Scientific
Research
Publishing

Film and Television Website Scores Authenticity Verification Based on the Emotional Analysis

Weiyu Tong

Aberdeen Institute of Data Science and Artificial Intelligence, South China Normal University, Foshan, China
Email: Tongweiyu666@gmail.com

Abstract

Sentiment analysis is a method to identify and understand the emotion in the text through NLP and text analysis. In the era of information technology, there is often a certain error between the comments on the movie website and the actual score of the movie, and sentiment analysis technology provides a new way to solve this problem. In this paper, Python is used to obtain the movie review data from the Douban platform, and the model is constructed and trained by using naive Bayes and Bi-LSTM. According to the index, a better Bi-LSTM model is selected to classify the emotion of users' movie reviews, and the classification results are scored according to the classification results, and compared with the real ratings on the website. According to the error of the final comparison results, the feasibility of this technology in the scoring direction of film reviews is being verified. By applying this technology, the phenomenon of film rating distortion in the information age can be prevented and the rights and interests of film and television works can be safeguarded.

Keywords

Bi-LSTM Model, Film Review Emotion Analysis, Naive Bayes, Python, Data Crawl

1. Introduction

In the information technology era, film and television websites are widespread. Some argue that data companies provide standardized outsourced services, commonly known as a "water army", to boost the popularity of movies and TV dramas. As a result, evaluating the quality of these works has become a notable chal-

lenge for today's audiences. Presently, advancements in natural language development and the evolution of sentiment analysis technology offer a viable solution to this issue.

Sentiment analysis, a methodology for identifying and comprehending emotions within text through natural language processing and text analysis, has emerged as an effective means to evaluate the quality of content. The primary aim of sentiment analysis is to ascertain the emotional orientation of text, typically categorized as positive, negative, or neutral. This categorization serves to aid social media platforms in content filtering or providing personalized recommendations, functioning as a branch of Natural Language Processing (NLP). The core of text sentiment analysis involves the extraction and understanding of emotions, attitudes, or sentiments from textual data.

Numerous scholars have delved into the sentiment analysis of film reviews, exploring ways to enhance the accuracy of results. Substantial research findings have been amassed both domestically and internationally in recent years.

Guo *et al*. have proposed the role of the LDA model in text sentiment analysis. They emphasize the importance of determining the optimal value for the topic quantity K during training. A nuanced balance is required; if K is too large, it may result in overly fine-grained topics, while if K is too small, the main content may be divided into overly coarse topics [1]. Li *et al*. proposed an improved algorithm, which combined with hierarchical sampling technology of data, can greatly improve the accuracy of text data with heterogeneous structure [2]. Chen *et al*. introduced the RoBERTa-PWCN-GTRU emotion analysis model, showcasing advancements in sentiment analysis accuracy specifically applied to Chinese film review datasets. Their model's impact on improving sentiment analysis accuracy in a Chinese context is noteworthy [3]. Chen *et al*. extracted the features of TF-IDF and Chi-square statistics of jieba by combining the whole word and the two-word method, and after comparing various algorithms, BP neural network is proved to have the best effect in emotion classification [4]. Wang *et al*. conducted a study on the MLP model, examining its suitability for emotion analysis in film reviews. Notably, they observed that enhancing the dictionary capacity or data list length contributes to an increased accuracy in the model [5]. Jiang *et al*. introduced a comprehensive index recommendation calculation method that considers variations in users' individual needs. This approach enables the derivation of distinct recommendation indexes tailored to different users [6]. Deng *et al*. proposed a sentiment classification model based on naive Bayes, leveraging the Scrapy framework for extracting movie review data. Their work provides valuable insights into the effectiveness of this algorithm, particularly in the context of sentiment analysis for short film reviews [7]. By studying the CNN model, Zhou *et al*. proposed a Bagging algorithm using the CNN model as a weak classifier, and the final classification result is determined by voting method. It is verified that the classification accuracy of this integrated method is improved compared with Bi-LSTM model and single CNN model [8]. Zhu *et al*.

