

## PAPER

## Heterogeneous Graph Contrastive Learning for Stance Prediction

Yang LI<sup>†a)</sup> and Rui QI<sup>†</sup>, Nonmembers

**SUMMARY** Stance prediction on social media aims to infer the stances of users towards a specific topic or event, which are not expressed explicitly. It is of great significance for public opinion analysis to extract and determine users' stances using user-generated content on social media. Existing research makes use of various signals, ranging from text content to online network connections of users on these platforms. However, it lacks joint modeling of the heterogeneous information for stance prediction. In this paper, we propose a self-supervised heterogeneous graph contrastive learning framework for stance prediction in online debate forums. Firstly, we perform data augmentation on the original heterogeneous information network to generate an augmented view. The original view and augmented view are learned from a meta-path based graph encoder respectively. Then, the contrastive learning among the two views is conducted to obtain high-quality representations of users and issues. Finally, the stance prediction is accomplished by matrix factorization between users and issues. The experimental results on an online debate forum dataset show that our model outperforms other competitive baseline methods significantly.

**key words:** stance prediction, heterogeneous information network, contrastive learning, matrix factorization

## 1. Introduction

In recent years, online social media has become an important platform for users to express and discuss their personal opinions about events and issues in the real world [1]. Mining users' real feedback from the text published by users is of great significance for understanding public opinion and analyzing social dynamics [2]. Stance detection is an important step in analyzing social media discussions, assisting public opinion analysis and decision making [3]. Given a target topic, the task of stance detection aims to determine whether a user or his/her text is in favor (agree), against (disagree), or neutral towards it. Stance detection is a key component in downstream tasks like fake news detection, argumentation systems and recommendation systems, and has become an important research direction in the field of Natural Language Processing and Social Media Analysis.

Existing research of stance detection mainly relies on large number of human annotation data to train machine learning models, and identifies user's posts towards a given issue through a classification task, in which the text of on-topic posts are used as the features. However, it is expensive and difficult to extend to new issues [4]–[6]. In addition,

users usually do not express their stance on specific topics explicitly [7]. Instead, they may show endorsements of their support or opposition to the current issue by agreeing with some other issues or other users. Therefore, inferring the stance of "silent users" has attracted much attention in recent years [8]–[10] and the task of stance prediction has emerged.

Different from traditional stance detection task, *stance prediction* aims at inferring the stances of users on certain targets, which are not expressed or implied on social media. The stance in social media can be inferred from a mixture of signals that might reflect user's beliefs such as posts and online interactions. This is similar to recommending new items to users based on their purchase history. Therefore, we regard this research problem as a recommendation problem. In this paper, to capture these key factors for stance prediction, we propose a Heterogeneous Graph Contrastive Learning model (HGCL-SP), which leverages the rich heterogeneous information of users and correlation between issues [11]. Firstly, we build a heterogeneous information network (HIN) [12] to connect four types of entities (user, topic, issue and argument), and utilize a HIN to describe users, issues and the corresponding heterogeneous relationships between them. Next, motivated by the recent success of graph contrastive learning, we perform data augmentation on the constructed original graph structure to generate an augmented view. The meta-path is a relation sequence connecting two entities, which is widely used to obtain the high-order structure of HIN and the semantic relationship between entities. By passing messages along different meta-paths, high-quality node embedding can be learned. Based on this idea, a meta-path based GCN model is used to learn the two views and the final representation learning is obtained by contrasting the same node from different views. Finally, the stance prediction is completed by the matrix factorization of users and issues. We conduct comprehensive experiments on a public dataset and show that our model outperforms the SOTA methods significantly.

The contributions of this paper are summarized as follows:

- We model the task of stance prediction as a recommendation problem, and propose a Heterogeneous Graph Contrastive Learning model (HGCL-SP) that jointly models users, topics, issues and arguments through heterogeneous graph learning.
- In order to enrich the interactions of the heterogeneous

Manuscript received April 23, 2022.

Manuscript revised June 13, 2022.

Manuscript publicized July 25, 2022.

<sup>†</sup>The authors are with the College of Information and Computer Engineering, Northeast Forestry University, Harbin, 150040, China.

a) E-mail: yli@nefu.edu.cn (Corresponding author)

DOI: 10.1587/transinf.2022EDP7065

information network, we conduct contrastive learning to maximize the similarity of representations of the same node from the original graph and the augmented view. To the best of our knowledge, this is the first attempt to utilize a heterogeneous graph contrastive learning framework for the task of stance prediction.

- We conduct comprehensive comparison experiments on a public dataset, and verify the effectiveness of each part of our approach through ablation experiments. Experimental results show that the proposed approach outperforms the state-of-the-art approaches.

## 2. Related Work

In this part, we introduce the existing related work on stance detection, heterogeneous graph neural network and contrastive learning.

### 2.1 Stance Detection

According to whether the stance is expressed in the text, existing research work related to stance detection can be divided into two categories: one is stance classification, to detect the stances of given texts, and the other is stance prediction, inferring the stances of users on certain targets, which are not expressed or implied on social media. For stance classification, most studies mainly learn the representations of the text and the target content, and then perform the classification through different deep learning models [13]–[16]. Neural models for stance detection include a few informative categories like attention-based, convolution-based and pre-training models. Yang et al. [4] proposed a two-stage deep attention neural network for target specific stance detection. The model employed densely connected BI-LSTM to encode tweet tokens and traditional bidirectional LSTM to encode target tokens. Sun et al. [13] proposed a hierarchical attention network to learn the representation of documents with different language features and adjust the weights of different feature sets. Ghosh et al. [15] explored the reproducibility of several existing stance detection models, including both neural models and classical classifier-based models. Sun et al. [16] developed a new method that applies the recently developed BERT model for stance detection. Furthermore, Li et al. [6] proposed the BERTtoCNN model combining the classic distillation loss with similarity-preserving loss to improve the performance of stance detection. In order to enrich semantics of the user text, some work tried to introduce external knowledge [5], [17], [18]. Zhang et al. [5] used external knowledge such as semantic and emotion lexicons to supplement the target information, and put forward a knowledge transfer model that can be used for cross-domain stance detection. In addition, Liu et al. [17] proposed a graph representation learning model to capture the relationship between targets and introduced common knowledge to further improve the stance detection task.

