International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 5 May 2022, pp: 726-731 www.ijaem.net ISSN: 2395-5252

Sentiment analysis of product reviews using Deep Learning techniques

[1]Ms.M.P.Geetha, [2]P.Prethika, [3]J.Valentine Nithish, [4]P.S.Vignesh

[1]*Assistant Professor,Sri Ramakrishna Institute of Technology,Tamil Nadu.*
[2]*Student,Sri Ramakrishna Institute of Technology,Tamil Nadu.*
[3]*Student,Sri Ramakrishna Institute of Technology,Tamil Nadu.*
[4]*Student,Sri Ramakrishna Institute of Technology,Tamil Nadu.*

Submitted: 05-05-2022 Revised: 10-05-2022 Accepted: 13-05-2022

ABSTRACT:Sentiment analysis is one of the fastest growing research place, which enables customers to make better-informed purchase choices through right understanding and analysis of collective sentiments from the net and social media. It also presents organizations the capability to measure the impact in their social advertising strategies by figuring out the public emotions in the direction of the product or the occasions associated to them. Maximum of the research done thus far have targeted on obtaining sentiment features by analyzing syntactic and lexical features that had been explicitly expressed via sentiment words, emoticons and other special symbols. An approach is proposed to perform the sentiment analysis of product reviews using Deep Learning. Unlike traditional machine studying methods, Deep Learning models do not depend upon feature extractors as these features are learned directly throughout the training procedure. The main idea on this work is to use word2vec to research phrase embedding and recurrent neural networks to train and classify the sentiment lessons of the product evaluations. This blended word2vec- Long Short Term Memory version may be used to predict the sentiment of new product reviews. The proposed work ambitions to measure the accuracy of the sentiment analysis class version the usage of deep getting to know and neural networks.

KEYWORDS: NLP(Natural Language Processing), Sentiment Analysis,LSTM(Long Short Term Memory),Word vector model, Deep Learning, Word Embedding.

I. INTRODUCTION

A Sentiment analysis or opinion mining is the computational study of people's opinions, sentiments, emotions, appraisals, and attitudes closer to entities along with products, services, companies, individuals, issues, occasions, subjects, and their attributes. The inception and speedy increase of the sphere coincide with the ones of the social media on the internet, for example, critiques, forum discussions, blogs, micro-blogs, Twitter, and social networks, due to the fact for the first time in human records, we've got a big volume of opinionated data recorded in virtual paperwork. Due to the fact early 2000, sentiment analysis has grown to be one of the maximum active studies areas in natural language processing (NLP). it's also broadly studied in facts mining, web mining, textual content mining, and informtion retrieval.

The rapid increase in the rate of internet users every day leads to in evolution of e-commerce and social media websites like Amazon, Flipkart, facebook, Twitter etc. "In 2017, around 1.66 billion humans worldwide purchased items online and it became visible that during the same period, international e-retail sales amounted to 2.3 trillion U.S dollars and as consistent with projections, it may double via 2021 . Now a days, reviews and ratings have become source of information for consumer decisions. Sentiment analysis is one of the main research type of NLP(Natural Language Processing) for tracking the opinion in the direction of a specific product as postive or negative. Sentiment analysis, additionally referred to as opinion mining, studies peoples sentiments towards unique products or entities like mobile

International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 5 May 2022, pp: 726-731 www.ijaem.net ISSN: 2395-5252

telephones, resorts, airways, and so on. The opinion/sentiment is an mindset, notion, or judgment induced through the sensation of the author (consumer who is writing a review) . Now, the first issue someone does when she or he needs to buy a product, is to look the form of reviews and opinions that humans have written on social media such as Facebook, Twitter and diverse other blogs or product review sites.

Nearly 95 percent of the customers consult customer reviews prior to making a purchase decision . This has paved way for new research studies areas like sentiment evaluation or opinion mining. "Sentiment analysis gives insight to organizations via giving them immediately comments on products, and measuring the effect in their social marketing strategies . This helps the manufacturer to perceive new possibilities and control their reputations. maximum of the studies finished in sentiment analysis are primarily based on traditional device learning algorithms like SVM(aid Vector system), Naive bayes and so on. but recent research use Deep Learning models to gain higher overall performance because it has completed high-quality effects in other areas like image classification, laptop vision and speech recognition.

