Aspect-Based Sentiment Analysis with Explicit Sentiment Augmentations

Jihong Ouyang[1,2*], Zhiyao Yang[1,2*], Silong Liang[1,2], Bing Wang[1,2], Yimeng Wang[1], Ximing Li[1,2†]

[1]College of Computer Science and Technology, Jilin University, China
[2]Key Laboratory of Symbolic Computation and Knowledge Engineering of MOE, Jilin University, China
{ouyj@, zhiyaoy20@mails., liangsl23@mails.}jlu.edu.cn, {wangbing1416, wangyimeng116, liximing86}@gmail.com

Abstract

Aspect-based sentiment analysis (ABSA), a fine-grained sentiment classification task, has received much attention recently. Many works investigate sentiment information through opinion words, such as "good" and "bad". However, **implicit sentiment** data widely exists in the ABSA dataset, whose sentiment polarity is hard to determine due to the lack of distinct opinion words. To deal with implicit sentiment, this paper proposes an ABSA method that integrates explicit sentiment augmentations (ABSA-ESA) to add more sentiment clues. We propose an ABSA-specific explicit sentiment generation method to create such augmentations. Specifically, we post-train T5 by rule-based data and employ three strategies to constrain the sentiment polarity and aspect term of the generated augmentations. We employ Syntax Distance Weighting and Unlikelihood Contrastive Regularization in the training procedure to guide the model to generate the explicit opinion words with the same polarity as the input sentence. Meanwhile, we utilize the Constrained Beam Search to ensure the augmentations are aspect-related. We test ABSA-ESA on two ABSA benchmarks. The results show that ABSA-ESA outperforms the SOTA baselines on implicit and explicit sentiment accuracy.

Introduction

Aspect-**b**ased **S**entiment **A**nalysis (**ABSA**) aims to induce predictive models over manually annotated sentences to identify the sentiment polarity towards each specific aspect term (Wang et al. 2022a; Li et al. 2022). Taking the second sentence in Fig. 1 (a) as an example, the task aims to automatically identify the sentiment polarities of its aspect terms "*outside*" (`Negative`) and "*atmosphere*" (`Positive`) potentially with the corresponding opinion words "crushed" and "nice". Due to its popularity, ABSA has been widely applied in many real-world scenarios, and accordingly, it is one of the most significant tasks in the natural language processing community (Yang et al. 2023; Ouyang et al. 2023).

To handle the task of ABSA, many studies have been investigated during the past decade. Broadly speaking, the focus of recent work is on how to generate more discriminative representations for aspect terms to enhance the identification performance of sentiment polarity. Some early studies generate strong aspect term representations by directly employing deep neural encoders, such as LSTM (Tang et al. 2016; Wang et al. 2016; Cheng et al. 2017) and pre-trained language models (Xu et al. 2020; Dai et al. 2021). Beyond them, to further link the aspect terms and opinion words, some studies build dependency trees of sentences and then generate aspect term representations by employing graph convolution networks (GCN) (Sun et al. 2019; Wang et al. 2020; Chen, Teng, and Zhang 2020; Li et al. 2021a).

(a) 1. The *fried rice* (Positive) is amazing here .
2. It is crushed at *outside* (Negative), but the minute you walk inside, it has a nice *atmosphere* (Positive).

(b) 1. Our server checked on us maybe twice during the entire *meal* (Negative).
2. All the money went into the *interior decoration* (Positive), none of it went to the *chefs* (Negative).

Figure 1: Examples of the explicit sentiment sentences (a) vs. implicit sentiment sentences (b) from Rest.14 dataset. The words with are the opinion words.

The success of the GCN-based approach underscores the pivotal role that opinion words play in the realm of ABSA. However, recent research has highlighted a complex scenario characterized by a lack of distinct opinion words, termed "implicit sentiment" (Li et al. 2021b; Wang et al. 2022b). To delve into this phenomenon, we select four examples from the Rest.14 to compare the implicit and explicit sentiment sentences. In the context of Fig.1(a), the sentiment is discernible due to distinct opinion words. In contrast, as shown in Fig.1(b), unraveling the sentiment associated with aspect terms such as "meal," "interior decoration," and "chefs" is challenging. Implicit sentiment is a prevalent occurrence within ABSA datasets and it is hard to deal with (Li et al. 2021b).

To tackle the challenge mentioned above, in the paper, we design a novel **ABSA** method by integrating **E**xplicit **S**entiment **A**ugmentations (**ABSA-ESA**). Such augmentations provide more sentiment clues for predicting sentiment polarity. We add them after the corresponding input sen-

[*]These authors contributed equally.
[†]Corresponding author

tence, forming new ABSA training data. To obtain the augmentations, we design an **ABSA-specific explicit sentiment generation** method. We aim to generate the sentences **explicitly** conveying the **same sentiment polarity** as their corresponding input sentences, targeting **the same (or similar) aspect terms**. We post-train the generation model T5 (Raffel et al. 2020) by the rule-based data selected in the ABSA dataset, making the generated augmentations comply with the above requirements. Furthermore, we introduce three strategies to confine the generated augmentations about their sentiment polarity and aspect terms. Specifically, in the training procedure, we employ the **Syntax Distance Weighting** and **Unlikelihood Contrastive Regularization** to lead the model to generate explicit opinion words with the same polarity as the input sentence. Subsequently, when engendering the augmentations, we employ the **Constrained Beam Search** to ensure the augmentations are aspect-related.

To sum up, our contributions can be listed as follows:

- We propose a novel ABSA framework named ABSA-ESA, which focuses on solving the implicit sentiment issue by generating explicit sentiment augmentations.
- We propose an ABSA-specific explicit sentiment generation method that generates augmentations with distinct opinion words for specific aspect terms.
- Empirical results on two ABSA benchmarks show that ABSA-ESA outperforms other methods on both explicit and implicit accuracy.