For stance prediction, most studies mainly predict the stances of users from micro and macro levels in social media platforms [4], [9] or online debate forums [8], [10]. From the microscopic perspective, stance prediction can be viewed as a recommendation problem. Qiu et al. [19] proposed a latent factor model, which integrated the arguments, interactions and attributes of users into a collaborative filtering framework, and completed the stance prediction task on CreateDebate. In addition, Sasaki et al. [1] used Factorization Machine (FM) to model the preferences of social media users. By analyzing the experimental results, it can be seen that the historical behavior of users is helpful to predict the stance of “silent users”. Huang et al. [10] proposed a Heterogeneous Argument Attention Network, which jointly learned stance prediction and persuasiveness prediction in multi-round of dialogues. The advantage of this model lies in that it makes use of argument structure information through the GNN module. The macro level stance prediction is usually regarded as a collection of micro level predictions, to infer the opinion of public on an event in some specific cases [8].

### 2.2 Heterogeneous Graph Neural Network

Graph neural networks [20] have become a research hotspot in the field of deep learning. According to the number of entity types contained in the graph, it can be divided into homogeneous graphs [21] and heterogeneous graphs [22]. Wang et al. [23] proposed a heterogeneous graph neural network based on node-level and semantic-level attention mechanism. The importance of nodes and meta-paths can be fully taken into account by using the hierarchical attention mechanism. However, this model only considers the start node and the end node of meta-paths, resulting in information loss. Therefore, Fu et al. [24] proposed two information aggregation strategies, which aggregate neighbor information from within and between meta-paths to generate node embedding, thus effectively solving the problem of missing information. Gong et al. [22] proposed an attention-based heterogeneous graph neural network, which learns the representation of entities by using content features and heterogeneous context features. Because different meta-paths have different meanings, this model can learn users’ actual preferences under the guidance of attention mechanism.

### 2.3 Contrastive Learning

In recent years, graph representation learning based on contrastive learning has achieved good results [25], [26]. Zhu et al. [25] proposed a new heterogeneous graph contrastive learning framework, which generates different views through data augmentation, and learns node representations by maximizing the consistency of the same node representations in different views. Wang et al. [27] conducted a contrastive learning between the network schema view and the meta-path view of a HIN. In addition, contrastive learning can also be used for multi-task learning. Wu et al. [26] re-

garded the traditional supervised recommendation task as the main task, and constructed an auxiliary self-supervised task to improve graph representation learning. Experiments show that this method can effectively solve the long-tail recommendation problem.

In this paper, we model the traditional stance prediction problem as a recommendation problem, and propose a self-supervised heterogeneous graph contrastive learning method to predict the user's stance. Our model can be easily extended to other recommendation applications.

### 3. Model

In this section, we first formulate our task and then introduce the details of our proposed model.

#### 3.1 Task Definition

We use CreateDebate as our data source, which is a widely used public online debate forum. As shown in Fig. 1, given a topic in CreateDebate, it contains a number of debate issues, where each debate issue describes a specific question

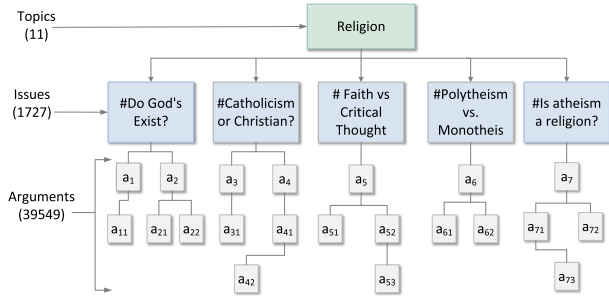


Fig. 1 Structure of the CreateDebate corpus.

such as “Catholicism or Christian?”. For each issue, users express their views by publishing arguments, such as “I am Christian, not anti catholic, just based off what I’ve learned, Catholicism is not right.” The task of stance prediction is to predict the stance (support, oppose or neutral) of a user  $i$  on a target debate issue  $j$  before the user expresses his opinions.

We propose a Heterogeneous Graph Contrastive Learning model for Stance Prediction (HGCL-SP) on online debate forums. The overall framework of HGCL-SP is shown in Fig. 2. It consists of four main parts: (1) HIN Construction and Data Augmentation, (2) Meta-path Guided Graph Encoder, (3) Contrastive Learning, and (4) Matrix Factorization for Stance Prediction.

#### 3.2 HIN Construction

*Heterogeneous Information Network*  $\mathcal{G} = (\mathcal{V}, \mathcal{E})$  is a directed graph consisting of a node set  $\mathcal{V}$  and an edge set  $\mathcal{E}$ .  $\mathcal{G}$  is related to a node type mapping function  $\phi: \mathcal{V} \rightarrow \mathcal{N}$  and an edge type mapping function  $\varphi: \mathcal{E} \rightarrow \mathcal{R}$ .  $\mathcal{N}$  and  $\mathcal{R}$  represent sets of node and edge types, where  $|\mathcal{N} + \mathcal{R}| > 2$ .