proposed a collaborative filtering recommendation algorithm ABFR based on ALBERT-BiLSTM model, and it was confirmed that the accuracy and recall rate of this algorithm were improved compared with other algorithms [9]. Lei *et al.* improved the traditional collaborative filtering algorithm, improved the accuracy of the measurement of interest similarity between users, and confirmed that the algorithm can obtain better recommendation quality [10]. Yang *et al.* propose a joint neural network model (CNN + LSTM) with attention mechanism, and compared with the traditional single model, the results show that it can reduce noise interference and make full use of article context information, thus improving the classification accuracy [11]. Zhou *et al.* proposed a TF-IDF algorithm to extract subject terms and determine weights, calculate the emotion value of online comments, and then subdivide the interval to determine the emotion degree. A PLTS-VIKOR method is proposed to rank the advantages and disadvantages of each film, and the validity and superiority of this method are verified by real review data. At the same time, it is pointed out that this method can also be applied to other fields such as music life [12]. Zhang *et al.* pointed out that the traditional recommendation algorithm has problems such as sparse data and lack of attention to the diversity of recommendation results. The LDA model is used to extract the theme of movie reviews and identify the emotional tendency related to the theme, which improves the content-based recommendation algorithm and confirms that the diversity of recommendation list has been significantly improved [13]. Razia Sulthana *et al.* proposed a method for classifying comments using a classifier-based Bagging algorithm on a newly constructed SVM classifier, and compared the results of the proposed method with similar existing work, concluding that the method achieves better results compared to existing systems [14]. Prabu *et al.* proposed a sentiment analysis technique based on Corporate-Integrated Autoencoder Convolutional Neural Networks (CI-AECNN). It realizes the generation of emotion dictionary based on corpus. This method is used to calculate the examinee's emotional tendency, add the identified emotional words to the seed dictionary, and remove the emotionally incorrect words from the seed dictionary. Use Long Short-Term Memory (LSTM) to perform word sense disambiguation. Conditional random fields are used to extract features. Finally, auto-encoder and convolutional neural networks are used for classification. In the MATLAB simulation environment, the overall analysis of the research work shows that the proposed technology produces better results than the existing technology [15]. Zhang *et al.* based on the principles of word vector and deep learning, a film review sentiment analysis technique based on deep learning and machine learning word stitching is proposed. The correct rate of 83.13% is achieved, which proves the effectiveness of the method [16]. Sentiment analysis of text based on LSTM and Bert, Si *et al.* combined with Bert pre-trained language model and Long Short-Term Memory (LSTM). The network builds a text sentiment analysis model, uses Bert pre-trained word vectors instead of traditional word vectors, generates semantic

vectors dynamically according to the word context, and inputs the semantic vectors into LSTM to capture semantic dependencies, thereby enhancing the ability to extract effective information. According to the comparison of accuracy rate, accuracy rate, recall rate and F-Measure, it is found that the proposed method is significantly better than the comparison method in text sentiment analysis [17]. Adepu *et al.* proposed a multi-channel CNN model based on bidirectional LSTM and attention mechanism, focusing on those parts of sentences that have an important impact on judging the emotion of the sentence. Finally, it is concluded that this algorithm is superior to traditional CNN, LSTM + CNN and other machine learning algorithms [18]. Yang *et al.* proposed a Collaborative Filtering Algorithm (CF-SA) based on sentiment analysis of comment text. LDA is used to calculate the similarity of user comments. In addition, ALBERT model and BiLSTM neural network are used to mine the emotional tendency of users in the project review text, and the similarity of user comments and user ratings are combined to obtain the final user similarity and predict the user's score on the project. The comparison with classical recommendation algorithms shows that the proposed algorithm has good recommendation performance [19]. Yang *et al.* proposed a Multi-Feature Fusion Sentiment Analysis (MBEA) method that combines capsule network, multilayer bidirectional short-term memory network and residual network. This article generates word vectors using Word2vec model, and then inputs word vectors into residual network, capsule network and multilayer bidirectional long short-term memory network to obtain their vector feature representations respectively. Finally, the fully connected layer is input, and the emotion discrimination is carried out by softmax activation function. The accuracy and effectiveness of the model are further verified by experiments on data sets [20]. He *et al.* proposed a bidirectional LSTM model with a trapezoidal structure. Compared with single-layer configuration, the bidirectional LSTM model with trapezoidal structure can better extract high-dimensional features of text by using multi-layer structure, and the parameters of the model can be reduced by using trapezoidal structure. According to the experimental results, it is concluded that the trapezoidal structure model and the ordinary structure model have similar performance. However, the parameters of trapezoidal structure model are 35. 75% less than those of ordinary structure model [21]. Kumaresan *et al.* proposed an integrated technique for analyzing tweets containing different languages through collaborative generative adversarial networks (Gans) and Self-Attention Networks (SANs). The GAN layer arranges pre-processed tweets into positive and negative parts while the SAN layer identifies neutral tweets. Comparing the proposed technique with existing parallel neural networks (CNNS) and Self-Attention Networks (SANs) on Tanglish tweets, it is concluded that the proposed model has better accuracy than the existing models [22].

In this paper, text sentiment analysis technology is used to verify the authenticity of film and television website scores to help audiences identify the quality of movies. In order to comprehensively analyze the emotions of users' film re-

views, this paper uses Python as a programming language to crawl the film review data from Douban movie website by calling Requests library and BeautifulSoup library. lstm and naive Bayes models were constructed to train the data, and the classification effect of the training model was evaluated. The optimal model was selected to conduct sentiment analysis and text classification of Chinese film review texts, and score them. The comparison was made with the real Douban score to verify the authenticity of the website score.

2. Experimental Design

2.1. Data Source

This paper utilizes movie review data from Douban.com (https://movie.douban.com/chart) for the test dataset. The data includes movie name, short review content, and review score, serving as the target data for analysis. The training dataset, consisting of 30,000 film reviews and corresponding scores, is obtained from an open dataset on CSDN. The reviews are categorized based on scores, with ratings 1 - 2 classified as negative, 3 as neutral, and 4 - 5 as positive. Figure 1 and Figure 2 present the statistical graphs illustrating sentence length and occurrence frequency, as well as the cumulative distribution function of sentence length for the training dataset, respectively.

