Research on Dynamic Political Sentiment Polarity Analysis of Specific Group Twitter Based on Deep Learning Method

Chengyang Chang¹, Xiaodong Wang*¹

¹Department of computer science and technology, National University of Defense Technology, Changsha, Hunan, 410000, China

*Corresponding author’s e-mail: xdwang@nudt.edu.cn

Abstract. The article introduce a method to analyzing the dynamic political sentiment polarity of American politicians’ tweets within a fixed period of time to assist analysts in judging the direction of American politics and the future trend of China-US relations. This research is applied in the field of think tanks or intelligence analysis. The service targets are intelligent analysts, and the data is the tweet text data sent by a specific group in a fixed time period. We propose an architecture that combines multiple deep learning models and use a dedicated tweet data set to construct a specific group to obtain an sentiment polarity multi-classifier, and then introduce the time characteristics of tweets, and finally obtain the dynamic political sentiment polarity of politicians. The US politician tweets data set proposed in this article verifies that the proposed comprehensive architecture is better than traditional deep learning methods in this task, and the accuracy of the classifier verification set reaches 80.66%. According to the sentiment polarity judgments of 20 US governors and senators, the success rate is 75%. The analysis of individual dynamic political sentiment polarity can provide effective help and intelligence support for analysts. The method in this paper effectively uses a variety of deep learning techniques to assist analysts to obtain more accurate dynamic political sentiment polarity from massive Twitter text data.

1. Introduction

Twitter is an important Internet social network and micro-blogging service website in the United States. The number of monetizable active users of Twitter is 166 million (as of March 2020) [1]. The Trump team began to widely use social media to publish campaign information during the 2016 U.S. election. After he was elected president, Trump announced many important decisions on Twitter, so that he was awarded the title of "Twitter governing the country". Trump himself also claimed that using Twitter to publish news is his only defense against the media. Since 2016, American politicians have generally begun to use Twitter on a large scale as a channel for simultaneous release of official news. As a micro blog, Twitter has the characteristics of immediacy, fragmentation, interactivity, user core and shallowness. Due to the political system in the United States, domestic politicians in the United States, whether they are senators, congressmen, or state governors and mayors, have a large number of supporters and voters in their campaign places, so the Twitter messages they publish always reflects their truly political opinions and information that politicians want voters to know, and these messages are often quickly passed on the Internet. Therefore, the tweets issued by American politicians can be an effective channel for us to deeply judge the camp of the politicians and whether they are satisfied with the current Trump administration.
Sentiment polarity analysis is the process of analyzing, processing, inducing and reasoning about subjective texts with emotional colors. Sentiment analysis is a popular branch of text classification. Its purpose is to analyze people's opinions in text data (such as product reviews, movie reviews, and tweets) and extract their polarity and opinions. This article discusses the classification problem of sentiment polarity differences caused by various subjective and objective factors within a specific group, which is still different from traditional sentiment analysis tasks. The research content of this article is an analysis of the dynamic political sentiment polarity of the Trump administration for the American politicians. However, due to the complexity of the political circle, the sentiment polarity of politicians' tweets often changes due to time or the occurrence of certain special events. Traditional multi-emotional classification methods alone cannot meet the needs of solving this problem. This paper compares the effects of multiple deep learning models on the multi-classifiers trained on the constructed data set, and proposes a framework based on four deep learning methods to analyze the dynamic political sentiment polarity of American politicians, and according to The framework proposes a time-varying emotion dynamic analysis method based on deep learning, and provides a feasible solution to the problem of dynamic political emotion polarity analysis.

2. Related Work
The essence of sentiment polarity analysis of tweets is a downstream task of automatic text classification, the basic task of natural language processing. The main methods at this stage are the same as those of automatic text classification, mainly based on traditional machine learning algorithms and neural network based methods. There are a few researches in the field of dynamic sentiment analysis.

2.1. Text Sentiment Polarity Analysis Based on Traditional Machine Learning Algorithms
This type of method mainly uses traditional machine learning algorithms such as support vector machine (SVM), naive Bayes (NB), n-gram language model algorithms, etc., and selects a certain feature item weight calculation method to construct a vector space model, such as Correct the word frequency feature weight (TF-IDF), then use the corresponding feature selection method to select the feature, and finally use the machine learning algorithm to train the classification model. Bermingham et al. [2] designed a model that integrates twitter traffic and sentiment analysis to predict election results. The author selects TF-IDF to construct a vector space model, trains SVM and MNB classifiers to classify the emotions of Twitter English text, the accuracy of the classifier is 62.94% and 64.82%, respectively. Liu Zhiming et al. [3] used three machine learning algorithms, three feature selection algorithms, and three feature item weight calculation methods to carry out an empirical study on sentiment classification of Weibo. The TF-IDF weight calculation method was used, and the classification algorithm used SVM. The classification effect obtained by the feature selection of information gain is the best, and the classification accuracy of 87.07% can be achieved for the two-class classification of micro blog data.

