Multi-Step Cross-Domain Aspect-Based Sentiment Generation with Error Correction Mechanism

Ningning Mao[1], Xuanliang Zhu[2], Yadi Xu[3]*
School of Digital Business, Zhejiang Dongfang Polytechnic, Wenzhou, China[1, 2]
School of Information Engineering, Wenzhou Business College, Wenzhou, China[3]
School of Computer Science, Universiti Sains Malaysia, Penang, Malaysia[3]

***Abstract*—With the rapid growth of social media and user-generated content, cross-domain aspect-level sentiment analysis has become an important research direction in sentiment computing. In this study, a cross-domain sentiment analysis method based on the T5 model is proposed. This method integrates a multi-step generative training mechanism with a correction mechanism to improve the model's generalization ability and sentiment classification accuracy when processing texts from different domains. First, domain-invariant sentiment features are extracted through training on texts and their associated aspect vocabularies from both the source and target domains. This process effectively reduces inter-domain discrepancies. Unlike other methods, the generative task is formulated in the source domain to produce both aspect and sentiment element pairs, which improves the model's reasoning ability through multi-step generation. Finally, a correction mechanism is used to detect the aspect labels in the generated labels of the target domain and regenerate the sentiment predictions when errors are detected, which improves the model's robustness. Experimental results show that the proposed method performs well in several cross-domain sentiment analysis tasks and significantly outperforms traditional methods in sentiment classification accuracy. The study provides an innovative solution for cross-domain sentiment analysis with broad application potential.**

***Keywords*—Cross-domain aspect-based sentiment analysis; multi-step generation; correction mechanism; domain-invariant feature learning**

I. Introduction

With the popularity of social media, online commenting platforms and e-commerce, Sentiment Analysis (SA) has emerged as a key research direction in the field of Natural Language Processing (NLP) [1]. Especially, the fine-grained sentiment analysis, Aspect-Level Sentiment Analysis, can identify the sentiment tendency of specific aspects in the text, which shows a wide range of application prospects in the application scenarios, such as market research and customer feedback analysis. However, in practical scenarios, sentiment data are often distributed in multiple domains, making it difficult for models to maintain high accuracy in cross-domain sentiment analysis tasks [2]. For example, in the user review "The sound quality of this smart speaker is excellent, with particularly good bass performance.", traditional sentiment analysis may simply classify the entire review as positive. In contrast, Aspect-Based Sentiment Analysis (ABSA) can further identify "sound quality" as a specific aspect and correctly determine that the sentiment expressed by "excellent" is positive as shown in Fig. 1. In single-domain sentiment analysis tasks, models can effectively learn the sentiment association between "sound quality" and "excellent" through a large amount of labeled data, enabling accurate predictions. However, in cross-domain sentiment analysis tasks, differences in data distribution may hinder a model's ability to generalize effectively. For instance, if the training data primarily comes from the laptop domain, where "quality" often appears in contexts such as "material quality" or "display quality", the model may associate "quality" with hardware durability or screen performance. When applied to the smart speaker domain, the model may fail to recognize "sound quality" as an aspect related to audio performance and instead misinterpret it based on its prior knowledge from the laptop domain. Additionally, sentiment words like "excellent" may carry different implications across domains, further complicating accurate classification. Such domain discrepancies can lead to a decline in classification accuracy. Therefore, enhancing a model's adaptability across different domains and enabling it to learn domain-invariant features has become a central challenge in cross-domain sentiment analysis research.

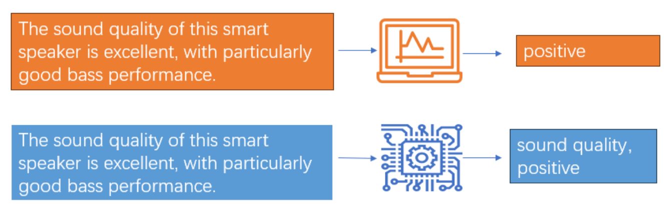

Fig. 1. Traditional Sentiment Analysis versus Aspect-Based Sentiment Analysis

Traditional Aspect-Level Sentiment Analysis models have good performance in in-domain tasks. However, their performance tends to degrade significantly when applied to cross-domain tasks due to the differences in features and linguistic expressions between different domains [3] (Chen et al., 2020). These inter-domain differences are mainly reflected in two aspects. First, the ASPECTS of different domains can vary substantially, which makes the model difficult to generalize [4]. Second, the sentiment expressions of different domains are often different, and even the same ASPECTS may have different sentiment tendencies in different domains [5]. Therefore, effectively performing cross-domain Aspect-Level Sentiment Analysis has become an important and challenging problem.

*Corresponding Author.

Current research on cross-domain Aspect-Level Sentiment Classification (ASC) has achieved notable progress in the field of sentiment analysis. However, several critical challenges persist. Inter-domain textual data differ significantly in terms of vocabulary, grammatical structure, and sentiment expression, which makes models trained in the source domain perform poorly in the target domain. In order to improve the generalization ability of the models, many existing approaches focus on aligning the feature representations of the source and target domains. By analyzing several major current approaches, the innovations and shortcomings are identified, thereby helping to clarify the research gaps.

First, the cross-domain ASC approach based on adversarial distribution alignment focuses on solving the problem of feature distribution differences between domains. Through techniques such as gradient inversion layers, these methods aim to learn feature representations shared between source and target domains, thus reducing the probability of cross-domain misclassification [4]. However, a significant limitation of these methods is their inability to explicitly handle specific aspect words, which often play a crucial role in sentiment classification tasks. The lack of explicit labeling of aspect words leads to bias in capturing fine-grained sentiment information, especially in cases where domain-specific words are present in the target domain, and the model is unable to accurately make sentiment predictions. Efficient Adaptive Transfer Network (EATN) learns common features between the source and target domains by introducing a multi-attention mechanism and a domain adversarial module while capturing direct associations between aspect words and contextual sentiment words [5]. Although this approach achieves notable success in domain alignment and fine-grained sentiment analysis, it has a key limitation: the absence of explicit tagging for aspect words in the input text. Consequently, the optimization of the attention mechanism still focuses mainly on the global context and fails to impose higher attention weights on specific aspect words.

Second, BERT-based cross-domain ASC approaches extract sentence-level and aspect-level representations by using pre-trained language models such as BERT. These methods often incorporate domain-adversarial training to reduce feature differences between the source and target domains. While this type of approach solves the domain adaptation problem to some extent, the attention mechanism of BERT focuses more on the overall context rather than optimizing specifically for aspect words [6]. Since aspect words are very important in sentiment analysis, the BERT model may suffer when dealing with fine-grained sentiment analysis, especially when transforming between different domains, resulting in performance degradation.

