Event Feature Pre-training Model Based on Public Opinion Evolution

WANG Nan[1], TAN Shu-Ru[2], XIE Xiao-Lan[3] , LI Hai-Rong[4], JIANG Jia-Hui[5]
School of Management Science and Information Engineering, Jilin University of Finance and Economics, Changchun, China[1, 5]
School of Information Science and Engineering, Guilin University of Technology, Guilin, China[2, 3]
School of Information Engineering, Xinjiang Institute of Technology, Xin Jiang, China[4]

***Abstract*—The comments in the evolution of network public opinion events not only reflect the attitude of netizens towards the event itself, but also are the key basis for mastering the dynamics of public opinion. According to the comment data in the event evolution process, an event feature vector pre-training model NL2ER-Transformer is constructed to realize the real-time automatic extraction of event features. Firstly, a semi-supervised multi-label curriculum learning model is proposed to generate comment words, event word vectors, event words, and event sentences, so that a public opinion event is mapped into a sequence similar to vectorized natural language. Secondly, based on the Transformer structure, a training method is proposed to simulate the evolution process of events, so that the event vector generation model can learn the evolution law and the characteristics of reversal events. Finally, the event vectors generated by the presented NL2ER-Transformer model are compared with the event vectors generated by the current mainstream models such as XLNet and RoBerta. This paper tests the pre-trained model NL2ER-Transformer and three pre-trained benchmark models on four downstream classification models. The experimental results show that using the vectors generated by NL2ER-Transformer to train downstream models compared to using the vectors generated by other pre-trained benchmark models to train downstream models, the accuracy, recall, and F1 values are 16.66%, 44.44%, and 19% higher than the best downstream model. At the same time, in the evolutionary capability analysis test, only four events show partial errors. In terms of performance of semi-supervised model, the proposed semi-supervised multi-label curriculum learning model outperforms mainstream models in four indicators by 6%, 23%, 8%, and 15%, respectively.**

***Keywords*—Event vectorization; NL2ER-transformer model; public opinion reversal prediction; evolution of public opinion event; multi label semi supervised learning**

I. Introduction

Network public opinion refers to the public's opinions on an event on the network media. When an event causes heated discussion among netizens and forms a certain scale of network public opinion, this work call it a network public opinion event, hereinafter also referred to as an event. Network public opinion often leads to various public opinion phenomena, such as public opinion reversal, network violence, secondary derived events, etc. Network public opinion prediction from the perspective of public opinion phenomenon recognition is one of the current research directions of public opinion prediction: by analyzing the influencing factors of public opinion phenomena, designing event feature vectors, and further constructing machine learning classification models to predict various phenomena of public opinion events. This paper focuses on the upstream event pre-training model, while the downstream task takes the prediction of public opinion reversal [1] as an example to verify the model.

This paper maps the public opinion event vector construction problem to the pre-train model based natural language feature vector generation problem. Based on the Transformer structure, we build an event vector generation model NL2ER-Transformer (Natural Language to Event Representation from Transformer) with the ability to predict the event evolution. The generated event vector is more in line with the evolution characteristics of public opinion reversal event. It solves the problem that the current event vectorization models only consider the static subjective characteristics of events and have not learned the dynamic evolution characteristics. The specific contributions of this paper are summarized as follows:

- Combining manually designing features and event evolution analysis, we put forward the concepts of comment word, event word and event sentence. Based on the three concepts, this work transforms an event into a sequence, which is similar to vectorized natural language. This kind of representation form of event can enable the model to better learn the evolution characteristics of the event. In order to obtain comment words automatically, a semi-supervised multi-label curriculum learning model is proposed, which can automatically convert comments into discriminative feature vectors, and then construct event vectors based on them.
- Based on the sequence data proposed in this paper, by the training mode of simulating event evolution, an event vector generation model NL2ER-Transformer with event evolution prediction ability is constructed based on Transformer. It solves the problem that the current event vectorization models only consider the static subjective characteristics of events and have not learned the dynamic evolution characteristics. This paper has carried out several experiments to verify the effectiveness of the model. The experimental results show that using the event vectors constructed by NL2ER-Transformer to train the public opinion reversal prediction model are better than the classifiers trained

by the event vectors constructed by other event vectorization models.

II. Related Work

Feature extraction of event representation vector is the most important step to build a network public opinion prediction model based on machine learning. At present, the feature extraction of event vectors mainly includes three methods: manually designing features, pre-training model and knowledge graph.

A. Manually Designing Features

At first, scholars generated event vectors by manually defining or extracting qualitative and quantitative event features from social media. Research [2] uses 33 artificial features to assign values to public opinion events; study [3] made a secondary improvement on the basis of [2], reducing 33 features to 30 features; author in [4] combines Twitter and Weibo, and proposes some characteristics to analyze public opinion based on their commonalities; researcher in [5] combines time features and proposes new event features for rumor analysis. The manually defined event features are very dependent on expert knowledge which usually can only be obtained by observing the influencing factors of one period or even the whole period of the event. Although the manually defined event features contain some event evolution information and the prediction model based on them can obtain better prediction accuracy, this method requires that the event has evolved for a period of time before the event features are obtained, and it is difficult to determine the event features at the early stage of the event [6][7]. Therefore, the model trained with manually designed event features has poor prediction ability for events with short occurrence time, and the prediction accuracy is not high before the occurrence of public opinion phenomena such as public opinion reversal.