The accuracy of text emotion classification is difficult to reach the level of ordinary text classification, which is mainly caused by the complex emotion expression and a large amount of emotion ambiguity in the emotion text. Especially in the multi-classification problem of emotion polarity, in the emotion classification algorithm based on machine learning, each article is converted into a corresponding feature vector to represent. The quality of feature selection will directly affect the performance of sentiment analysis tasks, and the direct use of TF-IDF to calculate the weights of feature words under multi-classification is not good for vectorization. Therefore, in recent years, this type of method has gradually become the mainstream of solving sentiment classification. method. The political sentiment polarity classification involved in this article is more complicated than ordinary sentiment classification, and more emotional ambiguities. At the same time, we hope to use models to automatically identify deeper emotional features in different categories. This is We have chosen a variety of deep neural network methods to study the reasons.
2.2. Text Sentiment Polarity Analysis Based on Deep Learning

In text sentiment polarity analysis, the application of deep neural network can use the network structure to automatically obtain text features to solve classification problems, effectively avoiding complex artificial feature engineering. In the training phase, the weights are adjusted through multiple algorithms such as forward propagation algorithm and reverse correction algorithm, so that the test text can learn accurately according to the adjusted weights, so as to obtain different neural network models, and then combine one. The texts of unknown categories are sequentially passed through these neural network models to obtain different output values, and the text category is finally determined by comparison.

There are many mature models used in text classification, such as HAN[4] based on hierarchical attention network model, TextCNN[5], DCNN[6], MVCNN[7], etc., based on recurrent neural network, etc. Network Tree-LSTM[8], densely connected two-way LSTM model (DC-Bi-LSTM[9]), C-LSTM[10], RCNN[11] and other models based on the combination of the above two. Among the above methods, the densely connected two-way LSTM DC-Bi-LSTM model performed the best on the sentiment polarity test data set with fewer classification categories, reaching 89.7% in the two-class data set SST-2 [9]. The accuracy rate is close to the best results on the five-category SST-1[12] and six-category TREC[13] data sets, but only an accuracy rate of up to 50% [14].

For the problem of political sentiment polarity analysis based on tweets, the biggest difference from traditional sentiment polarity classification is that the data set is much smaller than traditional text classification. The data set that can be collected for different event views is very small. Limited. At this stage, there are some specific studies on the classification of tweets mainly in the agricultural and medical fields. Reddy et al. [15] believe that users’ opinions can be discovered from tweets or re-posts. The author divides tweets into positive sentiment and negative sentiment, and uses long-term short-term memory (LSTM) networks and convolutions for a dataset of 150,000 small tweets. The neural network (CNN) training classifier, compared with the naive Bayes classifier, obtained 81% of the binary classification effect on the long short-term memory (LSTM) network. Dunnmon et al. [16] used direct learning on small-scale agriculture-related tweet data sets, and then migrated to learn larger sentiment data sets in other fields to accurately predict the potential of agricultural sentiment. The performance of the convolutional neural network in classification is better than that of the recurrent neural network, and its performance on the ternary sentiment classification problem is much lower than that of the binary classification.

Social media provides a low-cost alternative resource for public health surveillance, and health-related classification plays an important role in identifying useful information. Dai et al. [17] proposed a clustering method based on word embeddings, Perform vector representation and clustering to classify tweets as relevant or irrelevant to the topic (such as influenza), and the best accuracy can reach 87.1%. Gencoglu et al. [18] proposed a deep convolutional neural autoencoder model for learning the information representation of medical and health-related tweets, and then combined clustering methods to extract hot topics and events.

It is found that the convolutional neural network and the recurrent neural network are better or worse for the sentiment polarity classification effect of Twitter short texts. It depends more on the quality of the data set. The other two neural network methods are more complicated. The effect on the classification problem is not ideal, but it can have a good classification effect in a simple two-class classification problem, so it still has strong application value.

In recent years, the large-scale pre-training language model proposed in the sentiment classification research of tweets has better proposed solutions for multi-classification, multi-label and complex scenes. There are two main ideas. One is to fine-tune training on existing pre-trained language models. This method is suitable for problems with small data sets. Biseda et al. [19] explored the use of drug reviews and social media as potential alternative sources for the discovery of adverse drug reactions. They performed three models of BERT [20] and its two variants, BioBERT [21] and Clinical BERT [22]. Fine-tuning the training, achieved the best accuracy rate of 88.9% in the three categories of drug emotional reviews. The second is to directly train a pre-training model of a specific target domain
through large-scale related domain corpus. The performance effect of this method is greatly improved. The model can perform well for various natural language tasks in a specific target domain, but the required hardware resources are high and the amount of data is large. Müller et al. [23] aimed at the recent COVID-19 In the hot spot of the epidemic, the COVID-Twitter-BERT (CT-BERT) pre-training model was trained on a large data set of 160 million tweets. This model is used to complete natural language processing tasks related to the COVID-19 epidemic text from social media. 10-30% improvement over the basic model BERT-LARGE.