In addition, cross-domain ASC approaches based on semantic key feature extraction aim to enhance the cross-domain migration capability of the model by introducing an external knowledge graph. Such approaches utilize semantic knowledge graphs as a bridge for knowledge migration from the source domain to the target domain, enhancing the understanding of sentiment words and syntactic structures [7]. However, while external knowledge graphs can enrich semantic information, their reliance on external resources rather than in-depth analysis of the input text may limit the model's ability to capture the diversity of sentiment expressions in certain domains. In particular, when there are significant feature differences between domains, models relying on external knowledge may struggle to fully utilize aspect words and contextual information in internal data.

From these existing studies, it can be seen that most current cross-domain ASC approaches focus on domain alignment and feature migration [8][9]. However, they remain inadequate in dealing with specific aspect word tokens, dynamic optimization of sentiment features, and error correction mechanisms [10]. Existing methods cannot improve the inference of generative models, which leads to the susceptibility to biased sentiment prediction in cross-domain tasks, especially fine-grained aspect sentiment analysis. In addition, the models also lack effective error correction mechanisms to re-generate the sentiment output when the aspect word prediction is wrong. Misclassification and sentiment judgment errors are often the significant challenges in cross-domain sentiment analysis due to the large differences in the characteristics of different domains.

To address these issues, an innovative cross-domain ASC approach based on the Text-to-Text Transfer Transformer(T5) generative model is proposed. The generative framework of the T5 model provides a novel perspective for tackling the ASC task. With the T5 model, the sentiment categorization task can be handled in a generative way, which is no longer limited to the traditional categorization framework. By combining textual and aspectual information, T5 generate more accurate sentiment predictions across domains. Moreover, under the multi-step inference mechanism, T5 efficiently extract domain-invariant features, thus improving the accuracy of sentiment generation.

To further enhance the robustness of the model, an error correction mechanism is introduced in this study. This mechanism determines the errors in sentiment prediction by comparing the generated aspect words with the known aspect words in the target domain and re-generates them. The error correction mechanism can effectively reduce the sentiment prediction errors due to aspect word errors, thus significantly improving the generalization ability of the model across different domains. Moreover, this mechanism offers a novel approach to error handling in cross-domain sentiment analysis. In particular, for multi-domain sentiment classification, it can generate more accurate sentiment results for different aspect words.

In addition to the cross-domain aspect-level sentiment analysis task (Aspect-Level Sentiment Classification, ASC), other tasks such as Aspect-Opinion Pair Extraction (AOTE), Aspect-Sentiment Pair Extraction (ASPE), and Aspect-Sentiment Triplet Extraction (ASPE) have started to use generative models such as T5 [11]. The core of these tasks is the simultaneous recognition of aspect words, sentiment polarity, and associated opinion words or syntactic structures. Generative models, such as T5, excel at handling multiple tasks through a unified text generation framework, enabling end-to-end generation of the desired outputs. In these tasks, generative models demonstrate a high degree of flexibility and reasoning ability to adapt to different input formats and task goals. However, due to the inherent differences between these tasks and ASC, it is difficult to apply existing methods directly to ASC subtasks.

Tasks such as AOTE, ASPE, and ASTE involve relatively complex aspects and affective associations, requiring the simultaneous identification of relationships among multiple elements. Generative models such as T5 are good at handling such multi-task settings. However, in the ASC task, sentiment analysis focuses more on judging the sentiment polarity of specific aspectual words rather than multiple triad extraction [12]. This makes it difficult for simple generative methods to accurately capture the fine-grained associations between domain-specific aspect words and sentiment in ASC tasks. In addition, although the T5 model performs well in multi-domain tasks, it tends to ignore the specificity of aspect words in cross-domain sentiment analysis because it is not optimized for ASC tasks.

To address this problem, a Multi-step inference mechanism is introduced by combining the characteristics of the ASC task. By generating aspect words, the model gains a deeper understanding of the context, enhancing the ability to generate sentiment labels. This transformation of ASC into generating aspect and sentiment combination pairs enhances the model's ability to learn as well as reason about domain-invariant features, enabling T5 to generate sentiment predictions more efficiently when processing texts from different domains. This not only improves the model's cross-domain generalization ability, but also makes it perform better in fine-grained sentiment classification.

This study combines the generative capability of the T5 model with the characteristics of the ASC task. Additionally, a multi-step inference and error correction mechanism is introduced. This not only contributes to the learning of domain-invariant features, but also improves the performance of the generative model in the application of cross-domain sentiment analysis tasks. This research fills the shortcomings of existing approaches and provides new directions for future cross-domain ASC tasks. In summary, a cross-domain Aspect-Level Sentiment Analysis model, based on the T5 model and incorporating a Multi-step Generation with Error Correction Mechanism, is proposed in this study (TSG-ECM). The main contributions are as follows:

- Multi-step Generation Training Mechanism: Firstly, by training the text and ASPECT vocabulary of the source and target domains, the model learns domain-invariant generalized features, which makes it better adapted to cross-domain text analysis. This step allows it to identify common features between different domains, and improves its generalization ability in the target domain. Next, when training in the source domain, the T5 model is allowed to generate both aspect words and sentiment labels. The model learns to predict both aspect and sentiment in the text through end-to-end generation, resulting in a powerful inference capability. In cross-domain tasks, the simultaneous output of aspect words and sentiment polarity facilitates capturing features that are common across domains and enhances its generalization ability in the target domain.
- Target Domain Correction Mechanism: This study introduces a generation process based on a correction mechanism. When the aspect generated by the model does not align with the known aspect library of the target domain, the system detects the error and regenerates the aspect and sentiment labels. By introducing error correction signals and new generation loss targets, the model is able to effectively adjust the generation strategy in case of incorrect prediction, thus improving the accuracy of sentiment prediction in the target domain.

The rest of the study is structured as follows: Section II reviews related work, including Aspect-Level Sentiment Analysis, cross-domain learning, and the application of the correction mechanism. Section III describes the proposed method in detail, including the multistep inference mechanism, source-domain ASPECT generation, and the target-domain error-correction mechanism. Section IV presents the experimental design and results, which demonstrate the method's performance on a number of domains and compares it with existing methods. Section V presents the discussion. Finally, Section VI summarizes the main contributions and discusses future research directions.

