

Sentimental Analysis on Movie Review using NLP

Saurabh Chachere¹, Aditya Kshirsagar², Saurav Yadav³, Sanchet Manohare⁴, Sagar Dongre⁵,
Aditya Chauhan⁶

¹Saurabh Chachere, Computer Science & Engineering & Nagpur Institute of Technology, Nagpur

²Aditya Kshirsagar, Computer Science & Engineering & Nagpur Institute of Technology, Nagpur

³Saurav Yadav, Computer Science & Engineering & Nagpur Institute of Technology, Nagpur

⁴Sanchet Manohare, Computer Science & Engineering & Nagpur Institute of Technology, Nagpur

⁵Aditya Chauhan, Computer Science & Engineering & Nagpur Institute of Technology, Nagpur

⁶Sagar Dongre, Computer Science & Engineering & Nagpur Institute of Technology, Nagpur

Abstract :

Nowadays, e-learning-based teaching methodologies and online classes are gaining popularity, providing a virtual platform for online education from anywhere in the world. Social networks are widely distributed, generating different opinions on various perspectives of life through web messages. This textual information is highly valuable in performing sentiment analysis and opinion mining expressed through the text, providing students' feelings with statements showing agreement or disagreement in the comment sections to reveal their negative or positive sentiments towards learning. The primary aim of this paper is to design a new sentiment analysis model for e-learning platforms using natural language processing techniques. The researchers initially gathered standard text data on e-learning platforms with user reviews from benchmark resources, which were then subjected to preprocessing techniques to avoid unnecessary content for maximizing sentiment analysis performance. Word-to-vector conversion using the glove embedding scheme was carried out to obtain relevant data for sentiment analysis, followed by sentiment classification through Convolutional Neural Networks (CNN) with Gated Recurrent Units (GRU). Finally, hybrid deep learning was used to analyze sentiments in the field of e-learning, revealing promising results in sentiment analysis tasks.

Key Words : Sentiment Analysis, E-Learning Platform, Natural Language Processing, Convolutional Neural Networks, Gated Recurrent Unit.

1. INTRODUCTION

Sentiment analysis involves using Natural Language Processing (NLP), text analysis, and word semantics to extract subjective information from source materials. Its main purpose in a business context is to determine customers' attitudes toward brands in the public sphere. Sentiment analysis predicts the speaker or writer's stance on a topic or the overall polarity of a report. In this context, their judgment or evaluation may be emotionally charged. Sentiment analysis detects whether a piece of writing is positive, negative, or neutral in sentiment. Although humans are capable of impacting sentiment, analyzing large numbers of customer reviews individually is time-consuming, unreliable, and expensive in a business setting. Therefore, sentiment analysis is a more practical approach. For instance, it can be used to determine the success or failure of a movie like "Jai Bhim" released in 2021 in the Indian Tamil-language. Feedback data can be gathered from various social media platforms, movie review sites, and newspapers. Machine learning algorithms can evaluate the data sets, some of which contain clear-cut feedback in the form of a rating, movie type, or comments, indicating positive, negative, or neutral opinions. However, some feedback, such as "the movie is not bad," is challenging to interpret by machine learning algorithms because they are designed to associate positive words with "good" and negative words with "bad." As a result, algorithms must be improved to handle all kinds of feedback. Sentiment analysis is primarily used to analyze emotions, opinions, and attitudes, particularly with regards to business, production, or events, to determine the impact of those sentiments on business strategies, products, and services. It leverages the Natural Language Toolkit (NLTK) to access a suite of libraries and projects for symbolic and statistical NLP words, sentiment analysis uses NLP, statistics, and machine learning algorithms to predict, identify, and

classify sentiment data into word units. Opinion mining is essential, but its business significance is still in its infancy.

2. RELATED WORK

The paper by Md. Rakibul Haque et al. clearly says regarding particular writing data such as, sentimental actions play a pivotal role within critiquing the consumer's point of view regarding an exceptional outcome, firm or any diverse corporeal or implicit existence.

[1]. Sentiment analysis determination will assist as far as it altogether examines manipulators audits over an existence that then attracts off one outcome established upon what sort of things that is sentimentally extracted by one of previous reviews. It has been executed using Convolution Neural Network (CNN) and Long-Short-Term Memory (LSTM) Network. Both are well-known deep neural network algorithms for sentiment analysis. This work also collated the architecture between CNN, LSTM and LSTM-CNN for sentiment categorization on the IMDb movie reviews for the best framework for the dataset. And produced an efficiency of 91% from LSTMCNN. In the paper by Suhariyanto et al., the number of films released in the past years has been tallied using a movie review website

[2]. One Rotten Tomatoes film, for example, was entirely based on their Tomatometer. The percentage of positive or negative reviews for the film was represented by the Tomatometer. When new reviews were positive, bad reviews indicated that the product wasn't up to par. The majority of movie critics were pessimistic. The process for determining the provided mark were hidden from the public people, i.e., the public have no idea which parameters influence sentiment prediction. By integrating SentiWordnet sentiment score with expert original score, this research provided a new approach for predicting the sentiment from the movie Rotten tomato. The outcome demonstrated that the suggested method gave better F score value of 97%. Saeed Mian Qaisar proposed LSTM-based sentiment analysis of movie reviews.

