Knowledge-Enhanced Hierarchical Heterogeneous Graph for Personality Identification with Limited Training Data

Yuxuan Song[1, 2], Qiudan Li[1*], Yilin Wu[2, 1], David Jingjun Xu[3], Daniel Dajun Zeng[1, 2]

[1]The State Key Laboratory of Multimodal Artificial Intelligence Systems, Institute of Automation, Chinese Academy of Sciences, Beijing, China

[2]The School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing, China

[3]Department of Information Systems, College of Business, City University of Hong Kong, Hong Kong, China

songyuxuan2023, qiudan.li, wuyilin2022@ia.ac.cn, davidxu@cityu.edu.hk, zengdaniel@outlook.com

Abstract

Personality identification plays important roles in understanding user behavior and offering foresight ability for downstream applications. The key challenge is how to address the scarcity of labeled personality data. Recently, some studies have adopted data augmentation and prompt learning to perform personality identification. However, they still heavily require a large amount of labeled data to learn an appropriate distance strategy, which limits the generalization and flexibility of the model. This study proposes a knowledge-enhanced hierarchical heterogeneous graph model, which adopts a global multi-view graph node encoding to acquire comprehensive personality features and their inherent associations, where three types of knowledge including part-of-speech (POS) tag, entity, and Linguistic Inquiry and Word Count (LIWC) are introduced. Then, a hierarchical heterogeneous graph with a "post-word-diverse knowledge" structure is constructed for each post to obtain enhanced representation. Finally, a relation guided representation optimization that considers intra-user relationships and inter-label relationships is further developed to learn more discriminative semantic representation. Experimental results on three widely used datasets demonstrate that the model outperforms state-of-the-art methods when training with only 100 samples (approximately 1% of the total data set).

Introduction

With the rapid development of the Internet, social media has provided a quick channel for users to express their opinions on events and products. The voice of the customer (VOC) contains personality data, which reflects life choices, health, and happiness. Automatic personality identification from such data can enhance the explainability and foresight for applications, such as proactive policy prediction (Yang, Lau, and Abbasi 2023) and personalized recommendations (Yang et al. 2022), thus attracting widespread attention in the fields of artificial intelligence, cognitive psychology, and business intelligence.

Most personality identification methods mainly focus on using pre-training methods (Reimers and Gurevych 2019; Jiang, Zhang, and Choi 2020; Leonardi et al. 2020) and graphical models (Wang et al. 2020; Ramezani, Feizi-Derakhshi, and Balafar 2022; Li et al. 2024) to exploit all labeled data to learn text representations. Although these methods have achieved success, personality labels are limited in real-world scenarios. To address the data scarcity problem, some studies have recently adopted data augmentation and prompt learning to perform personality identification. Zhu et al. (2024) and Zhu et al. (2022) construct heterogeneous graphs on the basis of user text and Linguistic Inquiry and Word Count (LIWC) lexicon. Different contrastive loss strategies between unlabeled data (Zhu et al. 2024) and between text graphs (Zhu et al. 2022) are designed for data enhancement. Moreover, Wen et al. (2023) uses the T5 model for fine-tuning and prompt learning. Hu et al. (2024) enhances the pre-trained model using the knowledge generated by the large model. These strategies may be affected by the illusion of large models.

	User 1	User 2
Pair 1	oh nice! I love what you did with the wires :D	Woooooow I love what you did to love song!!
Pair 2	Were you able to get rid of it? Or did you make peace with it?	Oh well then sir what are you listening to?
Traits	EXT AGR **OPN**	NEU **OPN**

Table 1: An example of two pairs of posts from two users with different personality traits.

The above methods fail to model external information that embodies grammatical and semantic content, such as part-of-speech (POS) tag, entity, and writing styles. Moreover, their utilization of the psychological knowledge embedded in LIWC also needs to be strengthened. During the graph learning process, node representations are updated by the external information solely at a global level, resulting in low update efficiency. Furthermore, they overlook the dynamic interaction between users and the generated representation. A large amount of labeled data is still needed to determine an appropriate distance measurement strategy.

Table 1 shows two pairs of posts from two different users. For text pair 1, these two users all express admiration and affection, and this similarity between posts reflects the shared

*Corresponding author

openness trait exhibited by both users. However, it is difficult to distinguish the personality differences only based on semantics, while style information, such as POS tag (e.g. the adjective "nice") and emoticons (e.g. ":D"), reflects the personality characteristics of extraversion and agreeableness. As shown in text pair 2, the interrogative sentence like "what are you listening to?" highlights user's dissatisfaction and neuroticism, indicating that strengthening the association between posts from the same user can further help eliminate confusion caused by similar texts. Therefore, how to obtain more robust and comprehensive feature representations by designing fine data augmentation and feature extraction strategies, while efficiently utilizing user-text associations to reduce the effect of the posts with low discriminative capacity is the key technical challenge.

