Prompt Learning for Low-Resource Multi-Domain Fake News Detection

Zitao Peng
Guangdong University of Foreign Studies
Guangzhou, China
pengzt318@163.com

Nankai Lin
Guangdong University of Technology
Guangzhou, China
neakail@outlook.com

Yongmei Zhou
Guangdong University of Foreign Studies
Guangzhou, China
yongmeizhou@163.com

Dong Zhou *
Guangdong University of Foreign Studies
Guangzhou, China
dongzhou@gdufs.edu.cn

Aimin Yang *
Lingnan Normal University
Zhanjiang, China
amyang@gdut.edu.cn

***Abstract*—The spread of fake news has caused severe damage to people's lives and society nowadays. Social media is inundated with fake news from multiple domains. Previously, the methods used to detect fake news have tended to be limited to single domains and have performed inadequacy in other or multiple domains. Therefore, the detection of multi-domain fake news has garnered significant interest. However, multi-domain fake news detection approaches rely more on sufficient training samples. In the real world, the low-resource problem has become a significant challenge that restricts the detection of fake news in multi-domain. In that case, prompt learning approaches have significant advantages in low-resource scenarios, but the existing fake news detection approaches based on prompt learning differ greatly in performance across different domains. In addition, the verbalizer in the prompt learning framework is the key module for mapping label words to classification labels, and the performance of existing methods is also limited by simply designed verbalizer modules, which makes the label words coverage small and the label words prediction inaccurate, especially in zero-shot scenarios. In this paper, we propose prompt learning for low-resource multi-domain fake news detection(PLDFEND). We incorporate domain-aware and relational learning into the prompt learning framework to improve the prompting effect. In addition, the verbalizer is optimized to adapt to different scenarios and map the label words to classification labels, thus achieving multi-domain news classification detection. After conducting comprehensive experiments in settings with limited resources and abundant data, we have confirmed the effectiveness of PLDFEND.**

***Index Terms*—multi-domain fake news detection, low-resource, prompt learning**

I. Introduction

Social media platforms have become one of the main channels for people to receive news, and fake news has propagated strongly on social media. People will undoubtedly be misled and public opinions manipulated, which is extremely detrimental to society. Therefore, automated detection of false news is both necessary and challenging.

Now that news in the real world covers various domains, misleading the facts can seriously cause national economic losses and the safety of human lives. Most existing methods focus on detecting fake news in a single domain [1], [2] and often perform poorly in other domains. Therefore, multi-domain fake news detection methods [3]–[5] have gradually gained attention.

However, all of these approaches using pre-trained language models require sufficient data samples to train the model and performance tends to be poor in low-resource scenarios [6]. In the real world, data distribution across domains is very unbalanced, with certain domains having very limited data resources [3]. The scarcity of data resources in certain domains may lead to biased models and unbalanced performance. These methods continue to perform poorly on domains with scarce training samples. During the early stage of fake news detection, there was very limited information available, and labeled data was far from enough [7], making it quite difficult to allow the models to learn sufficient features and patterns from the limited data. Its performance and generalization ability got restrained. Recently, the prompt learning paradigm has achieved good performance in low-resource tasks, and prompt learning for low-resource fake news detection tasks has also garnered attention. In [6], the authors propose a fake news detection method based on prompt learning for knowledge entities, but it has a high correlation with its focus domain, and the performance varies widely across domains. In this way, a low-resource multi-domain scenario is a difficult problem that exists for fake news detection at this stage. Another problem with existing instant prompt learning-based fake news detection approaches is that simply designed verbalizer have very limited performance and there is no filtering of the label words, resulting in low label words coverage and insufficiently accurate label words predictions, especially in zero-shot conditions [6]. The verbalizer module in prompt learning is the bridge between the predicted label terms and the classification label space, and the label words corresponding to the classification labels are in great numbers and have a certain degree of ambiguity. Therefore the construction of the verbalizer greatly affects the performance of prompt learning

Dong Zhou and Aimin Yang are the co-corresponding authors.

979-8-3503-3078-6/23/$31.00 ©2023 IEEE

[8].