Related Work

Aspect-based Sentiment Analysis

Aspect-Based Sentiment Analysis (ABSA) methods primarily focus on integrating sentiment information from contextual words into aspect terms. In earlier approaches, this was often achieved by utilizing LSTM or Bi-LSTM as encoders (Tang et al. 2016; Wang et al. 2016; Cheng et al. 2017). Consequently, recent advancements have embraced the Attention mechanism as the preferred encoder (Tang et al. 2016; Wang et al. 2016; Cheng et al. 2017). Notably, leveraging pre-trained language models has emerged as the prevailing trend in ABSA (Xu et al. 2020; Dai et al. 2021). Furthermore, to establish stronger connections between aspect terms and opinion words, numerous studies have delved into constructing dependency trees within sentences and refining aspect term representations using Graph Convolutional Networks (GCNs) (Sun et al. 2019; Wang et al. 2020; Li et al. 2021a).

Concurrently, alongside developing robust encoders, researchers have explored the enrichment of training data to provide external sentiment information for the model (He et al. 2019; Wang et al. 2022a; Yang et al. 2023). These additional data often lack fine-grained annotations and necessitate subsequent data processing. Addressing this, this paper integrates ABSA-specific augmentations into ABSA models, bypassing the need for extensive reprocessing.

Implicit Sentiment Analysis

Implicit sentiment classification, a pivotal subfield within sentiment analysis, was pioneered by Liu (2012), drawing significant scholarly interest. Initial works revolved around implicit sentiment at the sentence level (Deng, Wiebe, and Choi 2014; Choi, Wiebe, and Mihalcea 2017; Zhou et al. 2021a; Xu et al. 2022). Recent endeavors have shifted towards tackling implicit aspect-based sentiment classification (Li et al. 2021b; Wang et al. 2022b; Fei et al. 2023). A prevailing approach involves incorporating external knowledge to capture sentiment expression patterns. For instance, Xu et al. (2022) integrates external sentiment-related knowledge into sentence features, enhancing the model's sentiment comprehension. Similarly, Li et al. (2021b) employs a post-training strategy with BERT, leveraging contrastive learning on expansive sentiment-annotated corpora. ABSA-ESA utilizes the data generated by the model instead of obtaining external knowledge.

Data Augmentation

Within NLP, the data augmentation technique has gained substantial traction to expand the pool of available training instances. This approach finds widespread application across diverse domains, including text classification (Wu et al. 2022; Liu et al. 2022; Ouyang et al. 2022), neural machine translation (Lam, Schamoni, and Riezler 2022; Kambhatla, Born, and Sarkar 2022; Gao et al. 2019), and text generation (Bi, Li, and Huang 2021; Xu et al. 2021). Notably, recent strides in ABSA have similarly leveraged data augmentation (Chen, Faulkner, and Badyal 2022; Wang et al. 2022a; Hsu et al. 2021). However, their augmentation techniques tend to be relatively simple, *e.g.,* token replacement, masked aspect prediction, and polarity reversal, limiting the semantic diversity of the enhanced samples. The augmentation method in this paper is based on the language model, which generates augmentations with rich sentiment information.

Our Proposed ABSA-ESA Model

In this section, we introduce the proposed ABSA method named **ABSA-ESA**.

Overall Framework

Generally speaking, ABSA methods take the review sentence $\mathbf{s} = \{s_j\}_{j=1}^{M}$ and its corresponding aspect term $\mathbf{a} = \{a_j\}_{j=1}^{|\mathbf{a}|}$ as the model input, M denotes the length of all sentence. And output the sentiment polarity $y \in \mathcal{Y} = \{\texttt{Positive}, \texttt{Negative}, \texttt{Neutral}\}$ for $\mathbf{a}$. To deal with the sentences containing implicit sentiment, we extend this paradigm by introducing an augmented sentence $\mathbf{s}'$ following the initial input $\mathbf{s}$. This augmented sentence contains explicit sentiment tied to the aspect term $\mathbf{a}$. For clarity, we present the comprehensive framework of ABSA-ESA in Figure 2.

To generate the augmented sentence $\mathbf{s}'$, we propose an ABSA-specific explicit sentiment generation method. We post-train T5 by utilizing $\hat{\mathbf{s}}'$ as generation targets selected from the dataset. $\hat{\mathbf{s}}'$ has the same (or similar) aspect terms

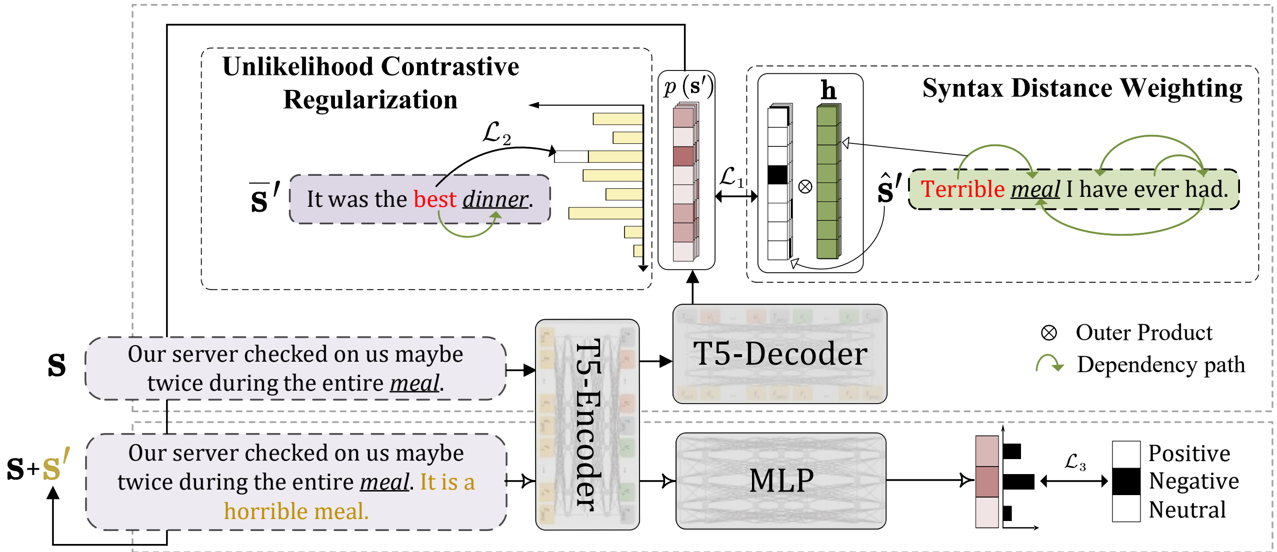

Figure 2: Overall framework of ABSA-ESA.

