SynPrompt: Syntax-aware Enhanced Prompt Engineering forAspect-based Sentiment Analysis

University of Electronic Science and Technology of China, Chengdu China

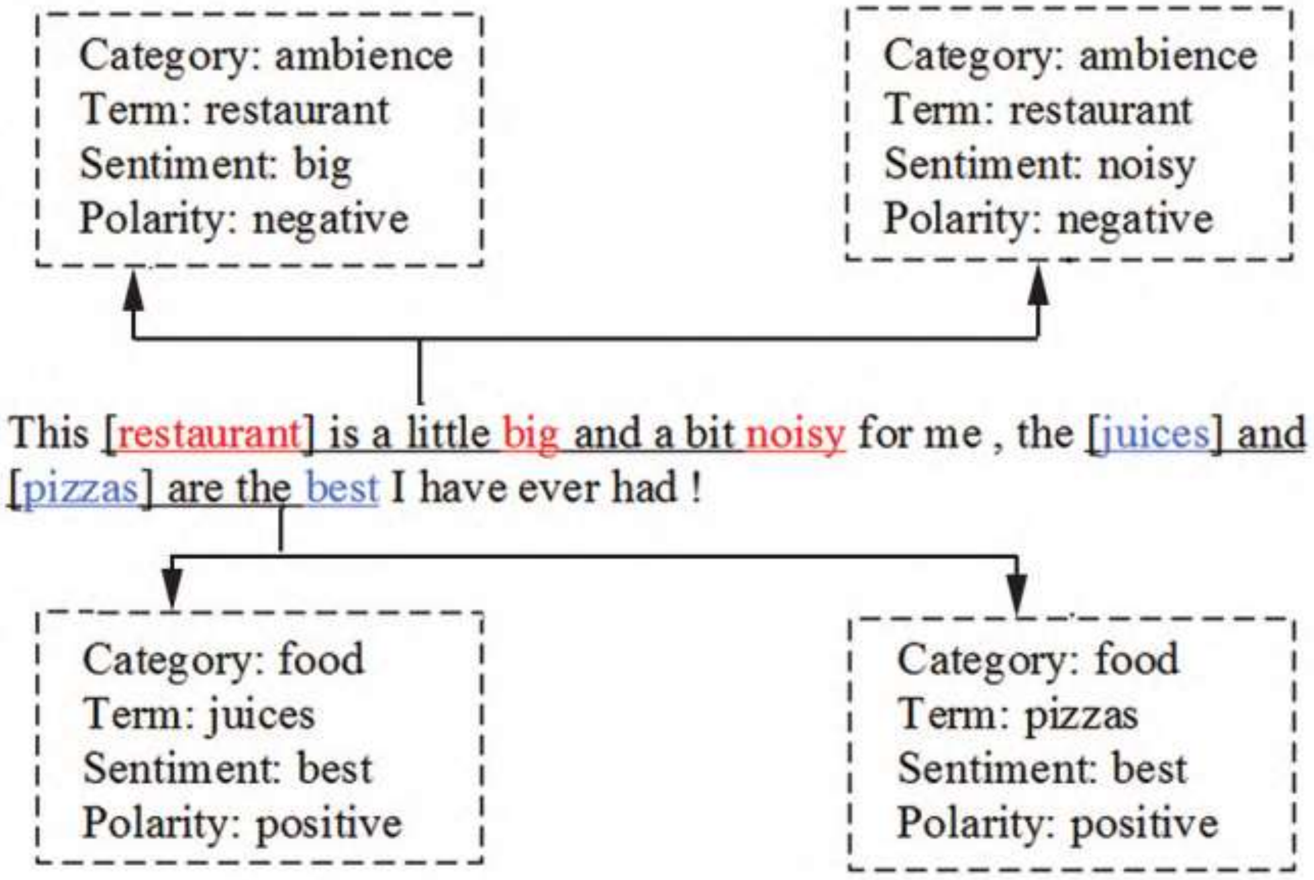
Wen Yin

LREC-COLING 2024

Introduction

Introduce the starting point and main idea.

Aspect-based sentiment analysis(ABSA)



Basic sub-tasks:

Aspect Sentiment Classification(ASC)

Aspect Term Extraction (ATE)

Critical problems

Rough prompts

Recent studies suggest that rough engineering is difficult to adapt to specific task scenarios.

