SentKB-BERT: Sentiment-filtered Knowledge-based Stance Detection

Hongzhou Chen[1], Ke Yan[1,2,*], Mustafa Raad Kadhim[1], Kui Wu[1], Ling Tian[1,2]
[1] School of Computer Science and Engineering, University of Electronic Science and Technology of China, Chengdu, China
[2] KASH Institute of Electronics and Information Industry, Kashi, China
202221081122@std.uestc.edu.cn, {kyan, kadhim_mustafa}@uestc.edu.cn, 202312081605@std.uestc.edu.cn, lingtian@uestc.edu.cn

***Abstract*—Stance detection plays an important role in detecting users thoughts and opinions through their comments or feedback. However, few-shot and zero-shot stance detection is not yet ready to analyze a textual document based on various topic aspects. To address the issues, we present a novel sentiment-filtered knowledge-based stance detection based on BERT (SentKB-BERT). By encoding Wikipedia summaries and knowledge graphs between external entities together, the proposed text filtering technique increases the presentation of external knowledge and decreases the lack of sentiment expressions. Additionally, the sentiment and external knowledge are encoded jointly in a unified framework to construct the stance representation. The experimental results demonstrated that our proposed method significantly improves the performance of few-shot and zero-shot stance detection, where the improvements were up to 10.3% in different indicators. The source code is available at https://github.com/FFFcapatin328/SentKB.**

***Index Terms*—Stance Detection, Multi-modal Learning, Sentiment Analysis, External Knowledge**

I. Introduction

Nowadays, the global Internet is continuing to grow at a staggering scale. There are 5.16 billion internet users in the world today, meaning that 64.4 percent of the world's total population is now online [1]. A significant open issue in the field of research refers to comprehending people's thoughts and opinions in their daily postings [2]. Thus, stance detection is one of the main research problems regarding the automatic analysis of such contents [3].

Stance detection is a task in natural language processing that involves classifying text to discover the author's position towards a certain target, which can be either in favor, against, or neutral [4]. The traditional approach to stance detection is topic-specific, which develops classifier-topic pairs for a specific topic [5]. In order to develop a classifier for a new topic, it is necessary to have a large number of well-labeled training examples, which is impractical for the thousands of possible topics. To formalize the description of the issue, the researcher defined two variations of stance detection [6]: zero-shot stance detection, where the classifier is tested on a multitude of entirely new topics, and few-shot stance detection, where the classifier is tested on a large set of topics with minimal training examples. Aiming to overcome the lack of labeled data in zero-shot and few-shot scenarios, some pioneer studies [7–12]

* Ke Yan is the corresponding author.

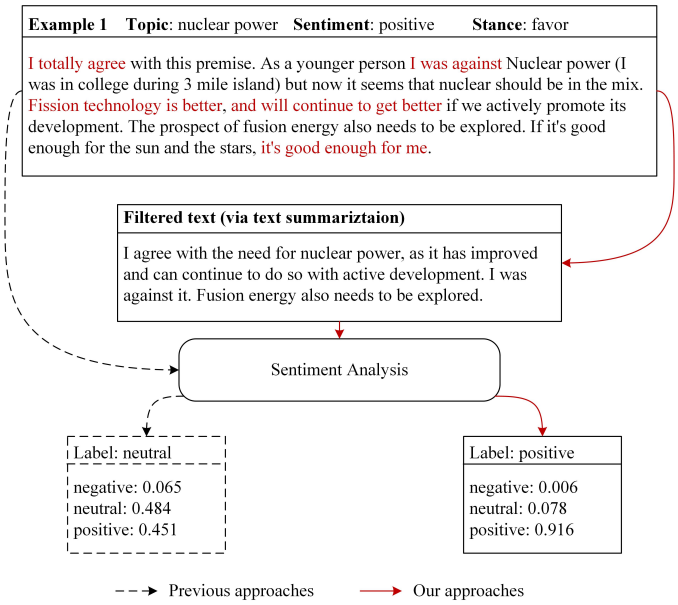


Fig. 1: An interpretive example of the importance of text filtering. We use a pre-trained model [19] as the sentiment analyzer in the figure.

focus on mining the deep semantics of the text to enrich the representation of the raw text. Some recent studies [13–15] seek external knowledge to supplement missing background information in the training data. Furthermore, to make the stance detection model better at generalization and reasoning, some researchers try to combine both deep semantics and external knowledge [16].

Recent works have proved that sentiment has a close relationship with stance [16–18]. From a linguistic perspective, the basic task in sentiment analysis is to have enough information to extract the characteristics [20]. For instance, if we base our evaluation on human evaluation, we will easily evaluate the given example in Fig. 1 as a positive sentiment (favor stance). However, the example expression caused bias when applying sentiment analysis to the whole text content and led to misclassifying the sentiment, as shown by the dashed line in Fig. 1. This issue also happens due to casual writing status, which leads baseline models to suffer from misclassifying issues.