To solve the problem, we propose prompt learning for low-resource multi-domain fake news detection (PLDFEND). Specifically, we build a template for the fake news detection task using domain-enhanced prompt learning by injecting domain information into customizable learnable tokens in the prompt template. In this way, the model can perceive information from different domains and learn the knowledge of the relationship between domains and texts, as well as the differences and relevance of fake news in different domains, thus achieving the purpose of enhancing the guidance of prompt learning. Then the optimized verbalizer can map the label words predicted by the model to the corresponding real and fake classification labels. Unlike existing methods, we optimize the verbalizer using contextual prior and learnable weights, which leads to more comprehensive coverage of label words and more accurate label word prediction. This model takes advantage of the good performance of prompt learning in low-resource scenarios and integrates domain information to focus on multiple domains and possess good generalization performance. It also optimizes the verbalizer to give full play to the performance of prompt learning, performing well in fake news detection across various domains.

We have conducted experimental evaluations under the Weibo21 dataset [3]. Our model is optimal in both zero-shot settings and few-shot settings and also has good performance in full-scale settings, which validates the effectiveness of our model on multi-domain fake news in scenarios with both low-resource and abundant data.

Our major contributions can be summarised as follows:
(1) In this paper, we propose a model to effectively solve the problem of multi-domain fake news detection in low-resource scenarios using domain information-enhanced prompt learning.
(2) We optimize the verbalizer module for prompt learning, which can be adapted to different low-resource scenarios and further improve the model performance.
(3) We conduct experiments on a dataset containing nine different domains, and the experimental results demonstrate the effectiveness and superiority of PLDFEND in scenarios with both low-resource and abundant data.

II. Related Work

Fake news seriously impacts people's lives, and the automatic detection of fake news has also received attention from researchers. Methods based on news content are mainly detected by extracting features of news content, and there are differences between real news and fake news in writing style, vocabulary and syntax [9]. It is essential to integrate additional information to improve the model's performance and reliability. For example, fake news content is often inflammatory, and in [10], the authors use emotional information to improve model performance. Combining external knowledge is also an effective means, and integrating external knowledge into news content can enhance the model's understanding of the semantic information present in the news. [11]. It also contains many objective facts, and some approaches have formed entity information of news content and external knowledge to detect fake news through entity comparison networks [1]. Social context-based methods concern the interaction between publishers and users (comments, retweets, etc.) and the propagation mode (Propagation threads and structures). More recently, hypergraph neural networks are being increasingly studied for their ability to capture group interactions among news clips, as news articles are typically shared among users who exhibit similar interests [2]. The authors in [12] utilize graph attention networks to integrate news spreading and user social information in order to detect fake news.

There are differences in news from different domains, such as writing style, vocabulary usage and distribution methods. The difference in data distribution across domains is called domain shift [13]. In [3], the authors present the first multi-domain fake news detection model. Research [5] explores the acquisition of interactive information through multi-view modeling and effectively reinforces domain-specific features in the context of multi-domain fake news detection using a domain memory bank.

In the real world, the distributions of data across different domains are also different, and some of the data in specific domains is also insufficient. Meanwhile, in the early stages of news propagation, there was limited information and a lack of labeled data [7]. Previous methods have struggled to perform well in these circumstances [6]. The method in this paper differs from the existing methods in that domain-enhanced prompt learning can still perform well in low-resource multi-domain conditions.

III. Methodology

In this section, we start with a problem statement, which is then divided into 4 modules to introduce the modeling approach. First, we introduce the prompt template construction module, using a prompt template to wrap the input news text. The domain-enhanced prompt representation module is then elucidated, which injects learnable domain tokens into the prompt representation. A module of optimized verbalizer construction, optimized with contextual prior distributions and learnable weights, is then defined. Finally, a fake news detection module is introduced to enable the authenticity detection of news texts.