2.3. Research on Dynamic Sentiment Analysis Method at the Present Stage
At this stage, there is still room for improvement in the accuracy of static multi-sentiment classification, and there are relatively few related researches in various fields of dynamic sentiment analysis. However, there are still some cutting-edge research and exploration results at home and abroad. Yin et al. [24] proposed a framework of emotion recognition method (VADAR[25]) based on thesaurus and grammar rules to analyze the emotions (positive and negative emotions) of a large number of tweets related to the COVID-19 pneumonia epidemic. The dynamic change law of time; Li Hui et al. [26] used the short text emotional topic model (SSTM[27]) model to obtain the emotional distribution of the review document, and proposed a dynamic emotional topic model to collect the mobile phone review data of a certain brand according to time Segmentation to obtain the trend of the content and emotion of the subject over time; Gupta et al. [28] used a variety of machine learning classifiers such as random forest, logistic regression, and three-layer perceptron to maliciously treat the Amazon product review data. The extreme tendency of review text is classified, and dynamic analysis of the behavior of this type of group is carried out to better understand brand-level dynamic opinion fraud and to improve the existing Amazon product scoring mechanism.

At this stage, due to the low accuracy of deep neural network models for traditional emotional multi-classification problems, the classifier models used in related researches on dynamic sentiment analysis are more traditional machine learning models or emotion recognition based on thesaurus and grammar rules. Method, and all the research is two-category dynamic emotion. However, starting from our problems, dynamic political emotions show more multipolar tendencies. The political emotions of a certain politician fluctuate between different polarities, especially before and after certain events (such as general elections). , The COVID-19 epidemic, etc.), the political sentiment of politicians may undergo drastic changes, and the static multi-category analysis method of sentiment for a single tweet simply cannot meet the relevant requirements. Therefore, this paper introduces the external feature-time of the tweet text, improves the traditional deep learning-based emotional multi-classification model, and proposes a time-varying political sentiment dynamic analysis model based on multiple deep learning.

The problem scenario used in this article is to further classify the portrait of the person based on the tweet, and assist the analyst in the fine-grained portrait of the target person. It requires the model used to have a higher classification accuracy rate, and at the same time due to the political emotion of the person. Changes over time also require the model to be easy to update and train quickly, and the volume and parameters should not be too large, otherwise it is not conducive to the deployment of work application scenarios.

In summary, we choose to use three traditional neural network models and the BERT-based pre-training model to solve the problem.

3. Research Framework and Model Introduction
The sentiment polarity discussed in this article is not the sentiment polarity in the traditional sense, but in a specific context, it classifies the difference in sentiment polarity that occurs within a specific group due to various subjective and objective factors. This question has strong application value, because it is different from the problem of sentiment polarity classification in the traditional sense. Such problems are often very specific to the sentiment polarity within a specific group under a specific social background, not Commendatory, derogatory, or neutral. The research object of this article
focuses on the specific group of American politicians, in view of the internal division of this group since the Trump administration took office in 2016, and judges the dynamic changes of the political sentiment polarity among individual politicians in the US government over time.

3.1. Research Framework

The main users are oriented to a large number of analysts, and provide them with a tweet-person classifier based on the Twitter platform that is convenient to deploy, has high accuracy, and the model can be updated regularly, and describes the specific content of the tweets. The system mainly realizes regular targeted tweet collection, tweet identification of specific groups of people, and provides a variety of optional classifiers to analyze the political sentiment polarity of target groups. The groups targeted by this article are members of the U.S. Senate and House of Representatives and politicians at the level of state governors. Among them are 39 governors and 75 senators.

Based on the previous investigation and analysis, this article divides the sentiment polarity of American politicians into the following four categories:

1. The pro-Trump faction is characterized by unreserved support for the various policies of the Trump administration. This type of personnel mainly represents the existing ultra-right forces in American politics.

2. Moderate opposition, whose main characteristic is its opposition to the Trump administration's unreasonable policies. Such personnel are widely present in the Democratic and Republican parties, and their attitudes are often vacillating.

3. Direct competitors. This type of personnel is a direct competitor of the presidential election and Trump's team. However, the characteristics of these personnel are also to attack China as a means to attract votes.

4. Objective moderates, whose main characteristic is their sympathetic attitude towards Chinese and Asian Americans, resolutely opposing most of the Trump administration's policies, and actively fighting against the COVID-19 epidemic. This part of politicians should be actively striving for us.
According to the application background, this article specifically crawls the Twitter text of a specific group, cleans and annotates the data after obtaining the data, and then selects and annotates the tweets of some people with known sentiment polarity to form annotated Training data set. Then, the labeled data is divided into training set, test set, and verification set according to a certain proportion, and then input into different types of classifiers for training. Due to differences in data types, text lengths, and classification methods, the effects of classifiers based on different deep learning methods are also different. The main indicators are the accuracy and F1 value of the trained classifier on the validation set data. At the same time, due to the application Scenario needs, but also needs to consider the size of the model and the speed of training.

3.2. Time-varying dynamic sentiment analysis method based on deep learning (TVD-SAM)

In the dynamic political sentiment polarity analysis problem, the problem is abstractly expressed through the four-tuple relationship as

\[ F : \{g, s, h, t\} \]

Where \( g \) represents the emotional object, \( s \) represents the emotional orientation, \( h \) represents the opinion holder, and \( t \) represents the time. Here we have also defined the label categories of all tweets separately, and the corresponding four-tuple relationships are: the emotional object is the text of the tweet; the emotional orientation selected 4 types of political sentiment polarity Label based on our previous research; the opinion holder is the user who posted the tweet; the time is the date the tweet was issued.

