IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 1, April 2024

Implementation of Sentiment Analysis using Deep Learning

Siddhesh Bhor[1], Sanskruti Bhor[2], Borhade Rani[3], Fodse Aniket[4], Prof. Dipali Dube[5]
Students, Department of Computer Engineering[1,2,3,4]
Guide, Department of Computer Engineering[5]
Vidya Niketan College of Engineering, Ahmednagar, Maharashtra, India

Abstract: *The usage of internet as well as online platforms is booming day by day. Understanding the public opinions can be beneficial for business and political entities in making strategic decisions. In light of this, sentiment analysis plays an important role to understand the polarity of the public opinions. Today, this ocean of data can be used for the fruitful purposes. Analysis of sentiment textual posts can supply knowledge and information that can be used in citizen opinion polling, business intelligence, social contexts, and Internet of Things (IOT)-mood triggered devices. The main focus is the sentiment analysis based on Emotional Recognition (ER). We are going to implement the model for the prediction of sentiment on the basis of general words length words and emoji over any online platform. We are also going to compare the traditional concept related to sentiment analysis. Using the sentiment analysis we can try to control some illegal activities that are post on online platforms such as Movies Reviews Social media etc.*
In this model we divide the process in the six phases, first is Data overview, second is Data preprocessing third is Feature Engineering fourth is Model selection, fifth is Model Evaluation and last is Model Deployment We can try to take result in two class those are positive and negative or Good and Bad Sentiment. We can try to make two columns first one for data sample and second for the result. Therefore, the above process conclude that the accuracy of the sentiment analysis will be increases by using the length words and emoji from the data.

Keywords: Emotional Recognition

I. INTRODUCTION

The rapidly increasing data on the online platform can be use for the various analysis process for use them to achieve some goals or the targets for the various fields. The data can be used to predict the future of the decision that will be make. As well the data can be use for analysis of sentiment of the user. online platform users can post the some information about the any thing with some emotions now a day.so their emotions or sentiment can be analysis possible.
Analysis of online platform's (movie reviews, social media etc.) sentiment textual posts can supply knowledge and information that can be used in citizen opinion polling, business intelligence, social contexts Sometime the Bad or Good sentiments can be useful for prediction of the persons mental health or the behavior about the some thing. Online movie reviews can be use to predict the movie is good or bad. Reviews can be used to other peoples to judge the movie, some reviews can be make more impact on the other users or clients.
Reviews analysis can be used at other sector also for predicting their performance will be good or bad some death cases can be related to the person's online platform activities they can deliver the some specific note on the online platforms. Sentiment Analysis can be used to improve the content quality and used to monitoring the world or specific activities.

II. LITERATURE REVIEW

This chapter will provide research of different paper, including author requirements, research perspective, and overview of project idea, general constraints. In addition, it will also provide the information of algorithm and functionality needed for this system - such as algorithm, techniques, functional requirements and performance requirements. With the event of Information and Communication Technology, various varieties of information security threats may be seen.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 1, April 2024

In this paper, Kian Long Tan et al. [1].In Sentiment Analysis With Ensemble Hybrid Deep Learning Model paper the compare and combine the Robustly optimized Bidirectional Encoder Representations from Transformers approach (Roberta), Long Short-Term Memory (LSTM) approach can be done.

A.Athar et al. [2] In Sentimental analysis of movie reviews using soft voting ensemble- based machine learning paper proposed thisstudy of sentiment analysis on the movie reviews using voting by use machine learning. In the paper natural language processing can be consider for analysis the reviews.

Ashima kukkar et al. 131. Sentiment analysis can be done on the basis oflengthy words in this paper. That was be achieve on the such process the happpyyy can be present in the text of such review that can be stem into happy and the extended character score can be calculated. Length word can be use for the increase the accuracy of sentiment analysis.

Ilic et al. [4] This method introduced a richer word representation and applies it to a variety of NLP(Natural language processing) tasks, but it still trained the core task model from the start and treats pretraining embedding as a fixed parameter, limiting its impact.

Andrew L. Mans et al. [5] By researchers at Stanford University wherein they used unsupervised learning to cluster the words with close semantics and created word vectors. They ranvarious classification models on these wond vectors to understand the polarity of the reviews. This approach is particularly useful in cases when the data has rich sentiment content and is prone to subjectivity in the semantic affinity of the words and their intended meanings

Pang et al. [6] polarity detection of movie reviews and product reviews. They have also worked on creating a multi-class classification of the review and predicting the reviewer rating of the movie/product. These works discussed the use of Random Forest classifier and SVMs for the classification of reviews and on the use of various feature extraction techniques. One major point to be noted in these papers was exclusion of a neutral category in classification under the assumption that neutral texts lie close to the boundary of the binary classifiers and are disproportionately hard to classify.