2.2. Emotion Classification Model Based on Naive Bayes Algorithm

2.2.1. Naive Bayes

Naive Bayes is a generative model commonly used in classifiers, which transforms joint probability into conditional probability based on Bayes theorem, and
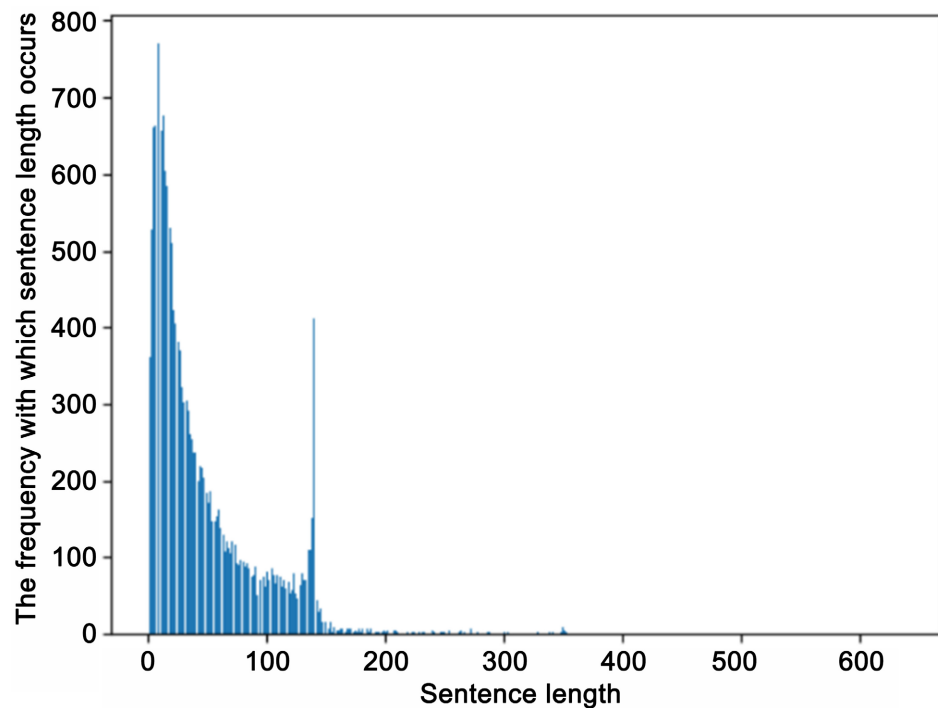

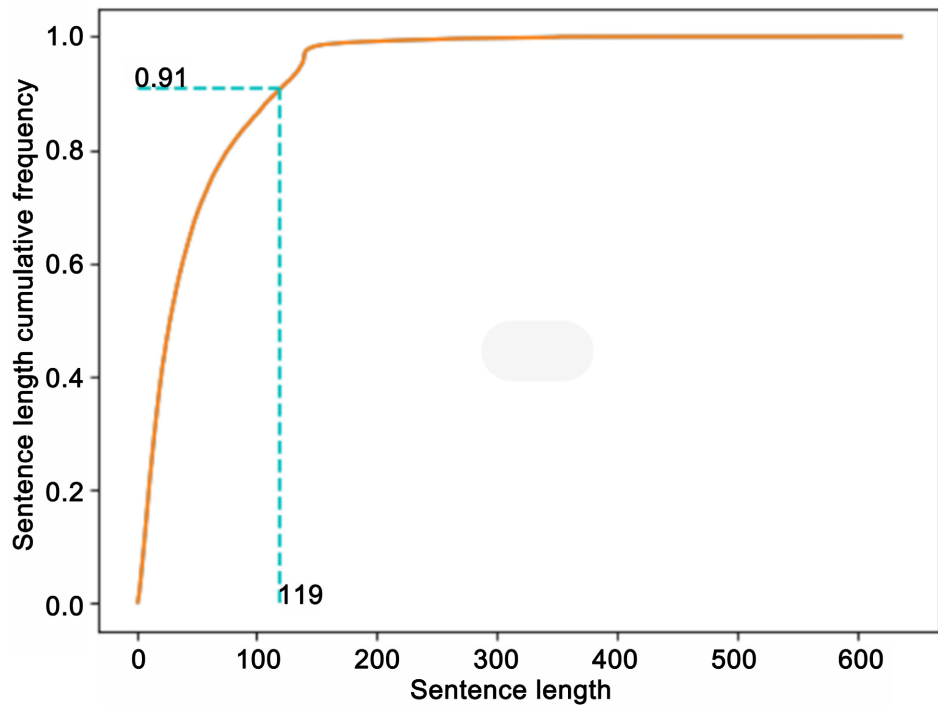**Figure 1.** Sentence length and frequency statistics chart.