B. Pre-Training Model

With the continuous development of computer hardware technology, people have made great progress in natural language vectorization, from the initial word2vec [8]-[10] to ELMo [8] and GPT [11] with context awareness as the core. With the proposal of Transformer [12] in 2017, natural language vectorization has made amazing progress again, and language transfer models have begun to flourish. At the same time, the vectorization of public opinion events, which is closely related to natural language, has also been correspondingly improved. For example, studies [13] and [14] transform event description information into event vectors based on Transformer structure; Mohammadreza Samadi[15]used RoBERTa[16], XLNET[17], BERT[18] and other transfer learning models to transform events into vectors, and tested them on multiple groups of classification models; SZU Yin Lin[19] and Guillermo Blanco[20] used Bert to transform event information into vectors for subsequent prediction. Such methods are the mainstream methods for vectorization of news events or public opinion events at present, but they only vectorize according to the event description text, and do not consider the dynamic factors of event evolution, such as public sentiment and attitude in the manually designing features. The method generated event vectors do not have enough information to express the change and development of events and it is difficult to accurately predict the reversal phenomenon in the process of public opinion evolution.

C. Knowledge Graph

As the mainstream method of public opinion research, knowledge graph has also been used by many scholars to construct public opinion event vectors. For example, research [21] proposed a method to construct a knowledge graph of public opinion with online news comments, which is used for the decision-maker master the online public opinion quickly and directly; study [22] constructed an NPOKG based knowledge graph for extracting features of public opinion events; study [23] based on ELECTRA and REDP methods, entity extraction and relationship extraction are performed on public opinion text information respectively. Network public opinion knowledge graphs are constructed for Weibo platform and short video platform, and comparative analysis is conducted on each network public opinion knowledge graph. However, from the experimental results of the research, it can be seen that the knowledge items contained in Internet public opinion are relatively sparse, and a large number of entities extracted through text can only construct a very small number of nodes, so it is not suitable to be used as a vectorization means of public opinion events.

In view of the problems of the above event vectorization methods, on the basis of the influence factors of the reversal event and dynamic temporality of public opinion evolution, this paper maps the public opinion event vector construction problem to the pre-train model based natural language feature vector generation problem. Based on the Transformer structure, we build an event vector generation model NL2ER-Transformer (Natural Language to Event Representation from Transformer) with the ability to predict the event evolution. The generated event vector is more in line with the evolution characteristics of public opinion reversal event. The classifier can accurately predict the events that may be reversed only relying on the information before the event reversal.

III. Event Vectorization and Training Method of Simulating Event Evolution

In this section, we show the model for event vectorization and training method of simulating event evolution. The model consists of two parts. The first part transforms an abstract event that has occurred for *m* days into a sequence based on netizens' comments, by defining comment word (automatically obtained through the proposed semi-supervised multi-label curriculum learning model in Section Ⅲ B), event word vector, event word and event sentence; The second part uses the proposed NL2ER-Transformer model to generate event vector that contain the characteristics of the event itself and the evolution characteristics of the event. The model framework is shown in Fig. 1.

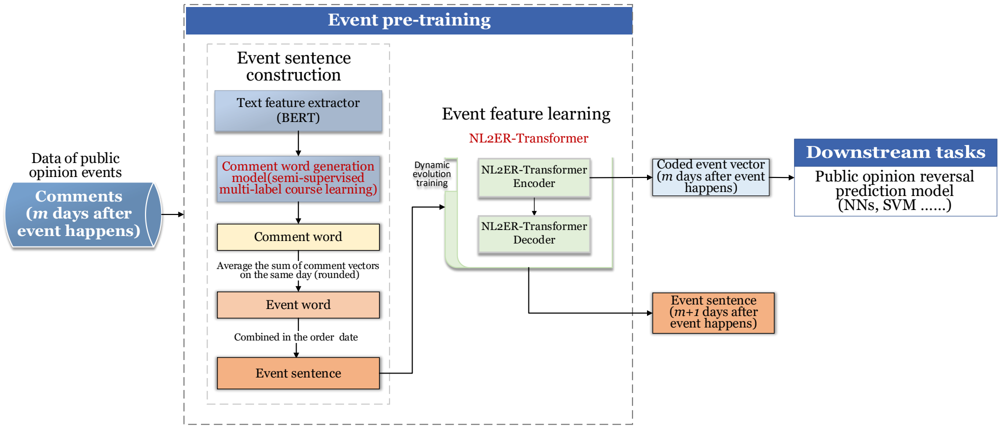

Fig. 1. Model framework.

A. Mapping Public Opinion Event into Sequence Like vectorized Natural Language

Definition 1. Comment word

A comment word is a vector constructed by the corresponding comment text based on the characteristics of it. The specific form is shown in Formula (1).

$$C = [c_1, c_2, \dots c_7] \qquad c_i \in \{0,1\}\ i \in \{0,7\} \tag{1}$$

We use vector C to describe comments and c_i is the ith feature. The value of c_i and design criteria will be discussed in Section III.1(a).