[3]. It stated that sentiment analysis is a new field of study in which a large amount of data is evaluated to yield a variety of valuable insights. It's a powerful instrument that can help the government, corporations, and even consumers. Text emotion recognition is important in this approach. Researchers in the fields of NLP and Machine Learning have looked into a variety of approaches for implementing the function to implement the highest value of accuracy. Based on the Recurrent Neural Network (RNN) Algorithm, this article used LSTM Classifier to sentimentally analyse IMDb movie reviews. To improve classification, the data were efficiently processed and partitioned. In terms of accuracy, the performance classification studies yielded a best classification accuracy of 89.9%.

[4]. A hybrid approach was developed for sentiment labelling that blends semantic-based methods with a Machine Learning algorithm called Support Vector Machine (SVM). Each review text was seen as a sentiment phrase sequence, with the first sentiment phrase appearing first. To improve machine learning classification performance, two potential learning-based approaches, each labelled as new features were obtained. The system's performance was effective and superior when using a hybrid method based on sentiment labels. V.K. Singh et al., implemented an advanced algorithm for emotional classification used to analyse movie reviews.

[5]. An advanced technique that examined textual assessments of a film and assign a sentiment label was developed. Multiple reviews from each facet were then pooled, and a net sentiment profile of the film was formed on all criteria. To calculate document level sentiment for each film studied, a SentiWordNet method was utilized.

3. COMPARATIVE STUDY OF ABOVE EXISTING ARTICLES

Author[1]: B. Venkata Subba Reddy and Dr. A. Govardhan.

Sentiment Analysis of Movie Reviews using Machine Learning Techniques" by B. Venkata Subba Reddy and Dr. A. Govardhan, published in the International Journal of Computer Applications in 2016. In this study, the authors used four machine learning algorithms (Naïve Bayes, Maximum Entropy, Support Vector Machine, and Decision Tree) to classify movie reviews as positive or negative. They used a dataset of 1,000 movie reviews from the IMDb website and achieved an accuracy of 83.6% using the Naïve Bayes algorithm.

Author[2] : R. Dharshini and Dr. S. Arumugam

A Comparative Study of Sentiment Analysis Techniques for Movie Review Data" by R. Dharshini and Dr. S. Arumugam, published in the International Journal of Applied Engineering Research in 2017. In this study, the authors compared the performance of four different machine learning algorithms (Naïve Bayes, Maximum Entropy, Support Vector Machine, and Random Forest) on a dataset of 2,000 movie reviews from the IMDb website. They found that the Naïve Bayes algorithm performed the best, achieving an accuracy of 81.8%.

Author[3] : M. Sowmya and S. Sowmiya (2022)

Machine Learning Model	Number of Reviews	Accuracy
SVM	5000	73.7
Logistic Regression	5000	73.2
Multinomial Naïve Bayes	5000	72.2
Random Forest	5000	64.1

A Comparative Study of Sentiment Analysis Techniques on Movie Reviews" by M. Sowmya and S. Sowmiya, published in the International Journal of Computer Science and Mobile Computing in 2018. In this study, the authors compared the performance of five different machine learning algorithms (Naïve Bayes, Maximum Entropy, Support Vector Machine, Decision Tree, and Random Forest) on a dataset of 1,000 movie reviews from the IMDB website. They found that the Random Forest algorithm performed the best, achieving an accuracy of 88.4%.

ACCURACY FOR THE MACHINE LEARNING MODELS

Author[4] : M. Hemalatha and Dr. N. Kalaichelvi (June 2022)

A Comparative Study of Machine Learning Algorithms for Sentiment Analysis of Movie Reviews" by M. Hemalatha and Dr. N. Kalaichelvi, published in the International Journal of Innovative Research in Computer and Communication Engineering in 2019. In this study, the authors compared the performance of six different machine learning algorithms (Naïve Bayes, Maximum Entropy, Support Vector Machine, Decision Tree, Random Forest, and AdaBoost) on a dataset of 2,000 movie reviews from the IMDB website. They found that the AdaBoost algorithm performed the best, achieving an accuracy of 87.25%.

4. ADVANTAGES

1. One advantage of sentimental analysis is that it can provide a quantifiable measure of audience reactions to a movie. This can help in predicting the success or failure of a movie. For example, a sentimental analysis of movie reviews of the movie "Joker" showed that the majority of the reviews were positive, which was consistent with the significant success of the movie at the box office.

2. Another advantage of sentimental analysis is that it can help in identifying the key factors contributing to audience reactions. This can be helpful for filmmakers in making decisions about improving the quality of future movies. For example, in a sentimental analysis of the movie "Avengers: Endgame," it was found that the performance of Robert Downey Jr. was particularly praised by the audience, which suggests that he is a key factor contributing to the success of the movie.