We define a HIN with four types of entities: user (U), topic (T), issue (I), and argument (A). The following five heterogeneous relationships are considered to enrich the user and issue information:

- **R1:** To model the relation between a user and his participated issues, we construct the *user-participate-issue* matrix  $\mathbf{A}_1$ , where each element  $t_{i,j} \in \{0, 1\}$  represents if user  $i$  participates in issue  $j$ .
- **R2:** We construct a *user-post-argument* matrix  $\mathbf{A}_2$  to model the interaction between a user and his posted arguments, where each element  $s_{i,j} \in \{0, 1\}$  indicates whether argument  $j$  is posted by user  $i$ .
- **R3:** To model the relationship between a user and his

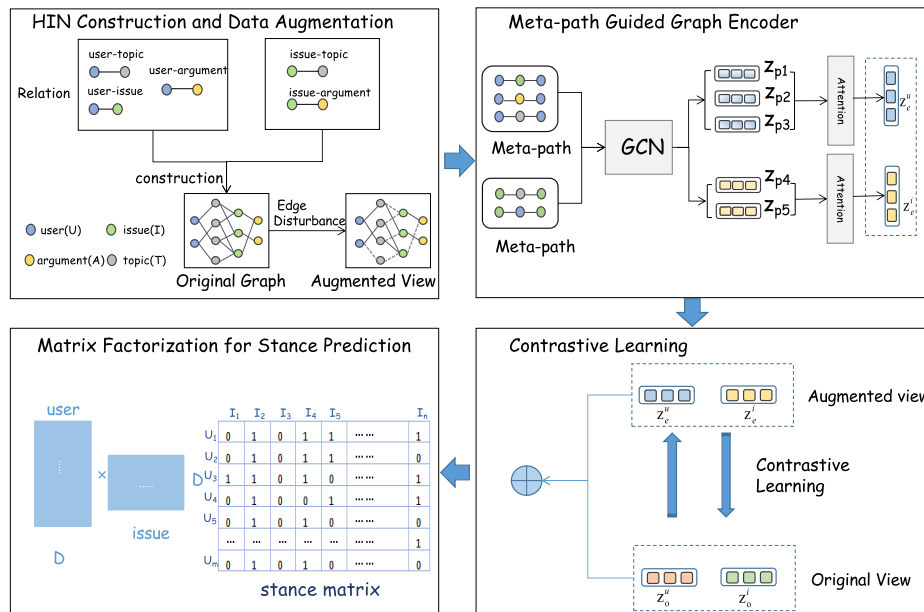


Fig. 2 The overall architecture of our proposed model HGCL-SP.

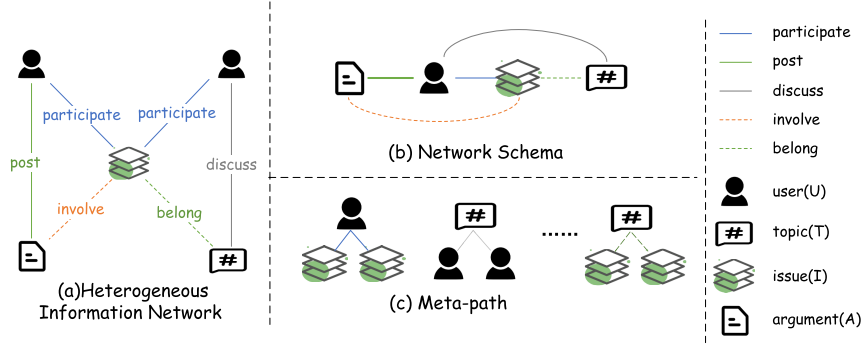


Fig. 3 An example of HIN and the description of the corresponding network schema and meta-path.

discussed topics, we construct a *user-discuss-topic* matrix  $A_3$ , where each element  $d_{i,j} \in \{0, 1\}$ . It represents if user  $i$  participates in the discussion of topic  $j$ .

- **R4:** To model the relationship between an issue and its topic, we construct an *issue-belong-topic* matrix  $A_4$ , where each element  $b_{i,j} \in \{0, 1\}$  denotes whether issue  $i$  belongs to topic  $j$ .
- **R5:** We construct an *issue-involve-argument* matrix  $A_5$  to model the relationship between an issue and an argument, where each element  $l_{i,j} \in \{0, 1\}$  denotes whether issue  $i$  involves argument  $j$ .

Based on the above relationships, we construct a HIN as the original graph  $\mathcal{G}_o = (\mathcal{V}, \mathcal{E})$  to learn the representation of the users and issues. In our case, there are four types of entity (i.e., user, issue, argument and topic) and five types of relationship as aforementioned in  $\mathcal{G}_o$ , as shown in Fig. 3.

### 3.3 Data Augmentation

The core idea of contrastive learning is to get high-quality representation of entities by making similar samples closer and different samples far away from each other. Inspired by the recent success of contrastive learning, some recent approaches leverage contrastive learning losses to learn node representations by maximizing the similarity between two randomly perturbed versions of the intrinsic features and link structure of the same node's local subgraph [27]. Therefore, we first perform data augmentation to generate multiple views, and then conduct contrastive learning based on the generated representation. In order to improve the effectiveness of graph data augmentation, we propose a joint *Edge Disturbance* method, which includes *Edge Drop* and *Edge Add*. For *Edge Drop*, some edges in  $\mathcal{G}_o$  are randomly discarded at a certain dropout ratio  $r$ . While for *Edge Add*, some edges are randomly increased with a certain ratio. We first perform the *Edge Drop* on the original graph, and then add edges randomly. Experiments show that the proposed *Edge Disturbance* method is the most effective.

### 3.4 Meta-Path Guided Graph Encoder

A meta-path  $P$ , in the form of  $N_1 \xrightarrow{R_1} N_2 \xrightarrow{R_2} \dots \xrightarrow{R_l} N_{l+1}$ ,

which describes the composite relationship  $R = R_1 \circ R_2 \circ \dots \circ R_l$  between node  $N_1$  and  $N_{l+1}$ , where  $\circ$  represents the composite operator acting on relations.