Tumasjan et al. [7] There are many sentiment analysistools and software existing today that are available for free or under commercial liceme. With the advent of microblogging. sentiment analysis is being widely used to analyze the public sentiments and draw inferences out of these. One famous application was use of Twitter to understand the political sentiment of the people in context of German Federal elections.

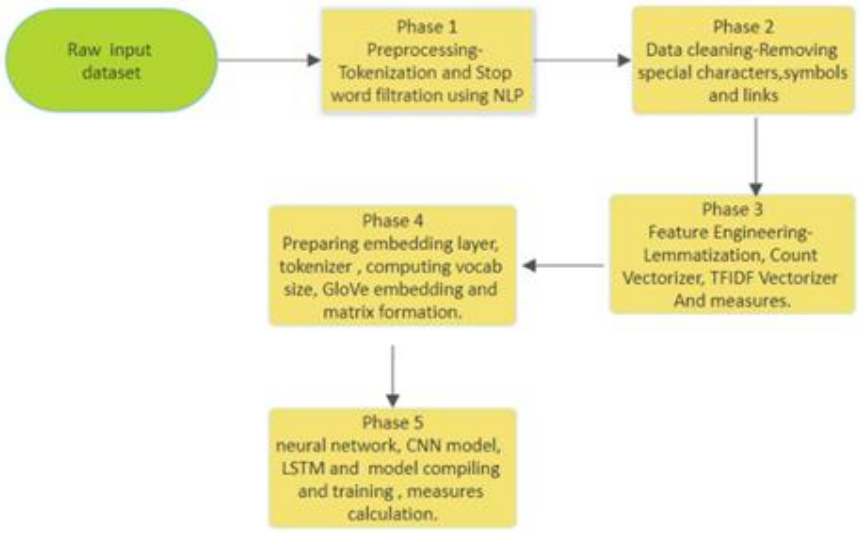
Chen et al. [8] suggested a strategy that combines the benefits of machine learning and information retrieval approaches in their paper, Authors used semantic orientation indexes to feed a buck-propagation several network and discovered that the proposed method improves sentiment categorization performance while also saving time during training

Zhang and Wallace et al. [9] developed a model utilizing convolution neural network approach that also includes pre-trained word embeddings to categories texts at the sentence level, demonstrating that a basic convolution neural network with a modest number of hyper parameters, coupled with static visual words can obtain excellent outcomes on a variety of benchmarks

Howard and Ruder et al. (10) Tn 2018 Universal Language Model Fine-tuning (ULLMET) method established by Howard and Rider that can be applied to natural language processing for the text classification that can be proposed in under Universal language model tine-tuning forte classification paper.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 1, April 2024

III. SYSTEM ARCHITECTURE



IV. EXPERIMENTALMETHODOLOGY

The methodology of the proposed system involves the phases described below and in Figure 5. The raw data from the different sources, which is in the unstructured format, is taken as input to the Tokenization phase.

PHASE I: TOKENIZATION

In this phase, the data is divided into tokens by using some delimiters. Mainly, Space, Comma, Hash (#), @, are used as delimiters.

PHASE II: STOP WORD ANALYSIS

The tokens formed in Tokenization are fed into this phase in which Stop word removal will be performed. Stop word is the commonly used word that does not have any meaning in itself; such as ‘‘the’’. These words do not play any role in determining the sentiment. So, they are removed. Further emoji’s are also removed from the text.

PHASE III: NORMALIZATION AND SENTI-SCORE GENERATION

The remaining tokens are sent to this phase in which the normalization of tokens will take place. In parallel, the same tokens will be sent to phase IV. Every content word has some Senti-score in the Senti WordNet. The normalized word’s Senti-score will be extracted from the Senti-WordNet in the Senti-Score generation.