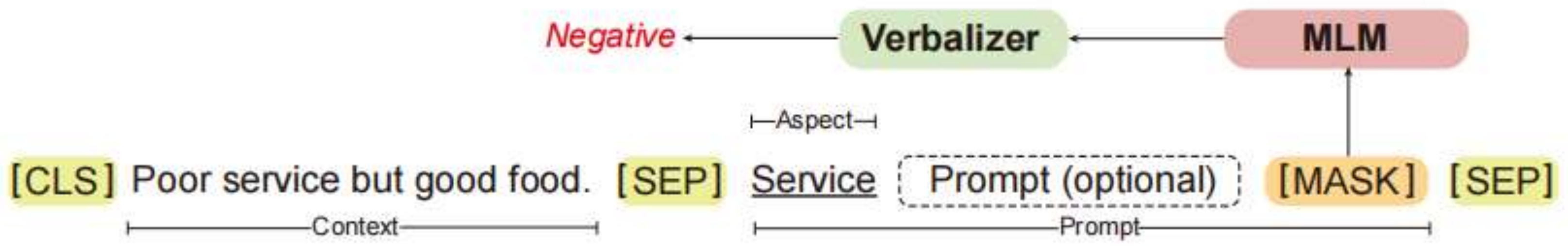


Figure 1: A general template based on prompt tuning for the ABSA task.

Solution

More robust and professional **Prompt Engineering**

Syntax-aware Enhanced Prompt method

Constructs our prompt by looking for words that are related to aspect words syntactic

Apply

Adapt

Adaptive prompt frameworks

Utilizing discrete prompt or continuous prompt structure

Methodology

Introduce the technical details of PFDualBERT.

Overall architecture of PFDualBERT

Prompt-based Tuning

Prompt-based tuning uses the [MASK] token as a predictor instead of the [CLS] token.

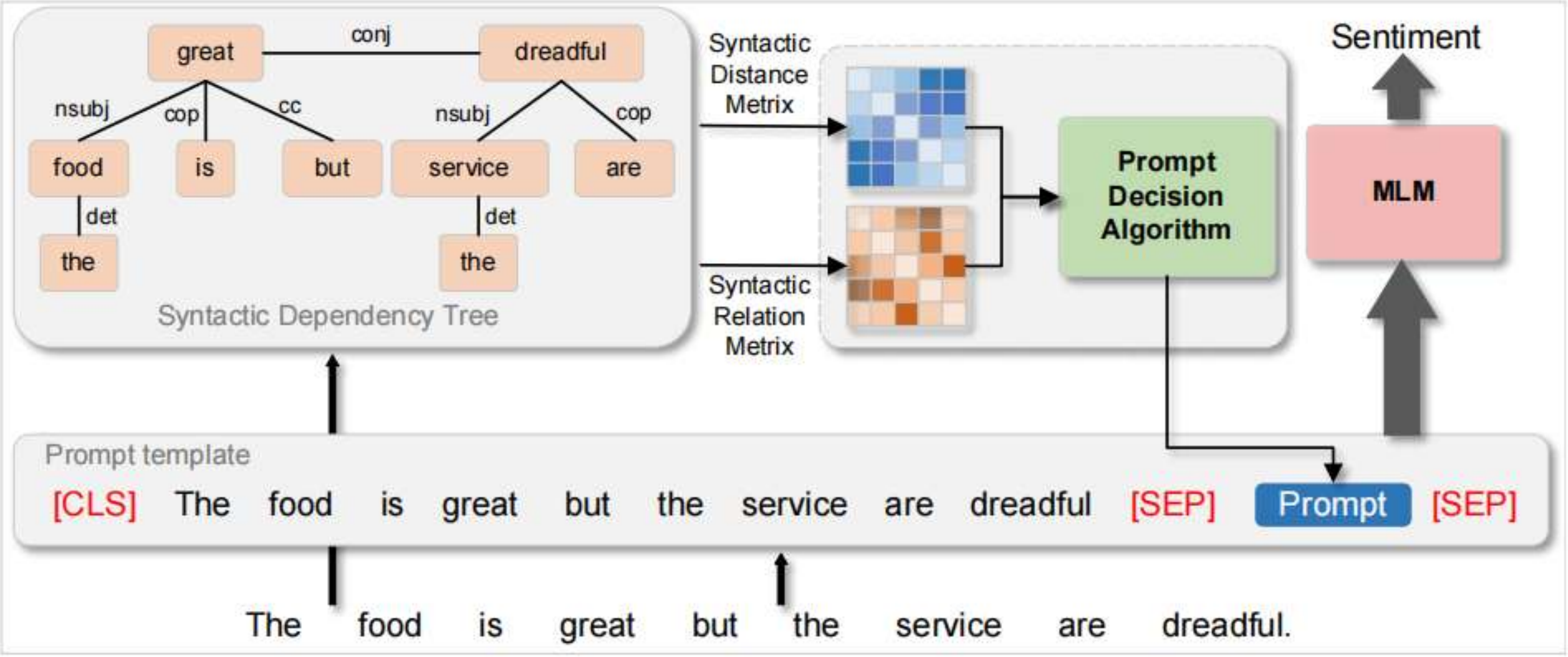


Figure 2: The construction process of Syntactic Distance Matrix and Syntactic Relation Matrix targets the example sentence.