Figure 2. Graph of sentence length cumulative distribution function.

computes the probability of characteristic conditions and simplified conditions based on independent hypothesis. The expression of naive Bayes is as follows:

$$P(A|B) = P(B)P(B|A) \cdot P(A) \tag{1}$$

Through the total probability formula, this article can see:

$$
\begin{aligned}
P(\text{feature}) = {}& P(\text{category1}) * P(\text{feature}|\text{category1}) \\
& + P(\text{category2}) * P(\text{feature}|\text{category2})
\end{aligned}
\tag{2}
$$

Whether it is a positive probability or a negative probability, it will eventually be divided by P(feature), so P(feature) does not change for these two probabilities, so it can be ignored and directly compared.

2.2.2. Data Preprocessing of Naive Bayes

Data preprocessing plays a pivotal role in data analytics and machine learning endeavors, aiming to cleanse, transform, and ready raw data for the more effective application of models. The primary objectives of data preprocessing encompass Chinese word segmentation, stop word removal, dictionary construction, and the segmentation of training sets. In this paper, the stop word table is sourced from CSDN, and the model processes the data in the following manner:

1) Construction dictionaries are used to store words that have already occurred.

2) Defines the stop word table, reads the stop word in the stop word file and stores it in the stop word table.

3) Define the cleaned data set.

4) Iterate over each comment in the raw dataset.

5) Use jieba to segment comments in Chinese to get a list of words.

6) For each word after the participle, if it does not belong to the stop word list, it is added to the words.

7) Adds words that do not appear in a dictionary to a dictionary.

8) Add the processed comments (the list of words after the stop word) and corresponding flags to the clear_dataset.

9) For each comment data, the comment text information after word segmentation is extracted, and the dictionary is used to transform each comment text information into a word vector, the dimension of which is the length of the dictionary.

10) The word vector for each review is added to the X list, and the label for each review is stored in the Y list.

11) Converts lists X and Y to numpy arrays.

12) Use the train_test_split function to divide the data set into the training set and the test set on a 3:1 ratio.

2.2.3. Algorithm Flow of Naive Bays

The implementation process of sentiment analysis based on naive Bayes is depicted in Figure 3. Initially, the movie review is segmented using Jieba for word tokenization. Subsequently, a classification model is constructed and trained using a designated training set. Simultaneously, the effectiveness of the model is evaluated using a separate test set. Finally, the trained classifier categorizes the emotions conveyed in the classified text.
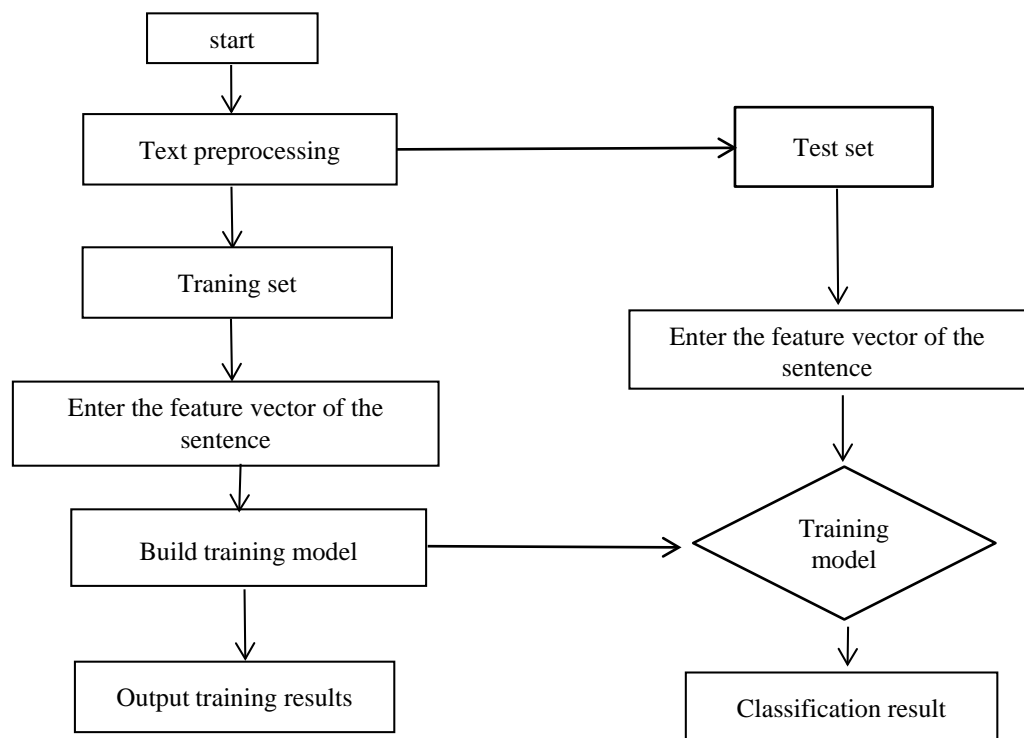

Figure 3. Naive Bayes algorithm flow.

2.2.4. Naive Bayes Classifier

The experiment in this paper selects a polynomial Bayesian classifier, that is, naive Bayes with polynomial distribution prior. It assumes that the features are generated by a simple polynomial distribution that can describe the probability of occurrence of various types of samples, so polynomial naive Bayes is suitable for describing the features of occurrence or proportion of occurrence. This model is often used for text classification, where features represent the number of times, such as the occurrence of a certain word.