To realize more generalized and effective personality identification under limited training data environment, this study proposes a knowledge-enhanced hierarchical heterogeneous graph model, which provides a mechanism that combines global and local perspectives. It introduces diverse external knowledge to update nodes to obtain a more flexible node representation, then generates and optimizes an efficient representation associated with the user and personality label. Specifically, three types of external knowledge, including POS tag, entity and LIWC psychological information, are encoded by a global multi-view graph to acquire comprehensive personality features and their inherent associations. A hierarchical heterogeneous graph consisting of post-word-diverse knowledge layers is constructed locally for each post. A graph convolutional network (GCN) is adopted to obtain enhanced text representations. Finally, a relation guided representation optimization is developed to learn more discriminative semantic representation, which mines intra-user relationships and inter-label relationships by a self-attention mechanism and a hybrid multi-label contrastive learning, respectively. The main contributions of this work are as follows:

- It is the first attempt to achieve personality identification from a global-local perspective, which requires limited personality labeled data.
- A knowledge-enhanced hierarchical heterogeneous graph approach is proposed. This approach not only introduces rich personality clues but also designs the structure of "post-word-diverse knowledge" and a relation guided representation optimization mechanism, which makes the model more flexible and generalized.
- The experiments are conducted on three widely used datasets. With the same training size of 100 samples (approximately 1% of the total dataset), the proposed model outperforms state-of-the-art methods.

Related Work

Personality Detection

Traditional personality detection methods rely on the use of psycholinguistic features, such as psycholinguistic features from LIWC (Tausczik and Pennebaker 2010) or statistical text features from the bag-of-words model (Zhang, Jin, and Zhou 2010). With the rapid development of deep learning, various deep neural networks, including CNN (Xue et al. 2018), LSTM (Tandera et al. 2017), and Transformer (Leonardi et al. 2020), have been used for personality detection and have achieved significant success. Given the benefit from large-scale pre-trained language models, a series of pre-training fine-tuning paradigms are also being explored for this task. For example, Jiang, Zhang, and Choi (2020) concatenated all posts from a user into a document and encoded it using BERT (Devlin et al. 2019) and RoBERTA (Liu et al. 2019). Some works leveraged the contextual information and external psycholinguistic knowledge of LIWC to improve performance. Transformer-MD (Yang et al. 2021a) stored the hidden state of the post in the memory to avoid introducing order bias. TrigNet (Yang et al. 2021b) constructed a heterogeneous graph between posts for each user based on LIWC and used GAT to aggregate useful information. D-DGCN (Yang et al. 2023) developed a dynamic graph that enables the model to learn the connections between posts and adopted DGCN to integrate information. Semi-PerGCN (Zhu et al. 2024) proposed a semi-supervised approach to predict personality, which constructed a supervised personality GNN by combining LIWC with an attention mechanism to enhance multi-view graph, and further introduced unlabeled data to perform unsupervised consistency constraints on parameters.

Contrastive Learning

Contrastive learning proposed by Hadsell, Chopra, and LeCun (2006) has been widely used in self-supervised representation learning. It mainly uses data augmentation strategies to construct semantically related paired samples (Li et al. 2024). For example, Yan et al. (2021) and Gao, Yao, and Chen (2021) proposed different data augmentation strategies for contrastive learning using unlabeled data. DeCLUTR (Giorgi et al. 2021) regarded sentences extracted from the same article as positive samples; otherwise, they were considered negative samples. Shen et al. (2020) proposed an augmentation method based on truncated data. PairSCL (Li et al. 2022) considered sentence category information. For personality prediction tasks, researchers have also introduced contrastive learning methods. Zhu et al. (2022) proposed a graph enhancement method based on LIWC and text semantics and used contrastive self-supervised learning on the enhanced graph. Li et al. (2024) exploited user representation based on historical semantic information and psychological knowledge for cross-view contrastive learning.

Different from the above work, in addition to LIWC, diverse external knowledge, such as POS tag, entity, and writing styles, is introduced, and the dynamic interaction between post and these types of knowledge is learned to generate a more comprehensive semantic representation. Furthermore, an attention mechanism is designed to introduce user text related information to enhance the effectiveness of multi-label contrastive learning in capturing relationships between users and across labels.