Regarding how the CNN[5], Bi-LSTM[29], C-LSTM[10] and BERT[20] pre-training models used in this article realize the process and principle of emotional multi-classification, this article will not repeat them. The main introduction is to get a lot of push How to realize the combination of external time variables of tweets after the classification results of the posts.

There are two main manifestations of the time-varying characteristics of emotional dynamic analysis.

The dynamic political sentiment analysis method based on special events divides the designated research time period according to a specific time node, and at this specific time node we can select the time window in which major events have occurred, such as the outbreak of COVID-19 pandemic, the United States Domestic demonstrations, etc., are used to analyze whether politicians’ political sentiment polarity changes significantly before and after these special events.

We assume that the time point when the special event \( E \) occurs in \( T \), and if its duration is \( t_1 \), then the time windows selected in the dynamic analysis are respectively as

\[ [T - t_1, T], [T, T + t_1] \]

According to the time point \( T \), input the optimal classifier to classify the tweets sent by the research object during the time period to obtain the classification results of tweets \( M \) before the special event and tweets \( N \) after the special event.

\[
S_{before} = \{S_{b_0}, S_{b_1}, ..., S_{b_m} | S \in [0,1,2,3], m = card(F_{t \in [T-t_1,T]}) \}
\]

\[
S_{after} = \{S_{a_0}, S_{a_1}, ..., S_{a_n} | S \in [0,1,2,3], n = card(F_{t \in [T,T+t_1]}) \}
\]

Time-window-based dynamic political sentiment analysis performs window interception of tweets on the time axis according to a specific time window, and then analyzes the political sentiment polarity tendency of the tweets within each fixed window time, and the interception in the window completely covers the delineation After the research period of time, obtain the dynamic picture of a certain politician’s political sentiment tendency in a fixed period of time, and analyze the change of political sentiment polarity in a certain period of time from the perspective of the time axis.
We choose the size of the time window as $\Delta T$ and the starting time as $T_0$, then the time axis set $T$ of the dynamic sentiment polarity change is

$$T_n = \{T_0 + k_n \Delta T \mid k \in [0,1,2,\ldots,n]\}$$

In such a time discrete space, the tweets of the research object in different time periods are input into the optimal classifier for classification, and the normalized classification result $S_{Tn}$ of the sentiment classification of the tweets in the time window is obtained.

$$S_{Tn} = \{S_{T_n,0}, S_{T_n,1}, \ldots, S_{T_n,p} \mid S \in [0,1,2,3], p = \text{card}(F_{t \in [T_{n-1}, T_n]}))\}$$

The dynamic political sentiment analysis method based on special events can be regarded as the special situation based on the time method. Although the two methods have similarities in the model expression, the two methods are actually cut from two different perspectives in the analysis process. For analysis, the method based on special events is to analyze whether a politician's political sentiment tends to change strongly before and after the special event, and analyze and judge his views on the special event. The time-window-based method is to intercept and analyze the political sentiment polarity of tweets within a fixed time period according to a fixed time window to find the time period when the research object undergoes greater political sentiment polarity changes, and then time period to explore the reasons for political swings in it.

4. Experiment and result analysis

4.1. Dataset selection and construction

As a social media platform recognized by American politicians today, Twitter is a social media platform. The tweets posted by politicians on Twitter are directly pushed to their main voters, so their tweets can truly reflect the politician’s macro-political perspective. Therefore, based on the above information, this article first selects 21 American politicians with strong sentiment polarity classification to establish a data set. These 21 American politicians are shown in Table 1.

| Politicians Name that Dataset chosen | Politicians Name |
|-------------------------------------|------------------|
| 1. Pro-Trump faction                | John Bolton, Donald Trump, Mike Pence, RoBert O’brien, Mike Pompeo, Steven Mnuchin |
| 2. Moderate opposition              | Congressmen Frank Palone, Congressman Eric Swawell, Senator Richard Blumenthal |
| 3. Direct competitors               | Joe Biden, Congressmen Adam Schiff, Senator Bernie Sanders, Speakers Nancy Pelosi |
| 4. Objective moderates              | Governor Gretchen Whitmer, Senator Kamala Harris, Lawrence H. Summers, Governor Andrew Cuomo, Sally Yates, Senator Maria Cantwell, Senator Edward Markey, Senator Elizabeth Warren |

After determining the crawl target, this article obtained the following 21 U.S. politicians' tweets from May 25, 2019 to May 25, 2020 through the targeted crawling method based on Twitter id. After cleaning, a total of 53292 tweets are used as a data set, and the ratio of training set, validation set, and test set is set to 5:1:1. Examples of data sets are shown in Table 2.
Twitter can truly reflect the politician’s attitude towards his voters and the current government, and we speculate that they have the same political sentiment.

