interpreting mood. The architecture of the AWA-GCN model explicitly tackles these issues by improving the identification of lexical items, particles, and punctuation marks, which play a vital role in comprehending sentence structure and mood in Chinese. The attention processes of the model excel in detecting nuanced linguistic signals and merging contextual information over various segments of a sentence, which is crucial for languages like Chinese with flexible word order.

In addition, the AWA-GCN model proficiently manages idiomatic terms and phrases often found in Chinese, guaranteeing that these emotionally charged aspects are not disregarded. By prioritizing word-level characteristics instead of individual characters, the model preserves the integrity of Chinese morphological structures, resulting in more accurate identification of boundaries and categorization of mood. This strategy boosts the model's capacity to analyze intricate sentence patterns and improves its resilience in handling the complexities of Chinese syntax and grammar.

The usefulness of the AWA-GCN model is confirmed by its better metrics in both triplet and quadruplet sentiment extractions. Compared to baseline models like EMC-GCN[32] and other GTS[3] configurations, the AWA-GCN exhibits enhanced accuracy and recall and proves its thorough comprehension of the complex nature of emotional expression in Chinese text.

Overall, the AWA-GCN model establishes a higher benchmark for sentiment analysis in the Chinese language, presenting significant implications for future uses.

Appendices

For those who are interested in viewing or need to reproduce the improvements, the source code of the specific AWA-GCN model and dimABSA dataset can be found on GitHub at the following URL: *https://github.com/SpikeShaun/AWA-GCN-for-MLA2024*. We sincerely welcome your valuable suggestions and hope you will gain innovative inspiration.

References

[1] Maria Pontiki, Dimitris Galanis, Haris Papageorgiou, Ion Androutsopoulos, Suresh Manandhar, Mohammed AL-Smadi, Mahmoud Al-Ayyoub, Yanyan Zhao, Bing Qin, Orphée De Clercq, et al. Semeval-2016 task 5: Aspect based sentiment analysis. In *ProWorkshop on Semantic Evaluation (SemEval-2016)*, pages 19–30. Association for Computational Linguistics, 2016.

[2] Caroline Brun and Vassilina Nikoulina. Aspect based sentiment analysis into the wild. In *Proceedings of the 9th workshop on computational approaches to subjectivity, sentiment and social media analysis*, pages 116–122, 2018.

[3] Zhen Wu, Chengcan Ying, Fei Zhao, Zhifang Fan, Xinyu Dai, and Rui Xia. Grid tagging scheme for aspect-oriented fine-grained opinion extraction. *arXiv preprint arXiv:2010.04640*, 2020.

[4] Bo Pang, Lillian Lee, et al. Opinion mining and sentiment analysis. *Foundations and Trends® in information retrieval*, 2(1–2):1–135, 2008.

[5] Minqing Hu and Bing Liu. Mining and summarizing customer reviews. In *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 168–177, 2004.

[6] Lu Xu, Hao Li, Wei Lu, and Lidong Bing. Position-aware tagging for aspect sentiment triplet extraction. *arXiv preprint arXiv:2010.02609*, 2020.

[7] Haiyun Peng, Lu Xu, Lidong Bing, Fei Huang, Wei Lu, and Luo Si. Knowing what, how and why: A near complete solution for aspect-based sentiment analysis. In *Proceedings of the AAAI conference on artificial intelligence*, volume 34, pages 8600–8607, 2020.

[8] Lianzhe Huang, Peiyi Wang, Sujian Li, Tianyu Liu, Xiaodong Zhang, Zhicong Cheng, Dawei Yin, and Houfeng Wang. First target and opinion then polarity: Enhancing target-opinion correlation for aspect sentiment triplet extraction. *arXiv preprint arXiv:2102.08549*, 2021.

[9] Hongjiang Jing, Zuchao Li, Hai Zhao, and Shu Jiang. Seeking common but distinguishing difference, a joint aspect-based sentiment analysis model. *arXiv preprint arXiv:2111.09634*, 2021.

[10] Zhexue Chen, Hong Huang, Bang Liu, Xuanhua Shi, and Hai Jin. Semantic and syntactic enhanced aspect sentiment triplet extraction. *arXiv preprint arXiv:2106.03315*, 2021.

[11] Hai Wan, Yufei Yang, Jianfeng Du, Yanan Liu, Kunxun Qi, and Jeff Z Pan. Target-aspect-sentiment joint detection for aspect-based sentiment analysis. In *Proceedings of the AAAI conference on artificial intelligence*, volume 34, pages 9122–9129, 2020.

[12] Mengting Hu, Yike Wu, Hang Gao, Yinhao Bai, and Shiwan Zhao. Improving aspect sentiment quad prediction via template-order data augmentation. *arXiv preprint arXiv:2210.10291*, 2022.