To solve this issue, this paper proposes a novel Sentiment-

filtered Knowledge-based Stance Detection BERT (SentKB-BERT) for few-shot and zero-shot stance detection. The proposed approach involves the integration of i) text embedding via the process of fine-tuning a pre-trained model; ii) incorporating external knowledge into the original text embedding; and iii) applying a text filter to highlight the sentiment expression. In comparison to the previous approaches, the proposed model focuses on the sparse sentiment expression about the topic in the text by by the text filter from the raw text.

To be more specific, the main idea of the proposed model is described in Fig. 1, where the generated filtered text would be more obvious to both the sentiment model and humans to obtain the correct sentiment label. Finally, the contributions of this work is to propose a novel model to solve the issue of casual writing, the limited presence of sentimental expressions, and lack of relationships between external entities.

The rest of this paper is arranged as follows: Section II reviews background and related work. Section III provides the details of the proposed model. Section IV is the experimental and results discussion. Finally, section V details the conclusion and future work.

II. Related Work

Before going deeper into the details of the proposed method, we first review recent works on few-shot and zero-shot stance detection via internal semantics mining and external knowledge infusion methods.

The pioneering studies proposed some approaches to solve the problem of few-shot and zero-shot stance detection by mining the deep semantics between the text and the topics. BiCond [9] was introduced to encode the topics with a BiLSTM first and encode the text with another BiLSTM based on the encodding of the topics. Another study [10] proposed a model called CrossNet, which added an aspect-specific attention module after the same BiLSTM as BiCond before the classification. Based on BiLSTM, SEKT [11] presented a method using a semantic-emotion heterogeneous graph to enhance BiLSTM for cross-target stance detection. Simultaneously, a topic-grouped attention network [6] is constructed to encode the relationship between the training and evaluation topics in an unsupervised way. The TSDA multi-task learning framework [12] has been proposed to identify relevant contents for different topics, which can improve the quality of text representation.

However, those approaches only relied on the training and evaluation data, which could not fill in the implicit background knowledge in the text. To solve the issue, other studies introduced external knowledge into the stance detection models based on different techniques. CKE-Net [13] extracted a subgraph from common sense knowledge graph with the seed entities in the raw text. By concatenating the subgraph embedding with the text embedding, the external knowledge of the common sense knowledge graph can be infused into the model. However, a substantial amount of knowledge that is not captured by such relationships in the knowledge graph will be ignored by the subgraph embedding. Thus, Wikipedia Stance Detection BERT (WS-BERT) [14] was introduced to encode the knowledge from Wikipedia in the form of raw text and achieve an inspiring result. Furthermore, based on Wikipedia knowledge, a recent work [15] crawled summarizations of top-k reltaed targets on Wikipedia and infused all of the crawled wiki summaries into the text embedding. Inspired by these techniques, some researchers tried to integrate deep semantics and external knowledge. BS-RGCN [16] introduced an autograph encoder to represent external knowledge and utilized a pre-trained sentiment analysis model for encoding sentiment knowledge of the text, respectively.

III. Method

Following the inspiration of integrating external knowledge and internal deep semantics, the proposed method contributes in two aspects compared with the related work. i) For external knowledge embedding, we propose an external knowledge fusion module that integrates the relationships between entities and the attribute information of the entities in the text. ii) For mining internal deep semantics, we propose a sentiment-related text filter module that removes the neutral text to emphasize the sentiment expressions.

The proposed model for stance detection, depicted in Fig. 2, incorporates three complementary modules. i) The sentiment encoding module aims at filtering out expressions that are not relevant to sentiment. ii) The external knowledge infusing module represents external concepts and relationships. iii) This classification module merges the embedding results to identify the stance of the raw text.

A. Problem definition

Suppose D is a dataset that concludes N instances where each ith instance contains a textual document d^i toward a specific topic t^i. Every d^i and t^i are unimportant to be unique in the whole dataset because a single textual document could correspond to different stance labels based on different topic t^i. Thus, $D = \{(d^1, t^1), \cdots, (d^i, t^i), \cdots, (d^n, t^n)\}$, where $N = [1, 2, \cdots, n]$. Further, the dataset should contain the y set, which contains the correct stance labels, where the number of instances in the y set should be equal to D. Thus, $y = \{\gamma^1, \cdots, \gamma^i, \cdots, \gamma^n\}$ and $\gamma^i \in [Aginst, Favor, Neutral]$. Finally, the datasets used in this paper contain three types of data and ground truth labels, which are training, validation, and testing, respectively.

B. Encoding stance-related Sentiment

The stance-related sentiment encoding module is divided into two consistent processes, which contain a text filter and a sentiment encoder.