To effectively learn the potential interactions of entities and capture the high-order structure, we select several meta-paths for users and issues respectively. Typical meta-paths describing the interaction between users are as follows: the meta-path  $user \xrightarrow{post} argument \xrightarrow{involved} issue \xrightarrow{involved^{-1}}$   $argument \xrightarrow{post^{-1}} user$  means if two users post their arguments on the same issue, they are similar. The meta-path  $user \xrightarrow{participate} issue \xrightarrow{participate^{-1}} user$  denotes that two users can be related when they participate in the same issue. The meta-path  $user \xrightarrow{discuss} topic \xrightarrow{discuss^{-1}} user$  means two users are related if they have participated in discussions of the same topic.

We also select two meta-paths to describe the relationship between issues:  $issue \xrightarrow{participated} user \xrightarrow{participated^{-1}} issue$  and  $issue \xrightarrow{belong} topic \xrightarrow{belong^{-1}} issue$ . These two meta-paths indicate that two issues are related if they have been participated by the same user or belong to the same topic.

After selecting the meta-paths, we use two GCN models to learn representation of the original graph  $\mathcal{G}_o$  and the augmented graph  $\mathcal{G}_e$ . Taking the original graph  $\mathcal{G}_o$  as an example, we will introduce the model in detail. Given a HIN  $\mathcal{G}_o = (\mathcal{V}, \mathcal{E})$  with a group of meta-paths  $\mathcal{P} = \{P_1, P_2, \dots, P_M\}$  as well as the corresponding adjacency matrices  $\mathcal{A} = \{A_1, A_2, \dots, A_M\}$  ( $M$  represents the number of meta-paths), a multi-layer GCN is used to generate the representation of user based on each meta-path, as shown below:

$$\mathbf{z}_{p_i}^{(l+1)} = \sigma \left( \mathbf{D}_{p_i}^{-\frac{1}{2}} \left( \mathbf{A}_{p_i} + \mathbf{I}_{p_i} \right) \mathbf{D}_{p_i}^{-\frac{1}{2}} \mathbf{z}_{p_i}^{(l)} \mathbf{W}_{p_i}^{(l)} \right) \quad (1)$$

where  $\sigma(x)$  is a rectified linear activation function defined as  $RELU(x) = \max(0, x)$ .  $\mathbf{z}_{p_i}^{(l+1)}$  denotes the  $(l+1)$ -Layer the representation of user. Here  $p_i$  means the  $i$ -th meta-path.  $\mathbf{z}^{(0)}$  is the content feature. Specifically, for the content feature of issue, we use Word2vector to get the semantic embedding of the issue. We also use the similar method to generate the content feature of the user.  $\mathbf{A}_{p_i}$  and  $\mathbf{I}_{p_i}$  represent the

adjacency matrix and the corresponding identity matrix of the meta-path  $p_i$ , respectively.  $\mathbf{D}_{p_i}$  is a degree matrix, and only diagonal elements have numerical values. This matrix is used to normalize the adjacency matrix.  $\mathbf{W}_{p_i}^{(l)} \in \mathbb{R}^{d_{in} \times d_{out}}$  is the weight matrix of layer  $l$  under the guidance of meta-path  $p_i$ .  $d_{in}$  and  $d_{out}$  represent the input dimension and output dimension of each layer of GCN respectively.

However, the representation of different meta-paths contribute differently to the final representation of entities. Therefore, we use attention mechanism to solve this problem. Under the guidance of attention mechanism, the representations based on different meta-paths are fused to obtain the final representation under this view:

$$\mathbf{z}_0^u = \sum_{i=1}^M \beta_{p_i} \cdot \mathbf{z}_{p_i} \quad (2)$$

where  $\mathbf{z}_0^u$  is the final representation of users under this view.  $\beta_{p_i}$  denotes the weight of the representation from different meta-paths, it is calculated as follows:

$$\beta_{p_i} = \frac{\exp(\sigma(\mathbf{a}\mathbf{z}_{p_i}))}{\sum_{j \in M} \exp(\sigma(\mathbf{a}\mathbf{z}_{p_j}))} \quad (3)$$

$\mathbf{z}_{p_i}$  is the representation of user learned under the meta-path  $p_i$ .  $\mathbf{a}$  is an attention vector. For the augmented graph  $\mathcal{G}_e$ , we can also obtain the representation of users  $\mathbf{z}_e^u$  and the representation of debate issues  $\mathbf{z}_e^i$ .

### 3.5 Contrastive Learning

After obtaining the representations of the same node from the original graph and the augmented graph, we leverage a contrastive learning loss to maximize agreement between them. Specifically, we treat the representations under different views of the same node as positive samples and the representations of other nodes as negative samples. Inspired by [26], we maximize the agreement between positive samples and minimize the similarity between negative samples. The contrastive loss of user nodes can be defined as follows:

$$\mathcal{L}_u = \sum_{u \in m} -\log \frac{\exp(\text{sim}(\mathbf{z}_0^u, \mathbf{z}_e^u)/\tau)}{\sum_{v \in m} \exp(\text{sim}(\mathbf{z}_0^u, \mathbf{z}_e^v)/\tau)} \quad (4)$$

where  $\text{sim}(x, y)$  represents the cosine similarity between the two vectors  $x$  and  $y$ , and  $\tau$  denotes a temperature parameter. Similarly, we obtain the contrastive loss  $\mathcal{L}_i$  of debate issues.