Table 2. Dataset Example

| Tweet text                                                                 | Label |
|---------------------------------------------------------------------------|-------|
| Enjoyed talking davidgura at Select USA summit. Tax reform trade and regulation rollback are critical to serve hardworking Americans | 0     |
| Enforcers must stop scammers and bottom feeders from exploiting COVID-19 and endangering health. False pitches and sky-high price hikes should be halted and prosecuted. | 1     |
| President Trump may be a slick salesman who fooled many people in this country, but you didn't fool me and you didn't fool New Yorkers. | 2     |
| With respect, Mr. President, not sure we can rely on Mr. Manafort's lawyer to tell us whether there was collusion, as unbiased as he may be. | 3     |

4.2. Datasets Cleaning and Labeling

Because the data set constructed in this article has never existed before, it is mainly because of the particularity of the problem and the research population. In the process of data cleaning, first of all, because tweets will be mixed with a large number of emoji emoticons, URL addresses, Twitter ids, etc. when crawling. First of all, for emoji emoticons, which may represent part of the emotion in the tweet, this article chooses to convert them into text for preservation. Secondly, a series of non-English symbols such as mixed URL addresses have been deleted, but punctuation marks and stop words are retained due to the particularity of the Twitter short text. The third point is that there are two main types of processing for the Twitter ids that are mixed in this article, some of which appear frequently and contain the country, region, and specific person names, for example: @realDonaldTrump and @POTUS are converted to President Trump, and @Iran is converted to Iran, @HilaryClinton are converted to Hilary Clinton, etc., because I think this part of the id appears as the subject or object in the content of the tweet, and is a named entity that has an important influence on the sentiment polarity of the tweet, so the conversion and retention are carried out. The id with a lower frequency of occurrence is deleted in this article. The biggest difficulty in the cleaning process is the need to manually filter out the tweets that are not related to politics in the tweets sent by these politicians. For the reposting part, a large number of reposted tweets can be filtered by identifying keywords, but there are still a small number of For tweets that are irrelevant to our research field, this part mainly relies on manual cleaning of the data set.

In the labeling part, first, the tweets of different politicians are classified and merged according to the crawled id, and the tweets are unified, and then the content is too short (less than 10 characters) or pure retweets are filtered out. In order to ensure that the tweet data of the four categories in the training set, test set, and validation set are relatively balanced, the tweets of different categories are first shuffled and then separated at a ratio of 5:1:1, and then the four categories of tweets The training data, verification data, and training data are combined to obtain a training set, a verification set, and a test set. Table 3 shows the specific situation of each classification of the data set constructed in this paper.

Table 3. Datasets Specific Distribution

|          | Label=0: Pro-Trump faction | Label=1: Moderate opposition | Label=2: Direct competitors | Label=3: Objective moderates | Total   |
|----------|-----------------------------|-------------------------------|-----------------------------|-------------------------------|---------|
| Training set | 9300                      | 6430                           | 7427                        | 14905                        | 38062   |
| Evaluation set | 1861                      | 1286                           | 1486                        | 2981                          | 7614    |
| Test set   | 1862                      | 1286                           | 1486                        | 2982                          | 7616    |
| Total      | 13023                     | 9002                           | 10399                       | 20868                        | 53292   |
4.3. Experiment results
After completing the construction of the data set, the overall experiment of this article is divided into three parts. One is based on the CNN[5] model of convolutional neural network, the Bi-LSTM[29] model based on recurrent neural network, and the combination of the two, The C-LSTM[10] model and the BERT[20] model based on the feedforward neural network are selected from the four models that can be used for sentiment polarity classification tasks as the classifier; the second is in On the trained classifier, the 1000 tweets of some American politicians with unknown sentiment polarity are classified and judged, and the possible sentiment polarity probability of this American politician during this time period is analyzed; the third is showing the US Senator Jeanne Shaheen as an example, using dynamic political sentiment analysis methods based on special events and time windows to show the analysis results.

The experimental environment of this article is: operating system: Ubuntu18.04, CPU: Intel E5 12 core, graphics card: NVIDIA RTX2080ti, TensorFlow version: 1.15 .2. The training set, validation set, and test set data are placed in the data folder. The experimental results of the classification of the four models of CNN, Bi-LSTM, C-LSTM, and BERT are shown in Table 4.

Performance comparison of multiple classification models of different neural networks

4.3.1. Performance comparison of multiple classification models of different neural networks
In the process of constructing the data set in advance, this article pays special attention to the problem of data imbalance that often occurs in multi-classification problems. Although the data of the four classifications are not completely balanced, the overall ratio is 1.3: 0.9:1.1:2.1, except that the fourth type of data is more than the average, the other three types of data sets are balanced. Therefore, it is reasonable and effective to use the accuracy of the validation set as the criterion for the evaluation model.

| Model    | Eval set precision | Training loss | F1Score | Model Training time | Model Size |
|----------|--------------------|---------------|---------|---------------------|------------|
| CNN      | 63.79%             | 0.8271        | 62.35%  | 12min19s            | 54.3MB     |
| C-LSTM   | 67.56%             | 0.6645        | 66.87%  | 14min15s            | 69.3MB     |
| Bi-LSTM  | 72.59%             | 0.7365        | 71.83%  | 13min42s            | 55.1MB     |
| BERT     | 80.66%             | 0.6282        | 79.34%  | 52min23s            | 1.22GB     |

It can be found that the classification effect of the BERT[20] model is stronger than that of the other three classifiers. As a single tweet, the accuracy of the verification set is 80.66%, which is sufficient to determine the polarity of emotions. The second part of the experiment assists human judgments. The sentiment polarity of American politicians in a certain period of time, so this article chooses the classifier obtained after fine-tuning the BERT and the training model as our sentiment polarity discriminator. However, after analysis, this article believes that although BERT-based classifiers in this data set are better than classifiers based on convolutional neural networks and recurrent neural networks, as the data set continues to expand, the amount of data in the data set gradually Balanced, the effects of other classifiers may be improved.