II. Related Research

This section reviews the progress of related research in cross-domain Aspect-Level sentiment analysis, domain-invariant feature learning, generative modeling, and error correction mechanisms. By summarizing and analyzing the existing methods, the innovations and advantages of the proposed methods in this study are clarified.

A. *Aspect-Level Sentiment Analysis*

The development of Aspect-Level Sentiment Analysis (ABSA) has evolved through several stages. Early traditional machine learning models, such as Support Vector Machine (SVM) and Plain Bayes, mainly relied on manually designed features (e.g., bag-of-words model and TF-IDF) to distinguish sentiment categories. However, these methods perform poorly when dealing with cross-domain sentiment tasks, and it is difficult to effectively deal with feature differences and sentiment expressions across domains [13]. With the rise of deep learning, Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) are becoming mainstream. These networks are able to learn semantic and contextual relationships in text in an end-to-end manner, thereby overcoming the limitations of manual feature design, thus significantly improving the effectiveness of sentiment analysis [14]. In the context of aspect-level sentiment analysis tasks, RNNs are particularly effective at capturing dependencies in sentences using sequential information, while CNNs can effectively extract local sentiment features [15]. Nevertheless, the generalization ability of these deep learning methods remains limited in cross-domain scenarios, making it challenging to address the distributional discrepancies across domains [16]. In recent years, models based on the Transformer architecture (e.g., BERT and T5) have become the core technology for cross-domain sentiment analysis. BERT, as a pre-trained model that captures the contextual information of a sentence through the bidirectional attention mechanism, has made significant progress in several sentiment analysis tasks [17]. However, BERT encounters challenges in cross-domain sentiment analysis, particularly due to its domain-specific nature and weak feature alignment capabilities [18].

B. Cross-Domain Sentiment Analysis with Domain-Invariant Feature Learning

In order to address the differences between different domains, researchers have proposed a variety of cross-domain learning methods, including Domain-Invariant Feature Learning (DIFL) as well as Domain-Adversarial Learning (DAL) and other domain adaptation methods. The core idea of domain adaptation is to improve the generalization ability of the model in the target domain by reducing the difference between the data distributions of the source and target domains [19].

Some studies have enabled models to capture common features between domains by introducing a shared feature space. For example, Xue et al. (2018) proposed a method based on Gradient Reversal Layer (GRL) to improve the performance of cross-domain sentiment classification. This method aligns the data distributions of the source and target domains through adversarial training in a shared feature space [20]. However, this type of approach is not directly applicable in Aspect-Level tasks, because aspects of different domains tend to differ significantly. Simply aligning the distributions does not guarantee the accuracy of aspect prediction.

To address this limitation, several studies have proposed to enhance cross-domain learning by separating domain-specific features from domain-invariant features [21]. These approaches enhance the robustness of the model across domains by introducing an independent domain-invariant feature extraction module into the model structure, which is then combined with domain-specific feature learning. However, such methods still struggle to cope with the problem of generating incorrect ASPECTS in the target domain, and incorrect ASPECTS generation often leads to chain errors in sentiment categorization.

C. Application of Error Correction Mechanisms in Generative Tasks

Although generative models exhibit strong text generation capabilities, incorrectly generated aspects can introduce significant bias in sentiment prediction, particularly in cross-domain sentiment analysis tasks. Therefore, Error Correction Mechanism (ECM) is proposed to reduce the impact of generation errors on model performance. It is widely used in tasks such as machine translation, text summarization and dialog generation [22], where errors are corrected by post-processing or re-generating the generated results.

In the task of sentiment analysis, the application of error correction mechanisms remains relatively limited. Dong et al. (2019) proposed an error correction-based generation framework, where the model re-generates if it detects incorrect words or phrases during the generation process, thereby reducing the impact of errors on the final task [23]. Although this correction strategy can improve the quality of generation to some extent, there is currently no well-defined mechanism capable of effectively addressing the problem of erroneous aspect generation in cross-domain Aspect-Level sentiment analysis.

D. Research Innovations

Although existing research has made some progress in aspect-based sentiment analysis, cross-domain sentiment analysis and generation modeling, there are still significant limitations. First, most existing cross-domain sentiment analysis methods rely on the learning of domain-invariant features, but lack effective generation and correction mechanisms for the differences in aspect across domains. Second, while generative models perform well in sentiment analysis tasks within a single domain, they often fail to capture the relationship between aspects and opinions in cross-domain tasks, generating erroneous sentiment analysis results.

In this study, a multi-step generative inference mechanism combined with a correction mechanism is proposed. (Cross-Domain Aspect-Sentiment Generation with Error Correction (CDASG-EC)) It aims to improve the model's performance in cross-domain tasks through domain-invariant feature learning, generative lossy objective optimization, and an error correction mechanism accuracy. Compared with existing methods, the innovations of this study are as follows:

- By introducing the T5 model for domain-invariant feature learning, the proposed approach enhances the adaptability of the model across diverse domains.
- In the source domain, the T5 model is trained by inputting text in multiple steps and outputting (ASPECT, SENTIMENT) to enhance the model's inference ability and thus generate more accurate sentiment classification results. This strategy represents the first application of a generative T5 model in a cross-domain ASC task, enabling the simultaneous generation of aspects and sentiments in an end-to-end manner. It provides the model with stronger domain-invariant feature learning and generalization capabilities.
- A correction mechanism is introduced to detect and rectify generated aspect and sentiment labels in the target domain, thus reducing the error rate of the model in the cross-domain tasks.

III. METHODOLOGY

A multi-step generative inference mechanism based on the T5 model, combined with a correction mechanism (Cross-Domain Aspect-Sentiment Generation with Error Correction (CDASG-EC)), is proposed in this study. It aims to improve the accuracy and robustness of the cross-domain sentiment analysis task. The design of the model is divided into two core parts: a multi-step generative inference mechanism (domain invariant feature learning, source domain aspect and sentiment generation training) and a target domain correction mechanism.