Use the words predicted by the [MASK] as the input of the downstream task.

The class probability distribution is obtained by marginalizing the set of label tokens:

$$P(y|x_{\text{prompt}}) = \sum_{w \in \mathcal{V}_y} p([\text{MASK}] = w | x_{\text{prompt}})$$

Syntax-aware Enhanced Prompt

Syntactic Distance Matrix

The original distance matrix $D^*(i,j) = d(w_i, w_j)$

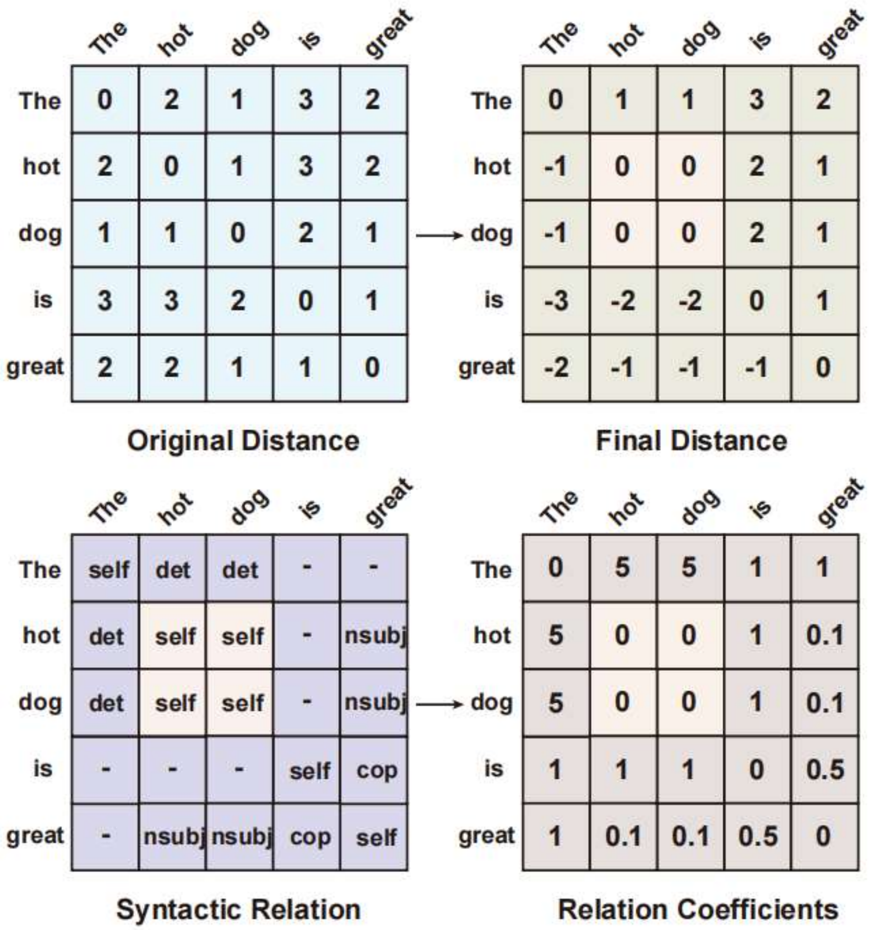
Determine the positive and negative of these distances

$$D(i,j) = \begin{cases} D^*(i,j), & i < j \\ -D^*(i,j), & i \geq j \end{cases}$$

Syntactic Relation Matrix

We obtain the final relation matrix R by mapping each relation to predefined coefficients

$$R(i,j) = \mathrm{MAP}(Relation(i,j))$$



Prompt Decision Algorithm

Algorithm 1: Prompt Decision

Data: Syntactic distance Matrix D and syntactic relation matrix R . The index of aspect words a in the sentence x . Decision threshold k.

Result: Decided prompt word list $\mathcal{P}$ of the given matrixes.