A. Problem Statement

Given $x = (x_1, x_2, \ldots, x_n)$ is a news text containing n words, and let M be a pre-trained language model, fake news detection can be considered a binary classification task in which the goal is to classify each news text into unique labels $y \in \{0, 1\}$ by the model M, where 0 represents Real news and 1 represents Fake news. The general approach is to train the model on dataset Q of multi-domain news from social media. However, low-resource is the scenario in the real world and in the early stages of fake news propagation, and we obtain a small portion of data from the dataset Q, forming a few-shot

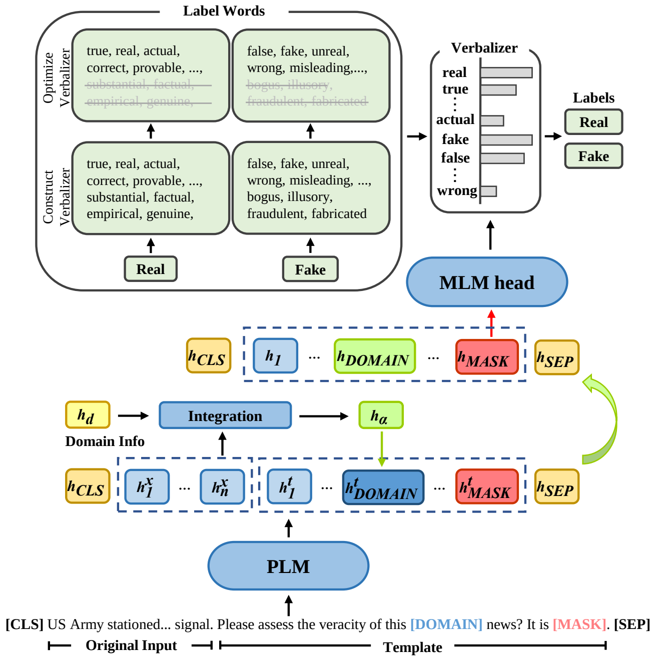

Fig. 1. The illustration of our domain information into prompt learning and optimizing the verbalizer. The above part is the general framework of PLDFEND.

dataset of Q_{few} for model training, to explore an effective method for low-resource multi-domain fake news detection.

B. Prompt Template Construction

The prompt learning paradigms formalize the fake news detection task as a masked language modeling task. Specifically, prompt learning wraps the news text input with a task-relevant template and masks the keywords in the template. For example, we classify the news text x =“US Army stationed... signal.” into real or fake labels and use a prompt template (e.g“Please assess the veracity of this news? It is [MASK].”) to wrap this text to get the prompt text x_p:

$$x_p = [CLS]\ x.\ Please\ assess\ the\ veracity\ of\ this\ news?\ It\ is\ [MASK].\ [SEP] \quad (1)$$

The model M is then allowed to predict the probability of each word v in the [MASK] token $P_M([MASK] = v|x_p)$, these words v constitute the set of label words V corresponding to the classification labels. Then, to map the probabilities of these label words to the labels' probabilities by the verbalizer we have designed to obtain the predicted classification results.

C. Domain-enhanced Prompt Representation

Several studies have shown that prompt templates with special and learnable tokens can improve prompt learning [14], [15]. So, in this paper, we add learnable tokens to prompt templates to improve prompt learning guidance. Specifically, as shown in Fig. 1, we add a special learnable token [DOMAIN] to the prompt template:

$$x_p = [CLS]\ x.\ Please\ assess\ the\ veracity\ of\ this\ [DOMAIN]\ news?\ It\ is\ [MASK].\ [SEP] \quad (2)$$

The embedding of this token is initialized and updated during training, and we use this token to enable the model to learn domain information.

Since domain labels have proven to be highly beneficial for multi-domain learning [5], [13], we integrated the domain information provided by domain labels into the prompt templates to improve prompt guidance.

We encode the original input news text x without template wrapping and the corresponding domain information through the model M to obtain the hidden original input news text word embedding vector $H^x = (h^x_{[CLS]}, h^x_1, ..., h^x_n, h^x_{[SEP]})$ and the hidden domain-embedding vector h_d, where $h^x_{[CLS]}$ integrates the semantic information of the whole text. Then We utilize the attention mechanism to integrate h_d and $h^x_{[CLS]}$ to obtain a learnable vector h_α with domain and text knowledge:

$$h_\alpha = softmax(\frac{h^x_{[CLS]} h^T_d}{\sqrt{d_k}}) h_d \quad (3)$$

where d_k is the dimension of the vectors $h^x_{[CLS]}$ and h_d.

We inject the integrated learnable embedding vector h_α into the [DOMAIN] position in the prompt template:

$$h_{DOMAIN} = h^t_{DOMAIN} + h_\alpha \quad (4)$$

Here h^t_{DOMAIN} denotes the hidden vector of [DOMAIN] token in the prompt template, and h_{DOMAIN} is the updated hidden vector of [DOMAIN] after injecting the learnable hidden vector h_α.