### 3.6 Matrix Factorization for Stance Prediction

Based on the representation of users and issues learned from previous step, we utilize matrix factorization, a widely used recommendation method to perform the stance prediction task.  $\hat{s}_{u,i}$  is user  $i$ 's stance (support or oppose) on issue  $j$ , and can be defined as follows:

$$\hat{s}_{u,i} = \alpha_u \cdot \mathbf{z}_0^u \mathbf{t}^i + \alpha_i \cdot \mathbf{t}^u \mathbf{z}_0^i + \mathbf{v}_u \mathbf{v}_i^T \quad (5)$$

where  $\mathbf{v}_u \in \mathbb{R}^{m \times D}$  and  $\mathbf{v}_i \in \mathbb{R}^{n \times D}$  are the latent factor of

users and issues, respectively.  $n$  and  $m$  are the number of the issues and users.  $D$  is the dimension of the latent factors.  $\mathbf{z}_0^u$  and  $\mathbf{z}_0^i$  are the representation of users and issues, respectively.  $\mathbf{t}^i \in \mathbb{R}^{O_a \times n}$  and  $\mathbf{t}^u \in \mathbb{R}^{m \times O_a}$  are transformation matrices that force  $\mathbf{z}_0^u$  and  $\mathbf{z}_0^i$  to the same space.  $O_a$  represents dimension of the attention vectors.  $\alpha_u$  and  $\alpha_i$  are tuning parameters.

In the setting of our problem, we need to predict the user's stance towards an issue is to support or oppose. In order to utilize the implicit feedback (user and issue pairs without interaction), stance  $s_{ij} = 0$  can be explained in two ways: user  $i$  is opposed to issue  $j$ , or user  $i$  has not expressed any stance on issue  $j$ . We introduce a confidence parameter  $c_{ij}$  to pose on the negative samples that the user and the issue have no interactions in the recommendation loss [28]. The larger  $c_{ij}$  is, the more we believe in  $s_{ij}$ :

$$c_{ij} = \begin{cases} a & s_{ij} = 1 \\ b & s_{ij} = 0 \end{cases} \quad (6)$$

where  $a$  and  $b$  are tuning parameters contented  $a > b > 0$ .

The object function of the MF part can be defined as:

$$\mathcal{L}_{MF} = (\mathbf{s}_{u,i} - \hat{s}_{u,i})^2 \cdot c_{ij} + \lambda(\|\mathbf{v}_u\|_2 + \|\mathbf{v}_i\|_2) \quad (7)$$

where  $\lambda$  denotes the regularization parameter.

Finally, we combine the loss of MF and the loss of contrastive learning to obtain the final objective function of our model:

$$\mathcal{L}_{final} = \mathcal{L}_{MF} + \mathcal{L}_u + \mathcal{L}_i \quad (8)$$

where  $\mathcal{L}_u$  and  $\mathcal{L}_i$  are the contrastive losses of users and issues respectively.

## 4. Experiments

### 4.1 Dataset

The dataset we use is from Qiu's work [19]. The data was crawled from the CreateDebate website, involving 4994 users and 11 topics. Each topic contains a number of issues. The debate issue describes a specific question such as "Does God exist?". The stances of these issues are usually divided into two categories, which are "support" or "oppose". In addition, each debate issue contains a number of arguments posted by users. Each argument is written by a user towards an issue, such as "I Well, I need to say that modern technology has exceeded some parts of the bible, coran, or other religious books." It can be either a separate post or a reply to the previous arguments. Table 1 shows the statistics of the CreateDebate dataset.

### 4.2 Baselines

As mentioned above, stance prediction problem can be regarded as a recommendation problem in a certain sense, we compare our proposed model with the following advanced recommendation methods:



**Table 1** Statistics of the CreateDebate dataset.

Type	Number	Description
# users	4,994	On average, each user has 8 arguments.
# topics	11	Religion, politics, technology and other topics.
# issues	1,727	On average, each topic contains 157 issues.
# arguments	39,549	An average of 23 arguments per topic.
# user stances	17,843	Total number of stances posted by users.
# support stances	8,601	Number of support stances posted by users.
# oppose stances	9,242	Number of oppose stances posted by users.

**Table 2** Results of our model and the baselines. \* indicates that the result is better than the method in the previous column at 5% significance level by Wilcoxon signed-rank test.

	Methods	$F1_{support}$	$F1_{oppose}$	$F1_{avg}$
Methods based on collaborative filtering	PMF	0.4694	0.5516	0.5105
	HFT	0.5238	0.5595	0.5417
	FM	0.6766	0.6410	0.6588
Methods based on deep learning	DCN	0.7322	0.6629	0.6976
	DeepFM	0.7430	0.6706	0.7068
	IGMC	0.6434	0.8015	0.7225
Our method	<b>HGCL-SP</b>	<b>0.8365*</b>	<b>0.8064*</b>	<b>0.8214*</b>

- **PMF**: Probabilistic matrix factorization [29]. The PMF model is widely used in ratings at first, which has a linear relationship with the number of observed values and works well on large, sparse and unbalanced datasets. We modified it to adapt to the one-class collaborative recommendation.
- **HFT**: Hidden factors and hidden topics models [30]. HFT performs numerical ratings by applying exponential transformation function to matrix factorization model. It combines the content filtering and the collaborative filtering by assuming that arguments are generated by user factors or issue factors.
- **FM**: Factorization Machine [1]. It has the advantages of both factorization models and Support Vector Machines (SVM). In stance prediction, FM is applied to model user preferences toward issues based on the historical interactions of stance arguments and users.
- **DCN**: Deep Cross Network [31], which can deal with a large number of dense and sparse features. The cross network makes up multiple layers, and in each layer, the interactions of the previous layer are retained to obtain high-order interactions.
- **DeepFM**: A neural network based on the combination of factorization machines and deep neural networks can model feature interaction from high-order and low-order aspects respectively [32].
- **IGMC**: An inductive matrix completion method independent of auxiliary information can effectively alleviate data sparsity [33]. This model uses graph neural networks to model the relationship between nodes.