Nowadays, if one wishes to buy a consumer product, one is now not restricted to asking one's friends and family for opinions due to the fact there are numerous consumer reviews and discussions approximately the product in public boards on the internet. For an organization, it can now not be necessary to conduct surveys, opinion polls, and recognition agencies for you to accumulate public evaluations because there may be an abundance of such facts publicly available. In current years, we've got witnessed that opinionated postings in social media have helped reshape groups, and sway public sentiments and emotions, which have profoundly impacted on our social and political structures. Such postings have additionally mobilized hundreds for political changes consisting of the ones befell in some Arab countries in 2011. It has for that reason become a necessity to accumulate and take a look at evaluations.

II LITERTURE SURVEY

[1] The approach includes use of collection of product based dataset from exclusive E-commerce web sites like amazon.com, epinion.com and many others. The reviews are accumulated on products like telephone, ipod etc. The goal of the work is to analyze and are expecting product based totally reviews by way of classifying them as high-quality, poor and impartial by way of the use of algorithms like naïve baye's and SVM. Due to the fact that enter is set product reviews that are unstructured, they perform pre-processing, extracts capabilities directly to which feedback are made, then calculates polarity of evaluations, and also plots graph for the end result. The effects also cowl managing negation part.For example- "the nokia smartphone isn't always awful" gives effective overview although it includes a negative phrase "now not".

[2] In this study, an attempt has been made to categorise sentiment analysis for movie reviews the use machine learning tchniques. Two distinctive algorithms specifically Naive Bayes (NB) and support Vector device (SVM) are implemented. These two algorithms have additionally been implemented earlier by distinctive researchers and results of all variations of implementation have been compared. It's observed that SVM classifier outperforms each other classifier in predicting the sentiment of a review.

[3]This paper proposes the deep learning model of Bert-BiGRU-Softmax to address the sentiment word disambiguation and the sentiment polarity problem, which makes use of the Bert model to extract capabilities selection from e-commerce evaluations at input layer, the hidden layer of BiGRU model with attention mechanism to acquire semantic codes together with the attention possibility of input layer, and Softmax model as the output layer to classify the sentiment tendency of the e-commerce reviews. The experiment analyses the massive-scale evaluation datasets from the Sunning, Taobao websites. The experimental consequences show that Bert-BiGRU-Softmax version has higher overall performance than RNN, BiGRU and Bert-BiLSTM, which improves the accuracy at least 3% on sentiment evaluation of the e-commerce product satisfactory reviews.

[4]This study has carried out extraordinary machine learning algorithms of SVM and Na¨ıve Bayes at the Amazon beauty products reviews. The consequences from the have a look at confirmed that during phrases of accuracy the SVM method achieves better effects than the Na¨ıve Bayes approach when the entire information set turned into used as schooling and testing facts set. As the

International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 5 May 2022, pp: 726-731 www.ijaem.net ISSN: 2395-5252

quantity of opinions reduced the Na¨ıve Bayes method executed higher performance than the SVM method.

[5] On this evaluation, they compare six unique sentiment category techniques, three supervised machine learning technique: SVM, Gradient Boosting, and LR algorithms and three supervised-based strategies: VADER, Pattern, and SentiWordNet lexicons to research Amazon reviews datasets. In addition they perform our experiments the usage of Amazon product critiques with diverse NLP strategies consisting of stopwords elimination, word lemmatization, and TF-IDF vectorization. Thier experimental methods studied the accuracy, precision, keep in mind, and F1 score of sentiment type algorithms. Mor eover, all our models were capable of classify bad and high quality critiques with distinctly precise accuracy and precision. The three supervised machine learning technique finished better than the lexicon-primarily based classifiers on all themetrics. This will be attributed to the fact that the Lexicon primarily based approaches uses a set listing of words to identify positive or bad sentiment. A number of the six fashions, the LR set of rules is the first-rate classifier typical with the highest accuracy, precision, bear in mind, and F1 score. a few of the 3 lexicon-based totally models, the VADER lexicon model has the very best ratings for all the metrics. each companies of algorithms executed better in time period of classifying effective class, and perform poorer in time period of classifying terrible elegance. The purpose for this can had been due to certain forestall words that might have a fantastic emotion related to it and additionally because of the inherent elegance imbalance problem due to the dataset having a big percentage of critiques that have a tremendous sentiment. In conclusion, thier device mastering outcomes are barely higher as compared to recent text sentiment machine studying works while our lexicon-primarily based result are worse compared to latest comparable lexicon-based works.

[6]. In this assignment various reviews and ratings of the distinctive on line product are taken from ecommerce websites are saved and used as a fundamental facts on which classifier is imposed on critiques and rankings from which distinctive bag of words are generated and the use of that bag of phrases polarity is decided. The sentiment generated from the internet site are proven inside the first-rate way in order that client can without problems understand the polarity of the opinions generated on that ecommerce internet site. This is how we get the quality suitable product overview.