However, the excellent effect of the classifier based on the BERT pre-training model will inevitably bring about the relatively slow training speed, and the model is much larger than the traditional neural network model. This is for our task to update the data set within a certain period of time, and at the same time fine-tune the training classifier requirements, the BERT-based classifier still has certain limitations. And in the architecture proposed in this paper, for different data sets, the best classifier among the three evaluation indicators of classification effect, training speed, and model size is selected as the preferred choice. When the hardware equipment allows and the time conditions are sufficient, the index of classifier accuracy must be preferred.
This article believes that there is still room for improvement of the model, and the author will also publish the data set and code to the Internet at the same time as it is released, hoping that more people of insight can work together to solve this problem.

4.3.2. Determine the static political sentiment polarity of unknown persons within a fixed period of time

The second part of the experiment is to input the tweets of some US politicians with unknown sentiment polarity in a certain period of time into the sentiment polarity discriminator, and then obtain the probability of the polarity of the tweet in a fixed period of time. This part of the experiment is to determine the sentiment polarity of the characters without manually interpreting a large amount of tweet information for American politicians who we do not know enough, and the sentiment polarity classification of their tweets can be quickly obtained through the discriminator, then judge the attitude of his character. We judged the tweets of 20 senators and governors in total, and their tweets were classified and normalized by the classifier we trained. The overall results of this part of the experiment are shown in Table 5. In order to show the effect of the discriminator in detail, the discriminating results and real political opinions of 5 digits are selected for concrete display, as shown in Table 6 and Table 7.

Table 5. Discrimination results for 20 senators and governors with unknown political sentiment polarity

|            | Total | Correction | Wrong |
|------------|-------|------------|-------|
| Senator    | 10    | 7          | 3     |
| Governor   | 10    | 8          | 2     |
| Total      | 20    | 15         | 5     |

Table 6. Judging the results of the tweets of 5 US politicians with unknown political orientation

| Name                | Time period     | Position       | Label=0 | Label=1 | Label=2 | Label=3 |
|---------------------|-----------------|----------------|---------|---------|---------|---------|
| Phil Murphy         | 2020-02-05 to 2020-05-24 | Governor of New Jersey | 0.098 | 0.454 | 0.019 | 0.428 |
| Richard Mike De Wine | 2019-12-20 to 2020-05-24 | Governor of Ohio | 0.357 | 0.034 | 0.019 | 0.590 |
| John Carney         | 2016-06-07 to 2015-05-24 | Governor of Delaware | 0.179 | 0.214 | 0.050 | 0.557 |
| Eric Brakey         | 2018-08-26 to 2020-05-23 | Senator in Maine | 0.347 | 0.131 | 0.220 | 0.302 |
| Laura Kelly(×)      | 2019-01-14 to 2020-05-22 | Governor of Kansas | 0.434 | 0.016 | 0.090 | 0.460 |

From the test results shown in the above table and manual verification, it can be found that the American politician sentiment polarity classification system proposed in this article is effective. The fine-tuned classifier based on the BERT pre-training model compares traditional convolutional neural networks with cyclic neural networks. The network has greatly improved and improved the classification effect of this problem. It shows that the pre-training model represented by BERT is effective and practical for solving various downstream tasks in natural language processing. However, limited by the accuracy of the classifier itself, the accuracy rate is 80.66%. In the second part of the experiment, the classification effect obtained by the classifier is still very helpful for us to judge the sentiment polarity of American politicians that we do not understand. The experiment verifies that the 20 American politicians in the United States, the discriminator’s accuracy rate was 75%. Among the five politicians shown, the American politician with the wrong discriminator’s judgment, Kansas Governor Laura Kelly, is a Democrat, but being in Kansas, a state with deep Republican traditions, her campaign success was also supported by a large number of former or current Republican
government officials. In the tweets for her judgment, although she strongly criticized the Trump administration's early anti-epidemic work, there were also many tweets thanking the Trump administration for supporting the large amount of medical supplies in Kansas. I think the above two points may be an important reason for disturbing the classifier to make an accurate judgment on this person. But judging from the overall effect of the second part of the experiment, it proves the effectiveness of this model.

