A Two-Stage Prompt Learning Method for Jointly Predicting Topic and Personality

Yilin Wu[12], Minjie Yuan[12], Yuxuan Song[12], Liping Chen[2], Chenyu Yuan[3], and Qiudan Li[2*]
[1] The School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing, China
[2] The State Key Laboratory of Multimodal Artificial Intelligence Systems, Institute of Automation,
Chinese Academy of Sciences, Beijing, China
[3] The School of Electronic and Information Engineering, Lanzhou Jiaotong University, Lanzhou, China
{wuyilin2022, yuanminjie2021, songyuxuan2023, liping.chen, qiudan.li}@ia.ac.cn, {22202209}@stu.lzjtu.edu.cn

2023 IEEE International Conference on Intelligence and Security Informatics (ISI) | 979-8-3503-3773-0/23/$31.00 ©2023 IEEE | DOI: 10.1109/ISI58743.2023.10297139

***Abstract*—Accurate personality prediction can help management departments analyze users' behaviors and make informed decisions effectively. Existing text-based personality prediction studies mainly rely on deep neural networks or pre-trained language models to extract semantic information and personality traits. However, the text's topic and label description may provide additional personality clues. This paper proposes a topic and personality prediction method based on a large language model (LLM), which utilizes a two-stage prompt strategy to mine the interaction between the topic and personality information. Additionally, labels' descriptions are incorporated to construct cue-based prompts, and a fine-tuning approach is adopted to optimize the model's performance. Experiments on two datasets show the efficacy of the proposed model.**

***Index Terms*—Personality, Topic, LLM, Prompt tuning**

I. Introduction

With the widespread usage and advancement of social platforms and self-media, users always express their opinions, attitudes, and stances on various products online. Mining users' personalities under specific topics or entities can provide clues for analyzing their interests and behavioral intentions [1].

Existing text-based personality prediction studies mainly rely on deep neural networks or pre-trained language models to extract semantic information and personality traits [2]–[4]. However, the performance of these methods can be further improved by considering other personality cues, such as the topic information and label description. The emerging abilities of LLM, such as multitask prediction and in-context learning, offers an opportunity to identify the topic and personality jointly [5]. For instance, when users evaluate the detail of a movie, their comments often like: "OMG, the starring is too cuteeeee, their acting is on fleek<3<3<3! The BGM? Totally perfection, the vibes are very good." These comments concentrate on specific details of the movie in which the language is lively and relaxed, exhibiting diverse expression styles. By deeply exploring the dynamic interaction among the topic, expression style, and personality, the model could gain a further understanding of the 'Openness' personality traits. This provides us with a perspective to model and analyze user personalities from relevant topics.

To mine the interaction between personality and topic information effectively while taking advantage of the learning and inference abilities of LLM, this paper proposes an LLM-based two-stage prompt learning mechanism for topic-personality prediction. The primary contributions of this paper are as follows:

- This paper proposes a research question of predicting topic and personality jointly.
- The proposed method mines the interaction between topic and personality by employing a topic-driven personality prediction model which combines LLM, prompt learning, and domain fine-tuning.
- We construct two Chinese datasets and conduct experiments to demonstrate the method's effectiveness.

*Corresponding author

II. Proposed Method

The overall framework of our method is shown in Fig. 1. The model consists of two modules, which are task-related prompt construction, topic and personality joint prediction based on multitask fine-tuning. The task-related prompt construction module introduces label description information into the two-stage prompt, which fuses the predicted topic information during the personality prediction stage. Then, the topic and personality joint prediction module freezes the pre-trained LLM parameters and adds trainable soft prompts before the task-related prompts, which could be optimized to predict the topic and personality accurately.

A. Task-related Prompt Construction

The user's cognitive load can be reduced to facilitate decision-making by introducing labels' descriptions [6]. Thus, we designed a two-stage prompt strategy to mine the interaction between topic and personality, and introduced corresponding labels' descriptions at different prediction stages to enhance the model's understanding of the labels, which could improve each stage's prediction accuracy.

Firstly, in the topic prediction stage, descriptions of the topic labels are incorporated into the following prompt: " According to the topic-label descriptions: [Descriptions]. Determine which topic the following comment belongs to: [Text]." Secondly, in the personality prediction stage, we use the predicted topic and its corresponding description to construct the interaction between the topic and personality information. Simultaneously, the descriptions of the personality labels are incorporated into the following prompt: "Considering the topic