and sentiment as $\mathbf{s}$ while also incorporating explicit sentiment expressions. Additionally, we utilize three strategies to guide the generation concerning sentiment polarity and aspect terms. During the training phase, a Syntax Distance Weighting strategy is implemented to prioritize context words closest to the aspect term in the dependency parse. Furthermore, we also gather $\bar{\mathbf{s}}'$, which has the opposite sentiment of $\mathbf{s}$, for Unlikelihood Contrastive Regularization. It instructs the model about undesirable word choices. When generating $\mathbf{s}'$, we employ Constrained Beam Search to ensure that the aspect term or its similar aspect is included in the augmentations and its context words are the most relevant to $\mathbf{a}$.

Next, we introduce the details of the ABSA-specific explicit sentiment generation method.

Training Data Collection

To train the explicit sentiment generation model, the initial step is to gather the training data. Given an input sentence $\mathbf{s}$ and its corresponding aspect term $\mathbf{a}$, the generating target $\hat{\mathbf{s}}'$ must satisfy the following rules:

- The target sentence should incorporate **the same (or similar) aspect term** as the input sentence.
- The target sentence should exhibit **identical sentiment polarity** to the input data.
- The target sentence must **contain explicit sentiment expressions**.

To obtain the target sentence $\hat{\mathbf{s}}'$ that satisfies the given rules, we begin by aggregating all aspect terms in dataset $\mathcal{D}$ to construct the aspect term set $\mathcal{A} = \{\bar{\mathbf{a}}_i\}_{i=1}^{|\mathcal{A}|}$. Each aspect term $\bar{\mathbf{a}}_i$ is associated with a representation $r_{\bar{\mathbf{a}}_i}$[1] acquired by consulting the GloVe embedding table (Pennington, Socher, and Manning 2014). Utilizing these representations $\mathcal{R} = \{r_{\bar{\mathbf{a}}_i}\}_{i=1}^{|\mathcal{A}|}$, we formulate a similarity matrix $\mathbf{C} = \{c_{ij}\}_{|\mathcal{A}|\times|\mathcal{A}|}$, where c_{ij} represents the similarity between aspect terms $\bar{\mathbf{a}}_i$ and $\bar{\mathbf{a}}_j$. c_{ij} is computed by the cosine distance:

$$c_{ij} = \cos(r_{\bar{\mathbf{a}}_i}, r_{\bar{\mathbf{a}}_j}). \tag{1}$$

With the similarity matrix $\mathbf{C}$ available, we proceed to the selection of $\hat{\mathbf{s}}'$. According to Li et al. (2021b), the dataset $\mathcal{D}$ can be divided into explicit subset $\mathcal{D}_e$ and implicit subset $\mathcal{D}_i$. As sentences in $\mathcal{D}_e$ contain explicit sentiment expressions, we choose the $\hat{\mathbf{s}}'$ from this subset to fulfill the third rule above.

We first select k_c aspect terms from $\mathcal{R}$, which are most similar to $\mathbf{a}$, thereby forming the set $\mathcal{A}'$. Subsequently, we extract sentences from $\mathcal{D}_e$ containing aspect terms from $\mathcal{A}'$ and share the same sentiment as $\mathbf{s}$. This forms the candidate target sentence set $\hat{\mathcal{S}}_t$. From this set, we randomly choose a target sentence $\hat{\mathbf{s}}' \in \hat{\mathcal{S}}_t$ to generate training data $(\mathbf{s}, \hat{\mathbf{s}}', \mathbf{a})$ with the input sentence and the corresponding aspect term. This process is iterated for all input sentences $\mathbf{s} \in \mathcal{D}$, resulting the final training dataset $\hat{\mathcal{D}} = \{(\mathbf{s}_i, \hat{\mathbf{s}}'_i, \mathbf{a}_i)\}_{i=1}^{N}$. Afterward, we begin to post-train T5 by $\hat{\mathcal{D}}$.

Sentiment Polarity Constraints

While our training pairs have been thoughtfully chosen, it's important to acknowledge that ABSA datasets often contain sentences with multiple aspect words and contrasting sentiment polarities. Take the following training data as an example:

Input sentence: I have been told to choose the *food* (`Positive`) in this restaurant several times.

Target sentence: Bad *service* (`Negative`) while the *food* (`Positive`) here in our dinner is worthy of being recommended.

[1] In the case where an aspect term $\mathbf{a}$ consists of multiple words ($\mathbf{a} = w_1, w_2, \cdots, w_{|\mathbf{a}_i|}$), we calculate the average word embedding $\bar{r}$ as the representation for that aspect.

In this instance, the target sentence involves two aspect terms, "service" and "food," associating with opposing sentiment polarities. Directly training the model with such pairs may lead the model to generate "bad" with "food" for positive polarity. To navigate this challenge, we integrate two constraints in training ABSA-ESA: Syntax Distance Weighting (SDW) and Unlikelihood Contrastive Regularization (UCR).