During the training process, the hidden vector $H = (h_{[CLS]}, h_1, ..., h_{[DOMAIN]}, ..., h_{[MASK]}, ..., h_{[SEP]})$ of the prompt text x_p is also updated with the domain information, and the updated [MASK] representation is used to predict the label words.

D. Optimized Verbalizer Construction

In this paper, we define a verbalizer, a critical component of prompt learning [16], to map the predicted set of label words V to the classification label space $\{0, 1\}$. We define the subset of V mapped to a particular label y as V_y. A good verbalizer has a significant impact on prompt learning performance. Most of the previous studies have used manual subjective vocabulary setting [17], which may have some bias in vocabulary coverage and some subjective preconceptions. This paper expanded the label words by collecting enough words that are semantically most similar to the classification labels and filtering the most effective words to form the verbalizer. Given CHATGPT's vast database and extensive knowledge coverage, it is highly adept at gathering diverse lexicons and information. We use it to collect 50 “Real” label words and 50 “Fake” label words to express the meaning and semantic features, as shown in Table I.

TABLE I
EXAMPLES OF THE LABEL WORDS.

Label	Label Words
Real	true, empirical, real, correct, actual, authentic, factual, fact, substantial, genuine, truth, veracity...
Fake	false, artificial, fake, fault, misleading, incorrect, bogus, wrong, virtual, unreal, illusory, disguised...

In the zero-shot scenario, excessive words may increase the ambiguity and vagueness of the model for words, and some words are rarely seen and unrepresentative. It will mislead the model prediction and make the model inaccurate in predicting probabilities, so it is necessary to filter and denoise these words. We use the contextual prior of the label words to solve this problem. Specifically, we denote the distribution of text x in the sample set as D, and each sentence in the distribution is wrapped as a prompt text x_p, and through the probability expectation of the whole sentence distribution, we obtain the prior distribution of the label words of the masked position as:

$$P_D(v) = E_{x \sim D} P_M([MASK] = v|x_p) \tag{5}$$

To estimate the expectation, we sample a small unlabeled support set C from the training set and assuming that the input samples $x \in C$ follow a uniform prior distribution, the context prior distribution of each label word is approximated as:

$$P_D(v) \approx \frac{1}{|C|} \sum_{x \in C} P_M([MASK] = v|x_p) \tag{6}$$

Then, we rank the prior probabilities and select the label words within the threshold.

In few-shot scenarios, training can be done with a small amount of annotated data. For each label word, we assign a learnable weight ω_v to minimize the effect of noise on the prediction. These weights are represented as vectors $\omega \in R^{|V|}$ while initialized to zero vectors and learned to weights in training. During the learning process, the weights of V_y within each label set are normalized, which is computed as follows:

$$\varepsilon_v = \frac{exp(\omega_v)}{\sum_{u \in V_y} exp(\omega_u)} \tag{7}$$

Where $\sum_{u \in V_y} exp(\omega_u)$ denotes the sum of the exponential terms of the weights of all label words in the label word set V_y.

E. Fake News Detection

The prompt text x_p encoded with domain information is fused to obtain the hidden vector H. The probability of each label word $v \in V$ of [MASK] is obtained after the Masked Language Modeling head prediction output:

$$P(v|x_p) = P_M([MASK] = v|x_p) \tag{8}$$

Then, through the verbalizer that we've defined, the predicted probability $P(v|x_p)$ for the label y can be calculated as:

$$P(y|x_p) = g(P(v|x_p)|v \in V_y) \tag{9}$$

where g is a function that converts the probability of a label word into the probability of the corresponding label.

In the zero-shot setting, no training of the parameters is required, and assuming that each label word has the same contribution to the class label y, the predicted probabilities of the label words are averaged so that the predicted probability of the label y is obtained as:

$$P(y|x_p) = softmax_{y \in \{0,1\}} \frac{\sum_{v \in Y_y} P_M([MASK] = v|x_p)}{|V_y|} \tag{10}$$

where V_y is a subset of V mapped to a specific label y.