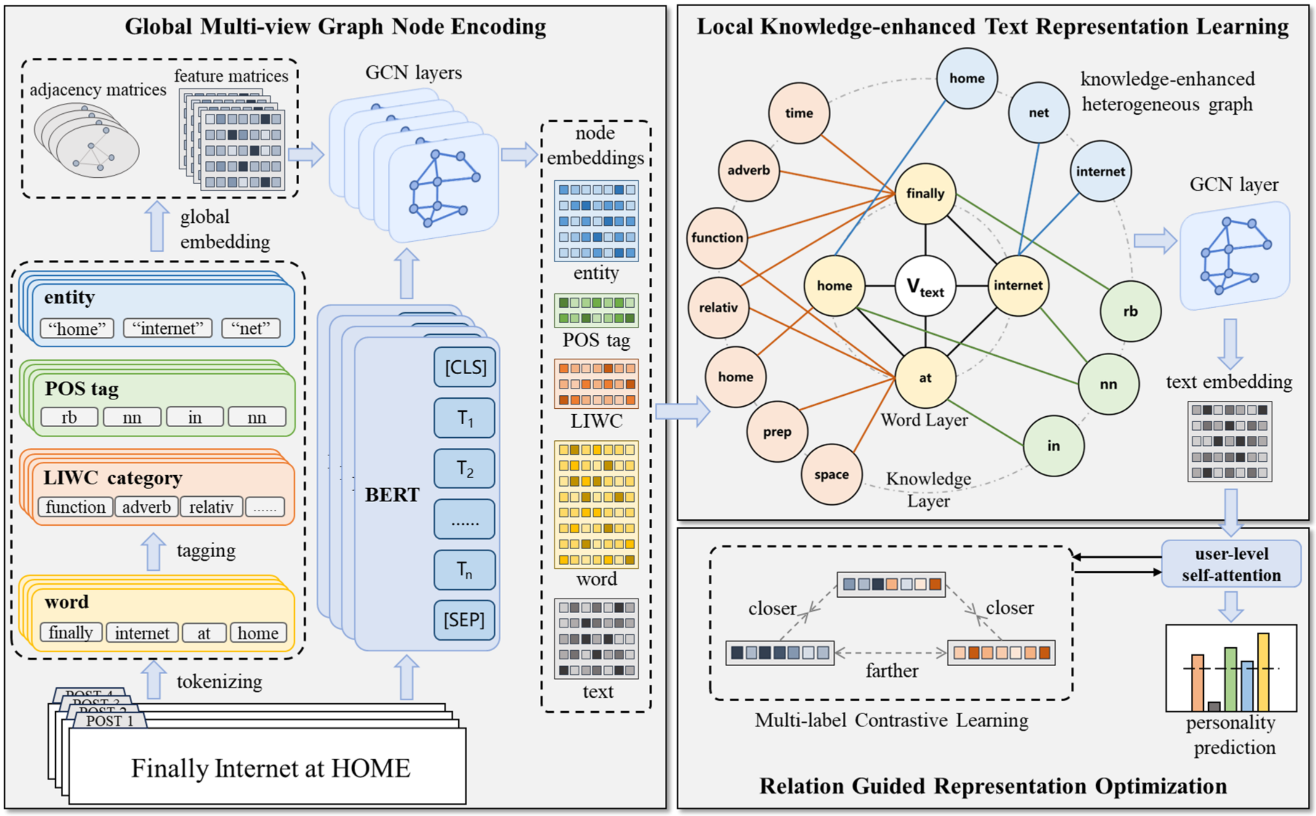

Figure 1: The overall architecture of the proposed model.

Preliminaries

Given an input dataset $S=\{x_i\}^N$ containing N posts and a label set L={Extraversion EXT, Neuroticism NEU, Agreeableness AGR, Conscientiousness CON, Openness OPN} defined in accordance with the Big Five personality theory, the goal of personality prediction task is to learn a mapping function f:$y_i=f\left(x_i, L\right)$ to obtain the corresponding personality label $y_i\in\{0,1\}^{|C|}$ based on the input post x_i, where |C| is the total number of labels, which is a fixed value of 5 in this task. Specifically, by mining the global feature representation of multiple knowledge in the dataset, this study constructs a knowledge-enhanced hierarchical heterogeneous graph for each post and uses a GCN to model the text feature vector. Then, we introduce user text association through user-level self-attention and predict the user's specific personality on the label set L by a multi-label contrastive learning method.

Methodology

The architecture of the proposed model, which consists of three main parts: global multi-view graph node encoding, local knowledge-enhanced text representation learning, and relation guided representation optimization, is shown in Figure 1. Specifically, the global multi-view graph node encoding part employs GCNs to capture features' inherent associations. It utilizes all available data to learn global node representations for each feature type, including semantic and external knowledge. The local knowledge-enhanced text representation learning part starts locally by constructing a knowledge-enhanced hierarchical heterogeneous graph for each post. Then, it uses GCN to learn text representations containing various personality-related external knowledge. The relation guided representation optimization part enhances global representation optimization by considering intra-user relationships through a user-level self-attention mechanism and inter-label relationships through a multi-label contrastive learning mechanism to achieve more accurate personality identification.

Global Multi-View Graph Node Encoding

Entities, POS tags, and LIWC categories reflect the personality information contained in a text from different perspectives. Entities reveal user interest themes, POS tags embody the syntactic structure and stylistic features of the text, and LIWC categories incorporate rich psycholinguistic prior knowledge. Fully mining these features and constructing internal connections allows for encoding comprehensive personality clues from a global perspective. Therefore, to encode these valuable features for personality identification, we modeled the three types of external information, words, and text representations from the entire dataset into five global graphs, which can be represented as $G_\tau=\{V_\tau,\ E_\tau\}$. Here, $\tau\in\{w, l, p, e,\ t\}$ represents word, LIWC classification, POS tag, entity and text representation, respectively. V_τ represents the set of all nodes of type τ, and E_τ represents the edges between nodes. Each feature's internal associations were then independently captured by two-layer GCNs, and local graph structure information was utilized to enrich the features contained therein, thereby building a more efficient global node representation $\mathbf{H}_\tau$:

$$\mathbf{H}_\tau = \widetilde{\mathbf{A}}_\tau \cdot ReLu\left(\widetilde{\mathbf{A}}_\tau \mathbf{X}_\tau \mathbf{W}_\tau^1\right)\mathbf{W}_\tau^2, \tag{1}$$

where $\widetilde{\mathbf{A}}_\tau = \mathbf{D}_\tau^{-\frac{1}{2}}(\mathbf{I} + \mathbf{A}_\tau)\mathbf{D}_\tau^{-\frac{1}{2}}$ is the normalized symmetric matrix of adjacency matrix $\mathbf{A}_\tau$, $\mathbf{X}_\tau$ is the initial node embedding, $\mathbf{W}_\tau^1$ and $\mathbf{W}_\tau^2$ are trainable parameters.