Definition 2. Event word vector

The event word vector Z_j represents the characteristics of an event on a certain day. The specific construction method is shown in formula (2)-(5).

$$\bar{C} = [\bar{c}_1, \bar{c}_2, \dots \bar{c}_7] \qquad \bar{c}_i \in \{0,1\}(1 \le i \le 7) \tag{2}$$

$$Z_j = [z_1, z_2, \dots z_7] \tag{3}$$

$$\bar{c}_i = \frac{\sum_{t=1}^{k_j} c_i^{(t)}}{k_j} \tag{4}$$

$$z_i = \begin{cases} 1 & \bar{c}_i > \omega \\ 0 & \bar{c}_i \le \omega \end{cases} \quad (1 \le i \le 7, 0 \le \omega \le 1) \tag{5}$$

Where k_j is the total number of comments collected on jth day of the event, $\bar{c}_i (1 \le i \le m)$ is the average value of the ith feature of all comments collected on the jth day, ω is a pre-defined threshold.

Definition 3. Event word

An event word is a decimal value corresponding to the event word vector.

The mapping relationship between an event word vector and an event word is represented by the event word dictionary in Table II. The construction process of a single event word is shown in Fig. 2.

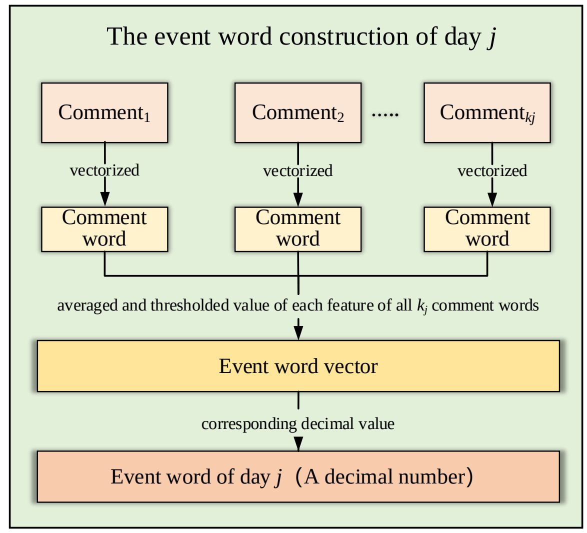

Fig. 2. Construction process of an event word for jth day.

Definition 4. Event sentence

Event sentence is a data form similar to vectorized natural language data, which is composed of event words according to the chronological order of events. The form of an event sentence is presented in Fig. 3.

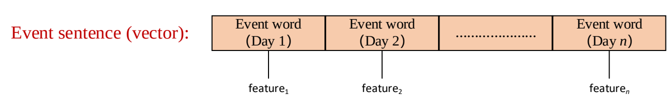

Fig. 3. The form of an event sentence.

According to definitions 1 to 4, we can interpret all the comments on the j day of an event as comment words one by one, construct the event word vector of that day with the average value of all the comment words of that day, map the event word vector to the event word of that day, and then arrange all the event words in order according to the time sequence of the event, so as to obtain the event sentence from the start of the event to the current time point.

Therefore, a public opinion event is represented as a description sentence (event sentence) with a certain meaning. The event words of each day in the event process constitute the vocabulary in the sentence, while the evolution time point of the public opinion event where the comment is located determines the grammatical order structure of the sentence. The event sentence obtained contains various specific information of the public opinion event from its start to the current time point including event evolution and change information.

1) Characteristics of comment words and construction of event word dictionary

a) Design of features about comment words: The comments are personal opinions that expressed by netizens on the situation occurred on the day of an event, usually in the form of short text or information combined with pictures and texts. In natural language, the number of words is limited, so each word can be uniquely encoded by constructing a dictionary, and sentences can be constructed using the codes of words. This paper regards an event as a sentence and the comments as words in the sentence. However, because the different lengths of comment texts are too complicated, and each comment is a unique mind expressed by different people, it is difficult to use comments as words to construct an upper-limited "dictionary" in natural language.

The comments not only reflect the current progress of the event, but also includes the netizens attitudes. It is including the deep excavation, questioning and disclosure of the event. Additionally, in many cases the comments will affect the evolution direction of the event and even reverse public opinion. Based on the important influencing factors of public opinion reversal analyzed in [2][3], this paper designs seven discriminative features for comments, and transforms text comments into comment words at a relatively abstract level. The features and values of a comment word are shown in Table I.