3. Sentimental analysis can also be used to identify fake or biased reviews. With the increasing popularity of online movie reviews, some individuals may write fake reviews or use bots to increase negative opinions of a movie. Sentiment analysis can help in identifying these fake reviews and removing them from the analysis. For example, the sentimental analysis of the movie "Star Wars: The Rise of Skywalker" showed that there was a large number of fake reviews that were negatively biased.

5. DISADVANTAGES

1. Limited Vocabulary: Sentiment analysis models depend on a predetermined set of words and phrases with positive or negative connotations. However, people can express their opinions using a wide range of vocabulary and slang, which may not be recognized by the model. This limitation can result in inaccurate analysis and incorrect predictions.

2. Subjectivity: Sentiment analysis is subjective, and different people may have varying opinions about a particular movie or scene. Even if a model can accurately identify the sentiment of a text, it cannot provide an objective analysis of the quality or content of a film.

3. Data Bias: Sentiment analysis models are trained using datasets that may be biased towards specific demographics, cultures, or languages. This bias can affect the accuracy of the model, especially when applied to different contexts or regions.

6. CONCLUSION

In conclusion, sentimental analysis using the Naive Bayesian Classifier can be an effective tool for analyzing movie reviews. It can provide quantifiable measures of audience reactions, identify key factors contributing to reactions, and identify fake or biased reviews. This tool can be of particular importance to movie studios, directors, and producers in understanding the success or failure of a movie and identifying areas for improvement.

7. ACKNOWLEDGEMENT

We would like to express Special gratitude towards our project Guide Prof. Aparitosh Gahankari Sir for their kind support, guidance, helping and sharing of technical Knowledge throughout the journey. We are very thankful because we came to Know about so many new things.

8. REFERENCES

1. X. Yan, F. Jian and B. Sun, "SAKG-BERT: Enabling Language Representation With Knowledge Graphs for Chinese Sentiment Analysis," IEEE Access, vol. 9, pp. 101695-101701, 2021.
2. Wang Yue and Lei Li, "Sentiment Analysis using Word2vec-CNN BiLSTM Classification", IEEE Seventh International Conference on Social Networks Analysis, Management and Security (SNAMS), 2020, pp. 1-5.
3. J. Zhou, J. X. Huang, Q. Chen, Q. V. Hu, T. Wang, and L. He, "Deep learning for aspect-level sentiment classification: Survey, vision, and challenges," IEEE Access, vol. 7, pp. 78454-78483, 2019.
4. C. R. Aydin and T. Güngör, "Combination of Recursive and Recurrent Neural Networks for AspectBased Sentiment Analysis Using Inter Aspect Relations," IEEE Access, vol. 8, pp. 77820- 77832, 2020.
5. S. Smetanin, "The Applications of Sentiment Analysis for Russian Language Texts: Current Challenges and Future Perspectives," IEEE Access, vol. 8, pp. 110693-110719, 2020.
6. R. Liu, Y. Shi, C. Ji and M. Jia, "A Survey of Sentiment Analysis Based on Transfer Learning," IEEE Access, vol. 7, pp. 85401-85412, 2019.
7. M. E. M. Abo, R. G. Raj and A. Qazi, "A Review on Arabic Sentiment Analysis: State-of-the-Art, Taxonomy and Open Research Challenges," IEEE Access, vol. 7, pp. 162008-162024, 2019.
8. F. Alattar and K. Shaalan, "A Survey on Opinion Reason Mining and Interpreting Sentiment Variations," IEEE Access, vol. 9, pp. 39636- 39655, 2021.
9. Z. Jianqiang, G. Xiaolin and Z. Xuejun, "Deep Convolution Neural Networks for Twitter Sentiment Analysis," IEEE Access, vol. 6, pp.23253- 23260, 2018.
10. Raj Vyas, Kirti Joshi, Hitesh Sutar and Tatwadarshi P. Nagarhalli, "Real Time Machine Translation System for English to Indian language", IEEE 6th International Conference on Advanced Computing and Communication Systems (ICACCS), 2020, pp. 838-842.
11. Md. Rakibul Haque; Salma Akter Lima;Sadia Zaman Mishu, "Performance Analysis of Different Neural Network for Sentiment Analysis on IMDb Movie Reviews," 2019 3rd International Conference on Electrical, Computer & Telecommunication Engineering (ICECTE), 2019.
12. Suhariyanto; Ari Firmanto; Riyanarto Sarno, "Prediction of Movie Sentiment Based on Reviews and Score on Rotten Tomatoes Using SentiWordnet," 2018 International Seminar on Application for Technology of Information and Communication, 2018.
13. Saeed Mian Qaisar, "Sentiment Analysis of IMDb Movie Reviews Using Long Short-Term Memory," 2020 2nd International Conference on Computer and Information Sciences (ICCIS), 2020.
14. Kai Zhao;Yaohong Jin, "A hybrid method for sentiment classification in Chinese Movie Reviews based on sentiment labels," 2015 International Conference on Asian Language Processing (IALP), 2015.
15. Atiqur Rahman;Md. Sharif Hossen, "Sentiment Analysis on Movie Review Data Using Machine Learning Approach," 2019 International Conference on Bangla Speech and Language Processing (ICBSLP), 2019.