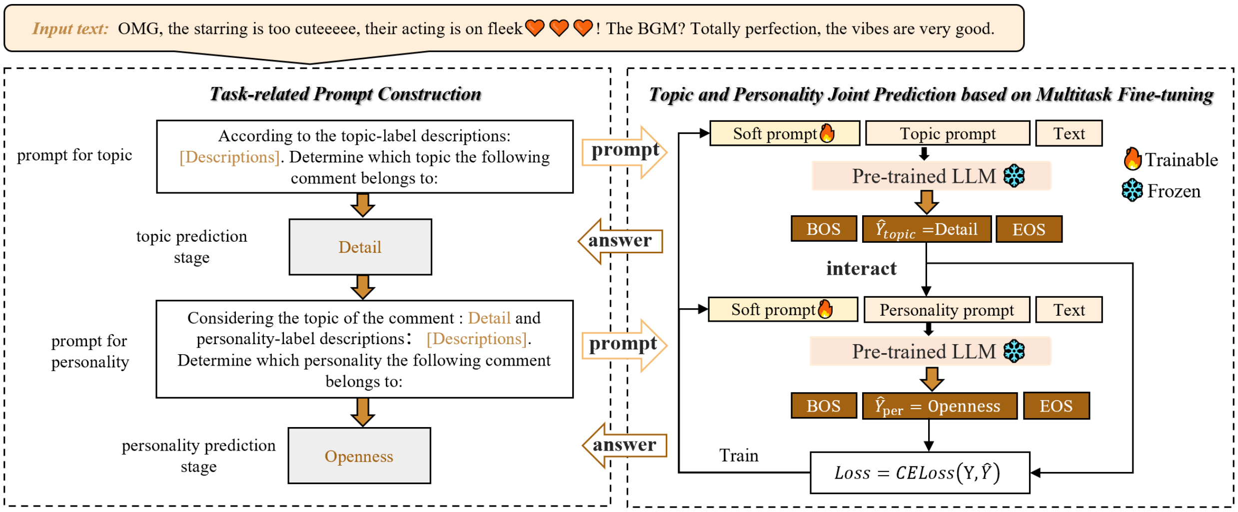

Fig. 1: The overall framework of the proposed method.

of the comment: [Topic] and personality-label descriptions: [Descriptions]. Determine which personality the following comment belongs to: [Text]." These descriptions of the topic and personality labels are shown in Table I and Table II, respectively.

B. Topic and Personality Joint Prediction based on Multitask Fine-tuning

To fine-tune the model on multi-tasks while leveraging the in-context learning capabilities of LLM, our method introduces the trainable soft prompt and uses the ptuning-v2 [7] to train the topic and personality prediction model jointly. First, the input X is consisted of soft prompts, task-related prompt, and text. The $soft\ prompts$ are added in front of the input sequence in the transformer structure of each layer, which will participate in the calculation of the self-attention mechanism for generating topic or personality label. Then, the LLM autoregressively generates predicted topic or personality labels $\hat{Y}$. Finally, the cross-entropy loss between each generated label $\hat{Y}_i$ and the true label Y_i is calculated to optimize the LLM. During the multitask fine-tuning process, only the attention weight corresponding to $soft\ prompts$ will be optimized, which are shared across the two stages. This reduction in trainable parameters improves the efficiency of fine-tuning and ensures the accuracy of topic and personality mining.

$$X = [Soft\ prompt, Task\ prompt, Text] \tag{1}$$

$$Loss\left(Y, \hat{Y}, \theta\right) = \sum_{i=1}^{t} CELoss\left(Y_i, \hat{Y_i}\right) \tag{2}$$

Where t is the number of tokens of the true label.

III. Experiment

A. Datasets

The movies named "Myth of Love" and "Be Somebody" were used to construct two Chinese datasets—AQSH and YMLW. The raw data was collected between November 2021 to October 2022 from platforms such as Douban Movie and Weibo. By preprocessing and removing the repeated and non-Chinese comments, AQSH with 2091 comments and YMLW with 2101 comments are finally obtained. According to the characteristics of the data, we summarized three types of topics, as shown in Table I, and four types of personalities, as shown in Table II. Two students familiar with the personality domain were invited to annotate the data. The Fleiss' kappa coefficient of the annotation result on the AQSH and YMLW are 0.71 and 0.73, respectively.

TABLE I: The Description Of Topic

Topic	Description
Detail	Evaluation of specific details of a movie such as content, structure, plot, actors, director, roles, lines, soundtrack, lighting, etc.
Overall	Evaluation of the overall aspects of movie production level, genre, theme, target audience, box office, etc.
Expectation	Expectations for the character and the future development of a movie.

TABLE II: The Description Of Personality

Personality	Description
Openness	Users' feeling of watching a movie from a novel, imaginative, and aesthetic perspective through various writing styles, such as emoji, punctuation, word repetition, etc.
Conscientiousness	Users' opinions and suggestions on the further enrichment and development of a movie, such as the actor and script.
Agreeableness	Encouragement and recommendation for a movie by using positive, optimistic, and comfortable words.
Neuroticism	Users' dissatisfaction with a movie through intense language or unreasonable negative emotions, such as anger and anxiety.