TABLE I. FEATURE DESCRIPTION AND VALUE OF COMMENT WORD

Feature name	**Feature description**	**Value**
c_1	Whether the comment is mixed with some kind of social emotion	Yes 1, NO 0
c_2	Whether the comment's inclination is positive	Yes 1, NO 0
c_3	Whether the comment is in a declarative tone	Yes 1, NO 0
c_4	Whether the comment is cyberbullying	Yes 1, NO 0
c_6	Is the comment stereotypical	Yes 1, NO 0
c_7	Whethdeer the comment has agenda setting	Yes 1, NO 0
c_8	Whether there are emoticons and special characters in the comment	Yes 1, NO 0

2) Event word dictionary and event sentence: According to the 7-dimensional comment word vector designed in Table I, after averaging and thresholding the comment words, the event word vector can be obtained. An event word dictionary composed of event word vector and its corresponding decimal value is constructed, which is shown in Table II (all comment words are also derived from this dictionary).

TABLE II. EVENT WORD DICTIONARY

Features of event word vector / Decimal value	**Whether the comment is mixed with some kind of social emotion**	**Whether the comment's inclination is positive**	**Whether the comment is in a declarative tone**	**Whether the comment is cyberbullying**	**Is the comment stereotypical**	**Whether the comment has agenda setting**	**Whether there are emoticons and special characters in the comment**
0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	1
……	……	……	……	……	……	……	……
127	1	1	1	1	1	1	1

3) A model of comment word generation based on semi-supervised multi-Label curriculum learning: In order to automatically generate the vector representation of a comment in the NL2ER-Transformer model, this paper designs a semi-supervised multi-label curriculum learning model, the structure is shown in Fig. 4.

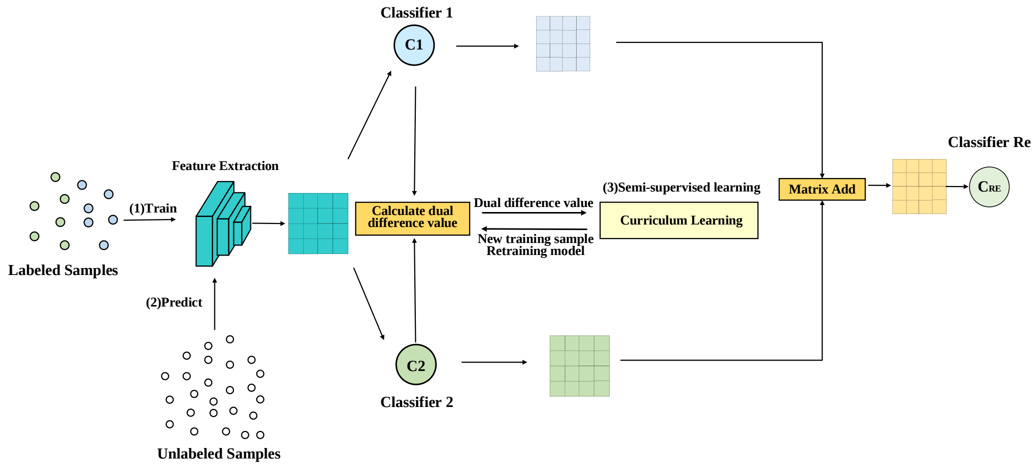

Fig. 4. Structure of the semi-supervised multi-label curriculum learning model.

The model contains four neural networks, one feature extractor $E(.)$, two multi label classifiers $C_1(.)$,$C_2(.)$, and one final classification $Cre(.)$. Specifically, $E(.)$ extracts features from samples; $C_1(.)$ and $C_2(.)$ are used to obtain the label prediction results(i.e., comment words); the label relation is abstracted by adding $C_2(.)$ and $C_1(.)$, and finally classify the samples through $Cre(.)$. During the curriculum learning, difference of $C_1(.)$ and $C_2(.)$ can be used to screen pseudo labeled samples added to training samples. The part of the semi-supervised curriculum learning is shown in Fig. 5.

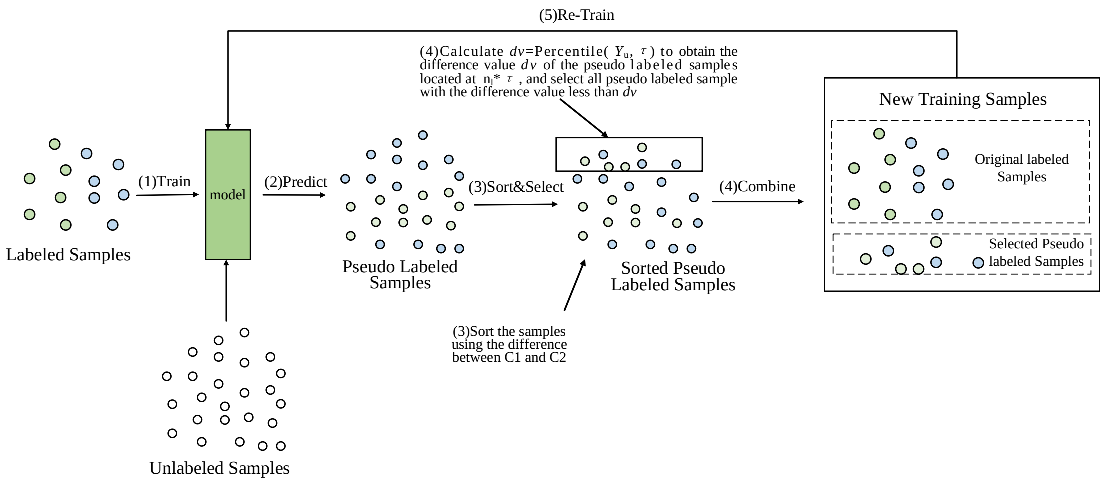

Fig. 5. Training method of the semi-supervised multi-label curriculum learning.