Stance-related text filter. The main idea of the text filter is to emphasize sentiment expressions related to stance by filtering out text that contains irrelevant background information or declarative sentences with neutral sentiment. We employ a text summarization model due to the fact that the text summarizer is designed to compress a given text into a more concise version that captures the essential information

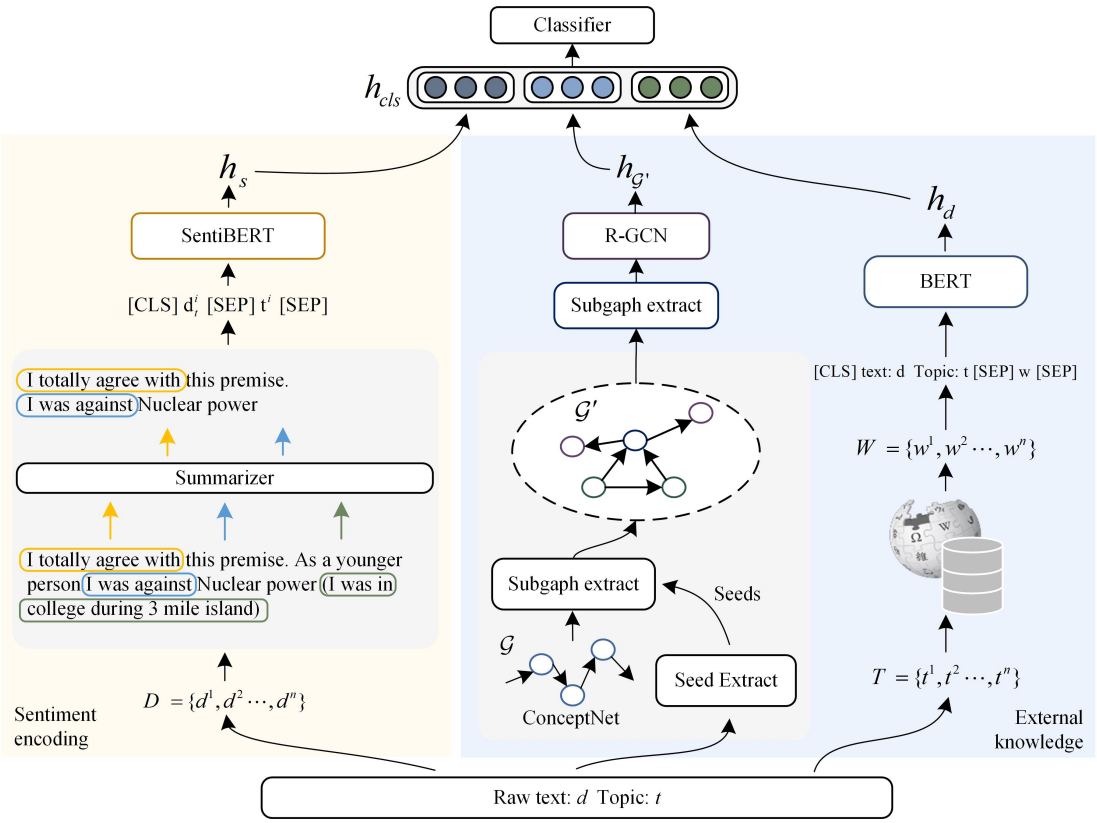


Fig. 2: The architecture of our model. The 'External knowledge' block indicates the external knowledge fusion modules. The 'Sentiment encoding' block indicates the stance-related sentiment encoding modules.

from the original document [21]. By summarizing the text, the sentiment encoder can remove most of the neutral sentiment sentences and get a summarization with non-neutral sentiment related to the topic. Every textual document d^i in D will be fed into the summarizer to obtain filtered d_f^i and store the reformed documents into $D_f = \{d_f^1, d_f^2, \cdots, d_f^n\}$.

Sentiment encoder. The sentiment encoder employs a BERT-based model trained with sentiment masks, denoted as SentiBERT. It takes the filtered sequence D_f as its input. As BERT requires data to be entered in the form of $[CLS]data[SEP]data[SEP]$, for instance d_f^i, the encoder concatenates the instance with the corresponding topic as $[CLS]d_f^i[SEP]t^i[SEP]$ and feeds it into the SentiBERT to obtain sentiment embedding $\boldsymbol{h_s^i}$ as follows:

$$\boldsymbol{h_s^i} = SentiBERT([CLS]d_f^i[SEP]t^i[SEP]) \quad (1)$$

C. Infusing External Knowledge

The external knowledge is integrated via the encoding of target-related text obtained from Wikipedia and the embedding of the target-related subgraphs of the commonsense knowledge graph. Therefore, the module can efficiently utilize the information from both raw text and the knowledge graph by mutually enhancing each other, thereby reducing the information loss during text embedding and knowledge graph construction.