The details for constructing the feature graphs are as follows. For the word graph, the pretrained model BERT generates the initial embedding matrix $\mathbf{X}_w$ and the adjacency matrix $\mathbf{A}_w$ is calculated by the PMI of two words. For the LIWC classification and the POS tag graphs, the embeddings are one-hot vectors, and the adjacency matrices are obtained using co-occurrence frequency and PMI. For the entities graph, entities are determined based on the entity set provided in the NELL knowledge base (Carlson et al. 2010), and the TransE method (Bordes et al. 2013) is used to learn $\mathbf{X}_e$. The adjacency matrix $\mathbf{A}_e$ is obtained using cosine similarity between nodes. For the text representation graph, the initial embedding is represented by the [CLS] vector output by BERT, and the adjacency matrix is also obtained using cosine similarity.

Local Knowledge-Enhanced Text Representation Learning

To learn an enhanced text representation incorporating more personality-related information, it is necessary to utilize the external information representations and establish connections with the text representation, facilitating dynamic interactions among them. For example, the post: "I diligently prepare my presentations and consistently seek feedback to improve." By integrating the analysis of work-related entities "presentations" and "feedback", adverbs "diligently" and "consistently" from POS tags, and the "improve" from the LIWC category of achievement, along with contextual semantic analysis, we can infer that the user has a high level of conscientiousness.

Therefore, we modeled a local knowledge-enhanced hierarchical heterogeneous graph for each post after obtaining the global representation $\mathbf{H}_\tau$ for each node to leverage rich external knowledge. As shown in Figure 1, the heterogeneous graph G_h corresponding to each post contains three layers: post–word–external knowledge. The feature vector $\mathbf{x}_h$ of each node comes from the previously obtained global embedding $\mathbf{H}_\tau$. The cross-layer edges between post–word and word–knowledge were constructed in an undirected manner, indicating that the word appears in the post or belongs to a certain category of external knowledge:

$$\mathbf{A}_{i,\,j}^{cross} = \mathbf{A}_{j,\,i}^{cross} = 1, \quad if\; contain\,(w_i, c_j)\,; \quad else\; 0, \tag{2}$$

where w_i represents a word and c_j represents a post or a type of external knowledge. The $contain(w_i,\ c_j)$ indicating that a word w_i is contained in a post c_j, or its corresponding POS tag, entity or LIWC classification belongs to a type of external knowledge c_j.

For word–word edges, we used the PMI method to find their co-occurrence relationships and connect them undirected in the order of appearance:

$$\mathbf{A}_{i,\,i+1}^{intra} = \mathbf{A}_{i+1,\,i}^{intra} = PMI\,(w_i, w_{i+1})\,, \tag{3}$$

where $PMI(w_i,\ w_{i+1})$ represents the pointwise mutual information of two adjacent words w_i and w_{i+1}.

For the hierarchical heterogeneous graphs constructed by the above method, we used a two-layer GCN to expand further the feature depth contained in the post nodes and merged the feature vectors $\mathbf{V}_{text}$ corresponding to all post nodes after calculation into a new text embedding matrix $\hat{\mathbf{X}}_t$.

Relation Guided Representation Optimization

Personality data on social media is characterized by short text lengths, casual language use, and semantically similar posts from users with different personalities. This presents challenges for representation learning methods due to interference from noise, such as low-quality and highly similar data. To address this issue, we designed a hybrid user-level and label-level relation-guided representation optimization method to obtain distinguished embeddings. It utilizes a user-level self-attention mechanism to enhance the similarity of posts from the same user at the representation and a label-level contrastive learning strategy to accentuate the differences between posts from users with different personalities, thereby obtaining distinguished representations.

Specifically, the user-level self-attention mechanism further enhances the connection between posts from the same user. The process of obtaining the text representation enhanced by user relations $\mathbf{X}_{user}$ is as follows:

In order to introduce user-text information and eliminate the interference caused by low-quality data, we introduced a user level self-attention based on the previously obtained text embedding matrix $\hat{\mathbf{X}}_t$ to further enhance the connection between posts from the same user on the basis of mining the correlation between texts. Specifically, we calculated $\mathbf{X}_{user}$ in the following way:

$$\mathbf{X}_{user} = (softmax(\frac{\mathbf{Q}\mathbf{K}^T}{\sqrt{d_k}}) + \mathbf{S})\mathbf{V}, \tag{4}$$

$$\mathbf{Q}, \mathbf{K}, \mathbf{V} = \mathbf{W}_q\hat{\mathbf{X}}_t, \mathbf{W}_k\hat{\mathbf{X}}_t, \mathbf{W}_v\hat{\mathbf{X}}_t, \tag{5}$$

$$\mathbf{S} = \begin{bmatrix} \mathbf{J}_{u_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \mathbf{J}_{u_n} \end{bmatrix}, \tag{6}$$

where $\mathbf{W}_q,\ \mathbf{W}_k,\ \mathbf{W}_v$ are trainable parameters, $\mathbf{J}_{u_i}$ represents a $u_i \times u_i$ matrix of all ones, and u_i is the number of posts belonging to the i-th user.

Then, the label-level relation-guided contrastive learning strategy to accentuate the differences between posts from

Datasets	User	Post	Training ratio
PANDORA	1458	7005	1.43%
PAN2015	278	9634	1.04%
MyPersonality	205	9799	1.02%

Table 2: Statistics of datasets.

users with different personalities. This strategy is implemented by designing a multi-label contrastive loss function as follows:

$$\mathcal{L}_d = BCE\left(y_d, y\right), \tag{7}$$

$$\mathcal{L}_c = \sum_{i,\ j \in B(i)} -\beta_{ij} log \frac{e^{-d(\mathbf{x}_i, \mathbf{x}_j)/\theta}}{\sum_{k \in B(i)} e^{-d(\mathbf{x}_i, \mathbf{x}_k)/\theta}}, \tag{8}$$

where y represents the ground-truth personality feature, $d(\cdot, \cdot)$ represents the Euclidean distance. $\mathbf{x}_i$ is the feature vector of the i-th text in $\mathbf{X}_{user}$, $B(i)$ denotes the batch to which instance i belongs. θ is the temperature hyperparameter of contrastive learning.