In the few-shot setting, since each label word is assigned a learnable weight, we weight and average the predicted probabilities of the label words to obtain the predicted probability of label y:

$$P(y|x_p) = softmax_{y \in \{0,1\}} \sum_{v \in V_y} \varepsilon_v P_M([MASK] = v|x_p) \tag{11}$$

We employ the binary cross-entropy loss function as the training objective function for the binary classification task of fake news detection:

$$L = -\sum_{i=1}^{N} y_i log P(\hat{y}_i|x_p) + (1 - y_i) log(1 - P(\hat{y}_i|x_p)) \tag{12}$$

where $\sum_{i=1}^{N}$ represents the summation of all training instances, y_i denotes the ground truth labels, and $\hat{y}_i$ is used to denote the predicted labels.

IV. EXPERIMENT

In this section, we perform experiments to evaluate the effectiveness of our method. We compare with other baselines on the dataset and analyze the experimental results of our method.

A. Dateset

We've conducted our evaluation on the Weibo21 dataset [5]. This dataset is a Chinese multi-domain fake news detection dataset that was gathered from Weibo. As shown in Table II, it contains news in nine domains: politics, society, health, finance, science, military, disaster, entertainment and education, with a total of 4,640 real news and 4,488 fake news.

B. Experiment Settings

We utilize $BERT_{base}$[1] [18] as a PLM with a fixed embedding vector dimension of 768. The maximum sentence length is 256. We used the RTX8000 as the GPU for our experiments. To simulate real-world scenarios with low resources and

[1]https://huggingface.co/bert-base-chinese

TABLE II
DATA STATISTICS OF WEIBO21.

Domain	Sci.	Mil.	Edu.	Dis.	Pol.	Hlth.	Fin.	Ent.	Soc.	All
Real	143	121	243	185	306	485	959	1000	1198	4640
Fake	93	222	248	591	546	515	362	440	1471	4488
All	236	343	491	776	852	1000	1321	1440	2669	9128

multiple domains of news, we've conducted experiments with zero-shot settings, few-shot settings, and full-scale settings for different domains and full domains, respectively. For the zero-shot setting experiments, we repeated multiple runs using 10 random seeds and calculated the average score as the final result after excluding the maximum and minimum values. For the few-shot setting experiments, we conducted experiments using 1, 2, and 4 shots. For each experiment, we selected k (1, 2, or 4) samples from each of the 9 domain categories in the original training set as the training set. We also used k samples from each category as the validation set. We repeated the experiments using 10 random seeds and calculated the average score by removing the highest and lowest values. We tuned the whole model for 10 epochs with a learning rate of 6e-5. For the optimized verbalizer, the unlabeled support set C was size 200. Finally, we performed ablation experiments, comparing the different configurations of our model separately. As for all the experiments, the F1 scores were used as an evaluation metric to assess the classification effectiveness.

C. Baselines

In this section, we present the baseline models we compare, starting with the traditional machine learning approaches, SVM [19]. We also compare the non-PLM deep learning approaches, TextCNN [20], BiGRU [21]. Then we compare the PLM-based deep learning approaches, FT [22], MDFEND [3], FuDFEND [4], M^{3}FEND [5]. We also compare a conventional prompt learning approach PT [17].

D. Main Results

In this section, the results of specific experiments are introduced, and from these, we can analyze the observations obtained for PLDFEND.

1) Few-shot: As shown in Table III, we conducted experiments with a few-shot setting. First, we observe that the conventional prompt learning approach PT outperforms the other PLM-based approaches in most cases, and the results are much better than those of the standard fine-tuned PLM approach FT. It proves that prompt learning can fully use the power of PLM to enable the model to possess better characterization and classification performance in low-resource conditions. It can also be shown that the PLM-based approaches, such as MDFEND, using domain information are more effective than the standard fine-tuned PLM approach FT in each domain, thus indicating the effectiveness of domain information in multi-domain fake news detection. Our method PLDFEND outperforms PT and other methods in various domains and in all domain, with relatively minor differences across domains. It utilizes domain-enhanced prompt learning and an optimized verbalizer, incorporating domain information and exhibiting good adaptability and generalization performance in different domains, even with limited training data. The above results also illustrate the superiority of our method in low-resource multi-domain scenarios.

TABLE III
RESULTS OF 1/2/4-SHOT MULTI-DOMAIN FAKE NEWS DETECTION.