Embedding the subgraph. The representation of relationship knowledge is enhanced by embedding the commonsense knowledge graph. We choose ConceptNet [22] as our base graph, which has 14 million edges with 34 kinds of relations. We denote ConceptNet as a directed graph $\mathcal{G} = \{\mathcal{V}, \mathcal{E}, \mathcal{R}\}$. To extract a related subgraph from the base graph $\mathcal{G}$, we extract all unique concepts that appeared in the original text and topic as seeds and denote them as C_d and C_t. Then we filter the base graph $\mathcal{G}$ with C_d and C_t to get the full subgraph $\mathcal{G}' = \{\mathcal{V}', \mathcal{E}', \mathcal{R}'\}$, which is related to all of the original text. After extracting the full subgraph $\mathcal{G}'$, we train a graph encoder following [23] to embed the subgraph as graph feature vectors, denoted as $\boldsymbol{h_{\mathcal{G}'}}$.

Target-related external text infusing. The summaries obtained from Wikipedia for the targets are regarded as an initial textual representation of external knowledge. Specifically, we search Wikipedia with the targets as key words and retrieve the summarization text from the pages. To infuse the retrieved text into the original text, we employ BERT as our text encoder. Since BERT requires data to be entered in the form of $[CLS]d^i[SEP]t^i[SEP]$, we concatenate the original text and the target and redesign the input form as $[CLS]text: d^i, traget: t^i[SEP]w^i[SEP]$, where w^i denotes the corresponding retrieved Wikipedia summaries. We obtain $\boldsymbol{h_d^i}$, the embedding of raw text and wiki summaries, as follows:

$$\boldsymbol{h_d^i} = BERT([CLS]text: d^i, target: t^i[SEP]w^i[SEP]) \quad (2)$$

D. Stance Classifier

The stance classifier is constructed with a multi-layer perception (MLP) to classify text-topic pairs (d^i, t^i) and obtain their stance label $\gamma^i \in \{aginst, favor, neutral\}$ through the incorporation of text embedding, external knowledge embedding, and sentiment embedding. In particular, after concatenating sentiment embedding $\boldsymbol{h_s}$, external knowledge graph embedding $\boldsymbol{h_{\mathcal{G}'}}$, and text embedding with wiki summaries $\boldsymbol{h_d}$, we feed the concatenated result into a fully-connected layer with a softmax layer as follows:

$$\boldsymbol{\gamma^i} = softmax(MLP([\boldsymbol{h_s^i}; \boldsymbol{h_{\mathcal{G}'}^i}; \boldsymbol{h_d^i}])) \tag{3}$$

in which the symbol ';' denotes the vector concatenation operation.

E. Learning objective

The graph encoder is first trained on the full subgraph $\mathcal{G}'$ and generates the corresponding graph encoding for each sentence and topic before training. Also, the summarization of each sentence is generated by the summarization model. During training, only the parameters of BERT and SentiBERT will be fine-tuned.

The final learning objective is to minimize the cross-entropy loss via the standard gradient descent algorithm. The learning objective is defined as follows:

$$\min_{\Theta} \mathcal{L} = -\sum_{i=1}^{N}\sum_{j=1}^{M} \gamma_i^j \log \hat{\gamma}_i^j + \lambda \|\Theta\|^2 \tag{4}$$

where M denotes the dimensionality of stance labels, γ_i denotes the ground truth label of the stance, $\hat{\gamma}_i$ denotes the prediceted label from our model, λ denotes the $L2$-regularization factor, and Θ denotes the trainable parameters of our model.

IV. Experiments

A. Datasets and Evaluation Metrics

We evaluated the proposed model on VAST [6] in 2020, which is released for few-shot and zero-shot stance detection. The statistics of the dataset are shown in Table I. We use an API to crawl the Wikipedia pages of the targets and save the summaries in string format as the raw text of the external knowledge.

Following previous works, we adopt the macro-average of the F1-score as the evaluation metric. Precisely, for VAST, which has three stance labels, $F_{\text{average}} = (F_{\text{against}} + F_{\text{favor}} + F_{\text{neutral}})/3$.

TABLE I: Data statistics of VAST

Statistics	***Train***	***Dev***	***Test***
Instances	13477	2062	3066
Unique comment	1845	682	786
Few-shot topics	638	114	159
zero-shot topics	4003	383	600

To evaluate our model in different scenarios, we calculated three different F1-scores for three different categories: *Few-Shot*, *Zero-Shot*, and *All*, which correspond to stance detection instances under few-shot, zero-shot, and all scenarios, respectively.