Syntax Distance Weighting. Syntax relationships are widely utilized in ABSA to establish a closer connection between aspect terms and their corresponding opinion words (Wang et al. 2020; Li et al. 2021a; Hou et al. 2021; Zhou et al. 2021b). To channel ABSA-ESA's focus towards generating explicit opinion words, we introduce varying learning weights to words within target sentences based on their syntax distance from the aspect term. The shorter the syntax distance between a context word and the aspect term, the higher the assigned learning weight. To start, we employ a dependency parser to construct the dependency tree of the target sentence $\hat{\mathbf{s}}_i'$. Subsequently, we calculate syntax distances between the aspect term and the context words, culminating in the formation of a distance vector $\mathbf{d}_i = \{d_{ij}\}_{j=1}^{M}$. Finally, syntax distance weights are computed using the ensuing formula:

$$\mathbf{h}_i = \mathbf{1} - \mathrm{Softmax}(\mathbf{d}_i), \tag{2}$$

where $\mathbf{1}$ denotes the matrix of ones with the shape same as $\mathbf{d}_i$. Then, we train the generation model with $(\mathbf{s}_i, \hat{\mathbf{s}}_i', \mathbf{a}_i)$ by minimize the following loss:

$$\mathcal{L}_1 = -\sum_{j=1}^{M} h_{ij} \log p(\hat{s}_{ij}' | \mathbf{a}_i; \mathbf{s}_i; \hat{\mathbf{s}}_{i;\leq j}'), \tag{3}$$

where $\hat{\mathbf{s}}_{i;\leq j}'$ denotes the sub-sentence from $\hat{s}_{i1}'$ to $\hat{s}_{ij}'$, and ; denotes the concatenate operation.

Unlikelihood Contrastive Regularization. To mitigate the adverse effect of unrelated words in the target sentence, we select the negative target sentence $\bar{\mathbf{s}}_i'$ from $\mathcal{D}_e$ that shares the same aspect term as the input sentence $\mathbf{s}_i$ yet has the opposite sentiment[2]. We obtain the dependency tree of $\bar{\mathbf{s}}_i$ and calculate the distance vector same as the SDW approach. Subsequently, we generate a negative word set $\bar{\mathbf{s}}'^{-} = \{\bar{s}_i'^{-}\}_{i=1}^{k_n}$ by selecting the top-k_n nearest words to the aspect term. Importantly, note that $\bar{\mathbf{s}}_i' \cap \hat{\mathbf{s}}_i' = \varnothing$, ensuring words from $\hat{\mathbf{s}}_i'$ are excluded when choosing negative words. Next, we minimize the Unlikelihood Contrastive Regularization loss as follows:

$$\mathcal{L}_2 = -\sum_{j=1}^{M} \log \frac{p(\hat{s}_{ij}')}{p(\hat{s}_{ij}') + \sum_{l=1}^{k} p(\bar{s}_k'^{-})}, \tag{4}$$

Aspect Term Constraint

Text generation with the language model is a Hidden-Markov-like process where each generated word is related to its front words. Inspired by this, we aim to have the generated sentence incorporate the input aspect term, resulting in context words that are intrinsically related to the aspect. Specifically, we utilize the Constrained Beam Search (CBS), which generates a set of sub-sentence $\mathcal{G} = \{\mathbf{g}_i\}_{i=1}^{V}$ at each generating step instead of just outputting one. At the t-th step, the previous sub-sentences $\mathcal{G}^{t-1}$ serve as input, and the subsequent word is generated for every sub-sentence, thereby forming candidate sub-sentences $\hat{\mathcal{G}}^t$:

$$\hat{\mathcal{G}}^t = \{(\mathbf{g}; w) | \mathbf{g} \in \mathcal{G}^{t-1}, w \in \mathcal{W}\}, \tag{5}$$

$$\mathcal{W} = \mathrm{argTopZ}(p(\mathbf{g}^{t-1})) \;\cup\; \bar{\mathcal{A}}' \tag{6}$$

where $\bar{\mathcal{A}}' = \mathbf{a} \;\cup\; \mathcal{A}'$ is the aspect terms set that ABSA-ESA need to generate. ; is the concatenate operation and argTopZ$(\cdot)$ output z words with the highest generation probability. The CBS process concludes when any sentence $\hat{\mathbf{g}} \in \hat{\mathcal{G}}^t$ contains the aspect term and the ending symbol </s>. CBS output $\hat{\mathbf{g}}$ as the generated sentence $\mathbf{s}'$. By contrast, if no sentence fulfills the above requirement, we randomly select V sentence to form $\mathcal{G}^t$ for the subsequent CBS step. For clarity, Algorithm 1 provides the i-th CBS step process.

Algorithm 1: t-th Step Constrained Beam Search

Input: Candidate sub-sentence $\mathcal{G}_{t-1}$, Model $\mathcal{F}$ Aspect term $\mathbf{a}$ and its similar aspects $\mathcal{A}'$
Output: Generated Sentence or $\mathcal{G}_t$

```
for g ∈ G_{t−1} do
    p ← Compute next word probability of F(g);
    W ← Select V words with top probability according p;
    W ← W ∪ a ∪ A′ ;
    for w ∈ W do
        g ← Add word w to g.
        Ĝ^t ← add g to Ĝ^t
    end
end
if ĝ ∈ Ĝ^t contains </s>and the aspect term then
    s′ ← Output ĝ as the generated sentences.
else
    G_t ← Randomly select V sub-sentences in Ĝ^t.
    Do the next constrained beam search step.
end
```

Overall Objective

Upon obtaining the augmented sentence $\mathbf{s}_i'$, we integrate it with the input sentence $\mathbf{s}_i$, yielding novel training triplets $(\mathbf{s}_i^c, \mathbf{a}_i, y_i)$, where $\mathbf{s}_i^c = \{s_{i1}, s_{i2}, \cdots, s_{i1}', \cdots, s_{i|\mathbf{s}'|}'\}$ represents the combined sentence. We then input $\mathbf{s}_i^c$ and the corresponding aspect term $\mathbf{a}_i$ into the T5-Encoder to obtain the aspect representation $\mathbf{H}_i^a$. Then we input $\mathbf{H}_i^a$ into a multi-layer perception (MLP) to obtain the sentiment prediction $\hat{y}_i$:

$$\hat{y}_i = \mathrm{MLP}(T5_e(\mathbf{s}_i^c, \mathbf{a}_i))), \tag{7}$$

Finally, we employ a Cross-Entropy loss to guide ABSA-ESA in sentiment prediction:

$$\mathcal{L}_3 = \ell_{\mathrm{CE}}(y_i, \hat{y}_i), \tag{8}$$

[2]Both `Positive` and `Negative` can be the opposite sentiment to `Neutral`.