Shot	Model	Sci.	Mil.	Edu.	Dis.	Pol.	Hlth.	Fin.	Ent.	Soc.	All
1	SVM	41.87	43.69	37.41	48.75	43.89	34.85	30.56	29.17	34.18	39.43
	TextCNN	31.56	34.36	29.32	31.47	32.29	22.43	22.57	20.85	24.06	27.71
	BiGRU	30.84	33.72	27.76	31.09	30.86	21.71	20.48	20.23	22.97	26.12
	FT	57.81	57.34	50.73	55.68	55.31	47.16	46.03	43.92	50.64	53.13
	MDFEND	61.28	59.42	52.53	58.34	56.14	49.56	49.19	44.75	52.63	55.48
	FuDFEND	58.33	57.79	51.08	57.37	59.79	47.75	46.41	44.56	53.26	54.35
	M^{3}FEND	59.38	60.51	52.64	56.65	56.36	51.38	48.68	45.04	53.12	55.79
	PT	62.09	62.23	57.63	59.85	61.35	53.19	50.73	46.19	56.89	56.83
	PLDFEND	**64.58**	**66.67**	**61.39**	**63.43**	**66.08**	**58.75**	**53.45**	**49.82**	**61.80**	**60.47**
2	SVM	53.62	57.80	51.18	52.89	56.93	48.16	47.08	44.25	51.97	53.28
	TextCNN	43.55	47.72	42.52	45.39	44.30	37.47	34.71	33.56	38.49	40.86
	BiGRU	44.31	47.38	41.16	43.64	42.66	37.16	32.27	33.84	37.57	39.62
	FT	59.49	62.74	58.65	60.73	61.56	57.62	54.27	52.69	58.93	59.48
	MDFEND	64.59	64.67	59.93	63.86	62.57	58.30	56.65	53.21	59.79	61.54
	FuDFEND	60.42	63.18	59.27	61.28	64.16	58.28	54.49	53.37	61.02	60.32
	M^{3}FEND	63.83	64.72	59.48	63.46	62.35	59.17	55.98	54.14	61.86	61.79
	PT	65.83	66.21	61.79	65.15	65.33	60.62	57.25	55.49	63.36	63.15
	PLDFEND	**68.75**	**68.83**	**64.63**	**67.95**	**69.74**	**64.51**	**59.47**	**57.70**	**66.48**	**65.72**
4	SVM	59.14	63.22	58.76	61.57	62.52	58.39	54.75	53.12	59.64	59.35
	TextCNN	58.63	64.57	59.64	62.54	62.29	60.27	53.71	54.48	60.13	60.78
	BiGRU	59.26	64.23	58.86	60.78	61.53	60.83	51.97	56.05	59.57	59.67
	FT	64.26	68.85	62.94	68.92	67.86	64.87	59.63	62.23	65.68	67.92
	MDFEND	67.71	70.49	65.16	70.36	71.05	66.31	61.65	63.49	66.37	69.53
	FuDFEND	64.89	69.70	64.53	69.57	69.59	65.47	60.32	62.59	68.72	68.78
	M^{3}FEND	68.25	70.62	65.79	69.39	70.83	67.35	62.58	64.12	67.23	69.86
	PT	69.05	71.13	66.41	71.39	72.34	67.83	63.19	64.74	70.25	70.32
	PLDFEND	**72.83**	**72.79**	**68.72**	**72.65**	**74.85**	**69.10**	**64.75**	**66.49**	**71.48**	**71.85**

2) Full-scale: As shown in Table IV, in the full-scale experiments, when there is an adequate amount of training data available, PLM already has a strong semantic understanding and feature extraction capabilities. We observe that the PLDFEND enhances prompting guidance by integrating domain information, and its performance turns out to be optimal in all domain as well as in most domains, especially in data-poor domains such as science, military, and education. Simultaneously, There is minimal variation in performance across and over all domains, indicating a strong generalization ability. We attribute this to our effective utilization of domain information, which enhances prompt learning and allows us to fully exploit the power of PLM.