The overall model is shown in Fig. 2. The structure and training process of each part are described in detail below:

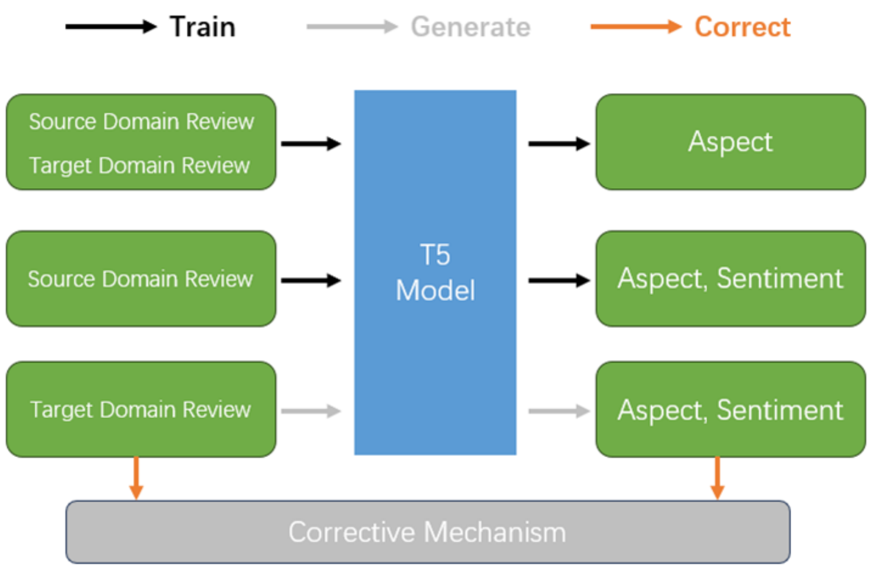

Fig. 2. Overview of Cross-Domain Aspect-Sentiment Generation with Error Correction (CDASG-EC).

A. Overview of the T5 Model

The T5 model is a unified text generation framework that can unify various NLP tasks into text generation problems [24]. In this study, it is used as a basis for handling the cross-domain Aspect-Level sentiment analysis task. Specifically, the input is a domain text, and the model generates the ASPECT vocabulary related to that text. Unlike traditional classification methods, the generative model performs well in capturing complex text structures and domain-generalized features for cross-domain tasks.

B. Multi-Step Generative Training Mechanism

1) Domain invariant feature learning: The main challenge of cross-domain sentiment analysis is the difference in data distribution between domains. To enable the transfer of sentiment analysis capabilities across different domains, domain-invariant feature learning is first performed. This process involves training the T5 model to learn transferable features by combining the text and ASPECT vocabularies from both the source and target domains. The specific steps are as follows:

a) Input data preparation: For texts from both the source and target domains, the input to the T5 model is formatted as [Text], where Text represents the raw text from either domain. The task of the T5 model is to generate the Aspect associated with Text, i.e., the specific vocabulary related to emotion in that text. The input format is [Text] and the generated output is [Aspect].

b) Loss function design: During training, the T5 model learns domain invariant features by minimizing the difference between the generated aspect and the true aspect. The cross-entropy loss function is employed to measure the difference between the generated aspect and the real aspect. The loss function is represented as follows:

$$L_{domian} = -\sum_{i=1}^{N}\left(y_i^{asp}\log\hat{y}_i^{asp} + \left(1-y_i^{asp}\right)\log\left(1-\hat{y}_i^{asp}\right)\right) \quad (1)$$

where, y_i^{asp} is the real ASPECT vocabulary, and $\hat{y}_i^{asp}$ is the probability distribution of the ASPECT vocabulary generated by the T5 model.

c) Domain-invariant feature extraction: During the training process, the model learns generic features from text data in the source and target domains. In this process, the T5 model captures not only the association between text and ASPECT, but also the common features between the source and target domains. This is achieved through a multi-task learning mechanism. This joint training ensures that the aspect generated by the model can be applied to texts in different domains.

2) Source domain aspect and sentiment generation training: After domain invariant feature learning is completed, the model is further trained on the source domain to generate aspect and corresponding sentiment labels. This step aims to ensure that the model learns to recognize aspect words in text and generate corresponding sentiment labels based on their context. In cross-domain tasks, this strategy enables the model to have stronger reasoning capabilities. The specific steps are as follows:

a) Source domain input and output format: During source domain training, the input of the model is the source domain text, denoted as [Text], and the generated output is [Aspect] and the corresponding sentiment label [Sentiment]. The model first generates the aspect and then predicts the sentiment label based on the generated aspect. The output is formatted as [Aspect Sentiment]. This strategy is the first of its kind to incorporate a generative T5 model in a cross-domain ASC task and generates both aspects and sentiments in an end-to-end manner, which is designed to provide the model with enhanced domain-invariant feature learning and generalization capabilities. In addition, this approach generates ASPECTS by inference in cross-domain contexts, which also reduces the dependence on a prior knowledge of the target domain and improves flexibility.

b) Source domain training objective: The source domain training objective is to minimize the difference between the generated sentiment labels and the true sentiment labels. Furthermore, if the ASPECT is wrong, the SENTIMENT loss is weighted. The loss function during training is:

$$L_{sentiment} = -\sum_{i=1}^{N} w_i^{asp} y_i^{sent} \log \hat{y}_i^{sent} \tag{2}$$

where, w_i^{asp} is a weight factor that depends on the correctness of the ASPECT. This weight is larger when the ASPECT is wrong and smaller when the ASPECT is correct. y_i^{sent} is the true sentiment label, and $\log \hat{y}_i^{sent}$ is the probability distribution of the sentiment label generated by the model.

C. Target Domain Correction Mechanism and Loss Objective Generation

When y_k^{asp} the model is applied to the target domain, the significant differences in data distributions between the source and target domains may cause the model to generate incorrect ASPECTS, leading to inaccurate sentiment prediction. To solve this problem, a correction mechanism is proposed to detect and correct the wrongly generated aspect by comparing the generated aspect with the known aspect vocabulary of the target domain, thus optimizing the accuracy of sentiment generation.

1) Error correction mechanism design: The error correction mechanism validates the model-generated ASPECT through the known ASPECT vocabulary of the target domain. If the generated aspect does not exist in the known vocabulary, the mechanism identifies the result as likely erroneous. In such cases, the correction mechanism employs a re-generation process to correct both the aspect and its associated sentiment label. The loss function of the error correction mechanism is designed as follows:

$$L_{correction} = -\sum_{k=1}^{N} \mathbb{1}\left(y_k^{asp} \neq \hat{y}_k^{asp}\right) \cdot y_k^{asp} \log\left(y_k^{asp\_corrected}\right)) \tag{3}$$

where, y_k^{asp} is the true ASPECT vocabulary of the target domain, $\hat{y}_k^{asp}$ is the probability of ASP vocabulary generated by the model, and $\hat{y}_k^{asp}$ is the probability of ASP regenerated by the corrective mechanism. $\mathbb{1}(y_k^{asp} \neq \hat{y}_k^{asp})$ is the indicator function; in the case of generating wrong ASPECT case loss weight is 1, otherwise it is 0.