The training method of curriculum learning in Fig. 5 is explained as following, the details are in [26].

- Train the model with labeled comments; (2) Predict all unlabeled comments; (3) Rank pseudo labeled data based on the difference between $C_1(.)$ and $C_2(.)$; (4) Obtain the difference value $dv = \text{Percentile}(Y_u, \tau)$ of the samples ranking at $n_l * \tau$, select the pseudo labeled samples with the difference value less than dv and combine it with the original training set to form a new training set; (5) Re-train the model, and repeat (2) - (5) until all the pseudo labeled data have all the marks. n_l represents the number of labeled samples; $Y_u \in R^{n_l * d_l}$ represents the unlabeled feature matrix and the label matrix.

The adaptive threshold dv is obtained in accordance with the following method: In the training process, all the unlabeled data X_u into the model are put to predict and calculate the difference df of each sample, and then rank it in ascending order according to the sample difference, after that we obtain sample difference value dv of location in $n_l * \tau$ and select the samples, which df is less than dv . Adding them to the labeled data set X_l, in which $\tau \in [0.2, 0.4, 0.6, 0.8, 1]$, that means after every cycle. It will select 20% or more pseudo labeled samples incorporated into the training set. The specific process of curriculum learning is shown in Fig. 5. The difference $df(x_i)$ is calculated as formula (6).

$$df = \| C_1(.) - C_2(.) \|_2^2 \quad (6)$$

B. Event Vectorization Model based on Transformer

Dynamic evolution is an important information of events. Learning the characteristics of the event evolution, which plays a significant role in predicting public opinion reversal from a developmental perspective. Inspired by training natural language, a training model NL2ER-Transformer is designed to simulated event evolution based on the event sentence structure mentioned above. The model is based on Transformer structure which is a classic NLP model proposed by Google in 2017 [12]. Transformer uses a self-attention mechanism and does not use the sequential structure of RNN, so that it can be trained in parallel and can have global information. In the model, the event sentence from the start to the *i*th day of the event is used as the training **feature**, and the event sentence from the start to the (i+1)th day of event is used as the **label**. This process is similar to text translation: the sentences to be translated are as features, and the corresponding translation results are as labels.

The whole training process of the event based on NL2ER-Transformer is shown in Fig. 6.

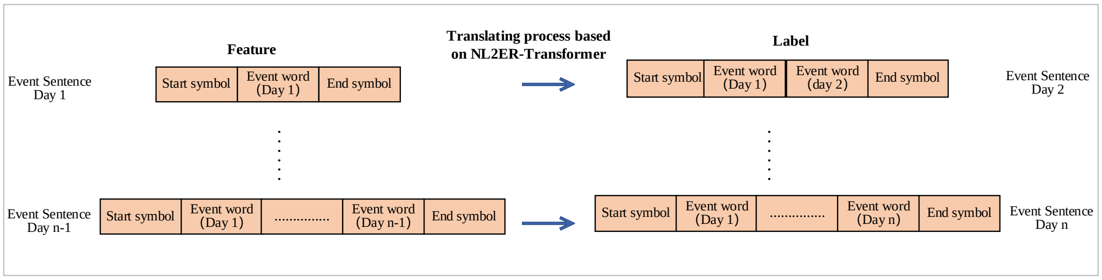

Fig. 6. Training method of simulating event evolution.

IV. Experiments

This paper verifies the effectiveness of the designed model from three aspects: (1) For the proposed comment word generating model based on semi-supervised multi-label curriculum learning, two multi-label classification benchmark models, Fast Tag [27] and Semantic Auto Encoder (SAE) [28], were selected for comparative experiments. At the same time, in the test data used for the downstream reversal prediction task, 50% of the comment words were automatically generated by the model, further verifying the effectiveness of the model; (2) For the event evolution analysis capability of the proposed NL2ER-Transformer model, we manually represent the event sentence value of each day of the event according to the definition, and then compare it with the event sentence value predicted by the model; (3) For downstream reversal prediction tasks, select different event vectorization baseline models and multiple classification models to verify the effectiveness of event vectors generated by NL2ER-Transformer model.

A. Datasets

We obtained 55 events in 2020-2022 from public opinion monitoring platforms such as Qingbo Index (https://www.gsdata.cn) and Zhiwei Shijian (https://ef.zhiweidata.com/library), including 26 public opinion reversal events, 29 public opinion non-reversal events, and a total of 38,091 comments on related events. And according to the assignment rules given above, seven discriminative features of 8041 comments were manually annotated. Data can be obtained from the following link: https://pan.baidu.com/s/1mEQxLdRyOHZd4T8Rmxz0fg. Extraction code: xl8m

In the data annotation, for reversal events, from start to the day of reversal, the label is 1 and from the day after reversal to the end of the event, the label is 0; for non-reversal events, from start to the end of the events, the label is 0.