Table 7. An analysis of the true political sentiment polarity of 5 American politicians

| Name          | Partisan     | Position            | Main political proposition                                                                                                                                                                                                 | Attitudes to the Trump administration | Analysis summary                                                                                           |
|---------------|--------------|---------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------|----------------------------------------------------------------------------------------------------------------|
| Phil Murphy   | Democrats    | Governor of New Jersey | Be sympathetic to Asians and people of color, support Trump on the fight against the epidemic, and thank the Trump administration for its massive support                                 | Neutral                               | This person belongs to the moderate opposition. Although he is different from Trump's party, he still follows the Trump administration and supports the Trump administration's decisions in most cases. The classifier also classifies this person into this category. |
| Richard Mike De Wine | Republican | Governor of Ohio | Unlike the traditional Republican Party that supports gun control, it opposes same-sex marriage and abortion, while actively preventing and controlling the new crown epidemic | Support                               | It can be seen that this person is very different from ordinary Republican far-right personnel. Although he supports the Trump administration, he deviates from many positions. Our classifier also shows this judgment. |
| John Carney   | Democrats    | Governor of Delaware | Support the protection of the rights and interests of people of color, emphasize racial equality, support gun control, and actively organize the fight against the epidemic | Strongly opposed                      | Since the Trump administration came to power, this person has objected to it, and it can be seen that the results of the classifier are also the same. |
| Eric Brokey   | Republican   | Senator in Maine   | Traditional Republican senators have a strong anti-China tendency                                                                                                                                                | Support                               | This person has a strong anti-China tendency, and he strongly agrees with the main political views of the Trump administration. The result of the classifier deviates from the actual result, but the label0 probability of 0.347 is already very high. |
| Laura Kelly(×) | Democrats   | Governor of Kansas | The traditional Democratic governor supports universal health insurance and actively fights against the new epidemic                                                                                            | Opposed                               | This person is a Democrat who strongly criticized the Trump administration on the issue of fighting the epidemic, but our |
crown epidemic, but his state is a traditional Republican vote warehouse and has received a lot of federal medical resources. discriminator has given the opposite conclusion. The reason may be that his tweet posted a lot of thanks to the Trump administration for its support materials, did not truly reflect its political orientation.

4.3.3. Analysis of dynamic political sentiment polarity
We first choose March 20, 2020 as the time node for the special event. The reason is that on this day, the number of people with new coronary pneumonia in the United States exceeded 10,000 for the first time. President Trump declared a state of emergency in the country. The level of attention has increased unprecedentedly. This day can be regarded as a replica of the first level of anti-epidemic status in all provinces of my country at that time when Wuhan broke out. We have selected an analysis to find that the senator's tweets in the fourth category increased by 43.1% before and after the incident, and the overall tweets increased by 34.2% after the outbreak of the COVID-19 pandemic. The following conclusions can be drawn from the analysis of dynamic political sentiment polarity at the time node of March 20: 1. Senator Shaheen has a high degree of concern about the COVID-19 pandemic; 2. Regardless of the outbreak of the COVID-19, she holds a negative attitude towards most of the Trump administration's policies.

| Tweet time | Tweet text |
|------------|------------|
| ['2020-05-21'] | By withdrawing from the Open Skies Treaty, Pres. Trump is barreling down a path that makes us less secure,... I urge the President to reverse this reckless decision. |
| ['2020-05-14'] | ...With limited supplies, I'm calling on the Trump administration to be transparent with the American people about how this drug will be distributed. |

Table 8. Senator Shaheen’s partial tweets

Figure 2. Dynamic political sentiment analysis based on special events
Based on the time window of the dynamic political sentiment analysis method, in the example, I choose seven days, one week as a time window. In the specific analysis, it can be changed according to different research objects and time periods. In the example, it can be seen that the number of tweets sent by the senator during the week of March 20-26, 2020, has surged, and most of the increments are objective and moderate tweets, and the two weeks of continuous attention After entering April, as the epidemic entered a period of steady increase, the total number of tweets and the number of tweets in category 4 also fell into a stable period. The dynamic changes of its political sentiment polarity can be well represented by the model proposed in this article, which can provide analysts with more intuitive intelligence based on the Twitter platform of the target person in real time.

5. Conclusion
Twitter, as a world-renowned social media platform, has now become an important information release channel for American politicians. It is of great significance to analyze the political sentiment of specific people’s tweets to determine the future trend of American politics and the future trend of China-US relations. It is an extremely meaningful exploration for the automatic acquisition and analysis of open source intelligence. Focusing on the introduced dynamic political sentiment polarity analysis, two analysis models based on special events and time windows are proposed. After comparing the classification effects of the traditional convolutional neural network model and the recurrent neural network model on the data set, this article chooses to use the BERT pre-training model to fine-tune the training to obtain the classifier, and then input the unknown American politician tweets texts into the classifier and analyzing the results, a relatively ideal effect was obtained at last. Using multi-classifiers combined with analysis models, the dynamic political sentiment polarity of individual tweets was analyzed and displayed, which provided a new method for the collection, acquisition and analysis of open source intelligence. There are also some problems in this article. One is that based on the experimental data, it can be found that because the amount of data of label3 is higher than that of the other three categories, the classifier's judgment for the third part is also higher after normalization. For the other three types of situations, this is also the main improvement point in the later stage. The data set is improved, supplementing the tweet data of the first three types of personnel, and further improving the balance and richness of the dataset.
References

[1] Twitter, Inc. Financial information 2020 annual report Twitter Inc. Financial Report 2020 [EB/OL]. https://investor.twitterinc.com/financial-information/default.aspx

[2] Bermingham A, Smeaton A. On using Twitter to monitor political sentiment and predict election results[C]//Proceedings of the Workshop on Sentiment Analysis where AI meets Psychology (SAAIP 2011). 2011: 2-10.