Dataset	Positive	Negative	Neutral	Implicit
Lap.14$_{\text{train}}$	987	460	866	715
Lap.14$_{\text{test}}$	341	169	128	174
Rest.14$_{\text{train}}$	2164	633	805	1030
Rest.14$_{\text{test}}$	728	196	196	267

Table 1: Statistics of the Rest.14 and Lap.14 dataset. The "Implicit" column denotes the number of sentences with implicit sentiment expression.

Overall, we train ABSA-ESA by minimizing the following objective:

$$\mathcal{L} = \mathcal{L}_1 + \mathcal{L}_2 + \mathcal{L}_3. \tag{9}$$

Experiments

Dataset and Implementation Details

In our experiments, we use the dataset released by Li et al. (2021b), which has already been labeled with explicit and implicit sentiment. The statistics of the dataset are presented in Table 1. In the training process, we set the learning rate as 5×10^{-5} (Restaurant) and 2×10^{-5} (Laptop). The batch size is 4 (Restaurant) and 8 (Laptop). We set the training epoch as 15 for both datasets. We set k_c=2 in data selection and k_n=4 for Unlikelihood Contrastive Regularization. For Constrained Beam Search, we select V=6 and z=3.

Baselines

Common ABSA Baselines. PT (Xu et al. 2019) used BERT, which post-trained on several ABSA sub-tasks, as the encoder. ADA (Rietzler et al. 2020) conducts ABSA by incorporating data from other domains. R-GAT (Wang et al. 2020) learns both aspect term embedding and dependency relation embedding to obtain more information. ASA-WD (Tian, Chen, and Song 2021) uses the key-value memory network to explore the word dependencies.

Implicit ABSA Baselines. BERTAsp (Li et al. 2021b)[3] utilizes contrastive learning to post-train BERT for more implicit sentiment knowledge. CLEAN (Wang et al. 2022b) explores implicit sentiment by studying the causal relations in each sentence. THOR (Fei et al. 2023) employs the chain-of-thought strategy to manually provide more sentiment clues for the implicit sentiment sentences.

Data Augmentation Baselines. Back Translation (BT) translates the training sentence into another language and then translates it back. EDA (Wei and Zou 2019) augments data at the token level by methods such as synonym replacement and random insertion. C^{3}DA (Wang et al. 2022a) generates augmentations by replacing the aspect term and reversing the sentiment polarity.

[3]We select the results of BERTAsp with the version not trained with external data. And the results of its variants, such as T5Asp, are also obtained in the same way.

Main Results

The main results are in Table 2 and Table 3, organized based on the backbone employed by each baseline. In general, ABSA-ESA consistently demonstrates superior performance compared to the baselines across all evaluation metrics. Remarkably, ABSA-ESA significantly enhances the classification accuracy of implicit sentiment data in contrast to the state-of-the-art CLEAN baseline (1.12↑ on Rest.14 and 1.71↑ on Lap.14). Additionally, ABSA-ESA exhibits notable advancements in the classification accuracy of explicit sentiment data, (average 1.27↑ on Rest.14 and 2.14↑ on Lap.14). In the subsequent analysis, we delve into a detailed comparison of results across various baseline categories.

Compared with Implicit ABSA Baselines. Our initial focus is comparing ABSA-ESA against implicit ABSA baselines closest to ABSA-ESA. We first focus on the BERTAsp and CLEAN. Even though these methods employ cutting-edge methods such as contrastive learning and causal reasoning, the results in Table 2 show that ABSA-ESA has the best implicit accuracy. This indicates that generating explicit sentiment augmentations is a sample and effective way to deal with implicit ABSA.

Subsequently, we compare ABSA-ESA with THOR. Although both ABSA-ESA and THOR share the core notion of enriching the input sentence with additional sentiment cues, there exists a fundamental distinction between them. THOR employs predefined prompts to guide model responses, whereas ABSA-ESA directly generates explicit sentiment tied to the given aspect. The results in Table 3 illustrate that ABSA-ESA outperforms THOR on both datasets, particularly excelling in the F1-score of implicit data. It's noteworthy that on the Lap.14 dataset, THOR's implicit data F1-score trails behind that of the basic prompt model (67.63 vs. 76.40). This underscores the vulnerability of methods built on the chain-of-thought strategy, where outcomes are heavily reliant on manually designed prompts. In contrast, ABSA-ESA's strength lies in its autonomous generation of explicit sentiment augmentations aligned with the provided aspect terms.

Compared with Common ABSA Baselines. Reviewing Table 2, it becomes apparent that the overall and explicit accuracy of the common ABSA methods and ABSA-ESA are relatively similar. However, a more pronounced disparity emerges when analyzing the accuracy of implicit data. ABSA-ESA outperforms these common ABSA methods by a considerable margin. This indicates that the most common ABSA methods struggle with implicit sentiment expression. Notably, among these methods, R-GAT presents a relatively competitive implicit performance by incorporating dependency relations within sentences to foster closer connections between aspect terms and opinion words. Nonetheless, there remains a substantial gap compared to ABSA-ESA's performance.

Compared with Data Augmentation Baselines. Analyzing Table 2, ABSA-ESA demonstrates a slight edge over the state-of-the-art data augmentation ABSA method (C^{3}DA)