2.3. Emotion Classification Model Based on LSTM

2.3.1. LSTM (Long Short-Term Memory) Neural Network

Long Short-Term Memory (LSTM) is a variant of Recurrent Neural Networks (RNNs) specifically designed to address the challenge of long dependencies in traditional RNNs. Due to issues like the vanishing or exploding gradient problem, traditional RNNs struggle to effectively capture long-term dependencies when processing extended sequence data. LSTM, however, excels with lengthy sequences of information by introducing a mechanism known as a "gate". This gate enables selective retention or discarding of information. LSTM has demonstrated remarkable performance in various fields, including natural language processing and time series analysis.

The complete LSTM architecture diagram is shown in Figure 2, which is mainly composed of 5 different parts:

1) Unit state, the internal unit state of LSTM is memory.

2) Hidden state, used to calculate the external hiding of the predicted result status.

3) The input gate determines how much of the current input will be sent to the cell status.

4) The forget door, which determines how many previous cell states, will be sent to the current cell state.

5) Output gate, which determines how much cell state, is output to hideStatus.

The Bi-LSTM model used in this paper is a variant of the LSTM model, addressing the limitation of the LSTM model, which can only rely on the forward sequence for learning. This is achieved by incorporating feedback from both forward and reverse direction sequences. Emotion analysis of film reviews often requires a comprehensive consideration of contextual content to assess the overall emotional tendency of the review. Consequently, the classification accuracy of the Bi-LSTM model surpasses that of the LSTM model. The structure of the Bi-LSTM model is depicted in Figure 4.

2.3.2. Data Preprocessing of LSTM

1) Read data from CSV files through the pandas library.

2) Use jieba for Chinese word segmentation, if the words after segmentation are in the stop word list, filter them.
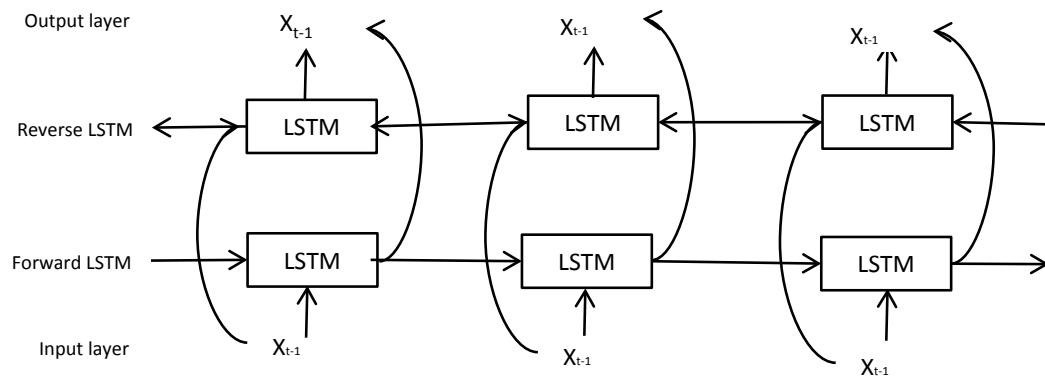
3) Get unique labels and terms in the data.

Figure 4. Bi-LSTM structure.

4) Build word-level features, form vocabularies, create dictionaries and tag dictionaries, and save them as pickle files.

5) Converts the text to the corresponding sequence of numbers and does the sequence padding to ensure that the input length is consistent.

6) The label is processed to convert it to a unique thermal encoding form

7) Finally, return the processed input sequence x, label y, output label mapping output_dictionary, vocabulary size vocab_size, label category number label_size, and reverse dictionary inverse_word_dictionary.

The processed data will be used for the subsequent training of the LSTM model.

2.3.3. Algorithm Flow of LSTM

The implementation process of sentiment analysis based on an LSTM model is illustrated in Figure 5. Initially, the stop word table is loaded, and the data is cleaned and filtered using Jieba word segmentation. Subsequently, a dictionary is constructed for data pre-processing, encompassing steps such as label processing and sequence filling. Next, a double-layer Bi-LSTM deep learning model is built, incorporating L2 regularization in the LSTM layer to reduce weight complexity. The model is compiled using the cross-entropy loss function and the Adam optimizer. The model's architecture is depicted in Figure 6. Following the completion of model training, the accuracy of the model on the test set is computed, and the trained classifier is employed to classify the emotion of the text.