Initialize $\mathcal{P} = \{x_a\}, \mathcal{C} = \{\}, t = 0$

Get adjacent matrix $M \leftarrow D \odot R$

Get List of shortest paths about aspect:

$\mathcal{C} \leftarrow Dijkstra(M, a)$

while $t \leq k$ **do**

- $i \leftarrow getIndexOfMin(\mathcal{C})$
- if $i < a$ then $\mathcal{P}.insert(\mathrm{x}_i)$
- if $i \geq a$ then $\mathcal{P}.append(\mathrm{x}_i)$
- Update $\mathcal{C}_i \leftarrow$ DBLMAX
- $t \leftarrow t + 1$

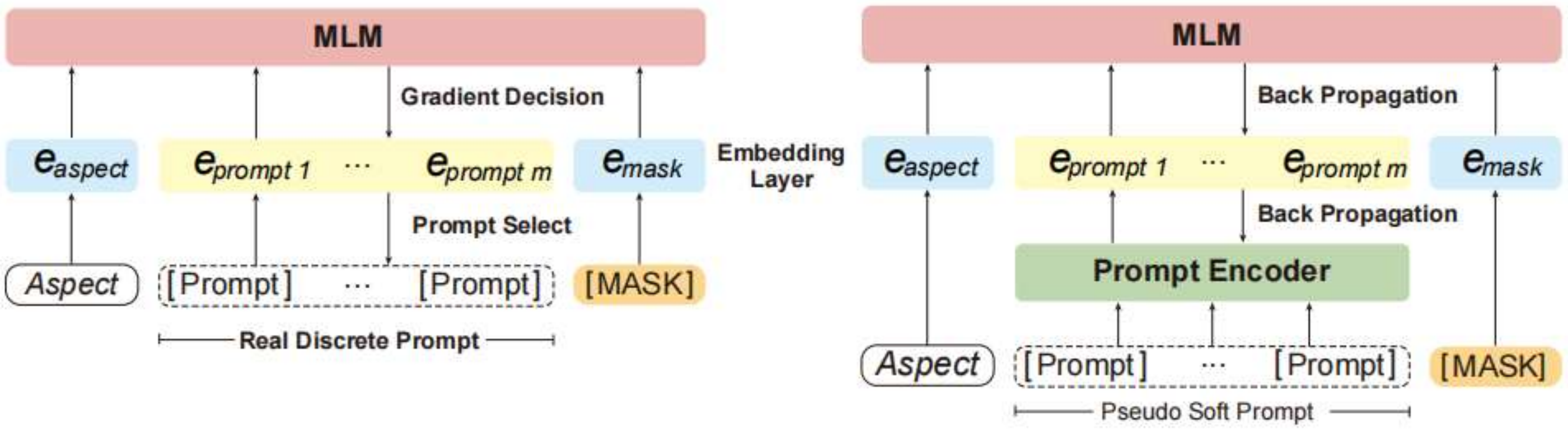
return $\mathcal{P}$;

We compute the result of the dot multiplication of matrix D and R.

Then, we compute the shortest distance from each node with respect to the aspect nodes using **Dijkstra**'s algorithm.

The word closest to the aspect is determined as the prompt word.

Adaptive Prompt Frameworks



(a) Auto Prompt

(b) Soft Prompt

Gradient-Based Prompt Selection

Identify a candidate set V_cand of the top-k tokens estimated to cause the greatest increase

$$\mathcal{V}_{cand} = \underset{w\in\mathcal{V}}{\text{top-}k}\left[\mathbf{e}_w^\top \nabla \log p\left(y|x_{prompt}\right)\right]$$

Prompt Encoder

Minimizing the discreteness among prompt words.

we initially initialized the [Prompt] tokens constructed an encoder comprising a LSTM

$$\mathbf{e}_i = \text{MLP}\left(\text{LSTM}\left[\overrightarrow{\mathbf{e}_i} : \overleftarrow{\mathbf{e}_i}\right]\right)$$

Datasets

SemEval 2014 Task 4 : Restaurant and Laptop reviews ,and Twitter posts

Baselines

BERT-SPC	**2019**
SCAPT	**2021**
SentiPrompt	**2021**
MRCOOL	**2022**

Main Results

Method		Restarant		Laptop		Twitter	
		Accuracy	F1-score	Accuracy	F1-score	Accuracy	F1-score
FT	BERT-SPC	84.46	76.98	78.99	75.03	74.13	72.73
	SCAPT	83.39	74.53	77.17	73.23	74.64	73.91
PT	SentiPrompt*	84.39	75.53	79.17	76.23	75.64	74.88
	MRCOOL	85.49	**79.14**	78.12	75.78	-	-
Our	SynPrompt	84.17	76.45	78.28	75.36	74.88	73.73
	SynPrompt+AutoP	85.67	78.37	80.79	76.09	75.84	**74.37**
	SynPrompt+SoftP	**85.96**	78.45	**81.28**	**77.19**	**76.23**	74.30

Table 3: Experimental results (%) comparison on three publicly available datasets. We underline the second best performed baseline. The results of our rerunning version are marked with *.