#### 4.3 Experiment Settings

In our experiment, we divide the data set into training set, validation set and test set with a ratio of 80%, 10% and 10%. We tuned all parameters in the validation set and report the results on the test set. We use GCN with three layers to gen-

erate the node representations. The dimensions of the three layers are set to 256, 128 and 64 respectively. For the number of user and issue latent factors, we tested several values in validation set and found that when it is equal to 50, the result is the best. We mainly use Precision, Recall, F1-score as the evaluation metrics. Based on the Recall and Precision, the harmonic average F1-score is calculated. Firstly, we can get the values of  $F1_{support}$  and  $F1_{oppose}$  by the following formulas, and then take the average  $F1_{avg}$  of them as the final result of the model.

$$F1_{support} = \frac{2 * Precision_{support} * Recall_{support}}{Precision_{support} + Recall_{support}} \quad (9)$$

$$F1_{oppose} = \frac{2 * Precision_{oppose} * Recall_{oppose}}{Precision_{oppose} + Recall_{oppose}} \quad (10)$$

$$F1_{avg} = \frac{F1_{support} + F1_{oppose}}{2} \quad (11)$$

#### 4.4 Experiment Results

In this part, we compare the results of our proposed model HGCL-SP with the baseline methods. From Table 2, we can come to the following conclusions: (i) Compared with other methods, the performance of collaborative filtering based methods PMF, HFT and FM are worse. Among the three methods, FM achieves the best results because of the integration of a variety of complex features. (ii) The results of DCN and DeepFM are significantly better than that of collaborative filtering based methods, demonstrating that deep learning based methods can effectively capture the deep interaction between users and issues. (iii) Among these methods, HGCL-SP has the best performance. Specifically, compared with the state-of-the-art recommendation model IGMC, HGCL-SP improves nearly 9.89% under the evaluation metric of  $F1_{avg}$ . All the above results show that HGCL-SP can effectively identify potential interactions and understand user preferences, by introducing heterogeneous rela-

**Table 4** Results of different meta-paths combinations.

Combination number	Combination mode	$F1_{support}$	$F1_{oppose}$	$F1_{avg}$
Combination of two meta-paths	$P_1 \& P_4$	0.7041	0.6881	0.6961
	$P_1 \& P_5$	0.6542	0.6611	0.6577
	$P_2 \& P_4$	0.7230	0.7001	0.7115
	$P_2 \& P_5$	0.6836	0.6754	0.6795
	$P_3 \& P_4$	0.6652	0.6670	0.6661
Combination of three meta-paths	$P_3 \& P_5$	0.6674	0.6660	0.6667
	$P_1 \& P_2 \& P_4$	0.7206	0.6944	0.7075
	$P_1 \& P_2 \& P_5$	0.7212	0.6861	0.7037
	$P_1 \& P_3 \& P_4$	0.7310	0.6888	0.7099
	$P_1 \& P_3 \& P_5$	0.6981	0.6635	0.6808
	$P_1 \& P_4 \& P_5$	0.7327	0.7235	0.7281
	$P_2 \& P_3 \& P_4$	0.7420	0.6947	0.7184
	$P_2 \& P_3 \& P_5$	0.7005	0.6579	0.6792
Combination of four meta-paths	$P_2 \& P_4 \& P_5$	0.7579	0.7367	0.7473
	$P_3 \& P_4 \& P_5$	0.7387	0.7286	0.7337
	$P_1 \& P_2 \& P_3 \& P_4$	0.7767	0.7355	0.7561
	$P_1 \& P_2 \& P_3 \& P_5$	0.7468	0.7117	0.7293
	$P_1 \& P_2 \& P_4 \& P_5$	0.8171	0.7882	0.8026
Our combination mode	$P_1 \& P_3 \& P_4 \& P_5$	0.8046	0.7724	0.7885
	$P_2 \& P_3 \& P_4 \& P_5$	0.8016	0.7658	0.7837
	$P_1 \& P_2 \& P_3 \& P_4 \& P_5$	<b>0.8365</b>	<b>0.8064</b>	<b>0.8214</b>

**Table 3** Comparison of HGCL-SP and the two variants.

Methods	$F1_{support}$	$F1_{oppose}$	$F1_{avg}$
<i>w/o</i> heterogeneous information	0.7230	0.7001	0.7115
<i>w/o</i> contrastive learning	0.8031	0.7638	0.7834
<b>HGCL-SP</b>	<b>0.8365</b>	<b>0.8064</b>	<b>0.8214</b>

tionships between nodes and obtaining high-quality representations through constrative learning.

#### 4.5 Ablation Experiment

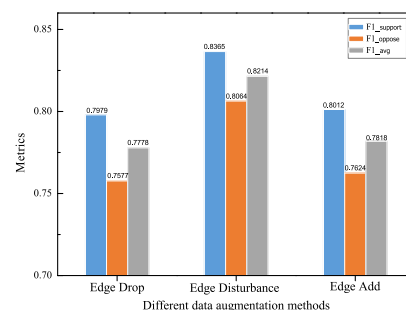
In this section, we get the variant models by removing different parts of HGCL-SP, and compare them with HGCL-SP to get the contribution degree of different parts to the final performance:

- **w/o heterogeneous information:** Our model removes the heterogeneous information and only keeps the interactions of users and issues.
- **w/o contrastive learning:** Our model without data augmentation and contrastive learning.

The comparison results of HGCL-SP and the two variants are shown in Table 3. Compared with the full model HGCL-SP, we find that the results of model *w/o heterogeneous information* drops by about 11% under all the metrics, which indicates that the integration of heterogeneous information network is very important for stance prediction in online debate forums. At the same time, we also find that contrastive learning can learn better representation of users and issues and improve the result of stance prediction.

#### 4.6 Detailed Analysis

In order to analyze the effects of different combination of meta-paths, different data augmentation methods and parameters on the experimental results, we carried out the following experiments.

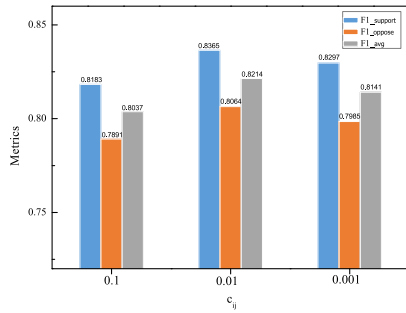
**Fig. 4** Performance of different data augmentation methods.