3. Experimental Design

3.1. Data Sets and Evaluation Criteria

This experiment utilized the Douban movie review dataset, which was published on CSDN and comprised 30,000 film reviews. Each review was labeled as positive, negative, or neutral based on user ratings, as outlined in Table 1. When constructing the LSTM model, performance evaluation indices were incorporated by utilizing the "scikit-learn" library. The evaluation criteria included accuracy rate, precision rate, recall rate, and F1 value for positive, negative, and neutral labels. The accuracy rate represents the ratio of samples correctly classified
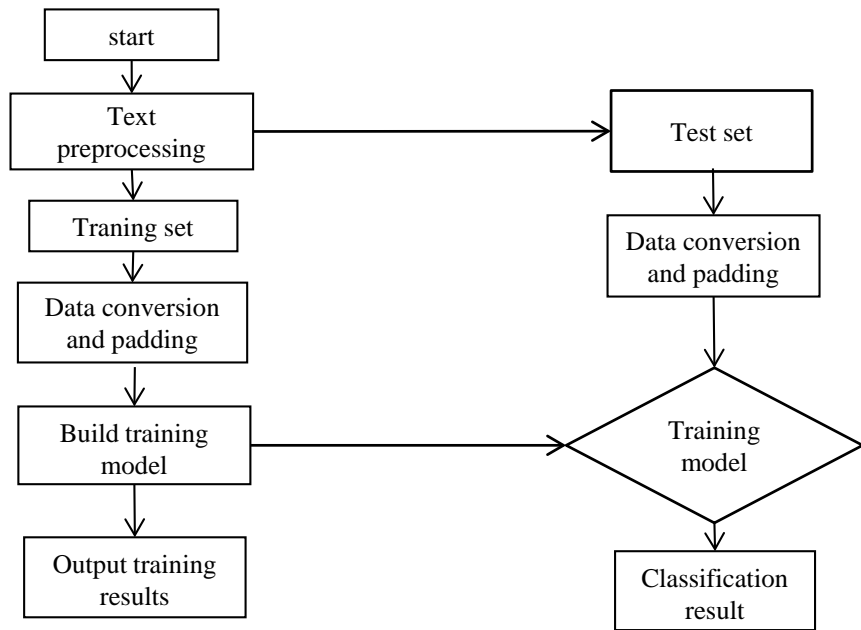
Figure 5. Bi-LSTM algorithm flow.

Embedding_input	input:	[(None,50)]
InputLayer	output：	[(None,50)]

embedding	input:	(None,50)
Embedding	output：	(None,50,200)

bidirectional(lstm)	input:	(None,50,200)
Bidirectional(LSTM)	output：	(None,50,64)

bidirectional_1(lstm_1)	input:	(None,50,64)
Bidirectional(LSTM)	output：	(None,64)

dropout	input:	(None,64)
dropout	output：	(None,64)

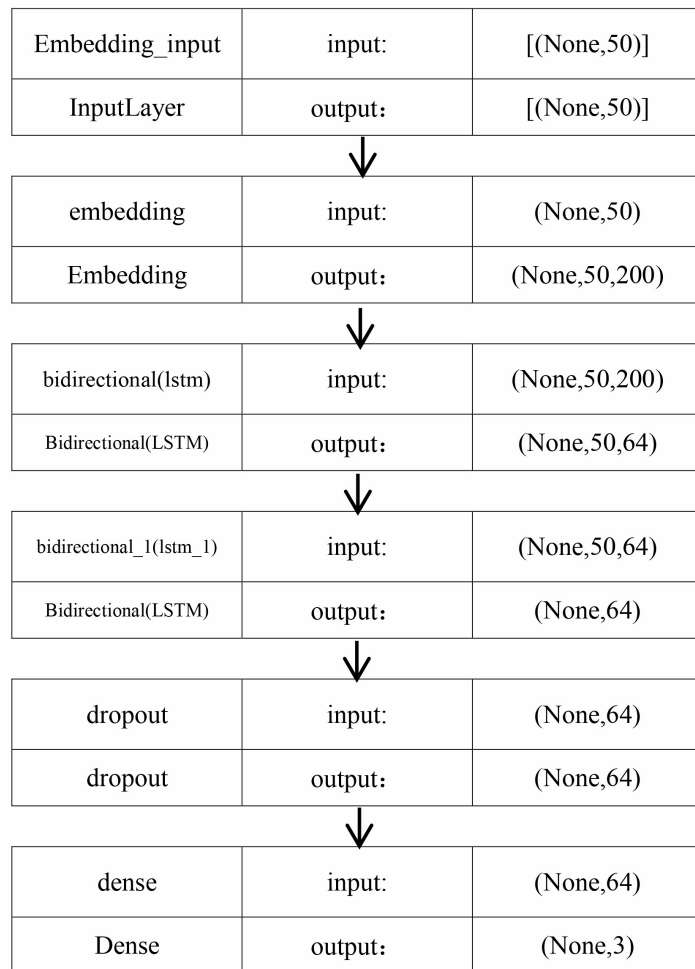| dense	input:	(None,64)
Dense	output：	(None,3)

Figure 6. Double-layer Bi-LSTM model structure.

Table 1. Classification label settings.

Type	Label
Positive	1
Negative	0
neutral	2

by the classifier to the total number of samples. The precision rate signifies the ratio of correct classifications to the total number of instances in a category. The recall rate denotes the ratio of correct classifications to the original total number in a category. Finally, the F1 value is the harmonic average of the precision rate and recall rate.

3.2. Experimental Parameter Settings

In order to evaluate the performance of LSTM model and naive Bayesian model, this article designed a set of comparison experiments. The classification accuracy, accuracy, recall and F1 values of the LSTM model were compared with those of the naive Bayes model. In order to control variables, the same data set was used for both models, and the jieba word segmentation was used for preprocessing.