Finally, after obtaining the predicted personality score y_d using the sigmoid function, we designed a multiple loss function that combines the traditional BCE loss $\mathcal{L}_d$ and the multilabel contrastive learning loss $\mathcal{L}_c$ based on the work of Wang, Dai et al. (2022) through weighted summation to train the model more effectively:

$$\beta_{ij} = \frac{{y^i}^T \cdot y^j}{\sum_{k \in B(i)} {y^i}^T \cdot y^k}, \tag{9}$$

$$\mathcal{L} = \mathcal{L}_d + \lambda \mathcal{L}_c, \tag{10}$$

where λ is a balanced hyperparameter.

Experiment

Experimental Settings

Datasets We conducted experiments on three commonly used datasets: PANDORA, PAN2015, and MyPersonality with Big Five personality labels. We randomly selected approximately 10,000 original posts with a length of more than five words in each dataset and cleaned the data by removing duplicate posts and posts that do not contain any personality labels to ensure fair comparisons on different datasets. A training set containing 100 posts and a validation set containing 100 posts were formed by a random algorithm based on depth-first search. The test set comprised the remaining data, and posts from the same user were ensured to appear in only one of the sets. The basic information of each dataset is shown in Table 2.

- PANDORA (Gjurković et al. 2021): This dataset is a large-scale collection of user-generated content sourced from the Reddit platform. It includes user comments alongside corresponding personality labels, which have been derived based on linguistic analysis and validated psychometric assessments. We randomly selected five posts from each labeled user for analysis.
- PAN2015 (Rangel Pardo et al. 2015): This dataset comes from the data science competition PAN2015 and contains a large corpus of posts from social media, including Twitter (now X) and blogs, in English, Spanish, Dutch and Italian. Our analysis mainly focuses on English data.
- MyPersonality (Xue et al. 2018): This dataset comes from an open-source project on Facebook allowing users to test their personality via online psychometric tests. It includes personality test results based on the Big Five personality traits, along with anonymized user-generated content.

Baselines We compare the proposed model with three groups of baseline methods: (1) deep neural network-based models, (2) graph-based models, and (3) metric learning-based models, which can be categorized as follows:

Deep neural network-based models

- textCNN (Kampman et al. 2018): This method is a trimodal architecture CNN for predicting personality from video clips with different audio, video, and text channels. We used the CNN structure designed for text in their study as a baseline.
- BERT (Devlin et al. 2019): This method uses a pretrained BERT-base-uncased model to encode the text, and then outputs the personality prediction results through a two-layer MLP.

Graph-based models

- BertGCN (Lin et al. 2021): This method constructs a heterogeneous graph containing word and text nodes, where the representation vectors of the nodes are extracted using the BERT model. Then GCN is used to obtain the feature representation of the text and perform predictions.
- SHINE (Wang et al. 2021): This method constructs a hierarchical heterogeneous graph containing words, part-of-speech tags, entities, and text, and enhances the propagation of labels across texts by capturing interactions within similar nodes.
- Semi-PerGCN (Zhu et al. 2024): This method constructs a heterogeneous graph containing posts, words, and LIWC classification, uses GCN and attention on the psychological path to predict personality, and uses large-scale unlabeled data for data augmentation.

Metric learning-based models

- Prototypical Network (Snell, Swersky, and Zemel 2017): This method generates class prototype vectors from training data, and determines the predicted label based on the distance between the query and the prototype.
- DCKPN (Zhang et al. 2023): This method constructs a dual-class knowledge propagation network with instance level and class level, propagates label information and feature structure in the latent space of the data, and constructs a prototype vector for prediction.
- IML-CL (Wang, Dai et al. 2022; Tian et al. 2024): This method designs a multi-label contrastive learning objective and combines it with the k-nearest neighbor method to obtain the final prediction result. We adopted BERT and applied this method to it as a baseline.

Datasets	Traits	Methods								
		CNN	BERT	BertGCN	SHINE	Semi-PerGCN	Prototypical Network	DCKPN	IML-CL	Ours
PANDORA	EXT	53.53	57.08	35.74	39.12	59.39	47.80	52.06	57.67	**63.97**
	NEU	53.76	34.94	34.23	32.41	**63.07**	46.14	47.34	46.35	62.21
	AGR	40.20	**57.22**	31.04	35.48	52.68	41.85	46.53	57.17	53.23
	CON	47.99	59.36	37.95	37.95	60.44	50.09	54.33	60.79	**65.56**
	OPN	45.80	42.58	42.30	42.30	50.38	21.07	56.40	44.04	**59.50**
	AVE	48.26	50.24	36.25	37.45	57.19	41.39	51.33	53.20	**60.89**
PAN2015	EXT	52.78	54.14	35.79	35.79	45.84	50.71	48.04	55.02	**55.24**
	NEU	38.18	53.66	39.71	39.71	58.96	52.45	52.32	54.93	**60.67**
	AGR	39.91	44.75	39.01	34.16	45.96	47.76	48.21	49.73	**50.53**
	CON	**55.84**	53.00	31.10	35.42	50.49	47.95	49.48	53.91	55.81
	OPN	51.43	36.57	34.86	34.86	41.69	46.58	48.42	48.49	**54.05**
	AVE	47.63	48.42	36.09	35.99	48.59	49.09	49.29	52.42	**55.26**
MyPersonality	EXT	54.20	47.07	29.76	29.76	51.66	36.56	44.46	42.52	**55.61**
	NEU	46.24	47.70	27.05	38.62	45.10	27.05	44.92	**49.08**	47.86
	AGR	47.10	47.09	34.72	34.72	48.75	31.88	46.86	**50.93**	49.41
	CON	**51.82**	43.95	31.43	31.42	46.21	35.14	46.87	43.69	46.52
	OPN	45.00	44.39	42.92	42.76	47.34	42.76	44.55	47.88	**50.66**
	AVE	48.87	46.04	33.18	35.46	47.81	34.68	45.53	46.82	**50.01**