[3] Liu Zhiming, Liu Lu. An Empirical Study on Sentiment Classification of Chinese Weibo Based on Machine Learning[J]. Computer Engineering and Applications, 2012, 48 (01): 1-4.

[4] Yang Z, Yang D, Dyer C, et al. Hierarchical attention networks for document classification[C]//Proceedings of the 2016 conference of the North American chapter of the association for computational linguistics: human language technologies. 2016: 1480-1489.

[5] Kim Y. Convolutional Neural Networks for Sentence Classification[C]//Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP). 2014: 1746-1751.

[6] Kalchbrenner N, Grefenstette E, Blunsom P. A Convolutional Neural Network for Modelling Sentences[C]//Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). 2014: 655-665.

[7] Yin W, Schütze H. Multichannel variable-size convolution for sentence classification[J]. arXiv preprint arXiv:1603.04513, 2016.

[8] Tai K S, Socher R, Manning C D. Improved semantic representations from tree-structured long short-term memory networks[J]. arXiv preprint arXiv:1503.00075, 2015.

[9] Ding Z, Xia R, Yu J, et al. Densely connected bidirectional lstm with applications to sentence classification[C]//CCF International Conference on Natural Language Processing and Chinese Computing. Springer, Cham, 2018: 278-287.

[10] Zhou C, Sun C, Liu Z, et al. A C-LSTM neural network for text classification[J]. arXiv preprint arXiv:1511.08630, 2015.

[11] Lai S, Xu L, Liu K, et al. Recurrent convolutional neural networks for text classification[C]//Twenty-ninth AAAI conference on artificial intelligence. 2015.

[12] Socher R, Perelygin A, Wu J, et al. Recursive deep models for semantic compositionality over a sentiment treebank[C]//Proceedings of the 2013 conference on empirical methods in natural language processing. 2013: 1631-1642.

[13] Li X, Roth D. Learning question classifiers[C]//COLING 2002: The 19th International Conference on Computational Linguistics. 2002.

[14] Wang Zhihui, Wang Xiaodong. Research on text classification method based on neural network[J]. Computer Engineering, 2020, 46(03): 11-17.

[15] Mahidhar Reddy D, Subba Reddy N V, Subba Reddy N V. Twitter Sentiment Analysis using Distributed Word and Sentence Representation[J]. arXiv, 2019: arXiv: 1904.12580.

[16] Dunnmon J, Ganguli S, Hau D, et al. Predicting State-Level Agricultural Sentiment with Tweets from Farming Communities[J]. 2017.

[17] Dai X, Bikdash M, Meyer B. From social media to public health surveillance: Word embedding based clustering method for twitter classification[C]//SoutheastCon 2017. IEEE, 2017: 1-7.

[18] Gencoglu O. Deep representation learning for clustering of health tweets[J]. arXiv preprint arXiv:1901.00439, 2018.

[19] Bisبدا B, Mo K. Enhancing Pharmacovigilance with Drug Reviews and Social Media[J]. arXiv preprint arXiv:2004.08731, 2020.

[20] Devlin J, Chang M W, Lee K, et al. BERT: pre-training of deep bidirectional transformers for language understanding[J]. Arxiv: Computation and Language, 2018,23(2):3-19.

[21] Lee Jinhuyuk, Wonjin Yoon, Sungdong Kim, Donghyeon Kim, Sunkyu Kim, Chan Ho So, Jaewoo Kang, BioBERT: a pre-trained biomedical language representation model for biomedical text mining, Bioinformatics. 2019.
[22] Azadeh Nikfarjam, Abeed Sarker, Karen O’connor, Rachel Ginn, and Graciela Gonzalez. Pharmacovigilance from social media: Mining adverse drug reaction mentions using sequence labeling with word embedding cluster features. Journal of the American Medical Informatics Association 22, no. 3: 671-681. 2015.

[23] Müller M, Salathé M, Kummervold P E. COVID-Twitter-BERT: A Natural Language Processing Model to Analyse COVID-19 Content on Twitter[J]. arXiv preprint arXiv:2005.07503, 2020.

[24] Yin H, Yang S, Li J. Detecting Topic and Sentiment Dynamics Due to COVID-19 Pandemic Using Social Media[J]. arXiv preprint arXiv:2007.02304, 2020.

[25] Hutto C J, Gilbert E. Vader: A parsimonious rule-based model for sentiment analysis of social media text[C]//Eighth international AAAI conference on weblogs and social media. 2014.

[26] Li Hui, Hu Yunfeng. Online comment analysis based on dynamic emotional topic model[J]. Data Analysis and Knowledge Discovery, 2017, 1(9).

[27] Xiong Shufeng, Ji Donghong. Short text sentiment topic model for product review analysis [J]. Acta Automatica Sinica, 2016, 42(008): 1227-1237

[28] Gupta V, Aggarwal A, Chakraborty T. Detecting and Characterizing Extremist Reviewer Groups in Online Product Reviews[J]. IEEE Transactions on Computational Social Systems, 2020.

[29] Minaee S, Azimi E, Abdolrashidi A A. Deep-sentiment: Sentiment analysis using ensemble of cnn and bi-lstm models[J]. arXiv preprint arXiv:1904.04206, 2019.