3.3. Analysis of Experimental Results

Table 2 presents the experiment results, focusing on the classification accuracy comparison between Naive Bayes and Bi-LSTM models. Clearly, the Bi-LSTM model demonstrates superior performance in classification accuracy. Therefore, utilizing the Bi-LSTM model for emotion classification in film review data substantially improves accuracy, and this improvement is not arbitrary.

Therefore, in the following score prediction experiment, this paper will use the trained Bi-LSTM model for training.

3.4. Film Review Emotion Analysis

3.4.1. Data Collection

When crawling movie reviews, this article employed a basic web crawler that iterates through the URLs of various pages, sends HTTP requests, and parses the HTML content using BeautifulSoup. Ratings and brief comments are extracted on each page by searching for HTML tags and class attributes within the comments section. Subsequently, this data is stored in a CSV file named "XXX.csv".

3.4.2. Results

According to the experimental results of 3.3, the Bi-LSTM model with better performance is selected to process the data. Select three movies with good ratings (more than 3.5 stars), three movies with average ratings (between 2.5 and 3.5 stars) and three movies with poor ratings (less than 2.5 stars) from the Douban website, and grab the review data for rating.

Table 2. Model performance.

Model	Index			
	Accuracy	Precision	Recall	F1 Score
Naive Bayes	0.546	0.664	0.546	0.523
Bi-LSTM	0.690	0.691	0.69	0.689

Table 3. Movie rating prediction results.

Movie	Index	
	True Score	Predictive Score
The Shawshank Redemption	4.85	4.0
Spirited Away	4.7	4.81
Titanic	4.75	4.5
Conan: Black Iron Submarine	3.3	2.85
Leave the World Behind	3	2.74
Napoleon	3.25	3.08
Lie Detector	1.4	2.2
League of Gods	1.5	2.45
Pure Hearts: Into Chinese Showbiz	1.1	2.40

If the emotion classification of the model output is 1, 5 points are recorded; if the output emotion is classified as 0, 1 score is recorded; if the output emotion is classified as 2, a score of 3 is recorded, and 220 movie reviews are collected for statistics for each movie. The experimental results are shown in Table 3.

3.4.3. Result Analysis

According to the experimental results, there is a correlation between lower movie scores and increased deviation between the model's predicted score and the actual movie score. Analysis of film reviews reveals that movies with low ratings often feature reviews characterized by irony, wherein the expressed emotion contradicts the literal meaning. This discrepancy leads to positive model classifications for movies that are, in reality, poorly rated. Additionally, the model demonstrates greater accuracy in predicting scores for high-rated and moderately-rated movies, confirming the practicality of the model in scoring applications.

The performance advantage of Bi-LSTM over naive Bayes may be related to its better capturing of long-term dependencies in film criticism. In a movie review, a reviewer may mention a particular emotion multiple times in the review, and Bi-LSTM is better able to remember this long-term contextual information. The performance advantage of Bi-LSTM over naive Bayes may be related to its better capturing of long-term dependencies in film criticism. In a movie review, a reviewer may mention a particular emotion multiple times in the review, and Bi-LSTM is better able to remember this long-term contextual information.

Compared with previous studies, the Bi-LSTM model adopted in this paper may be more robust in sentiment analysis of film criticism. Previous approaches may present challenges to long-term dependencies.

4. Closing Remarks

In this study, we use both the LSTM and naive Bayes models for deep learning on movie review data. We compare the models using various indicators and find that the LSTM model demonstrates higher accuracy in emotion classification compared to the naive Bayes model.

With the help of this technology, it can prevent the distortion of the true evaluation of the work such as grading, so as to avoid adverse effects on the publicity and revenue of the work. Film and television websites can increase users' confidence in their rating systems, allowing users to rely more heavily on these authentic ratings to make movie-watching decisions. This practice further consolidates the reputation and brand image of the site, enhances user loyalty to the site, and makes users more willing to continue to choose the site as the preferred platform to provide them with a quality viewing experience.

However, a few problems were identified in this study:

1) The data used in the training set has not been manually labeled, so the true emotions of some comments are inconsistent with the corresponding labels. This inconsistency makes it difficult to further improve the accuracy of the model.

2) In Chinese, there is a type of expression known as "irony", where the actual emotion is contrary to the apparent emotion in the text. This discrepancy poses a challenge for models to accurately classify such evaluations, thereby impacting the practical application's accuracy.

In the follow-up research, it is crucial not only to address the aforementioned two issues, but more importantly to acquire a more precise labeled dataset. This will significantly enhance the accuracy of the model.