Table 3: Overall results of the proposed model and baseline models in Macro-F1(%) score.

Implementation Details The model is implemented using PyTorch and trained by AdamW optimizer with an initial learning rate of 10^{-4} and a dropout ratio of 0.5. BERT-base-uncased is used to extract global text representations and word vectors. The hidden layer dimensions of all GCNs are set to 400, the batch size is 100, and the contrastive learning temperature θ is 10. All hyperparameters are tuned over the validation set to obtain the optimized results.

Overall Results

We compared the proposed model to all baseline models on three datasets, using Macro-F1 score as the evaluation metric. All experiments were repeated five times using random seeds, and the average results are presented in Table 3.

The proposed model demonstrated optimal performance on all datasets, outperforming the best baseline model Semi-PerGCN, IML-CL and CNN by 6.47%, 5.42%, and 2.33% on the three datasets, respectively. The experimental results demonstrate the effectiveness of the proposed model, showing that reliable predictive performance can be achieved even when trained with a limited number of data (only 100 samples in this article). The reasons may be two-fold: (1) With varied external knowledge introduced by the proposed model for enhancement, the generation of global node features and the construction of local hierarchical heterogeneous graph effectively enhance the predictive effectiveness of text feature vectors. (2) The strategy of enhancing contrastive learning with user-level self-attention effectively uncovers the correlation between data, eliminates the effect of low-quality samples within the dataset, and provides a basis for the propagation of labels from a few training samples to the entire dataset.

For different categories of methods, although the graph-based models can propagate labels across neighboring nodes using local structure, it requires node feature representations that contain sufficiently rich personality cues as a basis. Consequently, BertGCN and SHINE perform poorly in the personality domain. By contrast, Semi-PerGCN, which introduces psychological knowledge and unlabeled data for data augmentation, performs better and achieves the best result among all baselines with an F1 score of 57.19% on the PANDORA dataset. Both the deep neural network CNN and the pre-trained language model BERT have strong feature extraction capabilities, enabling them to extract sufficient knowledge from small samples for training and thus achieve good performance. CNN even achieves an F1 score of 48.87% on the MyPersonality dataset, which is the best among all baselines. Although the prototype network is developed for few-shot classification, it still faces certain limitations in terms of personality prediction and is considerably affected by the issue of imbalanced training data labels; thus, the overall performance is relatively poor. DCKPN and IML-CL combine strategies such as knowledge graph enhancement, contrastive learning, and unsupervised learning, which are commonly used in small dataset learning, thereby demonstrating good prediction performance.

Overall, all models demonstrated optimal performance on the PANDORA dataset, followed by PAN2015 and MyPersonality dataset. The PANDORA dataset contains more users, more diverse personality labels, and a more balanced number of posts for each user, which contributed to the improved prediction performance. In addition, in contrast to the PANDORA dataset, which consists of responses to specific topics and issues from Reddit, the PAN2015 and MyPersonality datasets are derived from posts by Twitter (now X) and Facebook users. These posts often contain shorter and more casually written texts, resulting in generally lower data quality, which in turn leads to poorer predictive performance on

Traits	F1 Score(%)			
	Full Model	-Knowledge Subgraph	-User-level Attention	-Contrastive Learning
EXT	63.97	63.40	63.00	64.24
NEU	62.21	54.39	60.02	60.75
AGR	53.23	51.21	54.44	52.66
CON	65.56	56.24	58.85	62.02
OPN	59.50	50.61	57.55	56.78
AVE	**60.89**	55.17 (↓9.40%)	58.77 (↓3.48%)	59.29 (↓2.63%)

Table 4: Results of ablation study of the proposed model on PANDORA, where "-" denotes the removal of a component.

Traits	F1 Score(%)			
	Full Model	-POS tag Nodes	-Entity Nodes	-LIWC Nodes
EXT	63.97	63.62	64.08	62.74
NEU	62.21	59.76	59.91	54.71
AGR	53.23	52.40	52.24	50.63
CON	65.56	64.20	64.98	66.03
OPN	59.50	57.45	51.09	54.57
AVE	**60.89**	59.49 (↓2.31%)	58.46 (↓4.00%)	57.74 (↓5.19%)

Table 5: Detailed results of ablation study for specific knowledge categories.

these two datasets. This observation also indicates that in the field of personality prediction, high-quality posts are crucial for achieving accurate results.

Experimental Analysis

This section analyzes the usefulness of each component used in the proposed model and the effect of using different strategies to generate initial word embeddings and the values of the contrastive learning weight λ.