3) Zero-shot: As shown in Fig. 2, we conducted experiments with zero-shot settings. We can see that all configuration cases of PLDFEND are superior to the PT baseline, illustrating the effectiveness of our approach. We also compared our approach to removing different modules -DM, -TP, -FO. For -DM, we remove the learnable [DOMAIN] tokens from the prompt templates, and there is no injection of domain information to enhance the guidance of the prompt templates; for -TP, we remove the manually designed discrete templates;

TABLE IV
RESULTS OF FULL-SCALE MULTI-DOMAIN FAKE NEWS DETECTION.

Model	Sci.	Mil.	Edu.	Dis.	Pol.	Hlth.	Fin.	Ent.	Soc.	All
SVM	69.43	84.51	79.59	76.25	81.73	84.38	78.64	79.17	81.93	80.79
TextCNN	72.54	88.39	83.62	82.22	85.61	87.68	86.38	84.56	85.40	86.86
BiGRU	72.69	87.24	81.38	79.35	83.56	88.68	82.91	86.29	84.85	85.95
FT	77.77	90.72	83.31	85.12	83.66	90.90	87.35	87.69	85.77	87.95
MDFEND	83.01	93.89	89.17	90.03	88.65	94.00	89.51	90.66	89.80	91.37
FuDFEND	81.33	94.68	89.17	90.59	**90.13**	94.17	89.01	91.61	**91.74**	92.13
M^{3}FEND	82.92	95.06	89.98	88.96	88.25	**94.60**	**90.09**	93.15	90.89	92.16
PT	87.23	91.46	86.57	88.74	87.82	91.53	87.12	89.97	88.05	89.49
PLDFEND	**91.22**	**95.75**	**90.39**	**91.67**	89.06	93.81	88.68	**93.40**	90.94	**92.45**

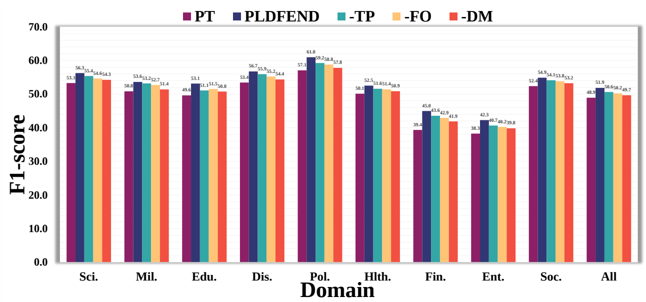

Fig. 2. Results of zero-shot multi-domain fake news detection.

for -FO, we do not perform contextual before optimizing the verbalizer. With the most significant performance degradation and the largest performance difference between domains for -DM, we speculate that without training and the injection of domain information, the model cannot understand and process texts from different domains accurately and it is unable to capture the features between the domains and the texts, which has the most significant effect on the performance of the model and reduces its ability to generalize. The above also illustrates the effectiveness of domain information enhanced prompt guidance and contextual prior optimized verbalizer.

V. CONCLUSION

In this paper, we proposed the PLDFEND method for multi-domain fake news detection in low-resource scenarios. This method is a prompt learning-based approach that uses domain information to improve the guidance of prompt templates and optimizes the verbalizer to improve model performance further. Experiments demonstrate the effectiveness of PLDFEND in scenarios with both low-resource and abundant data. In the future, we will further explore better methods for template construction. In addition, the method of incorporating domain information into prompt learning has broad applicability, and applying it to other tasks is also a future direction worth exploring.

ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China (No. 62376062), the Guangdong Basic and Applied Basic Research Foundation of China (No. 2023A1515012718), and the Philosophy and Social Sciences 14th Five-Year Plan Project of Guangdong Province (No. GD23CTS03).