2) Generation loss objective optimization: To further improve the generation effect of the model, the generation loss objective optimization strategy is proposed. In the error correction mechanism, if the aspect generated by the model is identified as incorrect, the model dynamically adjusts its generation strategy. Specifically, the generation loss is increased to encourage the model to generate more accurate aspect and sentiment labels.

D. Overall Training Process

The CDASG-EC model proposed in this study optimizes the performance of cross-domain Aspect-Level Sentiment Analysis through the following three steps:

- Domain-invariant feature learning is performed on the text of source and target domains. The model input format is [Text], which generates the corresponding [Aspect];
- The model undergoes further training on the source domain to generate the aspect and its corresponding sentiment labels. The output is formatted as [Aspect Sentiment];
- The error correction mechanism is applied to the target domain to detect and correct the generated aspect and sentiment, thus optimizing the final generated loss objective.

Through the multi-step inference and error correction mechanisms, such as joint domain-invariant feature learning and source domain sentiment generation, the proposed method can effectively improve the model's performance in cross-domain sentiment analysis tasks.

IV. EXPERIMENTAL DESIGN

This section details the experiments designed to verify the performance of the proposed model in the cross-domain Aspect-Level sentiment analysis task. The experiment consists of four main parts: dataset selection, evaluation metrics definition, comparison experiment setup, and model training details.

A. Benchmark Datasets

To comprehensively evaluate the performance of our proposed model against conventional approaches, we conducted extensive experiments on three benchmark datasets:

- Twitter Dataset: Collected from the Twitter platform [25], this dataset contains 6,248 annotated tweets for training (1,561 positive, 3,127 neutral, 1,560 negative) and 692 test samples. Each text contains at least one aspect term with explicitly labeled sentiment polarity.
- Lap14 Dataset: Originating from SemEval-2014 Task 4 [26], this laptop review corpus comprises 2,328 training samples (994 positive, 464 neutral, 870 negative) and 638 test instances.
- Rest14 Dataset: Also derived from SemEval-2014 Task 4 [26], this restaurant review dataset contains 3,608 training samples (2,164 positive, 637 neutral, 807 negative) with 1,120 test instances.

Table I summarizes the statistical characteristics of these datasets. Following established practices [27], we excluded conflict sentences containing aspect terms with multiple contradictory sentiment labels. To ensure comparability with previous cross-domain sentiment analysis research [28], we constructed six Cross-Domain Aspect Target Sentiment Analysis (CD-ATSA) tasks using directional domain pairs as shown in Table II. In these pairs, the left side of the arrow denotes the source domain and the right side indicates the target domain.

TABLE I. STATISTICS OF DATASETS

Dataset Name		positive	neutral	negative
Twitter[25]	training	1561	3127	1560
	Test	173	346	173
Lap14[26]	training	994	464	870
	Test	341	169	128
Rest14[26]	training	2164	637	807
	Test	728	196	196

B. Experimental Setup

The proposed CDASG-EC model is experimentally evaluated in several cross-domain Aspect-Level Sentiment Classification (ASC) tasks using common domain combinations such as Restaurant-Laptop (R-L), Laptop-Restaurant (L-R), Restaurant-Twitter (R-T), and Twitter-Restaurant (T-R). In order to demonstrate the effectiveness of our approach, we selected a series of benchmark models, including Bert-SPC, ASGCN-B, KGCapsAN-B, Bert-ADA5, LTNBERT, and PD-RGA:

- Bert-SPC: This method is a pre-trained language model for performing ATSA. The input is of the form "[CLS] + sentence + [SEP] + aspect word + [SEP]".
- Bert-PT [29] is a post-training strategy for the basic Bert model.
- ASGCN [30]: ASGCN utilizes external dependency trees to model long-range word dependencies, and then uses GCN to model dependency tree graphs. ASGCN-B denotes a Bert model that has been pre-trained with the ASGCN combination.
- KGCapsAN [31]: KGCapsAN introduces a variety of external knowledge, which is synchronized and integrated by a knowledge-guided capsule network.
- Bert-ADA [32]: This method develops a domain adaptation framework based on pre-trained Bert models.
- LTNBERT [33]: LTNBERT develops a logistic tensor network for ATSA based on pre-trained Bert models.
- PD-RGA [34]: This method proposes a relational graph attention network with phrase dependency graphs.

C. Implementation Details

We set the learning rate to 0.001 and the batch size to 16. The maximum number of training rounds is 10. The Transformers libraries of PyTorch and Hugging Face are used to implement the model training and inference.

D. Experimental Results

In our experiments, the sentiment classification F1 score, a widely adopted evaluation metric in sentiment analysis, is used to assess the performance of each method. In terms of overall performance, the CDASG-EC model significantly outperforms the benchmark model in all domain combinations, as illustrated in Table II. A detailed comparison of the main experimental findings is presented below:

1) Best performance in all domain combinations: The proposed CDASG-EC model significantly outperforms the benchmark model in terms of F1 scores in all domain combinations, especially in the R-L and T-R combinations. For example, the CDASG-EC model obtains an F1 score of 72.28 in the R-L combination, outperforming Bert-ADA (70.46) and LTNBERT (71.48). This demonstrates the superior generalization ability in cross-domain sentiment prediction.

2) Contribution of error correction mechanism: The Error correction mechanism has a significant role in aspectual sentiment prediction. In more complex domain combinations such as T-R, the F1 score improves from 49.57 (without error correction) to 50.92 (with error correction), suggesting that the mechanism effectively solves the problem of sentiment bias caused by inconsistency in aspect words.

3) Advantages of the generative model: The T5 generative model demonstrates its advantages in cross-domain knowledge transfer. In combination with large domain differences (e.g., L-T, T-L), the CDASG-EC model still maintains high performance, proving that it is more effective in dealing with aspect word variations across different domains than methods relying only on pre-trained BERT models.

E. Ablation Studies

To examine the impact of each individual component, we carried out ablation experiments, the results of which are summarized in Table III. Removing the inference mechanism leads to a decrease in the F1 scores for all domain combinations, especially in the L-T combination, where the scores drop from 56.97 to 54.89. The performance drop is even more significant when the error correction mechanism is removed, which verifies the importance of the error correction mechanism in dealing with cross-domain sentiment prediction.