Ablation Study We conducted ablation studies for the proposed model on the PANDORA dataset to investigate the effects of each part of the structure. As shown in Table 4, the performance of the model drops considerably after removing external knowledge, user-level information, or contrastive learning loss. When we do not use external knowledge enhancement and only construct the local graph using post and word nodes, the model's performance decreases by 9.40%, which indicates the necessity of introducing rich external knowledge for accurate personality prediction in limited data scenarios. Table 5 provides a more detailed breakdown of the proposed model's performance when a specific type of knowledge is removed. The LIWC classification nodes, representing psychological knowledge, are the most important, and the entity nodes reflecting language habits and target domains also have a great effect. The grammatical structure information contained in the POS tag nodes is affected by the informal nature of online posts and was included in the word vector generation process of BERT; therefore, its effect is minimal. Removing the user-level self-attention caused a drop of 3.48% on the performance, proving the importance of user information for personality prediction. The introduction of contrastive learning helps capture the differences between categories better, and also leads to a 2.63% performance drop after removal.

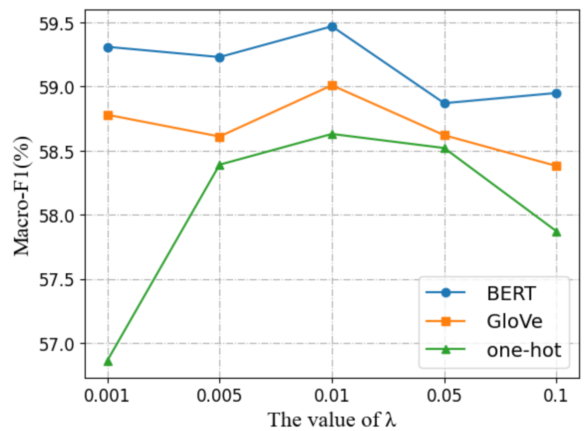

Figure 2: Performance curves of different initial word embedding methods and different hyperparameter λ values.

Effect of Different Initial Word Embedding Generation Methods

We investigated the effect of using different strategies, including BERT, GloVe, and one-hot encodings, to generate initial embeddings for word nodes before generating global features. The performance of the proposed model under different generation methods and different hyperparameter λ in Eq. 10 are shown in Figure 2. Compared with one-hot vectors, GloVe generates word vectors through global word frequency statistics, which contains more semantic information and therefore performs better. BERT, combined with a large-scale pre-training corpus, can generate more accurate semantic representations with the help of context, achieving the best results. Besides, as the value of λ increases, the performance improves initially but eventually starts to decline after λ reaches 0.01. This result indicates that cross entropy loss and contrastive learning loss should be considered simultaneously, and overemphasis on either side weakens the model's generalization ability.

Conclusion

This study proposes a knowledge-enhanced hierarchical heterogeneous graph method, which takes a step towards performing personality identification under limited training samples, addressing the key challenges of scare labels in real word. In the future, we will further analyze the impact of different ratios of training data on prediction performance. Besides, to solve the computational challenges faced with larger datasets or real-time applications, exploring the graph optimization methods and the combination strategy of offline storage and block parallel are valuable future research directions.

Acknowledgments

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

References

Bordes, A.; Usunier, N.; Garcia-Duran, A.; Weston, J.; and Yakhnenko, O. 2013. Translating embeddings for modeling multi-relational data. *Advances in neural information processing systems*, 26.

Carlson, A.; Betteridge, J.; Kisiel, B.; Settles, B.; Hruschka, E.; and Mitchell, T. 2010. Toward an architecture for never-ending language learning. In *Proceedings of the AAAI conference on artificial intelligence*, volume 24, 1306–1313.

Devlin, J.; Chang, M.-W.; Lee, K.; and Toutanova, K. 2019. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, 4171–4186.

Gao, T.; Yao, X.; and Chen, D. 2021. SimCSE: Simple Contrastive Learning of Sentence Embeddings. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, 6894–6910.

Giorgi, J.; Nitski, O.; Wang, B.; and Bader, G. 2021. DeCLUTR: Deep Contrastive Learning for Unsupervised Textual Representations. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, 879–895.

Gjurković, M.; Karan, M.; Vukojević, I.; Bošnjak, M.; and Šnajder, J. 2021. PANDORA Talks: Personality and Demographics on Reddit. In *9th International Workshop on Natural Language Processing for Social Media*, 138–152.

Hadsell, R.; Chopra, S.; and LeCun, Y. 2006. Dimensionality reduction by learning an invariant mapping. In *2006 IEEE computer society conference on computer vision and pattern recognition (CVPR'06)*, volume 2, 1735–1742. IEEE.

Hu, L.; He, H.; Wang, D.; Zhao, Z.; Shao, Y.; and Nie, L. 2024. LLM vs Small Model? Large Language Model Based Text Augmentation Enhanced Personality Detection Model. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, 18234–18242.

Jiang, H.; Zhang, X.; and Choi, J. D. 2020. Automatic text-based personality recognition on monologues and multiparty dialogues using attentive networks and contextual embeddings (student abstract). In *Proceedings of the AAAI conference on artificial intelligence*, volume 34, 13821–13822.

Kampman, O.; Barezi, E. J.; Bertero, D.; and Fung, P. 2018. Investigating Audio, Video, and Text Fusion Methods for End-to-End Automatic Personality Prediction. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, 606–611.

Leonardi, S.; Monti, D.; Rizzo, G.; Morisio, M.; et al. 2020. Multilingual Transformer-Based Personality Traits Estimation. *INFORMATION*, 11(4): 179–199.