PHASE IV: TAIL EXTRACTION AND INTEN-SIFIED TAILSCORE

Meanwhile, a separate table will be formed for each token saved in the previous phase. The first column will contain the unique characters of the respective tokens stored in the variables. For the second column of the table, it uses the respective normalized word which is the output of the normalization module. Count the frequency of each character by using the normalized word and store it corresponding to the character in the first column. In the third column, take the lengthened word which is saved in the variable in previous phase and the normalized word generated as the output of the normalization module, as the input. Compare these two words of the respective tokens and count the frequency of the character in the tail and insert it corresponding to the character in the first column. After the tables for all the tokens are generated, all the cells of the third column of the respective tables are added and stored in the Tail Score Database as each token’s tail score. In this phase, the tail score of each token is retrieved from the database and multiplied with 0.01 and stored in some variables namely ‘‘intensified tail score’’.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 1, April 2024

PHASE V: AGGREGATED INTENSIFIED SENTI-SCORE

After the completion of the fifth phase, the Intensified Senti-score is calculated by the aggregation of the output of Senti-score generation and intensified tail score. Further these scores are used to make the final prediction. The calculated scores are added to the score of sentences and the score for a review is calculated. The review classification threshold value is set to 0.5. Based on a threshold value, all reviews of a dataset are evaluated to determine whether they are positive, negative, or neutral. Anything below a score of - 0.05 is classified as negative, while Anything over 0.05 is classified as positive and anything between is classifies as neutral.

V. OUTPUT

Fig. Positive Sentiment

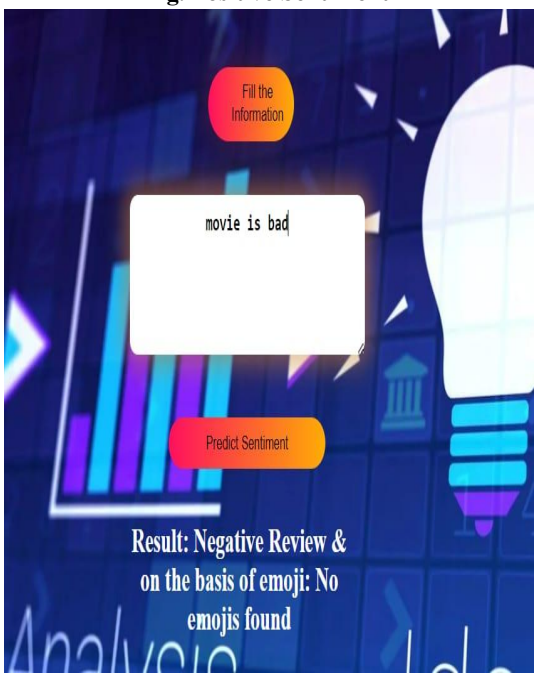


Fig. Negative Sentiment

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 1, April 2024

VI. CONCLUSION

In conclusion, this project creating a more adaptive and user-centric sentiment analysis sys- team. By focusing on functional requirements like natural language understanding, adaptability. and context awareness, it aims to provide a sentiment analysis system that can seamlessly un- Der stand and execute a wide range of tasks while learning and adapting to new challenges. Simultaneously, stringent non-functional requirements related to security, reliability, and privacy will ensure the system's trustworthiness and data protection. Through this project, the goal is to bridge the gap between user and business expectations and current sentiment analysis system capabilities, creating a system that can beuse to predict the sentiment based on user review more accurately. It can be used in business as well as social fields.

REFERENCES

[1] Kian Long Tan, Chin Poo Lee. "Sentiment analysis with ensemble hybrid deep learning model-2022
[2]A. AtharHC.Kim, "Sentiment analysis of movie reviews using soft voting ensemble based machine learning-2022
[3] Ashima kukkar Rajm mohana, "Improving sentiment analysis in social media by handling lengthened words", 2023
[4] S.Ilić, E. Marnese-Taylor, J. A. Balazs, and Y. Matsuo, "Deep contextualized word representation for detecting sarcasun and inamy 2018. arXiv: 1809.09795
[5]Andrewl. Mass, Raymond E. Daly, Peter Pham Dan Huang Andrew Y. Ng and Christo pher Potts, "Learning Word Vectors for Sentiment Analysis" 2021.
[6] Pang BocLec Lilian, Vaitbyanathan, Shivakumar. "Thumbs up Sentiment Classification wing Machine Learning Techniques. Proceedings of the Conference on Empirical Meth ods in Natural Language Processing (EMNLP) 2000
[7]Tumanjan, Andranik O.Sprenger. Tum G Sandner Philipp M. Welpe Isabell, Predicting Elections with Twitter: What 140 Characters Reveal about Political Sentiment Procondings of the Fourth International AAAI Conference on Weblogs and Social Media 2000
[8]L-S-Chen, C- Lin and H.-J. Chis. "A neural network-based approach for semimem classification in the blagophorn, 1. Informetrics, wil. 5. 2. pp. 313-322, 2019.
[9] Y. Zhang and B. Wallars, "A sensitivity analysis of fand peactitioners guide in).consoleGonal neural networks Sar semenceclassificatum 201 KarXiv:1510:03830
[10] Howard and S. Roder, "Universal language model fine-tuning for text classificabun", 201810001,0614.