	Model	Rest.14				Lap.14			
		All_A	All_F	ESE_A	ISE_A	All_A	All_F	ESE_A	ISE_A
BERT$_{base}$	BERT-$base$(110M)$^\flat$	83.57	77.16	89.21	65.54	78.22	73.45	81.47	69.71
	BT$^\natural$ (Fan et al. 2021)	86.47	80.29	93.37	64.41	79.67	75.79	82.60	72.21
	BERTAsp (Li et al. 2021b)	85.80	78.95	92.73	63.67	78.53	74.07	82.34	68.39
	EDA$^\natural$ (Wei and Zou 2019)	86.42	79.63	92.83	65.91	79.59	75.79	83.15	70.44
	PT$^\flat$ (Xu et al. 2019)	84.95	76.96	91.26	64.79	78.07	75.08	81.47	71.27
	ADA$^\flat$ (Rietzler et al. 2020)	87.14	80.05	**94.14**	65.92	78.96	74.18	82.76	70.11
	R-GAT$^\flat$ (Wang et al. 2020)	86.60	81.35	92.73	67.79	78.21	74.07	82.44	72.99
	CLEAN$^\flat$ (Wang et al. 2022b)	87.05	81.40	92.50	69.66	80.41	77.25	81.21	78.29
BERT$_{large}$	BERT-$large$(340M)†	86.89	79.62	93.45	66.06	79.77	76.64	82.73	72.02
	R-GAT†(Wang et al. 2020)	86.79	80.17	93.77	68.55	79.62	75.95	81.22	75.42
	BERTAsp† (Li et al. 2021b)	87.01	81.07	93.25	67.05	80.78	77.54	83.48	73.71
	ASA-WD (Tian, Chen, and Song 2021)	86.88	80.92	92.62	68.53	80.41	77.38	83.41	72.57
T5$_{base}$	T5-$base$(220M)†	86.60	79.24	93.08	65.92	80.25	76.13	83.17	72.60
	C^{3}DA$^\natural$ (Wang et al. 2022a)	86.93	81.23	93.59	65.54	80.61	77.11	82.68	73.57
	R-GAT† (Wang et al. 2020)	86.87	79.99	92.84	68.05	80.25	76.26	82.05	75.43
	T5Asp† (Li et al. 2021b)	87.11	80.95	93.27	67.42	79.49	77.46	81.92	73.14
	ABSA-ESA	**88.29**	**81.74**	93.77	**70.78**	**82.44**	**79.34**	**83.35**	**80.00**

Table 2: Main results of ABSA-ESA and each baselines. The result with Bold is the best result. The results with ♭ are from (Wang et al. 2022b) and the results with ♮ are from (Wang et al. 2022a). The results produced by ourselves are labeled with †. All, ESE and ISE denote the result obtained by using all data, explicit data and implicit data, respectively. The subscripts A and F represent the accuracy and macro-F1 score, respectively.

Model	Rest.14		Lap.14	
	All_F	ISE_F	All_F	ISE_F
Flan-T5(250M)	81.52	69.66	79.32	72.02
Flan-T5 Prompt ♡	81.50	70.91	79.02	76.40
THOR ♡(Fei et al. 2023)	82.98	71.70	79.75	67.63
ABSA-ESA $_{\text{Flan-T5}}$	**83.79**	**73.76**	**81.78**	**77.91**

Table 3: Results compared with THOR. All the baselines use Flan-T5 as the encoder. The results with ♡ are obtained from (Fei et al. 2023). The metrics are the same as Table 2.

Model	Rest.14		Lap.14	
	ESE_A	ISE_A	ESE_A	ISE_A
w/o SDW	92.78	68.16	82.64	77.58
w/o UCR	93.53	69.28	83.00	79.31
w/o SDW and UCR	92.67	68.39	82.41	77.90
w/o CBS	93.36	67.81	82.06	78.65
Full ABSA-ESA	**93.77**	**70.78**	**83.35**	**80.00**

Table 4: Ablation study of ABSA-ESA. w/o denotes the version without the specific component. The metrics are the same as Table 2.

concerning overall F1 scores. However, the advantage becomes more pronounced when considering implicit sentiment, with ABSA-ESA surpassing C^{3}DA by a significant margin. This divergence stems from ABSA-ESA's direct inclusion of sentiment clues within the input sentence and its ability to generate aspect-related opinion words autonomously. In contrast, C^{3}DA's augmentation quality heavily relies on seed spans, which produce sentiment polarity inversion with the input sentence.

Ablation Study

In this section, we analyze the influence of the three strategies on the final results, as outlined in Table 4. Notably, all strategies exhibit minimal impact on explicit sentiment data. Conversely, for implicit sentiment data, the omission of these strategies results in substantial drops in performance, particularly evident with the SDW strategy removed (2.62↓ on Rest.14 and 2.42↓ on Lap.14). Furthermore, a comparison between the version without SDW and the version without both strategies reveals no significant divergence in their

Aspect	Sentence	Prediction	Ground Truth
hostess	I have tried to make reservations, but both times, the *hostess* didn't have my name.	`Neutral`	`Negative`
	I have tried to make reservations, but both times, the *hostess* didn't have my name. **Hostess service was <u>awful</u>.**	`Negative`	
waitress	The *waitress* suggested glasses of wine that went very well with the food.	`Neutral`	`Positive`
	The *waitress* suggested glasses of wine that went very well with the food. **The waitress was very <u>nice and attentive</u>.**	`Positive`	
features	I was looking for a Mac which is portable and has all the *features* that I was looking for.	`Neutral`	`Positive`
	I was looking for a Mac which is portable and has all the *features* that I was looking for. **I <u>love</u> the features and the design.**	`Positive`	
service	Going to bring it to service today.	`Neutral`	`Neutral`
	Going to bring it to service today. **I have not seen a problem with the service.**	`Neutral`	

Table 5: Case study of ABSA-ESA. The sentence in Bold is the augmented sentence, and the word with underline is the generated explicit sentiment expression.

Model	Rest.14		Lap.14	
	All	ISE	All	ISE
T5-base	2.8936	2.5178	2.9123	3.0806
R-GAT	5.3242	5.4460	5.9510	5.5374
T5Asp	4.2844	4.5287	4.8858	4.9604
ABSA-ESA	**2.4267**	**2.3965**	**2.5218**	**2.6315**

Table 6: Average prediction entropy of each model. All models are based on T5.

results. This implies that relying on UCR offers limited enhancements to implicit accuracy and may even reduce model performance. Practical implementation suggests combining SDW and UCR for optimal outcomes.