[7] In this paper we endorse a topic sentiment mixture version i.e. Co-LDA model for topic sentiment evaluation. The CoLDA can model topics and sentiments concurrently. With the Co-LDA we will evaluation each opinion unit belong to which topic and what sentiment it is. So we are able to browse each opinion devices of every subject matter companion with wonderful and poor sentiments. We also can achieve which topics of the product are humans maximum interested in and which subjects are human beings glad. We compare our model on two product assessment sets; the effects show that the semi-supwevised Co-LDA is effective to topic sentiment evaluation.

[8] on this research, they suggest a deep getting to know approach to clear up a triumphing problem in e-commerce websites wherein the evaluation submitted by way of the user doesn't fit its score. critiques are a very vital supply of data for a ability consumer earlier than deciding to buy a product . They believe that such opinions with mismatched scores can create negative user revel in and consequently the inducement for this research.To assemble their version, they first employed paragraph vectors to study the syntactic and semantic relationship of a 'overview text'. They further grouped and looked after evaluate embedding to form a product series which is fed to a gated recurrent unit (GRU) to analyze product embedding. The concatenation of evaluate embedding generated from paragraph vectors and product embedding generated from GRU is used to teach a guide vector gadget (SVM) for sentiment classification. With best assessment embedding our classifier plays atan accuracy of 81.29%. Inclusion of product embedding increases the accuracy to 81.82%. This shows that product records is a effective characteristic that may be employed in sentiment evaluation. They later use this classifier via an internet service to predict rating of a assessment and examine it in opposition to given rating. This web carrier takes 'assessment text' and 'evaluation score' and affords a warning to the reviewer if there may be an inconsistency among the given score .

[9] In this paper they use word2vec representations to categorise more than four hundred,000 on-line customer critiques for diverse worldwide cell phone manufacturers acquired from Amazon. They first find functions most just like product aspects through word2vec and show that

International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 5 May 2022, pp: 726-731 www.ijaem.net ISSN: 2395-5252

word2vec is able to locate semantically similar words. Then we use CBOW and skipgram techniques with four extraordinary class algorithms: Naïve Bayes, SVM, Logistic regression and Random wooded area. results show that CBOW performs nicely as compared to bypass-gram, indicating that data may additionally consist often taking place equal phrases. Random wooded area outperforms all the algorithms when used with word2vec representations. for this reason, distributed phrase vector representations can be efficiently employed for the task of sentiment class by incorporating semantic word relations and contextual statistics.

[10]Sentiment evaluation the use of Recurrent Neural network ,Authors:Lilis Kurniasari1 and Arif Setyanto1, in this research they've proposed version for sentiment evaluation using RNN and word2vec. RNN in this model is applied the usage of framework Tensorflow. The studies results show that the model approach has better accuracy with other system learning fashions with result of accuracy is 91.98%. From this studies, additionally they have to pay attention to the opportunity of overfitting the version while sporting out the testing manner. within the future they are able to try to use RNN and LSTM to triumph over overfitting problems and to enhance the overall performance of the version

III PROPOSED METHODS

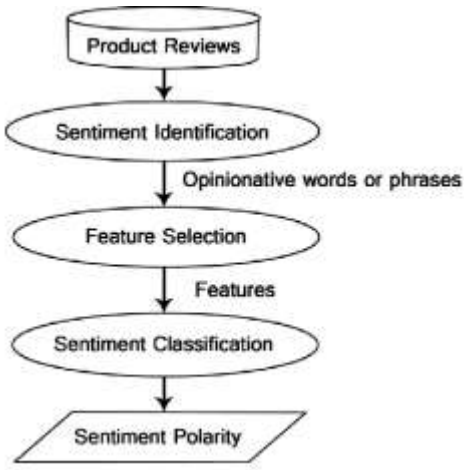

Fig 1. Block diagram

The first steps on this studies scraping and scrawling the reviews on the Ecommerce internet site. We use scraping and crawling strategies for data collection.We process the review data into word vectors and data sets. The data set in this study is split into components, first as a training data sets and second testing data sets.