In the training of the comment word generation model, we selected 30472 comments in total, including 10051 manually labeled comments and 20421 pseudo-labeled comments generated by the proposed model. In the construction of the downstream public opinion reversal prediction model, we randomly selected 12 events as the test set (including 6 reversal events and 6 non-reversal events. The 50% of comments used for the test set are pseudo-labeled data generated by the proposed semi-supervised multi-label curriculum learning model), and the remaining 43 events as the training set.

B. Verification of the Comment Word Generation Model based on Semi-Supervised Multi-Label Curriculum Learning

TABLE III. Comparison Results

Method	Absolute matching rate	Hamming loss	Accuracy	Precision	Recall	F1 score
FastTag	**0.31189**	0.14582	0.67537	0.2681	0.26729	0.2282
SAE	0.11021	0.2346	0.73342	0.49564	0.19515	0.24462
Our model	0.24489	0.16578	**0.79937**	**0.72863**	**0.34015**	**0.3908**

From Table III, it can see that excepting the absolute matching rate is 7% lower than FastTag, our model is much more than the traditional multi-label classification model in all indicators, This proves the effectiveness of the proposed multi-label curriculum learning model.

C. Evaluation for the Evolution Analysis Ability of NL2ER-Transformer Model

During training, the parameters of NL2ER-Transformer model are set as shown in Table IV.

TABLE IV. TRAINING PARAMETERS

Num_layers	d_model	num_heads	input_vocab_size	target_vocab_size	dropout_rate
2	64	4	128	128	0.1

For evolution analysis ability test, we train the NL2ER-Transformer model using the cross-validation method. One event is selected each time as the validation set, and the other events are used as the training set, and the method of error calculation for each learning result is shown in formula:

$$e = \frac{\sum_{i=1}^{n}(w_i - \widehat{w_i})^2}{n} \quad (7)$$

In formula (7), n is the number of days when the event occurs, the event sentence corresponding to the event is $S = [w_1, w_2, \dots, w_n]$, and the predicted result event sentence is $S = [\widehat{w}_1, \widehat{w}_2, \dots, \widehat{w}_n]$.

According to the experimental results, the model produced 4 prediction errors in total which happened in the 1st, 17th, 33th, and 34th experiments and the error values were respectively 3.6, 12.6667, 60, and 22.4285. The experimental results further show that the model does have the ability to analyze the evolution process of public opinion events, and can predict the evolution trend and changes of events. The event sentences with errors and their corresponding events are shown in Table V.

TABLE V. EVENT SENTENCES WITH ERRORS IN PREDICTION

Event	Type of data	Number of days	0	1	2	3	4	5	6	7
12-year-old girl lied about being raped by her teacher	Real event sentence	1	start	84	end	0	0	0	0	0
		2	start	84	82	end	0	0	0	0
		3	start	84	82	80	end	0	0	0
		4	start	84	82	80	82	end	0	0
		5	start	84	82	80	82	80	end	0
		6	start	84	82	80	82	80	80	end
	Prediction	1-6	start	84	82	80	82	80	**82**	end
The death of a female teacher in Jiangsu during the invigilation	Real event sentence	1	start	81	end	0	0	0	0	0
		2	start	81	81	end	0	0	0	0
	Prediction	1-2	start	81	**16**	end	0	0	0	0
Heilongjiang Xuexiang's 15 yuan per sausage attracts heated discussions	Real event sentence	1	start	16	end	0	0	0	0	0
		2	start	16	16	end	0	0	0	0
		3	start	16	16	16	end	0	0	0
		4	start	16	16	16	16	end	0	0
	Prediction	1-4	start	16	16	16	**17**	end	0	0
A 15-year-old girl in Hei Longjiang killed her mother and hid her corpse in a cold storage room	Real event sentence	1	start	16	end	0	0	0	0	0
		2	start	16	0	end	0	0	0	0
		3	start	16	0	80	end	0	0	0
		4	start	16	0	80	2	end	0	0
		5	start	16	0	80	2	2	end	0
		6	start	16	0	80	2	2	48	end
	Prediction	1-6	start	16	**16**	**17**	**17**	**1**	**17**	end

Table V show that three of the four erroneous events have only one day's prediction error, and the event with more errors have errors just from the second day. According to the analysis, the main reason is that of all events, only the error event contains ‘0’ event words. The model has never “seen” the ‘0’ event words during the training process, so the model prediction fails. Although the evolution of the event sentence is misplaced, we found that the label(reversal/non-reversal) of the event was predicted correct. So, we made a further comparison, and it turns out that the evolution process presented by the incorrectly generated event sentence is very similar to the event of ‘Female tambourine businessman in Dali, Yunnan scold tourists’ and the labels of these two events are also the same, i.e., non-reversal. This error obviously has no impact on the prediction results of public opinion reversal.