Li, M.; Zhu, Y.; Li, S.; and Wu, B. 2024. HG-PerCon: Cross-view contrastive learning for personality prediction. *Neural Networks*, 169: 542–554.

Li, S.; Hu, X.; Lin, L.; and Wen, L. 2022. Pair-level supervised contrastive learning for natural language inference. In *ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 8237–8241. IEEE.

Lin, Y.; Meng, Y.; Sun, X.; Han, Q.; Kuang, K.; Li, J.; and Wu, F. 2021. BertGCN: Transductive Text Classification by Combining GNN and BERT. In *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*, 1456–1462.

Liu, Y.; Ott, M.; Goyal, N.; Du, J.; Joshi, M.; Chen, D.; Levy, O.; Lewis, M.; Zettlemoyer, L.; and Stoyanov, V. 2019. Roberta: A robustly optimized bert pretraining approach. arXiv:1907.11692.

Ramezani, M.; Feizi-Derakhshi, M.-R.; and Balafar, M.-A. 2022. Knowledge Graph-Enabled Text-Based Automatic Personality Prediction. *Computational Intelligence and Neuroscience*, 2022(1): 3732351.

Rangel Pardo, F. M.; Celli, F.; Rosso, P.; Potthast, M.; Stein, B.; and Daelemans, W. 2015. Overview of the 3rd Author Profiling Task at PAN 2015. In *CLEF 2015 evaluation labs and workshop working notes papers*, 1–8.

Reimers, N.; and Gurevych, I. 2019. Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, 3982–3992.

Shen, D.; Zheng, M.; Shen, Y.; Qu, Y.; and Chen, W. 2020. A simple but tough-to-beat data augmentation approach for natural language understanding and generation. arXiv:2009.13818.

Snell, J.; Swersky, K.; and Zemel, R. 2017. Prototypical networks for few-shot learning. *Advances in neural information processing systems*, 30.

Tandera, T.; Suhartono, D.; Wongso, R.; Prasetio, Y. L.; et al. 2017. Personality prediction system from facebook users. *Procedia computer science*, 116: 604–611.

Tausczik, Y. R.; and Pennebaker, J. W. 2010. The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of language and social psychology*, 29(1): 24–54.

Tian, G.; Wang, J.; Wang, R.; Zhao, G.; and He, C. 2024. A multi-label social short text classification method based on contrastive learning and improved ml-KNN. *Expert Systems*, e13547.

Wang, R.; Dai, X.; et al. 2022. Contrastive learning-enhanced nearest neighbor mechanism for multi-label text classification. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, 672–679.

Wang, Y.; Wang, S.; Yao, Q.; and Dou, D. 2021. Hierarchical Heterogeneous Graph Representation Learning for Short

Text Classification. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, 3091–3101.

Wang, Z.; Wu, C.-H.; Li, Q.-B.; Yan, B.; and Zheng, K.-F. 2020. Encoding text information with graph convolutional networks for personality recognition. *Applied sciences*, 10(12): 4081.

Wen, Z.; Cao, J.; Yang, Y.; Wang, H.; Yang, R.; and Liu, S. 2023. DesPrompt: Personality-descriptive prompt tuning for few-shot personality recognition. *Information Processing & Management*, 60(5): 103422.

Xue, D.; Wu, L.; Hong, Z.; Guo, S.; Gao, L.; Wu, Z.; Zhong, X.; and Sun, J. 2018. Deep learning-based personality recognition from text posts of online social networks. *Applied Intelligence*, 48(11): 4232–4246.

Yan, Y.; Li, R.; Wang, S.; Zhang, F.; Wu, W.; and Xu, W. 2021. ConSERT: A Contrastive Framework for Self-Supervised Sentence Representation Transfer. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, 5065–5075.

Yang, F.; Quan, X.; Yang, Y.; and Yu, J. 2021a. Multi-document transformer for personality detection. In *Proceedings of the AAAI conference on artificial intelligence*, volume 35, 14221–14229.

Yang, K.; Lau, R. Y.; and Abbasi, A. 2023. Getting personal: A deep learning artifact for text-based measurement of personality. *Information Systems Research*, 34(1): 194–222.

Yang, Q.; Nikolenko, S.; Huang, A.; and Farseev, A. 2022. Personality-Driven Social Multimedia Content Recommendation. In *Proceedings of the 30th ACM International Conference on Multimedia*, 7290–7299.

Yang, T.; Deng, J.; Quan, X.; and Wang, Q. 2023. Orders are unwanted: dynamic deep graph convolutional network for personality detection. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 37, 13896–13904.

Yang, T.; Yang, F.; Ouyang, H.; and Quan, X. 2021b. Psycholinguistic Tripartite Graph Network for Personality Detection. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, 4229–4239.

Zhang, F.; Chen, W.; Ding, F.; and Wang, T. 2023. Dual class knowledge propagation network for multi-label few-shot intent detection. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 8605–8618.

Zhang, Y.; Jin, R.; and Zhou, Z.-H. 2010. Understanding bag-of-words model: a statistical framework. *International journal of machine learning and cybernetics*, 1: 43–52.

Zhu, Y.; Hu, L.; Ge, X.; Peng, W.; and Wu, B. 2022. Contrastive Graph Transformer Network for Personality Detection. In *IJCAI*, 4559–4565.

Zhu, Y.; Xia, Y.; Li, M.; Zhang, T.; and Wu, B. 2024. Data Augmented Graph Neural Networks for Personality Detection. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, 664–672.