B. Settings

We employ the base version of BERT as our text encoder. For our sentiment-related text filter, we chose a text summarization model as its backbone. Precisely, we employ the Flan T5 summarizer [24] and feed the summarized text to a RoBERTa-base model [19] trained on 124 million tweets from January 2018 to December 2021 and fine-tuned for sentiment analysis with the TweetEval benchmark. The graph encoder consists of two stacked RGCN [23]. We train our model on 1 GPU (Nvidia Geforce RTX 4090 24GB) using the Adam optimizer [25] with an initial learning rate of 8.142587485342115e-6 and batch size of 16 for training, which are the best hyperparameters automatically chosen by NNI [26]. The best checkpoints are selected according to the evaluation metrics on the evaluation set. In addition, we follow the standard train/validation/test splits of VAST in our experiments [6].

C. Results and Analysis

We compared our model with the baseline models introduced in section II including BiCond, Cross-Net, SEKT, TGA-Net, TSDA, CKE-Net, BS-RGCN, TarBK-BERT, and WS-BERT. Additionally, we also compared our model with BERT-sep and BERT-joint [27], which encode the text and topics separately and jointly only with BERT. The overall results of our model and baselines are shown in Table II. We observe that almost all of the external knowledge infusing methods outperform deep semantics mining methods except CKE-Net, which achieves a 1.2% lower F1-score than TSDA, which implies that infusing external knowledge is more effective in few-shot and zero-shot stance detection than mining deep semantics.

The performance of the proposed model on the *all* label is remarkable. Particularly, the macro F1-scores of three labels achieve 76.0%, which is 1.5% higher than the best baseline model. Compared with WS-BERT, TarBK-BERT, and CKE-Net, our model gets a 1.5%, 2.4%, and 5.9% higher F1-score on the *all* label, which indicates that the combination of sentiment and external knowledge can further improve the classification ability of the stance detection model. Compared with BS-RGCN, our model gets a 4.9% higher F1-score in the *zero-shot* scenario, which implies that only focusing on the high granularity of sentiment features will limit the transferability for the zero-shot task.

In addition, we also analyze the performance of the proposed model on three sub-labels. Our model achieves macro F1-scores of 95.1%, 93.6%, and 94.4% on the *neutral* label. The *neutral* instances are mostly unrelated to the topic, which indicates that our model has significantly improved the ability to distinguish whether the text is relevant to the given topic.

TABLE II: Performance comparisons on macro-average F1 scores of zero-shot and fewshot stance detection on VAST. The best performance is in bold font and the second-best results are underlined. All results are measured in percentages (%).The performance of CKE-Net†, WS-BERT† and TarBK-BERT† are obtained from the original papers, respectively.

Type	Method	Year	F1 Few-shot				F1 Zero-shot				F1 All			
			against	favor	neutral	all	against	favor	neutral	all	against	favor	neutral	all
Deep Semantics Mining Methods	BiCond	2016	46.3	45.4	25.9	39.2	47.5	45.9	34.9	42.7	46.8	45.7	30.6	41.0
	Cross-Net	2018	50.5	50.8	41.0	47.4	43.4	46.2	40.4	43.4	47.1	48.6	40.8	45.5
	SEKT	2020	47.9	51.0	21.5	47.4	44.2	50.4	30.8	41.8	46.2	50.7	26.3	41.1
	BERT-sep	2019	50.6	41.4	45.4	45.8	53.9	52.4	54.4	53.6	52.2	47.3	50.1	49.9
	BERT-joint	2019	58.4	54.6	85.3	66.0	59.7	54.3	79.6	64.6	59.1	54.5	82.3	65.3
	TGA-Net	2020	59.5	58.9	80.5	66.3	58.5	55.4	85.9	66.6	59.0	57.3	83.2	66.5
	TSDA	2023	65.2	60.1	88.6	71.3	63.8	63.6	86.0	71.2	64.5	62.0	87.3	71.3
External Knowledge Infusing Methods	CKE-Net†	2021	62.2	64.4	83.5	70.1	61.2	61.2	88.0	70.2	61.7	62.9	85.7	70.1
	BS-RGCN	2022	65.0	65.8	82.2	71.0	61.8	63.1	88.4	71.1	63.5	64.4	85.2	71.1
	TarBK-BERT†	2022	66.3	65.2	84.8	72.1	–	–	–	–	63.9	**65.7**	91.2	73.6
	WS-BERT†	2022	–	–	–	73.6	–	–	–	**75.3**	–	–	–	74.5
Ours	**(-)graph**	–	**67.4**	62.3	93.5	74.4	65.8	62.6	92.3	73.6	66.6	62.4	92.9	74.0
	(-)wiki	–	64.2	64.7	83.5	70.7	64.2	64.7	88.9	72.6	64.1	64.6	86.2	71.6
	(-)sentiment	–	61.8	63.3	94.6	73.2	58.6	62.0	93.5	71.4	60.3	62.6	94.1	72.4
	(-)summarization	–	63.4	65.9	91.7	73.7	**64.8**	62.3	92.3	73.2	64.2	64.1	92.0	73.4
	SentKB-BERT	–	66.7	**70.8**	**95.1**	**77.5**	63.1	**66.5**	**93.6**	74.4	**68.7**	65.0	**94.4**	**76.0**