D. Evaluation of Downstream Reversal Prediction Ability for The Generated Event Vectors

1) Experiment setup: We compare our experimental results with the results of three common models: RoBERTa[24], XLnet[25], Manual designing features[3]. We generate the vectors respectively by the above four models, the first three of which are trained with the blog texts of events, and our proposed NL2ER-Transformer Encoder. Then, through training four prediction models: Linear SVM(LSVM), Back Propagation Neural Network(BPN), nonlinear SVM and KNN, we further compare the prediction effects of the vectors generated by these four methods.

The neural network includes four hidden layers, the number of neurons in each layer is 16, 32, 64，and 128, the activation function is Relu, the output layer activation function is Sigmoid, the loss function is binary cross entropy function, the optimizer is Adam, and the training times are 100. In the linear SVM, C is set to 1, the loss function is hinge, and the training times are 100. The nonlinear SVM kernel function is set as sigmoid function, C is set as 1, and the number of iterations is 100. KNN algorithm sets the number of reference neighbors to 5.

2) Experimental results analysis: We use four indicators to evaluate the classification models: accuracy, precision, recall, F1-Score. We believe that as long as the model can give a reversal probability greater than 0.5 before the event reversal, the prediction is considered correct. From Table VI, we can see that the classification trained by our model is better than other vectorization methods in terms of ensuring high accuracy and precision. KNN and Nonlinear SVM trained in this way are not very good in the four indicators.

Public opinion reversal is a kind of phenomenon in the process of event evolution. Therefore, whether the vector representation of events includes the development and changes of public opinion determines whether the prediction model can learn the important knowledge of event reversal. Compared with netizens' comments, the factual description of events is relatively lacking in expressing the dynamics of public opinion evolution. So, the event vectors generated based on this factual description of events will affect the result of reversal prediction.

The presented classification models of reversal prediction trained with vectors generated by our proposed NL2ER-Transformer are better than or equal to those trained with vectors generated by other event vectorization methods. Therefore, we can infer that it is feasible to vectorize public opinion events by combining comments and event evolution.

TABLE VI. INDEX VALUE OF PREDICTION MODEL UNDER DIFFERENT EVENT VECTOR TRAINING

Representation	Method	Accuracy	Precision	Recall	F1-Score
RoBERTa	LSVM	66.66%	66.66%	66.66%	0.666
	BPN	66.66%	83.33%	62.5%	0.714
	Nonlinear SVM	50%	66.66%	50%	0.514
	KNN	41.66%	66.66%	44.44%	0.533
XLnet	LSVM	75%	**100%**	66.66%	0.8
	BPN	50%	83.33%	50%	0.625
	Nonlinear SVM	50%	**100%**	50%	0.666
	KNN	46.15%	**85.71%**	50%	0.5
Manual design features	LSVM	58.33%	16.66%	**100%**	0.2857
	BPN	58.33%	16.66%	**100%**	0.2857
	Nonlinear SVM	58.33%	16.66%	**100%**	0.2857
	KNN	58.33%	16.66%	**100%**	0.2857
NL2ER-Transformer	LSVM	**91.66%**	83.33%	**100%**	**0.909**
	BPN	**91.66%**	**100%**	85.71%	**0.923**
	Nonlinear SVM	**83.33%**	66.66%	**100%**	**0.8**
	KNN	**91.66%**	83.33%	**100%**	**0.909**

V. Conclusion and Future Work

Aiming at the problems existing in the current research on event representation, this paper proposes an abstract mapping of events to sequences, and combines subjective comments and event temporality to transform public opinion problem into natural language processing problem. The NL2ER-Transformer model is constructed to realize the extraction of event features based on comments. The experimental results show that, compared with current mainstream event vectorization models, the event vectorization scheme proposed in this paper has a better effect on public opinion reversal prediction, and the training method proposed in this paper can indeed enable the vector generation model to learn events evolution information.

There are still some limitations in this paper, and the future work will be carried out from the following aspects:

- Since the discriminative features of comment words are designed based on the prediction of public opinion reversal, the pre-training of the model is only for the prediction task of public opinion reversal, and the effectiveness of other public opinion prediction tasks has not been thoroughly discussed. Next, the team will design a more general feature representation and generation strategy, and construct an event pre-training model in the field of public opinion prediction by testing different types of downstream tasks based on event vectors.
- Event sentence is only generated based on comments, and the features of comment words are discriminative features with strong subjectivity, and the fusion of other objective data is not considered. In future research, we will try to use new methods to combine other unstructured data, such as event description text, with relevant structured data in the process of public opinion evolution, to achieve event pre-training with multi-source data fusion, and to improve the generalization performance of the model.

Acknowledgment

This work was supported by the National Social Science Foundation (No. 22BTQ048).

References

[1] Chang-bin Jiang, Yue-qi Zou, Hu Wang, et al. Research on public opinion reversal prediction of we-media based on SVM. Information Science, 2021, 39(04): 47-53, 61

[2] Nan Wang, Hai-rong Li, Shu-ru Tan. Research on public opinion reversal prediction based on improved SMOTE algorithm and integrated learning. Data Analysis and Knowledge Discovery, 2021, 5(04): 37-48

[3] Nan Wang, Hai-rong Li, Shu-ru Tan. Prediction of public opinion reversal based on evolution analysis of public opinion events and Improved KE-SMOTE algorithm, data analysis and knowledge discovery, 2022, 3(13): 1-20

[4] Yang, F., Liu, Y., Yu, X., & Yang, M. (2012). Automatic detection of rumor on Sina Weibo. *MDS '12*.

