International Journal of Scientific Research and Engineering Development-– Volume 5 Issue 5, Sep- Oct 2022

Available at www.ijsred.com

RESEARCH ARTICLE OPEN ACCESS

SENTIMENT CLASSIFICATION OF MOVIE REVIEW USING LONG SHORT-TERM MEMORY (LSTM) NETWORK

Ibiyomi M.A.[1], Ajinaja M.O.[2] and Adeola A.A.[3]

[1,2,3]Department of Computer Science, Federal Polytechnic Ile-Oluji, Ondo State, Nigeria

[1] micibiyomi@fedpolel.edu.ng
[2]micajinaja@fedpolel.edu.ng
[3] adeadeola@fedpolel.edu.ng

--***************************----------------------------------

Abstract:

With the massive growth of mining text, the need for sentiment analysis is gradually gaining ground. Companies, businesses, and brands constantly want to beat their nearest competitor to have greater profit and a large customer base. Sentiment analysis aids in the study of user evaluations of a subject so that conclusions can be drawn based on the sentiments gleaned from the reviews. A well-known deep neural networkin sentiment analysis is the Long-Short-Term Memory Networks (LSTM) which was used in this studyto classify movie reviews based on their sentiment and to generate confusion matrices for both training and test sets of data.

Keywords: *Sentiment analysis, lstm, movie, review*

--***************************----------------------------------

1.Introduction

With the advent of internet technologies and the effect on social media, more and more industries, social media platforms, and movie (film and television) platforms have emerged. A lot of these industries are heavily dependent on data and used to infer different analyses. People can easily filter out the products or movies they want to buy or see by reading other people's internet reviews of those products or films. Movie lovers go to read comments or reviews about a movie before deciding whether to watch the movie or not.Different company brands use a collection of opinions of their customers to determine how to produce or make decisions on sales [1]. On the platform side, they gather user feedback to assess user satisfaction and sentiment. For instance, news platforms want to examine netizens' thoughts and feelings, and e-commerce platforms want user feedback to improve their products. Sentiment analysis has developed into a very active field as a result of the high demand.

Sentiment analysis, often known as opinion mining, is a natural language processing (NLP) method for identifying the positivity, negativity, or neutrality of data [2]. Sentiment analysis on textual data is a popular practice among businesses to monitor how their brands and goods are perceived by consumers through online reviews and to better understand their target market.Deep learning and recurrent neural networks are two examples of methods that are used to evaluate the accuracy of sentiment analysis categorization models. An artificial neural network that employs sequential data or time series data is known as a recurrent neural network (RNN). They are implemented into well-known programs like Siri, voice search, and Google Translate [3]. These deep learning algorithms are frequently employed for ordinal or temporal issues, such as language translation, natural language processing (NLP), speech recognition, and picture captioning. It is a kind of neural network in which the output from one stage is fed into the next phase's input. Traditional neural networks have separate inputs and outputs, however, in some situations, such as when it's necessary to anticipatethe next word of a sentence, the previous words are required and hence there is a need to remember the previous words [4].

To solve issues requiring sequences of numbers or speech recognition, recurrent neural networks (RNN) are highly helpful. Long short-term memory (LSTM), a sort of recurrent neural network capable of learning order dependency in sequence prediction tasks, is a specific kind of RNN. Utilizing LSTM has several benefits, one of which is that it solves the Vanishing Gradient Problem. Network training is challenging for Words with a long succession of numbers because of the Vanishing Gradient issue. When using gradient-based techniques to train neural networks, the issue arises (for example, Back Propagation).The parameters of the network's previous layers are challenging to learn and fine-tune as a result of this issue. To get over this issue and collect the data, LSDM networks are used.

The dataset comprises 25,000 customer reviews in the training data and 25,000 customer reviews as test data and 500 set as the value to restrict the number of frequent words. The choice of the dataset was based on a popular movie review rating platform that contained huge collection of reviews about movies. Dataset was obtained from Keras [15] where each review is encoded as a sequence of word indexes. Then we splitthe dataset into 80:20 training and testing data. 20% of the training samples have been used as validation data. We have made the reviews the same in length by zero-padding the shorter reviews so that is easy to train the architectures.The general result from the research includes generating of a confusion matrix for both the training and test data.