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

[1] Guo, X.L., Niu, X.L. and He, L.Q. (2023) Emotion Analysis of Film Review Text Based on LDA Model. *Computer and Telecommunication*, **3**, 1-4. (In Chinese)

[2] Li, J.H. and Cheng, Y.Y. (2021) Text Sentiment Analysis Based on Douban Film Reviews. *The Han Teachers College Journal*, **6**, 80-86. (In Chinese)

[3] Chen, Y.J., Zheng, G.S. and Xiao, W. (2023) Emotion Analysis of Chinese Film Criticism Based on RoBERTa and Syntactic Information. *Science, Technology and Engineering*, **18**, 7844-7851. (In Chinese)

[4] Chen, P.P., Geng, X.R., Zou, M. and Tan, D.Y. (2020) Text Sentiment Orientation Analysis Based on Machine Learning. *Computer and Modernization*, **3**, 77-81+92. (In

Chinese)

[5] Wang, D., Gan, H., Huang, W.F., Huang, Z.L. and Chen, Z.J. (2021) Research on Emotion Analysis of Film Criticism Based on MLP Model. *Chinese Science and Technology Papers Online Quality Papers*, **2**, 208-211. (In Chinese)

[6] Jiang, L. and Zhang, Q.L. (2017) Research on Personalized Recommendation Strategies Based on Comment Sentiment Analysis: A Case Study of Douban Film Reviews. *Theory and Practice of Intelligence*, **8**, 99-104. (In Chinese)

[7] Deng, C.Y. and Yu, G.Q. (2023) Research on Sentiment Analysis of Film Criticism Based on Naive Bayes. *Intelligent Computers and Applications*, **2**, 210-212+217. (In Chinese)

[8] Zhou, J.Y., Guo, Y.D. and You, D. (2018) Emotion Analysis of Chinese Film Criticism Based on Deep Learning. *Journal of Shanghai University* (*Natural Science Edition*), **5**, 703-712. (In Chinese)

[9] Zhu, Q.Q., Lan, W.F., Sun, H., Zhang, Y.Q. and Tian, P. (2022) Recommendation Algorithm ABFR Based on Sentiment Classification of User Comments. *Journal of Central South University for Nationalities* (*Natural Science Edition*), **3**, 333-338. (In Chinese)

[10] Lei, L. and Zhu, M. (2016) Application of Sentiment Analysis in Movie Recommendation System. *Computer Engineering and Applications*, **10**, 59-63+107. (In Chinese)

[11] Yang, X.C., Mo, X.L. and Wang, C.D. (2021) Research on Emotion Analysis of Film Criticism Based on Neural Network Model. *Journal of Tianjin University of Technology*, **5**, 13-17. (In Chinese)

[12] Zhou, H., Ma, H.N. and Liu, J. (2020) Research on Film Recommendation Algorithm Combining Sentiment Analysis and Probabilistic Language. *Theory and Practice of Intelligence*, **6**, 180-186. (In Chinese)

[13] Zhang, Y.L. and Zhang, L.L. (2022) Movie Recommendation Algorithm Based on Sentiment Analysis and LDA. *Procedia Computer Science*, **199**, 871-878. https://doi.org/10.1016/j.procs.2022.01.109

[14] Razia Sulthana, A., Jaithunbi, A.K., Harikrishnan Haritha Varadarajan Vijayakumar. (2022) Sentiment Analysis on Movie Reviews Dataset Using Support Vector Machines and Ensemble Learning. *International Journal of Information Technology and Web Engineering* (*IJITWE*), **17**, 1-23. https://doi.org/10.4018/IJITWE.311428

[15] Prabu, P., Sivakumar, R. and Ramamurthy B. (2021) Corpus Based Sentimenal Movie Review Analysis Using Auto Encoder Convolutional Neural Network. *Journal of Discrete Mathematical Sciences and Cryptography*, **24**, 2323-2339. https://doi.org/10.1080/09720529.2021.2014139

[16] Zhang, F.Q., Zeng, Q.T., Lu, L.K. and Li, Y.L. (2021) Sentiment Analysis of Movie Reviews Based on Deep Learning. *Journal of Physics*: *Conference Series*, **1754**, Article ID: 012234. https://doi.org/10.1088/1742-6596/1754/1/012234

[17] Si, H.Y. and Wei, X.Y. (2023) Sentiment Analysis of Social Network Comment Text Based on LSTM and Bert. *Journal of Circuits, Systems and Computers*, **32**, Article ID: 2350292. https://doi.org/10.1142/S0218126623502924

[18] Adepu, R. and Tryambak, H. (2023) Sentiment Analysis from Textual Data Using Multiple Channels Deep Learning Models. *Journal of Electrical Systems and Information Technology*, **1**, 56.

[19] Yang, X. and Liu, N.A. (2023) A Collaborative Filtering Algorithm Based on Sentiment Analysis in Review Texts. *Automation and Machine Learning*, **2**, 68-75.

[20] Yang, C.C., Cai, C. and Ni, T.G. (2023) Sentiment Analysis of Medium and Long

Text Based on Feature Fusion Model. *Advances in Computer and Communication*, **4**, 74-79. https://doi.org/10.26855/acc.2023.04.002

[21] He, Z.F., Dumdumaya, C.E. and Machica, I.K.D. (2023) Text Sentiment Analysis Based on Multi-Layer Bi-Directional LSTM with a Trapezoidal Structure. *Intelligent Automation Soft Computing*, **37**, 639-654. https://doi.org/10.32604/iasc.2023.035352

[22] Kumaresan, C. and Thangaraju, P. (2023) ELSA: Ensemble Learning Based Sentiment Analysis for Diversified Text. *Measurement: Sensors*, **25**, Article ID: 100663. https://doi.org/10.1016/j.measen.2022.100663