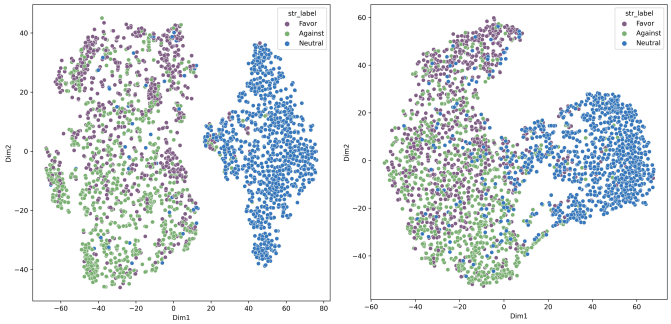
TABLE III: Accuracy on varying phenomena in VAST (%)

Method	Imp	mlT	mlS	Qte	Sarc
TGA-Net	59.4	60.5	53.2	66.1	63.7
CKE-Net	62.5	63.4	55.3	69.5	68.2
BS-RGCN	62.1	64.7	55.6	70.1	71.7
Zhang et. al.	61.0	65.9	55.8	67.7	70.8
WS-BERT	64.7	62.9	54.5	70.7	71.9
SentKB-BERT	**66.5**	**66.4**	**58.8**	**74.2**	**74.5**

Furthermore, due to the sentiment-related text filter, our model gets rid of those irrelavent contents with neutral sentiment and increases the F1-scores on the *against* label in three scenarios by 0.4%, 1.3%, and 4.8%.

We also visualize the results of the experiments to make the performance of the proposed model more intuitive. As shown in Fig. 3, we visualize the intermediate results before classification via t-SNE [28] and give a comparison between our model and pure BERT text encoding. The difference implies that our model performs better than BERT in dividing the data into different labels.

In order to make a full comparison with previous studies, we also examine our model on five challenging phenomena in VAST, as shown in Table III: (1) **Imp**, the text doesn't contain an explicit topic and the label is non-neutral; (2) **mlT**, there are other examples with the same text and different topics; (3) **mlS**, there are other examples with the same document and different, non-neutral, stance labels. (4) **Qte**, the text contains quotes and contains all three labels; (5) **Sarc**, the example contains sarcasm and contains all three labels. Our model outperforms all baseline models for all of the challenging phenomena. Imp, Qte, and Sarc examples require the model to recognize the topic-related concept in the text. The good performance on mlT and mlS examples proves that our model also fully learned the topic-to-topic and stance-to-stance patterns, which can enhance the model's ability to represent different aspects of the semantics of the same text.



(a) Our model (b) BERT

Fig. 3: Result of t-SNE visualization of intermediate output from our model and pure BERT.

D. Ablation Study

We conduct ablation studies on 4 variants of our model to evaluate the significance of the graph features, external knowledge, sentiment analyzer and summarizer modules, respectively. The modifications of these variants are shown below:

- **(-)graph**: removes the knowledge graph embeddings.
- **(-)wiki**: removes the external knowledge from Wikipedia.
- **(-)sentiment**: remove the whole sentiment encoding module.
- **(-)summarization**: remove the summarizer and keep the rest of the sentiment encoding module.

The results are also shown in Table II. First, the variant (-)graph lacks the relationship encoding between entities from the knowledge graph. It achieves macro F1-scores of 74.0% of all labels and it's 2.0% lower than the original model, which indicates that graph embeddings of common sense graphs can help improve the performance of few-shot and zero-shot stance detection tasks by introducing knowledge about relationships between entities in the text and topics.

TABLE IV: Cases selected for case study. Case 1, 2, 3 shows the effectiveness of sentiment analysis, external knowledge and text summarization, respectively. The **red bold text** indicates the expressions with non-neutral sentiment and the **black bold text** indicates the expressions with implicit background knowledge.