REFERENCES

[1] W. Xu, J. Wu, Q. Liu, S. Wu, and L. Wang, "Evidence-aware fake news detection with graph neural networks," in *Proceedings of the ACM Web Conference 2022*, 2022, pp. 2501–2510.
[2] U. Jeong, K. Ding, L. Cheng, R. Guo, K. Shu, and H. Liu, "Nothing stands alone: Relational fake news detection with hypergraph neural networks," in *2022 IEEE International Conference on Big Data (Big Data)*. IEEE, 2022, pp. 596–605.
[3] Q. Nan, J. Cao, Y. Zhu, Y. Wang, and J. Li, "Mdfend: Multi-domain fake news detection," in *Proceedings of the 30th ACM International Conference on Information & Knowledge Management*, 2021, pp. 3343–3347.
[4] C. Liang, Y. Zhang, X. Li, J. Zhang, and Y. Yu, "Fudfend: fuzzy-domain for multi-domain fake news detection," in *CCF International Conference on Natural Language Processing and Chinese Computing*. Springer, 2022, pp. 45–57.
[5] Y. Zhu, Q. Sheng, J. Cao, Q. Nan, K. Shu, M. Wu, J. Wang, and F. Zhuang, "Memory-guided multi-view multi-domain fake news detection," *IEEE Transactions on Knowledge and Data Engineering*, 2022.
[6] G. Jiang, S. Liu, Y. Zhao, Y. Sun, and M. Zhang, "Fake news detection via knowledgeable prompt learning," *Information Processing & Management*, vol. 59, no. 5, p. 103029, 2022.
[7] Y. Li, K. Lee, N. Kordzadeh, B. Faber, C. Fiddes, E. Chen, and K. Shu, "Multi-source domain adaptation with weak supervision for early fake news detection," in *2021 IEEE International Conference on Big Data (Big Data)*. IEEE, 2021, pp. 668–676.
[8] T. Gao, A. Fisch, and D. Chen, "Making pre-trained language models better few-shot learners," *arXiv preprint arXiv:2012.15723*, 2020.
[9] A. Mosallanezhad, M. Karami, K. Shu, M. V. Mancenido, and H. Liu, "Domain adaptive fake news detection via reinforcement learning," in *Proceedings of the ACM Web Conference 2022*, 2022, pp. 3632–3640.
[10] X. Zhang, J. Cao, X. Li, Q. Sheng, L. Zhong, and K. Shu, "Mining dual emotion for fake news detection," in *Proceedings of the web conference 2021*, 2021, pp. 3465–3476.
[11] Y. Wang, S. Qian, J. Hu, Q. Fang, and C. Xu, "Fake news detection via knowledge-driven multimodal graph convolutional networks," in *Proceedings of the 2020 international conference on multimedia retrieval*, 2020, pp. 540–547.
[12] E. Min, Y. Rong, Y. Bian, T. Xu, P. Zhao, J. Huang, and S. Ananiadou, "Divide-and-conquer: Post-user interaction network for fake news detection on social media," in *Proceedings of the ACM Web Conference 2022*, 2022, pp. 1148–1158.
[13] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Transactions on knowledge and data engineering*, vol. 22, no. 10, pp. 1345–1359, 2009.
[14] K. Hambardzumyan, H. Khachatrian, and J. May, "Warp: Word-level adversarial reprogramming," *arXiv preprint arXiv:2101.00121*, 2021.
[15] X. Han, W. Zhao, N. Ding, Z. Liu, and M. Sun, "Ptr: Prompt tuning with rules for text classification," *AI Open*, vol. 3, pp. 182–192, 2022.
[16] S. Hu, N. Ding, H. Wang, Z. Liu, J. Wang, J. Li, W. Wu, and M. Sun, "Knowledgeable prompt-tuning: Incorporating knowledge into prompt verbalizer for text classification," *arXiv preprint arXiv:2108.02035*, 2021.
[17] T. Schick and H. Schütze, "Exploiting cloze questions for few shot text classification and natural language inference," *arXiv preprint arXiv:2001.07676*, 2020.
[18] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "Bert: Pre-training of deep bidirectional transformers for language understanding," *arXiv preprint arXiv:1810.04805*, 2018.
[19] F. Yang, Y. Liu, X. Yu, and M. Yang, "Automatic detection of rumor on sina weibo," in *Proceedings of the ACM SIGKDD workshop on mining data semantics*, 2012, pp. 1–7.
[20] Y. Kim, "Convolutional neural networks for sentence classification," *arXiv preprint arXiv:1408.5882*, 2014.
[21] J. Ma, W. Gao, P. Mitra, S. Kwon, B. J. Jansen, K.-F. Wong, and M. Cha, "Detecting rumors from microblogs with recurrent neural networks," 2016.
[22] Y. Liu, M. Ott, N. Goyal, J. Du, M. Joshi, D. Chen, O. Levy, M. Lewis, L. Zettlemoyer, and V. Stoyanov, "Roberta: A robustly optimized bert pretraining approach," *arXiv preprint arXiv:1907.11692*, 2019.