##### 4.6.1 Different Combination of Meta-Paths

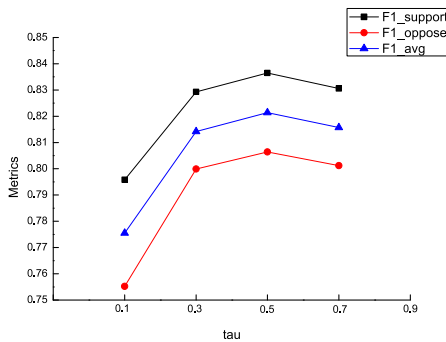
Due to different meta-paths have different semantics, it is important to choose high-quality meta-paths for the experimental results. We analyze the influence of different combinations of meta-paths on the experimental results. Specifically, for users, there are three meta-paths  $P_1: U \rightarrow A \rightarrow I \rightarrow A \rightarrow U$ ,  $P_2: U \rightarrow I \rightarrow U$ ,  $P_3: U \rightarrow T \rightarrow U$ . As for the issue, there are two meta-paths  $P_4: I \rightarrow U \rightarrow I$ ,  $P_5: I \rightarrow T \rightarrow I$ . We randomly select the above meta-paths to combine. As shown in Table 4, we can see that with the increase of the number of meta-paths, the results increase, and the best result is achieved when five meta-paths are combined. At the same time, different meta-paths have different effects on the experimental results. Therefore, it is reasonable for us to use attention mechanism to fuse different meta-paths.

##### 4.6.2 Different Data Augmentation Methods

Data augmentation directly affects the results of graph contrastive learning. Figure 4 shows the results of *Edge Drop*, *Edge Add* and *Edge Disturbance*. We can observe that the



**Fig. 5** Performance of different confidence parameter  $c_{ij}$ .



**Fig. 6** Performance of different temperature  $\tau$ .

*Edge Disturbance* method combined with *Edge Drop* and *Edge Add* achieves the best results. Through experiments, we prove that heterogeneous graph representation learning can be improved by contrastive learning of different views and different data augmentation methods.

#### 4.6.3 Parameter Sensitivity Analysis

Next, we will analyze the effects of different parameters on the experiment. In order to utilize the implicit feedback, we introduce a confidence parameter  $c_{ij}$  to pose on the negative samples that the user and the issue have no interactions in the recommendation loss. We vary the confidence parameter  $c_{ij}$  to 0.1, 0.01 and 0.001 in Fig. 5, and find that when  $c_{ij} = 0.01$ , the results are the best. Temperature  $\tau$  plays an important role in obtaining high-quality negative samples. We set different  $\tau$  values (i.e. 0.1, 0.3, 0.5 and 0.7) for the experiment. As shown in Fig. 6, we can see that with the increase of  $\tau$ , the results first increase and then decrease, and when  $\tau = 0.5$ , the best results are achieved.

## 5. Conclusion

In this paper, we propose a heterogeneous graph contrastive learning model for the task of stance prediction (HGCL-SP) on the online debate forums. HGCL-SP utilizes the rich heterogeneous information of users and correlation between issues to form a HIN and learns the graph representation via a meta-path guided GCN encoder. At the same time, HGCL-SP introduces the contrastive learning mechanism,

which generates an augmented graph view by *Edge Disturbance* and maximizes the agreement of node representations in these two views. Finally, the stance is predicted by the matrix factorization of the users and issues. Through experiments, we prove that the heterogeneous graph representation learning method proposed in this paper can effectively predict user stances, and heterogeneous graph representation learning can be improved by contrastive learning of different views.

In the future, we will try other data augmentation methods, such as counterfactual data augmentation, to enrich the heterogeneous graph learning and improve stance prediction.

## Acknowledgements

We thank the anonymous reviewers for their constructive suggestions. This work was supported by the National Natural Science Foundation of China [61806049] and the Heilongjiang Postdoctoral Science Foundation [LBH-Z20104]. Corresponding author: Yang Li, yli@nefu.edu.cn.

## References

- [1] A. Sasaki, K. Hanawa, N. Okazaki, and K. Inui, "Predicting stances from social media posts using factorization machines," Proceedings of the 27th International Conference on Computational Linguistics, pp.3381–3390, 2018.
- [2] A. AlDayel and W. Magdy, "Stance detection on social media: State of the art and trends," Information Processing & Management, vol.58, no.4, p.102597, 2021.
- [3] U.A. Siddiqua, A.N. Chy, and M. Aono, "Tweet stance detection using multi-kernel convolution and attentive lstm variants," IEICE Trans. Inf. & Syst., vol.102, no.12, pp.2493–2503, 2019.
- [4] Y. Yang, B. Wu, K. Zhao, and W. Guo, "Tweet stance detection: A two-stage dc-bilstm model based on semantic attention," 2020 IEEE Fifth International Conference on Data Science in Cyberspace (DSC), pp.22–29, IEEE, 2020.
- [5] B. Zhang, M. Yang, X. Li, Y. Ye, X. Xu, and K. Dai, "Enhancing cross-target stance detection with transferable semantic-emotion knowledge," Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, pp.3188–3197, 2020.
- [6] Y. Li, Y. Sun, N. Zhu, and W. Zhang, "BERTtoCNN: Similarity-preserving enhanced knowledge distillation for stance detection," Plos one, vol.16, no.9, p.e0257130, 2021.
- [7] A.I. Schein, A. Popescul, L.H. Ungar, and D.M. Pennock, "Methods and metrics for cold-start recommendations," Proceedings of the 25th annual international ACM SIGIR conference on Research and development in information retrieval, pp.253–260, 2002.
- [8] K. Darwish, W. Magdy, A. Rahimi, T. Baldwin, and N. Abokhodair, "Predicting online islamophobic behavior after #parisattacks," The Journal of Web Science, vol.4, no.3, pp.34–52, 2018.
- [9] L. Zhu, Y. He, and D. Zhou, "Neural opinion dynamics model for the prediction of user-level stance dynamics," Information Processing & Management, vol.57, no.2, p.102031, 2020.
- [10] K.Y. Huang, H.H. Huang, and H.H. Chen, "Hargan: heterogeneous argument attention network for persuasiveness prediction," Proceedings of the AAAI Conference on Artificial Intelligence, pp.13045–13054, 2021.
- [11] Y. Sun, J. Han, C.C. Aggarwal, and N.V. Chawla, "When will it happen?: relationship prediction in heterogeneous information networks," Proceedings of the fifth ACM international conference on Web search and data mining, pp.663–672, 2012.