Case Study

In this section, we present some cases in Table 5 to show the effectiveness of ABSA-ESA. We select totally four examples from two datasets. Firstly, it is evident that ABSA-ESA adeptly generates explicit sentiment expressions (highlighted in red within the sentences) in line with the input sentences, incorporating aspect-related opinion words. For instance, the opinion word "attentive" matches the aspect term "waitress" very well. Secondly, a discernible pattern emerges from the model's predictions: for implicit sentiment data, the model consistently outputs a `Neutral` sentiment polarity. After the augmentation, the model's predictions align accurately with the ground truth sentiment. Thirdly, in scenarios with `Neutral` ground truth sentiment, ABSA-ESA refrains from generating explicit sentiment. The augmentation for the final example has no opinion words, preserving the ABSA-ESA's original prediction.

Further Analysis

In this section, we introduce a novel perspective employing average entropy to dissect the ABSA results. The average entropy of ABSA model predictions is calculated as follows:

$$H(x) = -\frac{1}{n}\sum_{i=1}^{n}\sum_{c\in\mathcal{Y}} p_c(x_i)\log p_c(x_i), \tag{10}$$

where $p_j(x_i)$ denotes the probability of the sentence i on class c. As showcased in Table 6, ABSA-ESA boasts the lowest average entropy under the same backbone (T5). This outcome suggests that not only does ABSA-ESA accurately classify implicit sentiment data, but its classification tendencies are markedly distinct. This phenomenon arises primarily due to the integration of explicit sentiment augmentations, which endows sentences with pronounced sentiment inclinations, consequently enhancing discernibility. Furthermore, it is notable that both the overall and implicit entropy of ABSA-ESA is lower than those of the T5-base. This suggests that the incorporation of explicit sentiment augmentations not only enhances implicit classification tendencies but also brings about benefits to explicit sentiment data.

Conclusion

This paper proposes an ABSA method to deal with implicit sentiment expression by integrating explicit sentiment augmentations (ABSA-ESA). We design a novel ABSA-specific explicit sentiment generation method based on the T5 model and involve three strategies to constrain the generated augmentations on sentiment polarity and the aspect term. Specifically, Syntax Distance Weighting and Unlikelihood Contrastive Regularization are used to lead the model to generate explicit opinion words with the correct polarity. And Constrained Beam Search is used to ensure the augmentations are aspect-related. We test ABSA-ESA on two ABSA benchmarks, and empirical results demonstrate that ABSA-ESA can outperform existing ABSA baselines on both explicit and implicit sentiment accuracy.

Acknowledgements

This work is supported by the project from the National Key R&D Program of China (No.2021ZD0112500), the National Natural Science Foundation of China (NSFC) (No.62276113), the Project Funded by China Postdoctoral Science Foundation (No.2022M721321), the Energy Administration of Jilin Province (No.3D516L921421), and the Fundamental Research Funds for the Central Universities, JLU.

References

Bi, W.; Li, H.; and Huang, J. 2021. Data Augmentation for Text Generation Without Any Augmented Data. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics*, 2223–2237.

Chen, C.; Teng, Z.; and Zhang, Y. 2020. Inducing target-specific latent structures for aspect sentiment classification. In *Proceedings of the 2020 conference on empirical methods in natural language processing (EMNLP)*, 5596–5607.

Chen, D. Z.; Faulkner, A.; and Badyal, S. 2022. Unsupervised Data Augmentation for Aspect Based Sentiment Analysis. In *Proceedings of the 29th International Conference on Computational Linguistics*, 6746–6751.

Cheng, J.; Zhao, S.; Zhang, J.; King, I.; Zhang, X.; and Wang, H. 2017. Aspect-level sentiment classification with heat (hierarchical attention) network. In *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*, 97–106.

Choi, Y.; Wiebe, J.; and Mihalcea, R. 2017. Coarse-grained+/-effect word sense disambiguation for implicit sentiment analysis. *IEEE Transactions on Affective Computing*, 8(4): 471–479.

Dai, J.; Yan, H.; Sun, T.; Liu, P.; and Qiu, X. 2021. Does syntax matter? A strong baseline for Aspect-based Sentiment Analysis with RoBERTa. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics*, 1816–1829.

Deng, L.; Wiebe, J.; and Choi, Y. 2014. Joint inference and disambiguation of implicit sentiments via implicature constraints. In *Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers*, 79–88.

Fan, A.; Bhosale, S.; Schwenk, H.; Ma, Z.; El-Kishky, A.; Goyal, S.; Baines, M.; Celebi, O.; Wenzek, G.; Chaudhary, V.; et al. 2021. Beyond english-centric multilingual machine translation. *The Journal of Machine Learning Research*, 22(1): 4839–4886.

Fei, H.; Li, B.; Liu, Q.; Bing, L.; Li, F.; and Chua, T.-S. 2023. Reasoning Implicit Sentiment with Chain-of-Thought Prompting. *arXiv preprint arXiv:2305.11255*.

Gao, F.; Zhu, J.; Wu, L.; Xia, Y.; Qin, T.; Cheng, X.; Zhou, W.; and Liu, T.-Y. 2019. Soft contextual data augmentation for neural machine translation. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, 5539–5544.

He, R.; Lee, W. S.; Ng, H. T.; and Dahlmeier, D. 2019. An Interactive Multi-Task Learning Network for End-to-End Aspect-Based Sentiment Analysis. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, 504–515.

Hou, X.; Qi, P.; Wang, G.; Ying, R.; Huang, J.; He, X.; and Zhou, B. 2021. Graph Ensemble Learning over Multiple Dependency Trees for Aspect-level Sentiment Classification. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 2884–2894.

Hsu, T.-W.; Chen, C.-C.; Huang, H.-H.; and Chen, H.-H. 2021. Semantics-preserved data augmentation for aspect-based sentiment analysis. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, 4417–4422.

Kambhatla, N.; Born, L.; and Sarkar, A. 2022. CipherDAug: Ciphertext based Data Augmentation for Neural Machine Translation. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 201–218.

Lam, T. K.; Schamoni, S.; and Riezler, S. 2022. Sample, Translate, Recombine: Leveraging Audio Alignments for Data Augmentation in End-to-end Speech Translation. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, 245–254.