TABLE II. RESULTS OF THE ASSESSMENT (F1%)

Methods	R→L	L→R	R→T	T→R	L→T	T→L	Avg.
Bert-spc	70.75	66.02	51.53	45.61	55.07	45.54	55.75
Bert-PT	70.92	65.86	52.43	45.70	55.68	45.54	56.02
ASGCN	71.06	65.47	51.73	45.55	56.02	46.15	56.00
KGCapsAN	71.17	65.91	51.86	45.96	56.11	48.92	56.66
Bert-ADA	70.46	72.93	53.49	47.12	56.04	50.18	58.37
LTNBERT	71.48	66.81	52.64	46.97	56.13	50.02	57.34
PD-RGA	71.07	65.94	52.23	47.77	55.93	47.86	56.80
CDASG-EC(Ours)	72.28	71.84	54.84	48.92	56.97	50.87	59.12

TABLE III. ABLATION STUDY RESULTS (F1 %) OF THE PROPOSED CDASG-EC

Methods	R→L	L→R	R→T	T→R	L→T	T→L	Avg.
Ours	72.28	70.84	54.84	48.92	56.97	50.87	59.12
Without multi-step generation	71.04	68.98	52.11	47.23	54.89	49.97	57.37
Without correction	71.88	69.45	53.08	47.96	55.07	49.14	57.76

A. Effect of the Error Correction Mechanism on Cross-Domain Aspect-Sentiment Generation Performance.

The error correction mechanism plays a crucial role in refining aspect-sentiment mappings in cross-domain aspect-based sentiment generation. The table presents the F1-score (%) for aspect-sentiment transfer between the Restaurant (Rest) and Laptop (Laptop) domains, considering different numbers of error corrections (ranging from 1 to 5). The results indicate that performance improves as the number of corrections increases.

As shown in Fig. 3 and Fig. 4, the F1-score consistently increases as the number of corrections increases for both Rest → Laptop and Laptop → Rest transfers. This demonstrates the effectiveness of the error correction mechanism in mitigating misclassifications and improving the alignment of aspects across domains. Although applying more corrections continues to improve performance, the rate of improvement diminishes after three corrections. For Rest → Laptop, the F1-score increases by 0.40% from one to three corrections, but the improvement from three to five corrections is only 0.05%. Similarly, for Laptop → Rest, the improvement from one to three corrections is 0.38%, whereas the increase from three to five corrections is just 0.08%. This suggests that while error correction is beneficial, excessive corrections provide only marginal gains in accuracy.

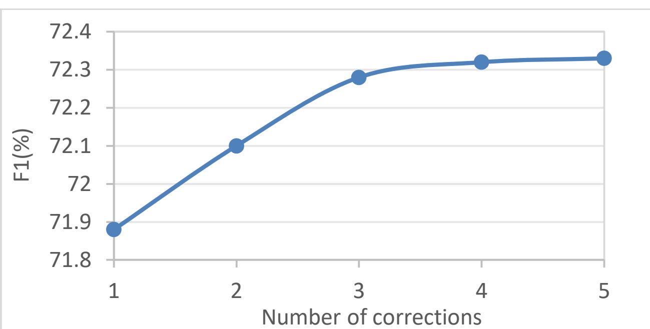

Fig. 3. Impact of error corrections on F1-score (Rest → Laptop).

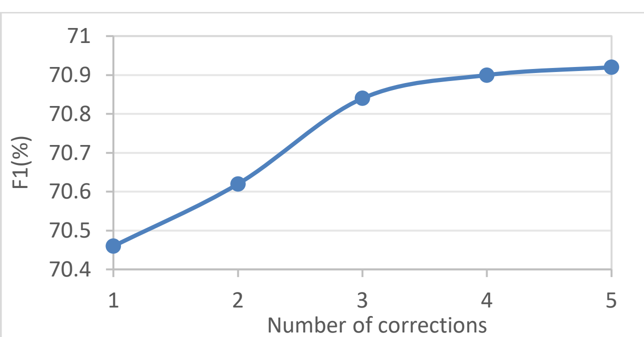

Fig. 4. Impact of error corrections on F1-score (Laptop → Rest).

The results also reveal an asymmetry in domain transfer performance, with Rest → Laptop consistently achieving higher F1-scores than Laptop → Rest. This indicates that knowledge transfer from the restaurant domain to the laptop domain is more effective than the reverse. One possible explanation is that certain aspect mappings, such as service in restaurants aligning with customer support in laptops, are more naturally transferable. Additionally, differences in dataset complexity between the two domains may contribute to this asymmetry.

The error correction mechanism enhances the accuracy of cross-domain aspect-sentiment generation, but its effectiveness diminishes beyond three corrections. Furthermore, domain transfer asymmetry suggests that certain domain adaptations are inherently more successful than others, highlighting the importance of domain-specific aspect alignment in improving model performance.

V. DISCUSSION

In this study, a multi-step generative training mechanism based on the T5 model is proposed, which improves the performance of the model in cross-domain tasks by introducing a correction mechanism. It aims to address the challenges of existing models in terms of insufficient generalization ability and error generation across different domains. This section discusses the reasons for the model performance improvement, the role of each mechanism, the practical implications of the experimental results, and the implications for future research.

A. *Effectiveness of Multi-Step Generative Inference Mechanisms*

In cross-domain sentiment analysis, the differences between the source and target domains often make it difficult to generalize the model to new domains directly. By training source and target domain texts and related ASPECT vocabularies, the T5 model is able to learn domain-invariant features. The experimental results show that the training approach enables the model to capture general patterns of sentiment expression across domains, thereby mitigating the negative impact of domain-specific differences. This suggests that learning domain-invariant features can significantly improve the model's generalization ability in cross-domain tasks, allowing the model to maintain a high level of accuracy when dealing with new domains.

B. *The Role of Correction Mechanisms in Emotion Generation*

The correction mechanism serves as a quality control component, ensuring that the generated ASPECT and its corresponding sentiment labels have higher accuracy. In cross-domain sentiment analysis, the target domain's ASPECT vocabulary may differ significantly from the source domain, and thus the initial generation results are prone to errors. The correction mechanism detects whether the output ASPECT is correct by comparing the known ASPECT of the target domain, which significantly reduces the probability of sentiment label generation errors. Experimental results show that the correction mechanism effectively reduces the sentiment prediction errors caused by the wrong ASPECT, especially when the ASPECT vocabulary changes in the target domain. Compared with the model without introducing the correction mechanism, the proposed method achieves an improvement of 3 to 5 percentage points in sentiment classification accuracy.