Section 2 presents the review of relevant literature. The methodology, data source, text pre-processing, and model architecture are presented in Section 3 while section 4 focuses on results and discussions. The conclusion drawn from the research is presented in Sect. 5

2.Related Works

To enable the model to explicitly learn the sentiment knowledge in Chinese text, [5] proposed a sentiment information-based network model (SINM). The transformer encoder and LSTM were employed as model building blocks. The researchers can automatically identify sentiment knowledge in Chinese text with the aid of a Chinese emotional vocabulary. To learn useful emotional expressions and forecast sentiment tendencies, they developed a hybrid task learning method for SINM. The sentiment knowledge in the text must first be learned through SINM. When emotional information is present, SINM will be more attentive to sentimental information than to pointless information. Studies using the ChnSentiCorp and ChnFoodReviews datasets have revealed that SINM outperforms most other approaches in terms of performance and generalizability.

A sentiment analysis pipeline is introduced in [6] and integrated into the continuing open-source cross-media analysis system. To conduct their research, they used the Apache-Hadoop infrastructure and the lexicon-based sentiment prediction approach from the Stanford coreNLP package, which includes the Recursive Neural Tensor Network (RNTN) model. While the lexicon-based one uses a sentiment dictionary with words annotated with sentiment labels and other essential lexical features, the latter one was trained using Sentiment Treebank with 215,154 phrases, tagged using Amazon Turk. According to the analysis of overall performance, RNTN outperforms the lexicon-based system by 9.88% on variable length positive, negative, and neutral comments. The lexicon-based approach, however, performs better when classifying compliments.Additionally, the team discovered that the F1-score values of the Lexicon-based are higher by 0.16 compared to the RNTN.

[7] examined how well vector representations performed on various text samples and assessed the value of domain-dependent vectors. To compute numerous vector-based attributes and carry out methodical studies to show their efficacy, vector representations were used. For text sentiment analysis of APP evaluations, they achieved F1 85.77%, recall 85.20%, and accuracy 86.35% using simple vector-based features.

To ascertain precise user views of the many elements of the two leading smartphone manufacturers in India, Vivo, and Oppo, in their most recent launches, [8] employed the sentiment analysis technique. The sentiment for each tweet was determined using a hybrid model that combines the Naive Bayes algorithms with the Lexicon Based Sentiment analysis to obtain improved accuracy. The tweets on the two smartphones, the Vivo Nex and the Oppo FindX, were used as individual user input. The next step is to compare the total ratings for each of the two smartphones, which provides a summary of how consumers feel about the two mobile devices.Further study was conducted to compare user opinions of the various features in the two smartphones, which provides businesses with significant input that they can use to either make immediate adjustments or enhance the design of their next models. Results of sentiment analysis submitted by UK energy customers on Twitter were provided in [9]. The first lexicon was picked by the researchers because it was efficient at identifying sentiment-bearing sentences and negative sentiments. Additionally, they merged functions from two sentiment lexica to improve the precision of the sentiment analysis results.They then used a second lexicon to categorize the remaining data. According to experimental data, the procedure improved the accuracy of the outcomes over the conventional approach of using only one lexicon.

The team demonstrated in [10] how sentiment structure and sentiment calculation criteria may be used to solve the issues. The dependency parsing with relationship migration and changed distance produced the sentiment structure, which proved useful in interpreting the sentiment of brief texts. According to the many relationships between the modifier and the emotion word and the contribution of each phrase to the computation of the sentiment of short text, the sentiment of short text was amassed. Results from experiments confirmed that the strategy was effective.