- [12] Y. Sun and J. Han, "Mining heterogeneous information networks: a structural analysis approach," *Acm Sigkdd Explorations Newsletter*, vol.14, no.2, pp.20–28, 2013.
- [13] Q. Sun, Z. Wang, Q. Zhu, and G. Zhou, "Stance detection with hierarchical attention network," *Proceedings of the 27th International Conference on Computational Linguistics*, pp.2399–2409, 2018.
- [14] J. Yuan, Y. Zhao, J. Xu, and B. Qin, "Exploring answer stance detection with recurrent conditional attention," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol.33, pp.7426–7433, 2019.
- [15] S. Ghosh, P. Singhania, S. Singh, K. Rudra, and S. Ghosh, "Stance detection in web and social media: a comparative study," *International Conference of the Cross-Language Evaluation Forum for European Languages*, pp.75–87, Springer, 2019.
- [16] Y. Sun and Y. Li, "Stance detection with knowledge enhanced bert," *CAAI International Conference on Artificial Intelligence*, pp.239–250, Springer, 2021.
- [17] R. Liu, Z. Lin, Y. Tan, and W. Wang, "Enhancing zero-shot and few-shot stance detection with commonsense knowledge graph," *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*, pp.3152–3157, 2021.
- [18] X. Zhang, J. Yuan, Y. Zhao, and B. Qin, "Knowledge enhanced target-aware stance detection on tweets," *China Conference on Knowledge Graph and Semantic Computing*, pp.171–184, Springer, 2021.
- [19] M. Qiu, Y. Sim, N.A. Smith, and J. Jiang, "Modeling user arguments, interactions, and attributes for stance prediction in online debate forums," *Proceedings of the 2015 SIAM international conference on data mining*, pp.855–863, SIAM, 2015.
- [20] R. Ying, R. He, K. Chen, P. Eksombatchai, W.L. Hamilton, and J. Leskovec, "Graph convolutional neural networks for web-scale recommender systems," *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pp.974–983, 2018.
- [21] X. Yu, X. Ren, Y. Sun, Q. Gu, B. Sturt, U. Khandelwal, B. Norick, and J. Han, "Personalized entity recommendation: A heterogeneous information network approach," *Proceedings of the 7th ACM international conference on Web search and data mining*, pp.283–292, 2014.
- [22] J. Gong, S. Wang, J. Wang, W. Feng, H. Peng, J. Tang, and P.S. Yu, "Attentional graph convolutional networks for knowledge concept recommendation in moocs in a heterogeneous view," *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp.79–88, 2020.
- [23] X. Wang, H. Ji, C. Shi, B. Wang, Y. Ye, P. Cui, and P.S. Yu, "Heterogeneous graph attention network," *The World Wide Web Conference*, pp.2022–2032, 2019.
- [24] X. Fu, J. Zhang, Z. Meng, and I. King, "MAGNN: Metapath aggregated graph neural network for heterogeneous graph embedding," *Proceedings of The Web Conference 2020*, pp.2331–2341, 2020.
- [25] Y. Zhu, Y. Xu, F. Yu, Q. Liu, S. Wu, and L. Wang, "Deep graph contrastive representation learning," *arXiv preprint arXiv:2006.04131*, 2020.
- [26] J. Wu, X. Wang, F. Feng, X. He, L. Chen, J. Lian, and X. Xie, "Self-supervised graph learning for recommendation," *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp.726–735, 2021.
- [27] X. Wang, N. Liu, H. Han, and C. Shi, "Self-supervised heterogeneous graph neural network with co-contrastive learning," *Proc. 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*, pp.1726–1736, 2021.
- [28] C. Wang and D.M. Blei, "Collaborative topic modeling for recommending scientific articles," *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp.448–456, 2011.
- [29] R. Salakhutdinov and A. Mnih, "Probabilistic matrix factorization," *Proceedings of the 20th International Conference on Neural Information Processing Systems*, pp.1257–1264, 2007.
- [30] J. McAuley and J. Leskovec, "Hidden factors and hidden topics: understanding rating dimensions with review text," *Proceedings of the 7th ACM conference on Recommender systems*, pp.165–172, 2013.
- [31] R. Wang, B. Fu, G. Fu, and M. Wang, "Deep & cross network for ad click predictions," in *Proceedings of the ADKDD'17*, pp.1–7, 2017.
- [32] H. Guo, R. Tang, Y. Ye, Z. Li, and X. He, "DeepFM: a factorization-machine based neural network for ctr prediction," *Proceedings of the 26th International Joint Conference on Artificial Intelligence*, pp.1725–1731, 2017.
- [33] M. Zhang and Y. Chen, "Inductive matrix completion based on graph neural networks," *International Conference on Learning Representations*, 2020.



**Yang Li** received her Ph.D. degree in July 2017 from the Department of Computer Science, Harbin Institute of Technology, Harbin, China. She is an associate professor in the College of Information and Computer Engineering at the Northeast Forestry University. Her current research interests include natural language processing, information retrieval, and social media analysis.



**Rui Qi** is currently pursuing a master's degree in electronic information with Northeast Forestry University in China. Her current research interests include natural language processing and stance detection.