Few-shot Study

Shot	Method	Restarant	Laptop	Twitter
1%	BERT-SPC	49.48	51.26	42.04
	SentiPrompt*	54.93 (+*5.45*)	54.06 (+*2.80*)	42.98 (+*0.94*)
	SynPrompt+AutoP	55.82 (+*6.34*)	55.12 (+*3.86*)	47.56 (+*5.52*)
	SynPrompt+SoftP	55.94 (+*6.46*)	54.53 (+*3.27*)	48.10 (+*6.06*)
2%	BERT-SPC	58.05	43.72	49.67
	SentiPrompt*	60.94 (+*2.49*)	49.93 (+*6.21*)	49.72 (+*0.05*)
	SynPrompt+AutoP	61.15 (+*3.10*)	53.06 (+*9.34*)	50.93 (+*1.30*)
	SynPrompt+SoftP	61.09 (+*3.04*)	53.96 (+*10.24*)	51.34 (+*1.67*)
4%	BERT-SPC	57.31	59.01	59.57
	SentiPrompt*	62.11 (+*4.80*)	61.65 (+*2.64*)	63.02 (+*3.45*)
	SynPrompt+AutoP	63.82 (+*6.51*)	62.65 (+*3.64*)	62.06 (+*2.49*)
	SynPrompt+SoftP	64.07 (+*6.76*)	60.64 (+*1.63*)	62.48 (+*2.91*)
8%	BERT-SPC	74.02	62.58	63.90
	SentiPrompt*	74.20 (+*0.18*)	66.67 (+*4.09*)	64.40 (+*0.50*)
	SynPrompt+AutoP	75.02 (+*1.00*)	66.92 (+*4.34*)	65.77 (+*1.87*)
	SynPrompt+SoftP	75.96 (+*1.94*)	67.03 (+*4.45*)	66.76 (+*2.86*)

Table 4: Experimental results (%) comparison of accuracy on three publicly available datasets through four levels of shot. The results of our rerunning version are marked with *.

Analysis of Key Variables Study

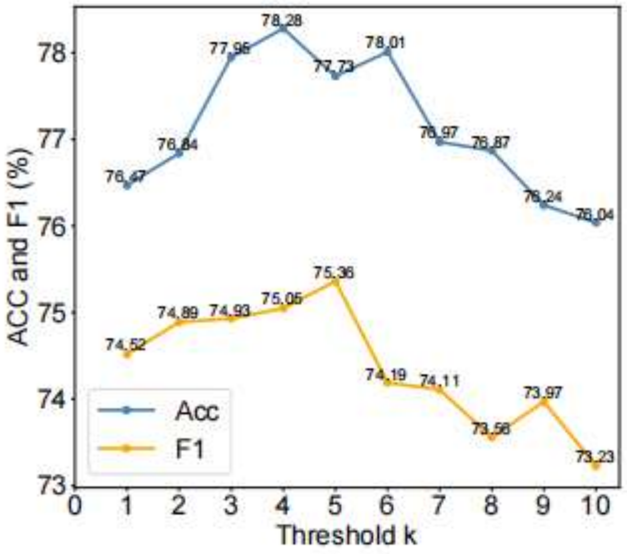


Figure 5: Effect of threshold k in SynPrompt. The results are produced under the full-shot setting on the Laptop dataset by vanilla SynPrompt.

Method	Template	Accuracy	F1-Score
AutoP	N=1	77.69	72.94
	N=2	77.82	73.05
	N=3	**78.14**	**73.29**
	N=4	77.95	72.91
	N=5	77.82	72.98
SoftP	N=1	77.98	72.98
	N=2	78.19	73.10
	N=3	**78.38**	73.19
	N=4	78.14	**73.29**
	N=5	78.17	73.07

Table 5: Effect of the number of prompt tokens N in the template. The results are produced under the full-shot setting on the Laptop dataset by vanilla AutoP and SoftP.

Conclusions

1.Propose the SynPrompt to conduct specifically tailored prompts

by fully exploiting the syntax information of the syntactic Dep.Tree.

2.Propose two prompt frameworks, which are adapted to our proposed SynPrompt

to improve the robustness and perception of SynPrompt.

3.The results demonstrate the effectiveness, rationality, and interpretability.

Thank You!

Yin wen

yinwenok@stu.uestc.edu.cn