Raw text	Summarization	Topic	Ground truth	Prediction
I totally agree with this premise. As a younger person **I was against Nuclear power** (**I was in college during 3 mile island**) but now it seems that nuclear should be in the mix. **Fission technology is better, and will continue to get better** if we actively promote its development. The prospect of fusion energy also needs to be explored. If it's good enough for the sun and the stars, **it's good enough for me**.	Fission technology is better, and will continue to get better if we actively promote its development. Fusion technology also needs to be explored.	nuclear power	favor	favor
The 'professor' **sidesteps the reality** of what is going on in our nation, with over 30,000 gun deaths per year, many many multiples of times of gun related deaths per capita of that in other developed nations. Even **in Dodge City of days old, where supposedly the Wild West** reigned, it was illegal to be wearing a firearm within the city limits, **why was that?**	The 'professor' sidesteps the reality of what is going on in our nation, with over 30,000 gun deaths per year, many many times of that in other developed nations.	gun death	against	against
Nuclear energy only provides 8% of U.S. energy. We could easily conserve 8% and then begin the decommissioning of every nuclear power plant. Then we could take the BILLIONS and BILLIONS of dollars thrown at the nuclear industry and invest it in Renewable Energies like Wind and Solar Energy. Germany is HUGELY successful with Wind and Solar Energy. Germany has produced the energy-equivalent of 20 nuclear power plants (yes 20!!) with Solar Energy alone! They are having the same success with Wind Energy. This success in Renewable Energy proves how obsolete and unnecessary nuclear energy is.	Nuclear Energy only provides 8% of U.S. energy. We could conserve 8% and begin the decommissioning of every nuclear power plant .	nuclear power	against	against

Second, the variant (-)wiki lacks the textual background knowledge about the topics. It achieves F1-scores of 71.6% of all labels, which is far lower than the original model. The result implies the essential role of external knowledge in the stance detection model. It also indicates that external knowledge in textual form can infuse background knowledge about topics with little information loss and enhance the reasoning and generation abilities in few-shot and zero-shot scenarios significantly.

Third, the variants (-)sentiment and (-)summarization remove the sentiment analyzer and summarizer, respectively. Specifically, variant (-)sentiment removes both the summarizer and SentiBERT, meaning that this variant only uses external knowledge in stance detection. The two variants get 72.4% and 73.4% of macro F1-score, which is 3.6% and 2.6% lower than the original model. Precisely, it proves that i) the sentiment of the text makes a significant enhancement to the classification ability of stance detection models due to the close relationship between the sentiment and the text. ii) By filtering out irrelevant text via text summarization, the sentiment analyzer can properly represent deep semantics and help the model perform better in few-shot and zero-shot stance detection.

E. Case Study

We also select some cases from the test data using the model trained on all the topics. The cases chosen for our case study are shown in Table IV. We selected three different cases that get the right prediction while they are predicted incorrectly by the best baseline model. Each case shows the effectiveness of sentiment analysis, external knowledge, and summarization, respectively.

The first case expresses the user's sentiment towards the topic with text in red, which we can use to classify the user's stance manually. However, we noticed that the model without a sentiment analyzer, for example, WS-BERT, could not take this into consideration.

The second case has little expression directly related to the sentiment, which makes the sentiment analyzer of little help for stance classification. Our model outputs the correct label thanks to the external knowledge introduced by the wiki summaries and commonsense knowledge graph. Our model can use extra knowledge to recognize the relationships between the topic and entities like 'wild west', 'Dodge City', and 'firearm'.

The third case is selected from the hardest kind of stance classification task, which is predicted incorrectly by the best baseline model without text summarization. The text contains explicit expressions about the topic and the majority of the text does not introduce entirely new and unexplained entities, which makes external knowledge hard to use. At the same time, due to the length of the text, the sentiment analyzer only returns neutral results most of the time, which is caused by the sparsity of the sentimental expressions. However, these issues can be solved by text summarization in the proposed model. As shown in Table IV, the summarized text can be unmistakably identified as non-neutral, which will simplify the task for the model to categorize the non-neutral stance of such lengthy texts.

V. Conclusions and Further Research Lines

This work aims to address the issue of sentiment bias resulting from the limited amount of sentiment expression within a sentence. We have proposed a model that incorporates the advantages of the text summarization technique to exclude irrelevant expressions, thereby assisting the sentiment analyzer in accurately representing the sentiment. This work summarizes the input to eliminate useless content, aiming to generate summaries categorized according to the attitude of each sentence. The experimental result proved the significance of our proposed model and detailed how the summarization technique severed sentiment analysis and stance detection. In future work, we are willing to improve performance by depending on extracting emotion expressions and considering them as external knowledge.

Acknowledgment

This research work was partly sponsored by Natural Science Foundation of Xinjiang Uygur Autonomous Region(No.2022D01B185), the Fundamental Research Funds for the Central Universities (No.ZYGX2021J019).

References

[1] Datareporta, Meltwater, and W. A. Social. Digital 2023: Global overview report. [Online]. https://datareportal.com/reports/digital-2023-global-overview-report.

[2] Y. Zhou, A. I. Cristea, and L. Shi, "Connecting targets to tweets: Semantic attention-based model for target-specific stance detection," in *Web Information Systems Engineering (WISE)*, 2017, pp. 18–32.

[3] D. Küçük and F. Can, "Stance detection: A survey," *ACM Computing Surveys*, vol. 53, no. 1, pp. 1–37, 2020.

[4] S. M. Mohammad, P. Sobhani, and S. Kiritchenko, "Stance and sentiment in tweets," *ACM Transactions on Internet Technology (TOIT)*, vol. 17, no. 3, pp. 1–23, 2017.