[5] Jing Ma, Wei Gao, Zhongyu Wei, Yueming Lu, and Kam-Fai Wong. 2015. Detect Rumors Using Time Series of Social Context Information on Microblogging Websites. In Proceedings of the 24th ACM International on Conference on Information and Knowledge Management (CIKM '15). Association for Computing Machinery, New York, NY, USA, 1751–1754. https://doi.org/10.1145/2806416.2806607

[6] Ting-rui Yan, Dong Wang, Li-zhong Xiao, Guang-zhong Wang. Simulation of public opinion trend prediction model based on AR-RBF. Computer Simulation, 2021, 38(04): 437-441

[7] Jian-feng Zhou. Research on popular microblog feature selection and prediction method based on FA-SVM. Computer Applications and Software, 2018, 35(12): 107-111

[8] Mikolov T , Chen K , Corrado G , et al. Efficient estimation of word representations in vector space. Computer Science, 2013.

[9] Pennington, J., Socher, R., & Manning, C. D. (2014). Glove: Global vectors for word representation. In Proceedings of the 2014 conference on empirical methods in natural language processing (pp. 1532–1543)

[10] Posadas-Durán, J.-P., Gomez-Adorno, H., Sidorov, G., & Escobar, J. J. M. (2019). Detection of fake news in a new corpus for the Spanish language. Journal of Intelligent & Fuzzy Systems, 36(5), 4869–4876.

[11] Peters M, Neumann M, Iyyer M, et al. Deep contextualized word representations: //Proceedings of NAACL, New Orleans American 2018. 2227–2237

[12] Vaswani A, Shazeer N, Parmar N, et al. Attention Is All You Need. arXiv, 2017.

[13] Yang Xu, Liang-qi Zhu, Bo Huang, et al. Public opinion analysis based on EEMD-Transformer model: COVID-19 public opinion as an example. Journal of Wuhan University of Science and Technology, 2020, 66(05): 418-424

[14] Khoo, L. M. S., Chieu, H. L., Qian, Z., & Jiang, J. (2020, April). Interpretable rumor detection in microblogs by attending to user interactions. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 34, No. 05, pp. 8783-8790).

[15] Samadi, M., Mousavian, M., & Momtazi, S. (2021). Deep contextualized text representation and learning for fake news detection. Information Processing & Management, 58(6), 102723.

[16] Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., et al. (2019). RoBERTa: A robustly optimized BERT pretraining approach. CoRR, abs/1907.11692.

[17] Yang, Z., Dai, Z., Yang, Y., Carbonell, J., Salakhutdinov, R. R., & Le, Q. V. (2019). XLNet: Generalized autoregressive pretraining for language understanding. In Advances in neural information processing systems 32 (pp. 5754–5764). Curran Associates, Inc.

[18] Devlin, J., Chang, M.-W., Lee, K., & 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) (pp. 4171–4186). Minneapolis, Minnesota: Association for Computational Linguistics.

[19] Lin, S. Y., Kung, Y. C., & Leu, F. Y. (2022). Predictive intelligence in harmful news identification by BERT-based ensemble learning model with text sentiment analysis. Information Processing & Management, 59(2), 102872.

[20] Blanco, G., & Lourenço, A. (2022). Optimism and pessimism analysis using deep learning on COVID-19 related twitter conversations. Information Processing & Management, 59(3), 102918.

[21] Zheng, M., Ma, Y., Zheng, A., & Wang, N. (2018, May). Constructing method of public opinion knowledge graph with online news comments. In 2018 International Conference on Robots & Intelligent System (ICRIS) (pp. 404-408). IEEE.

[22] Chen Xu (2022). Exploration of network public opinion management method based on knowledge map Industry and Technology Forum (02), 263-264

[23] Anning & Anlu (2022). Cross platform network public opinion knowledge map construction and comparative analysis Information Science (03), 159-165 doi:10.13833/j.issn. 1007-7634.2022.03.020.

[24] Cui, Y., Che, W., Liu, T., Qin, B., & Yang, Z. (2021). Pre-training with whole word masking for chinese bert. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 29, 3504-3514.

[25] Cui, Y., Che, W., Liu, T., Qin, B., Wang, S., & Hu, G. (2020). Revisiting pre-trained models for Chinese natural language processing. arXiv preprint arXiv:2004.13922.

[26] L. Wang, Y. Liu, C. Qin, G. Sun, and Y. Fu, “Dual relation semi-supervised multi-label learning,” in AAAI, 2020, pp. 6227–623

[27] M. Chen, A. Zheng, and K. Weinberger, “Fast image tagging,” in Proc. Int. Conf. Mach. Learn., 2013, pp. 1274–1282

[28] E. Kodirov, T. Xiang, and S. Gong, “Semantic autoencoder for zero-shot learning,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jul. 2017, pp. 3174–3183.