As a further step, [11] tried to determine the sentiment of tweets using sentiment analysis techniques from Logistic Regression, VADER, and BERT. While they may be adapted based on the domain, the suggested analytic approaches were more sensitive to sentiment expressions in social media situations. The sentiment analysis algorithm was not included in the preprocessing or subsequent processes, despite the use of 3 separate algorithms. It would be easier to compare the three suggested sentiment analysis algorithms if they underwent the same processing stages.Additionally, this research'snumerous practical implications included gathering public opinion for use by government authorities or even health officials so they may base their decisions on the findings. [12] created a sentiment analysis system for reviews of a picturesque location from customers. In the context of deep learning, it was based on CNNs constructed on LSTM for text feature extraction. To produce a robust text feature vector, the CNNs based on the LSTM model repeatedly operated their convolutional filters on the output matrix of LSTM. In the studies, the ideal parameter settings for each CNN and LSTM component are first determined independently. Following that, the optimal parameter configuration for the system's integration recognition frame was found to revolve around the best performance of each component. According to experimental findings, using CNNs based on LSTM models improves sentiment analysis accuracy over using only one CNN or LSTM model by 3.13% and 1.71%, respectively.

[13] used deep learning techniques to examine word embedding models (Word2Vec, Glove) in tweets. Here, they looked at how sentiment analysis may deal with long-term dependencies by including memory into a network model for prediction and visualization. They used recurrent neural networks (RNNs) and Long-Short Term Memory networks (LSTMs) for this. The findings demonstrated improved substantial classification accuracy taught at 80% for the training set and 20% for the testing set, demonstrating the dependability of our models for future forecasting. The Bidirectional LSTM Model (Bi-LSTM) is employed for additional inquiry investigations to enhance this performance.[14] described an approach for categorizing sentiment using a deep NN model called BiLSTM recurrent neural networks for a range of sentiment analysis tasks, including text categorization, cross-lingual issues, verbal and contextual evaluation, and industry news assessment, among others.

3.Methodology

This dataset is one of the largest IMDb movie reviewdatasets [16]. The dataset contains 50,000 movie reviewsbelonging to two categories either positive or negative. Wehave obtained this dataset from Keras [15] where each reviewis encoded as a sequence of word indexes. Then we splitthe dataset into 70:30 training and testing data respectively. 20% of thetraining samples have been used as validation data. We havemade the reviews the same in length by zero-padding the shorterreviews so that is easy to train the architectures

3.1.Long Short-Term Memory

An RNN modified to accommodate long-term reliance is called LTSM. While LSTM and RNN both include temporal loops inside their layers, the main distinction between the two is a memory cell that may store or update data based on the input phrases given [17]. There are three Gates in an LSTM cell. One is the Forget Gate layer, a Sigmoid layer that chooses which information needs to be erased from the memory cell. What new data will be stored in the memory cell is decided by the input gate layer, which is the next layer. A Sigmoid layer and a Tanh layer are the next two divisions.

Mathematically, the gate can be defined by

$$F_t = \sigma(W_f \,.\, [h_{t-1}, x_t] + b_f) \quad (1)$$

$$I_t = \sigma(W_i \,.\, [h_{t-1}, x_t] + b_f) \quad (2)$$

$$O_t = \sigma(W_o \,.\, [h_{t-1}, x_t] + b_o) \quad (3)$$

$$h_t = O_t * \tanh(C_t) \quad (4)$$

Here $F_t I_t$, and O_t represent the forget, input and output gate Layer. b_f ,b_o represents the bias factor of those layers. x_t and h_t are the input and output of the current unit and h_{t-1} is the output of the previous input $x_{t-1.].\sigma.}$σ represents the sigmoid layer with sigmoid activation function and tanh represents tanh layer with a tanh activation function[18] [19].

The input to our LSTM network is the same as the movie review. In the LSTM network, we have used only one LSTM layer after the Embedding Layer to avoid the complexity of the model. We have used 3 LSTM units in the LSTM layer. The output layer is also the same as in the movie reviews.

4. Results and Discussion

4.1The embedding layer

Each word may be turned into a fixed-length vector with a specified size using the embedding layer in an LSTM network. For incoming words, word vectors are produced using it. The output of the embedding layer serves as the input to the LSTM layer, placing it between the two.Fig. 1 shows the number of parameters used for the embedding layer which stands at 16,000 and 3 LSTM loops was used.