B. Baseline and Settings

To measure the performance of different models, F1 scores (represented by %) are used as evaluation metrics, and the baseline methods are as follows:

- BERT [8]: This method uses pre-trained bert-base-chinese model, which encodes each text into a vector of dimension 768. Two fully connected neural networks are used to reduce the dimensionality and extract topic and personality features from the vector, respectively.
- no prompt: This method employs ChatGLM-6B and ptuning-v2 to fine-tune directly without adding additional prompts.
- one-stage prompt: A variant of the two-stage prompt incorporates label description information through a template: "According to the topic label descriptions: [Description], and the personality label descriptions: [Description], determine which type of topic the following comments belong to, and determine which type of personality trait they exhibit: [Text].", which predicts the topic and personality simultaneously.

In the experiment, the datasets are divided into training and testing sets in a 7:3 ratio. The GLM-6B [9] is chosen as LLM. The trainable soft prompt length is set to 128, the max input length is set to 768, and the max output length is set to 32. In generating, the BOS marks the beginning, and the EOS marks the end of the generated output. The training is performed in a total of 3000 steps, and the batch size is 16. We optimized the model using the AdamW optimizer with an initial learning rate of 1e-2.

C. Experimental Results and Analysis

The overall experimental results are shown in Table III, where the best performing results are highlighted in bold. The two-stage prompt method achieves the best personality prediction results on the AQSH and YMLW datasets. Compared with the BERT baseline, the F1 scores improve from 63.80 to 65.87 and from 67.96 to 69.25, respectively.

TABLE III: OVERALL RESULT

Method	AQSH		YMLW	
	Topic	Personality	Topic	Personality
BERT	**71.61**	63.80	76.06	67.96
no prompt	64.77	62.64	73.70	66.90
one-stage prompt	70.88	63.30	77.02	65.96
two-stage prompt	69.90	**65.87**	**78.38**	**69.25**

On both datasets, BERT outperforms the no prompt method for topic and personality prediction, which shows that relying on trainable soft prompts alone is difficult to model multi-tasks simultaneously. Compared with no prompt, the one-stage prompt method that introduces labels' description improves the F1 scores in topic prediction from 64.77 to 70.88 and from 73.70 to 77.02 on the two datasets. This result indicates that designing task related prompts can help LLM better understand the semantic association between texts and labels. Based on the one-stage prompt, the two-stage prompt method constructs the interaction between the predicted topic and personality prompt, F1 scores for personality prediction increase from 63.30 to 65.87 on the AQSH dataset, and from 65.96 to 69.25 on the YMLW dataset. This performance improvement could be attributed to the capability of mining and constructing the interaction during the process of predicting personality.

IV. CONCLUSIONS

This paper proposes a topic and personality prediction method based on the two-stage prompt, which investigates the role of topic information in personality identification and the usefulness of labels' descriptions for topic and personality prediction. Experimental results demonstrate the efficacy of the proposed method. Future work will explore the reciprocal interaction between topic and personality information. To enhance the interpretability of the model, we will also investigate how to effectively integrate psychological knowledge in the framework, which not only allows LLM to gain a deeper understanding of personality but also helps improve the performance of some downstream tasks, such as user behavior modeling and demand analysis.

ACKNOWLEDGMENT

This work was partially supported by the National Natural Science Foundation of China under Grants (No. 62071467, No. 62141608, No. 72293575).

REFERENCES

[1] H. Yang and Z. Huang, "Mining personality traits from social messages for game recommender systems," *Knowledge-Based Systems*, vol. 165, pp. 157–168, 2019.

[2] K. Yang, R. Y. K. Lau, and A. Abbasi, "Getting Personal: A Deep Learning Artifact for Text-Based Measurement of Personality,“ *Information Systems Research*, vol. 34, no. 1, pp. 194–222, 2023.

[3] F. Yang, X. Quan, Y. Yang, and J. Yu, "Multi-Document Transformer for Personality Detection,“ *AAAI*, vol. 35, no. 16, pp. 14221–14229, 2021.

[4] A.-R. Feizi-Derakhshi et al., "Text-based automatic personality prediction: A bibliographic review,“ *Journal of Computational Social Science*, vol. 5, no. 2, pp. 1555–1593, 2022.

[5] B. Zoph et al., "Emergent Abilities of Large Language Models." *TMLR*, 2022.

[6] Q. Li, D. Zeng, D. Xu, R. Liu, and R. Yao, "Understanding and Predicting Users' Rating Behavior: A Cognitive Perspective," *INFORMS Journal on Computing*, vol. 32, no. 4, pp. 996-1011, 2020.

[7] X. Liu et al., "P-Tuning v2: Prompt Tuning Can Be Comparable to Fine-tuning Universally Across Scales and Tasks." in *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics*, vol. 2, pp. 61-68, 2022.

[8] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," in *Proceedings of NAACL-HLT 2019*, pp. 4171–4186, 2019.

[9] Z. Du et al., "GLM: General Language Model Pretraining with Autoregressive Blank Infilling.“ in *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics*, vol. 1, pp. 320-335, 2022.