[13] Hua Zhang, Xiawen Song, Xiaohui Jia, Cheng Yang, Zeqi Chen, Bi Chen, Bo Jiang, Ye Wang, and Rui Feng. Query-induced multi-task decomposition and enhanced learning for aspect-based sentiment quadruple prediction. *Engineering Applications of Artificial Intelligence*, 133:108609, 2024.

[14] Shen Zhou and Tieyun Qian. On the strength of sequence labeling and generative models for aspect sentiment triplet extraction. In *Findings of the Association for Computational Linguistics: ACL 2023*, pages 12038–12050, 2023.

[15] Li Yuan, Jin Wang, Liang-Chih Yu, and Xuejie Zhang. Encoding syntactic information into transformers for aspect-based sentiment triplet extraction. *IEEE Transactions on Affective Computing*, 2023.

[16] Chen Zhang, Qiuchi Li, Dawei Song, and Benyou Wang. A multi-task learning framework for opinion triplet extraction. *arXiv preprint arXiv:2010.01512*, 2020.

[17] Xuefeng Shi, Min Hu, Jiawen Deng, Fuji Ren, Piao Shi, and Jiaoyun Yang. Integration of multi-branch gcns enhancing aspect sentiment triplet extraction. *Applied Sciences*, 13(7):4345, 2023.

[18] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*, 2018.

[19] Chi Sun, Luyao Huang, and Xipeng Qiu. Utilizing bert for aspect-based sentiment analysis via constructing auxiliary sentence. *arXiv preprint arXiv:1903.09588*, 2019.

[20] Wenxuan Zhang, Yang Deng, Xin Li, Yifei Yuan, Lidong Bing, and Wai Lam. Aspect sentiment quad prediction as paraphrase generation. *arXiv preprint arXiv:2110.00796*, 2021.

[21] Hongjie Cai, Rui Xia, and Jianfei Yu. Aspect-category-opinion-sentiment quadruple extraction with implicit aspects and opinions. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 340–350, 2021.

[22] Qiao Sun, Liujia Yang, Minghao Ma, Nanyang Ye, and Qinying Gu. Rethinking aste: A minimalist tagging scheme alongside contrastive learning. *arXiv preprint arXiv:2403.07342*, 2024.

[23] Zhaoyang Niu, Guoqiang Zhong, and Hui Yu. A review on the attention mechanism of deep learning. *Neurocomputing*, 452:48–62, 2021.

[24] Dat Quoc Nguyen and Karin Verspoor. End-to-end neural relation extraction using deep biaffine attention. In *Advances in Information Retrieval: 41st European Conference on IR Research, ECIR 2019, Cologne, Germany, April 14–18, 2019, Proceedings, Part I 41*, pages 729–738. Springer, 2019.

[25] Seung-Hoon Na, Jinwoon Min, Kwanghyeon Park, Jong-Hun Shin, and Young-Gil Kim. Jbnu at mrp 2019: Multi-level biaffine attention for semantic dependency parsing. In *Proceedings of the shared task on cross-framework meaning representation parsing at the 2019 conference on natural language learning*, pages 95–103, 2019.

[26] Long-Shou GAO and Na-Na LI. Aspect sentiment triplet extraction based on aspect-aware attention enhancement. *Journal of Computer Applications*, page 0, 2023.

[27] Jie Hu, Li Shen, and Gang Sun. Squeeze-and-excitation networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 7132–7141, 2018.

[28] Jin-Seong Kim, Sung-Wook Park, Jun-Yeong Kim, Jun Park, Jun-Ho Huh, Se-Hoon Jung, and Chun-Bo Sim. E-hrnet: Enhanced semantic segmentation using squeeze and excitation. *Electronics*, 12(17):3619, 2023.

[29] Wei-Lin Chiang, Xuanqing Liu, Si Si, Yang Li, Samy Bengio, and Cho-Jui Hsieh. Cluster-gcn: An efficient algorithm for training deep and large graph convolutional networks. In *Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining*, pages 257–266, 2019.

[30] Pinlong Zhao, Linlin Hou, and Ou Wu. Modeling sentiment dependencies with graph convolutional networks for aspect-level sentiment classification. *Knowledge-Based Systems*, 193:105443, 2020.

[31] Jincheng Mei, Chenjun Xiao, Csaba Szepesvari, and Dale Schuurmans. On the global convergence rates of softmax policy gradient methods. In *International conference on machine learning*, pages 6820–6829. PMLR, 2020.

[32] Hao Chen, Zepeng Zhai, Fangxiang Feng, Ruifan Li, and Xiaojie Wang. Enhanced multi-channel graph convolutional network for aspect sentiment triplet extraction. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 2974–2985, 2022.

[33] Makoto Miwa and Yutaka Sasaki. Modeling joint entity and relation extraction with table representation. In *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*, pages 1858–1869, 2014.