Layer (type)	Output Shape	Param #
embedding_4 (Embedding)	(None, None, 32)	16000
lstm_3 (LSTM)	(None, None, 32)	8320
lstm_2 (LSTM)	(None, None, 32)	8320
lstm_1 (LSTM)	(None, 32)	8320
dense_4 (Dense)	(None, 1)	33

Fig. 1- layer type and corresponding parameter

4.2 Fit Model

We picked the number 25 as the epoch. The term "epoch" refers to how many times the neural network has been trained using all of the training data in a single cycle. The number of times the learning algorithm will go over the full training dataset is determined by a hyperparameter. We used batch size 128 to specify the number of samples to process before changing the internal settings. More samples and predictions were made to compare them to the anticipated output variables. Additionally, we used 80% of the data for training and 20% for testing when separating variation.

Figure 2 shows the plot of loss and accuracy. After 25 epochs, it was observed both the accuracy and loss were close to each other with overfitting hugely reduced. The closeness of both lines indicates the better performance of the model in general.

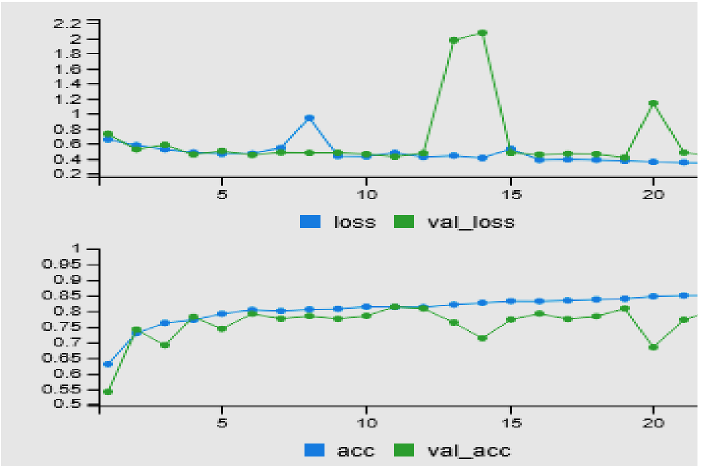

Fig. 2

Table 1 shows the confusion matrix for the trained data while table 2 shows the confusion matrix for the test data.It also explores the positive and negative sentiments for both the training and test data showing the sentiment classification performance of the model.

Table 1 – Confusion matrix for trained data

	Actual	
Predicted	**0**	**1**
0	10981	1850
1	1513	10650

Table 2 – Confusion matrix for test data

	Actual	
Predicted	**0**	**1**
0	10604	2084
1	1896	10416

5. Conclusion

Sentiment analysis is becoming very important as the amount of online data increases at a huge rate. For this reason, we need sentiment analysis on social media or online reviews for predicting and forecasting public opinion. We have found how to use the LSTM model and determine sentiment classification using a confusion matrix. For future work, we have decided to use the convolutional neural network in other fields of natural language processing and evaluate the performance of used methods in those fields

References

[1] G. Li, Q. Zheng, L. Zhang, S. Guo and L. Niu, "Sentiment Infomation based Model For Chinese text Sentiment Analysis," 2020 IEEE 3rd International Conference on Automation, Electronics and Electrical Engineering (AUTEEE), 2020, pp. 366-371, doi: 10.1109/AUTEEE50969.2020.9315668.

[2] Sentiment Analysis Guide. MonkeyLearn. (n.d.). Retrieved September 27, 2022, from https://monkeylearn.com/sentiment-analysis/

[3] IBM Cloud Education. (n.d.). What are recurrent neural networks? IBM. Retrieved September 27, 2022, from https://www.ibm.com/cloud/learn/recurrent-neural-networks

[4] IBM Cloud Education. (n.d.). What are recurrent neural networks? IBM. Retrieved September 27, 2022, from https://www.ibm.com/cloud/learn/recurrent-neural-networks

[5] G. Li, Q. Zheng, L. Zhang, S. Guo and L. Niu, "Sentiment Infomation based Model For Chinese text Sentiment Analysis," 2020 IEEE 3rd International Conference on Automation, Electronics and Electrical Engineering (AUTEEE), 2020, pp. 366-371, doi: 10.1109/AUTEEE50969.2020.9315668.