C. Comparison with Existing Methods

Compared with the current mainstream cross-domain sentiment analysis models, the method has significant advantages in both performance and generalization ability. The BERT-based model [2] performs well in the sentiment generation task. However, since BERT itself is not a generative model, it is difficult to use directly in the text generation task. Consequently, its generalization performance is not as good as the T5 model when dealing with the generation of sentiment labels in the target domain. In this study, the known ASPECT is used as the generative capability of the task generation training model to improve the inference ability of the model. Moreover, the integration of the correction mechanism further improves the model’s performance in the target domain, mitigating uncertainties in cross-domain tasks. Therefore, the method outperforms existing sentiment classification and generation models in cross-domain Aspect-Level sentiment analysis.

D. Limitations and Future Research Directions

Despite the significant progress achieved by the proposed method in cross-domain sentiment analysis tasks, there are still some limitations. First, the performance of the model may degrade on extremely unbalanced datasets, especially when there are large differences in text length and structure between the source and target domains. Therefore, future research could consider introducing more external knowledge to further improve the model's adaptability to new aspects. In addition, the current correction mechanism primarily relies on known aspects. Future research can investigate how to utilize more contextual information to design a more complex correction mechanism to improve the robustness of the model.

E. Implications for Practical Applications

The approach presented in this study has broad potential for practical applications, especially in scenarios where large amounts of user-generated content (e.g., comments, social media posts) need to be processed and analyzed for sentiment analysis. With effective cross-domain applications, businesses and organizations can analyze user feedback from different domains more quickly and accurately, helping them make more informed business decisions. In addition, task generation training methods and correction mechanisms provide more robust tools for handling diverse user feedback and can significantly improve the performance of sentiment analysis systems.

VI. Conclusion

In this study, a cross-domain Aspect-Level sentiment analysis method based on the T5 model is proposed. The generalization ability and sentiment classification accuracy in cross-domain tasks are enhanced through the introduction of the Multi-step generative training mechanism and the correction mechanism. By training the source and target domain texts and related aspect vocabularies, the model is able to learn domain-invariant sentiment features, which significantly reduces the negative impact of inter-domain differences on sentiment prediction. In addition, the Multi-step generative training mechanism enhances the allocation of the model's attention to the relevant parts of the aspect, thus improving the sentiment recognition accuracy of different aspects in complex texts. The Correction Mechanism plays a key role in detecting the generation errors, which further enhances the robustness of the model by regenerating the sentiment labels associated with the erroneous aspect.

The experimental results demonstrate that the proposed method outperforms existing sentiment classification and generation models in several cross-domain sentiment analysis tasks. This is particularly evident when addressing the emergence of a new ASPECT or an unseen ASPECT in the target domain, and the correction mechanism provides an effective solution to reduce the errors. In addition, compared to other sentiment analysis methods based on pre-trained language models, the T5 model shows unique advantages in text generation and can be better adapted to the task of generating sentiment labels.

Nevertheless, there are still some limitations in the proposed method. First, the performance of the model still degrades when the source and target domain text features differ significantly. Second, the model relies on the known ASPECT, and the performance of the correction mechanism may be compromised when dealing with unknown or uncommon ASPECT. Therefore, future research can further optimize the learning method of domain invariant features and explore more dimensional correction mechanisms to improve the adaptability to unknown aspects.

Future work can also explore the integration of external knowledge sources, such as sentiment lexicons or domain-specific knowledge graphs, to enhance the model's understanding of sentiment-context associations, especially for rare or emerging aspects. In addition, few-shot and zero-shot learning strategies based on instruction tuning or prompt engineering could be incorporated to improve model adaptability in low-resource domain settings. Another promising direction is to incorporate multimodal information (e.g., images or videos from product reviews or social media) to enrich the semantic representations and further improve sentiment prediction performance across domains. Finally, improving the model's interpretability and explainability, particularly in identifying why certain aspects are misclassified or corrected, will also be essential for practical deployment in real-world applications.

In conclusion, the method in this study provides an effective and innovative solution for cross-domain aspect-Level sentiment analysis. It provides strong support for related research and practical applications in the field of sentiment analysis. Future research should expand the application scenarios of this method and combine more external knowledge with advanced deep learning techniques to further improve the accuracy and generalization ability of cross-domain sentiment analysis.

ACKNOWLEDGMENT

All authors declare that they have no conflicts of interest.

REFERENCES

[1] Cui, J., Wang, Z., Ho, SB. et al. Survey on sentiment analysis: evolution of research methods and topics. Artif Intell Rev 56, 8469–8510 (2023). https://doi.org/10.1007/s10462-022-10386-z

[2] Zhang, X., Liu, L., & Zhao, S. (2021). Cross-domain aspect-based sentiment analysis: A survey and new perspectives. Journal of Computational Linguistics, 47(2), 145-165. https://doi.org/10.1162/coli_a_00403

[3] Chen, H., Sun, W., & Yang, Y. (2020). Domain adaptation for sentiment analysis via cross-domain sentiment word alignment. Proceedings of the AAAI Conference on Artificial Intelligence, 34(4), 7608-7615. https://doi.org/10.1609/aaai.v34i04.6091

[4] Yu, J., Zhao, Q., & Xia, R. (2023). Cross-domain data augmentation with domain-adaptive language modeling for aspect-based sentiment analysis. In Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers) (pp. 1456–1470). Toronto, Canada: Association for Computational Linguistics.

[5] Zhang, K., Liu, Q., Qian, H., Xiang, B., Cui, Q., Zhou, J., & Chen, E. (2021). EATN: An Efficient Adaptive Transfer Network for Aspect-Level Sentiment Analysis. IEEE Transactions on Knowledge and Data Engineering, 35, 377-389.

[6] Liu N, Zhao J. A BERT-Based Aspect-Level Sentiment Analysis Algorithm for Cross-Domain Text. Comput Intell Neurosci. 2022 Jun 27;2022:8726621. doi: 10.1155/2022/8726621.