Li, J.; Zhao, Y.; Jin, Z.; Li, G.; Shen, T.; Tao, Z.; and Tao, C. 2022. SK2: Integrating Implicit Sentiment Knowledge and Explicit Syntax Knowledge for Aspect-Based Sentiment Analysis. In *Proceedings of the 31st ACM International Conference on Information & Knowledge Management*, 1114–1123.

Li, R.; Chen, H.; Feng, F.; Ma, Z.; Wang, X.; and Hovy, E. 2021a. Dual graph convolutional networks for aspect-based sentiment analysis. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics*, 6319–6329.

Li, Z.; Zou, Y.; Zhang, C.; Zhang, Q.; and Wei, Z. 2021b. Learning Implicit Sentiment in Aspect-based Sentiment Analysis with Supervised Contrastive Pre-Training. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, 246–256.

Liu, B. 2012. Sentiment Analysis and Opinion Mining. *Synthesis Lectures on Human Language Technologies*, 5(1): 1–167.

Liu, G.; Yuan, O.; Jin, L.; and Zhou, J. 2022. Dynamic Augmentation Data Selection for Few-shot Text Classification. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, volume 2022, 4870–4881.

Ouyang, J.; Wang, Y.; Li, X.; and Li, C. 2022. Weakly-supervised text classification with wasserstein barycenters regularization. In *International joint conference on artificial intelligence*, 3373–3379.

Ouyang, J.; Yang, Z.; Xuan, C.; Wang, B.; Wang, Y.; and Li, X. 2023. Unsupervised Aspect Term Extraction by Integrating Sentence-level Curriculum Learning with Token-level

Self-paced Learning. In *Proceedings of the 32nd ACM International Conference on Information and Knowledge Management*, 1982–1991.

Pennington, J.; Socher, R.; and Manning, C. D. 2014. Glove: Global vectors for word representation. In *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*, 1532–1543.

Raffel, C.; Shazeer, N.; Roberts, A.; Lee, K.; Narang, S.; Matena, M.; Zhou, Y.; Li, W.; and Liu, P. J. 2020. Exploring the limits of transfer learning with a unified text-to-text transformer. *The Journal of Machine Learning Research*, 21(1): 5485–5551.

Rietzler, A.; Stabinger, S.; Opitz, P.; and Engl, S. 2020. Adapt or Get Left Behind: Domain Adaptation through BERT Language Model Finetuning for Aspect-Target Sentiment Classification. In *Proceedings of the 12th Language Resources and Evaluation Conference*, 4933–4941.

Sun, K.; Zhang, R.; Mensah, S.; Mao, Y.; and Liu, X. 2019. Aspect-level sentiment analysis via convolution over dependency tree. In *Proceedings of the 2019 conference on empirical methods in natural language processing*, 5679–5688.

Tang, D.; Qin, B.; Feng, X.; and Liu, T. 2016. Effective LSTMs for Target-Dependent Sentiment Classification. In *Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers*, 3298–3307.

Tian, Y.; Chen, G.; and Song, Y. 2021. Enhancing aspect-level sentiment analysis with word dependencies. In *Proceedings of the 16th conference of the European chapter of the association for computational linguistics: main volume*, 3726–3739.

Wang, B.; Ding, L.; Zhong, Q.; Li, X.; and Tao, D. 2022a. A Contrastive Cross-Channel Data Augmentation Framework for Aspect-Based Sentiment Analysis. In *Proceedings of the 29th International Conference on Computational Linguistics*, 6691–6704.

Wang, K.; Shen, W.; Yang, Y.; Quan, X.; and Wang, R. 2020. Relational Graph Attention Network for Aspect-based Sentiment Analysis. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 3229–3238.

Wang, S.; Zhou, J.; Sun, C.; Ye, J.; Gui, T.; Zhang, Q.; and Huang, X.-J. 2022b. Causal Intervention Improves Implicit Sentiment Analysis. In *Proceedings of the 29th International Conference on Computational Linguistics*, 6966–6977.

Wang, Y.; Huang, M.; Zhu, X.; and Zhao, L. 2016. Attention-based LSTM for aspect-level sentiment classification. In *Proceedings of the 2016 conference on empirical methods in natural language processing*, 606–615.

Wei, J.; and Zou, K. 2019. EDA: Easy Data Augmentation Techniques for Boosting Performance on Text Classification Tasks. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing*, 6382–6388.

Wu, X.; Gao, C.; Lin, M.; Zang, L.; and Hu, S. 2022. Text Smoothing: Enhance Various Data Augmentation Methods on Text Classification Tasks. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, 871–875.

Xu, H.; Liu, B.; Shu, L.; and Philip, S. Y. 2019. BERT Post-Training for Review Reading Comprehension and Aspect-based Sentiment Analysis. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics*, 2324–2335.

Xu, H.; Shu, L.; Philip, S. Y.; and Liu, B. 2020. Understanding Pre-trained BERT for Aspect-based Sentiment Analysis. In *Proceedings of the 28th International Conference on Computational Linguistics*, 244–250.

Xu, M.; Wang, D.; Feng, S.; Yang, Z.; and Zhang, Y. 2022. Kc-isa: An implicit sentiment analysis model combining knowledge enhancement and context features. In *Proceedings of the 29th International Conference on Computational Linguistics*, 6906–6915.

Xu, X.; Wang, G.; Kim, Y.-B.; and Lee, S. 2021. AugNLG: Few-shot Natural Language Generation using Self-trained Data Augmentation. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics*, 1183–1195.

Yang, Z.; Wang, B.; Li, X.; Wang, W.; and Ouyang, J. 2023. S3 map: Semisupervised aspect-based sentiment analysis with masked aspect prediction. *Knowledge-Based Systems*, 269: 110513.

Zhou, D.; Wang, J.; Zhang, L.; and He, Y. 2021a. Implicit sentiment analysis with event-centered text representation. In *Proceedings of the 2021 conference on empirical methods in natural language processing*, 6884–6893.

Zhou, Y.; Liao, L.; Gao, Y.; Jie, Z.; and Lu, W. 2021b. To be Closer: Learning to Link up Aspects with Opinions. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, 3899–3909.