[6] Y. Woldemariam, "Sentiment analysis in a cross-media analysis framework," 2016 IEEE International Conference on Big Data Analysis (ICBDA), 2016, pp. 1-5, doi: 10.1109/ICBDA.2016.7509790.

[7] X. Fan, X. Li, F. Du, X. Li and M. Wei, "Apply word vectors for sentiment analysis of APP reviews," 2016 3rd International Conference on Systems and Informatics (ICSAI), 2016, pp. 1062-1066, doi: 10.1109/ICSAI.2016.7811108.

[8] N. SrivatsAthindran, S. Manikandaraj and R. Kamaleshwar, "Comparative Analysis of Customer Sentiments on Competing Brands using Hybrid Model Approach," 2018 3rd International Conference on Inventive Computation Technologies (ICICT), 2018, pp. 348-353, doi: 10.1109/ICICT43934.2018.9034283.

[9] V. Ikoro, M. Sharmina, K. Malik and R. Batista-Navarro, "Analyzing Sentiments Expressed on Twitter by UK Energy Company Consumers," 2018 Fifth International Conference on Social Networks Analysis, Management and Security (SNAMS), 2018, pp. 95-98, doi: 10.1109/SNAMS.2018.8554619.

[10] J. Li and L. Qiu, "A Sentiment Analysis Method of Short Texts in Microblog," 2017 IEEE International Conference on Computational Science and Engineering (CSE) and IEEE International Conference on Embedded and Ubiquitous Computing (EUC), 2017, pp. 776-779, doi: 10.1109/CSE-EUC.2017.153.

[11] A. J. Nair, V. G and A. Vinayak, "Comparative study of Twitter Sentiment on COVID - 19 Tweets," 2021 5th International Conference on Computing Methodologies and Communication (ICCMC), 2021, pp. 1773-1778, doi: 10.1109/ICCMC51019.2021.9418320.

[12] J. Gao, R. Yao, H. Lai and T. -C. Chang, "Sentiment Analysis with CNNs Built on LSTM on Tourists Comments," 2019 IEEE Eurasia Conference on Biomedical Engineering, Healthcare and Sustainability (ECBIOS), 2019, pp. 108-111, doi: 10.1109/ECBIOS.2019.8807844.

[13] R. Monika, S. Deivalakshmi and B. Janet, "Sentiment Analysis of US Airlines Tweets Using LSTM/RNN," 2019 IEEE 9th International Conference on Advanced Computing (IACC), 2019, pp. 92-95, doi: 10.1109/IACC48062.2019.8971592.

[14] J. S. Vimali and S. Murugan, "A Text Based Sentiment Analysis Model using Bi-directional LSTM Networks," 2021 6th International Conference on Communication and Electronics Systems (ICCES), 2021, pp. 1652-1658, doi: 10.1109/ICCES51350.2021.9489129.

[15] F.Chollet, "Keras." (2015).

[16] Maas, A. L., Daly, R. E., Pham, P. T., Huang, D., Ng, A. Y., & Potts, C. , "Learning word vectors for sentiment analysis." Proceedings of the 49th annual meeting of the association for computational linguistics: Human language technologies-volume 1. Association for Computational Linguistics, 2011, pp. 142-150.

[17] M. R. Haque, S. Akter Lima and S. Z. Mishu, "Performance Analysis of Different Neural Networks for Sentiment Analysis on IMDb Movie Reviews," 2019 3rd International Conference on Electrical, Computer & Telecommunication Engineering (ICECTE), 2019, pp. 161-164, doi: 10.1109/ICECTE48615.2019.9303573.

[18] Hochreiter, S., &Schmidhuber, J., "Long short-term memory." Neural computation 9.8, 1997, pp. 1735-1780.

[19] Greff, K., Srivastava, R. K., Koutník, J., Steunebrink, B. R., &Schmidhuber, J. , "LSTM: A search space odyssey." IEEE transactions on neural networks and learning systems 28.10, 2016, pp. 2222-2232.