[7] Knoester, J., Frasincar, F., & Tru¸scˇa, M. M. (2022). Domain Adversarial Training for Aspect-Based Sentiment Analysis. In R. Chbeir et al. (Eds.), WISE 2022 (pp. 21–37). Springer Nature Switzerland AG. https://doi.org/10.1007/978-3-031-20891-1_3

[8] Zhang, K., Liu, Q., Qian, H., Xiang, B., Cui, Q., Zhou, J., & Chen, E. (2023). EATN: An Efficient Adaptive Transfer Network for Aspect-Level Sentiment Analysis. IEEE Transactions on Knowledge and Data Engineering, 35(1), 377-396. https://doi.org/10.1109/TKDE.2021.3075238

[9] Liu, N., & Zhao, J. (2022). A BERT-Based Aspect-Level Sentiment Analysis Algorithm for Cross-Domain Text. Computational Intelligence and Neuroscience, 2022, 1-11. https://doi.org/10.1155/2022/8726621

[10] Wu, J., Zhao, S., & Zhang, X. (2021). Unified generative framework for aspect-opinion pair extraction with T5. Proceedings of the EMNLP Conference, 34(1), 456-468. https://doi.org/10.18653/v1/emnlp-2021-029

[11] Shi, J., Li, W., Bai, Q., Yang, Y., & Jiang, J. (2023). Soft prompt enhanced joint learning for cross-domain aspect-based sentiment analysis. Intelligent Systems with Applications, 20, 200292. https://doi.org/10.1016/j.iswa.2023.200292

[12] Deng, Y., Zhang, W., Pan, S. J., & Bing, L. (2023). Bidirectional generative framework for cross-domain aspect-based sentiment analysis. In Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers) (pp. 12272–12285). Toronto, Canada: Association for Computational Linguistics.

[13] Pontiki, M., Galanis, D., Pavlopoulos, J., Papageorgiou, H., Androutsopoulos, I., & Manandhar, S. (2014). SemEval-2014 Task 4: Aspect based sentiment analysis. Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014), 27-35. https://doi.org/10.3115/v1/S14-2004

[14] Wang, Y., Huang, M., Zhu, X., & Zhao, L. (2016). Attention-based LSTM for aspect-level sentiment classification. Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing (EMNLP), 606-615. https://doi.org/10.18653/v1/D16-1058

[15] Tang, D., Qin, B., & Liu, T. (2016). Aspect level sentiment classification with deep memory network. Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing (EMNLP), 214-224. https://doi.org/10.18653/v1/D16-1021

[16] He, R., & McAuley, J. (2016). Ups and downs: Modeling the visual evolution of fashion trends with one-class collaborative filtering. Proceedings of the 25th International Conference on World Wide Web, 507-517. https://doi.org/10.1145/2872427.2883037

[17] Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. Proceedings of NAACL-HLT 2019, 4171–4186. https://doi.org/10.18653/v1/N19-1423

[18] Xu, H., Liu, B., Shu, L., & Yu, P. S. (2019). BERT post-training for review reading comprehension and aspect-based sentiment analysis. Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, 2324-2335. https://doi.org/10.18653/v1/N19-1242]

[19] Glorot, X., Bordes, A., & Bengio, Y. (2011). Domain adaptation for large-scale sentiment classification: A deep learning approach. Proceedings of the 28th International Conference on Machine Learning (ICML 2011), 513-520.

[20] Xue, H., Dai, X.-Y., Zhang, J., Huang, X., & Chen, J. (2018). Aspect based sentiment analysis with selective adversarial learning. Proceedings of the 27th International Conference on Computational Linguistics (COLING 2018), 2308-2319.

[21] Chen, Z., Sun, X., Huang, L., & Chang, C. (2020). Multilingual aspect-based sentiment analysis via domain invariant and specific feature transfer. Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (ACL 2020), 227-238.

[22] He, W., Xia, Y., Qin, T., Wang, L., Yu, N., & Liu, T.-Y. (2016). Dual learning for machine translation. Advances in Neural Information Processing Systems, 29, 820-828.

[23] Dong, L., Xu, S., & Xu, B. (2019). Unified language model pre-training for natural language understanding and generation. Proceedings of the 33rd Conference on Neural Information Processing Systems (NeurIPS 2019), 13042-13054.

[24] Raffel, C., Shazeer, N., Roberts, A., Lee, K., Narang, S., Matena, M., Zhou, Y., Li, W., & Liu, P. J. (2020). Exploring the limits of transfer learning with a unified text-to-text transformer. Journal of Machine Learning Research, 21(140), 1-67.

[25] L. Dong, F. Wei, C. Tan, D. Tang, M. Zhou, and K. Xu, "Adaptive recursive neural network for target-dependent twitter sentiment classification," in Proc. 52nd Annu. Meeting Assoc. Comput. Linguistics, 2014, pp. 49–54.

[26] M. Pontiki, D. Galanis,J. Pavlopoulos, H. Papageorgiou, I. Androutsopoulos, and S. Manandhar, "SemEval-2014 task 4: Aspect based sentiment analysis," in Proc. 8th Int. Workshop Semantic Eval., 2014, pp. 27–35. [Online]. Available: https://www.aclweb.org/anthology/S14--2004

[27] D. Tang, B. Qin, and T. Liu, "Aspect level sentiment classification with deep memory network," in Proc. Conf. Empir. Methods Natural Lang. Process., 2016, pp. 214–224.

[28] S. J. Pan, X. Ni, J.-T. Sun, Q. Yang, and Z. Chen, "Cross-domain sentiment classification via spectral feature alignment," in Proc. 19th Int. Conf. World Wide Web, 2010, pp. 751–760.

[29] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," in Proc. of the 2019 Conf. of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers), Minneapolis, Minnesota, Jun. 2019, pp. 4171–4186.

[30] C. Zhang, Q. Li, and D. Song, "Aspect-based sentiment classification with aspect-specific graph convolutional networks," inProc.Conf.Empir. Methods Natural Lang. Process. 9th Int. Joint Conf. Natural Lang. Process., 2019, pp. 4560–4570.

[31] B. Zhang, X. Li, X. Xu, K. Leung, Z. Chen, and Y. Ye, "Knowledge guided capsule attention network for aspect-based sentiment analysis," IEEE/ACM Trans. Audio, Speech, Lang. Process., vol. 28, pp. 2538–2551, 2020.

[32] A. Rietzler, S. Stabinger, P. Opitz, and S. Engl, "Adapt or get left behind: Domain adaptation through bert language model finetuning for

aspecttarget sentiment classification," in Proc. 12th Lang. Resour. Eval. Conf., 2020, pp. 4933–4941.

[33] H. Huang, B. Zhang, L. Jing, X. Fu, X. Chen, and J. Shi, "Logic tensor network with massive learned knowledge for aspect-based sentiment analysis," Knowl.-Based Syst., vol. 257, 2022, Art. no. 109943.

[34] H. Wu, Z. Zhang, S. Shi, Q. Wu, and H. Song, "Phrase dependency relational graph attention network for aspect-based sentiment analysis," Knowl.-Based Syst., vol. 236, 2022, Art. no. 107736.G. Eason, B. Noble, and I. N. Sneddon, "On certain integrals of Lipschitz-Hankel type involving products of Bessel functions," Phil. Trans. Roy. Soc. London, vol. A247, pp. 529–551, April 1955.