[5] K. S. Hasan and V. Ng, "Why are you taking this stance? identifying and classifying reasons in ideological debates," in *EMNLP*, 2014, pp. 751–762.

[6] E. Allaway and K. Mckeown, "Zero-shot stance detection: A dataset and model using generalized topic representations," in *EMNLP*, 2020, pp. 8913–8931.

[7] S. Jayaram and E. Allaway, "Human rationales as attribution priors for explainable stance detection," in *EMNLP*, 2021, pp. 5540–5554.

[8] Y. Jiang, J. Gao, H. Shen, and X. Cheng, "Zero-shot stance detection via multi-perspective contrastive learning with unlabeled data," *Information Processing & Management*, vol. 60, no. 4, p. 103361, 2023.

[9] I. Augenstein, T. Rocktäschel, A. Vlachos, and K. Bontcheva, "Stance detection with bidirectional conditional encoding," in *EMNLP*, 2016, pp. 876–885.

[10] C. Xu, C. Paris, S. Nepal, and R. Sparks, "Cross-target stance classification with self-attention networks," in *ACL*, 2018, pp. 778–783.

[11] B. Zhang, M. Yang, X. Li, Y. Ye, X. Xu, and K. Dai, "Enhancing cross-target stance detection with transferable semantic-emotion knowledge," in *ACL*, 2020, pp. 3188–3197.

[12] J. Zhang, S. Wu, X. Zhang, and Z. Feng, "Task-specific data augmentation for zero-shot and few-shot stance detection," in *WWW*, 2023, p. 160–163.

[13] R. Liu, Z. Lin, Y. Tan, and W. Wang, "Enhancing zero-shot and few-shot stance detection with commonsense knowledge graph," in *ACL/IJCNLP*, 2021, pp. 3152–3157.

[14] Z. He, N. Mokhberian, and K. Lerman, "Infusing knowledge from wikipedia to enhance stance detection," *WASSA 2022*, p. 71, 2022.

[15] Q. Zhu, B. Liang, J. Sun, J. Du, L. Zhou, and R. Xu, "Enhancing zero-shot stance detection via targeted background knowledge," in *SIGIR*, 2022, pp. 2070–2075.

[16] Y. Luo, Z. Liu, Y. Shi, S. Z. Li, and Y. Zhang, "Exploiting sentiment and common sense for zero-shot stance detection," in *COLING*, 2022, pp. 7112–7123.

[17] M. Hardalov, A. Arora, P. Nakov, and I. Augenstein, "Few-shot cross-lingual stance detection with sentiment-based pre-training," in *AAAI*, vol. 36, no. 10, 2022, pp. 10 729–10 737.

[18] Y. Li and C. Caragea, "Multi-task stance detection with sentiment and stance lexicons," in *EMNLP-IJCNLP*, Nov. 2019, pp. 6299–6305.

[19] D. Loureiro, F. Barbieri, L. Neves, L. Espinosa Anke, and J. Camacho-collados, "TimeLMs: Diachronic language models from Twitter," in *ACL*, 2022, pp. 251–260.

[20] M. Taboada, "Sentiment analysis: An overview from linguistics," *Annual Review of Linguistics*, vol. 2, pp. 325–347, 2016.

[21] R. Nallapati, B. Zhou, C. dos Santos, Ç. Gulçehre, and B. Xiang, "Abstractive text summarization using sequence-to-sequence rnns and beyond," in *CoNLL*, 2016, pp. 280–290.

[22] R. Speer, J. Chin, and C. Havasi, "Conceptnet 5.5: An open multilingual graph of general knowledge," in *AAAI*, vol. 31, no. 1, 2017.

[23] J. Chen, H. Hou, J. Gao, Y. Ji, and T. Bai, "Rgcn: recurrent graph convolutional networks for target-dependent sentiment analysis," in *International Conference on Knowledge Science, Engineering and Management*. Springer, 2019, pp. 667–675.

[24] H. W. Chung, L. Hou, S. Longpre, B. Zoph, Y. Tay, W. Fedus, E. Li, X. Wang, M. Dehghani, S. Brahma *et al.*, "Scaling instruction-finetuned language models," *arXiv e-prints*, pp. arXiv–2210, 2022.

[25] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *CoRR*, vol. abs/1412.6980, 2014.

[26] Microsoft, "Neural Network Intelligence," 1 2021. [Online]. Available: https://github.com/microsoft/nni

[27] J. D. M.-W. C. Kenton and L. K. Toutanova, "Bert: Pre-training of deep bidirectional transformers for language understanding," in *NAACL-HLT*, vol. 1, 2019, p. 2.

[28] L. Van der Maaten and G. Hinton, "Visualizing data using t-sne." *Journal of machine learning research*, vol. 9, no. 11, 2008.