A. Data collection:

Data collection is the manner to collect the simple data or file on which work is executed. Consumers explicit their sentiments approximately particular products on e-commerce web sites like amazons. Their sentiments and opinions are expressed in distinct way, with unique vocabulary, context of writing, utilization of quick paperwork and slang, making the records huge and disorganized. Manual analysis of sentiment records is in reality not possible. Therefore we uses the sentiment evaluation to make this effort smooth. The assessment is gathered inside the form of reviews which are given through the different purchasers of that product suggests their opinion approximately it.

B . Data Preprocessing:

The collected data can be either unstructured. There are numerous publically available datasets which offer the big collections of reviews which may be are utilized in sentiment analysis process. Every now and then the set of reviews consists of unwanted statistics together with HTML tags, URL statistics and so forth. Removing such vain information is achieved at the level of preprocessing of reviews.

C. Sentiment Classification

1) Document level analysis;

Document level sentiment evaluation determines the overall opinion of the report. Apart from the sentiment orientation of the character sentences it classifies the sentiment expressed through the entire document. The class is expressed by way of either effective or bad sentiment . This level analysis is beneficial best if the report related to a single entity. because it expresses opinion on a single entity (E.g. Product, person). hence it isn't always relevant to the documents which incorporate the comparison of the a couple of entities

2) Sentence level analysis:

The sentence level sentiment evaluation is also called as subjectivity classification. It distinguishes subjective information from the objective facts. It considers every sentence as a separate unit and it predicts that the sentence ought to incorporate most effective one opinion. It iterates each sentence and determines whether or not the sentiment orientation of the sentence is effective, negative or impartial.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 5 May 2022, pp: 726-731 www.ijaem.net ISSN: 2395-5252

3) Entity and aspect level analysis:

Entity and aspect level sentiment analysis captures the mixture of emotion from the review sentence. It plays pleasant-grained sentiment analysis and it differentiates what truely user wants and do now not-want. One most important characteristic of aspect level sentiment analysis is that, it without delay appears into the opinion rather than paragraph, sentences, terms and record. The aim is to find sentiments on entities and their components. Intended to mention that aspect level sentiment analysis focus on sentiment (positive or negative) and a goal (opinion).

D. Feature Extraction :

Features in reviews are extracted in order that it helps customer to recognise which feature has positive comment and which one has negative. For the reason that, overall conclusion about product is a whole lot wanted however there may be also situation where purchaser requirements come into the scenario. Use of adjectives is completed to categorise opinions as positive or terrible using unigram model.

Word vector model:

From the process of scraping and scrawling to get based information with string format and convert the records into vector form. We use word2vec to transform string phrases into vectors. Word2Vec is a device developed by Google . Word2vec will convert phrases into vectors through looking at the context of words with words that appear inside the sentence.

Fig 2. Word Vector Model

E. Long Short Term Memory

Long Short Term Memory could be a kind of recurrent neural network. In RNN output from the last step is fed as inpu twith in the current step. LSTM was designed by Hochreiter & Schmidhuber. LSTM is consists of three gates, i.e. the input (i), forget (f), output (o) gates, and a memory cell . This gates permits the LSTM to manipulate the contents stored within the memory cell. It tackled the problem of long-term dependencies of RNN in during which the RNN will not predict the word stored in the long-term memory but can give more accurate predictions from the recent information. As the gap length increases RNN does not give an efficient performance. LSTM will by default retain the information for a long period of time. It is used for processing, predicting, and classifying on the basis of time-series data.

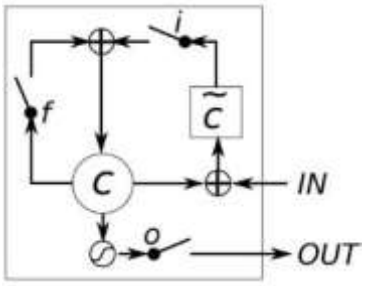

Fig 3. Long Short-Term Memory Unit

IV RESULTS AND DISCUSSION

This experiment uses Long Short Term Memory model and pre-trained word embedding layer FastText that contains three thousand vocabulary size. Training accuracy is mentioned in "Fig3". The program can initial split the document-level review by its punctuation into sentence-level punctuation and apply the pre-processing steps of data cleaning. Finally, the model will predict on the processed dataset by assignment every sentence level review a sentiment classification (Positive or Negative) and prediction confidence (0-100%).This section presents the experimental result of Deep Learning algorithms for sentiment analysis. The algorithmic rule was tested on amzon dataset obtained from kaggle. The dataset consists of four attributes and three thousand instances. Term frequency is calculated for each word with in the sentimnet and this frequency is used to train the model.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 5 May 2022, pp: 726-731 www.ijaem.net ISSN: 2395-5252

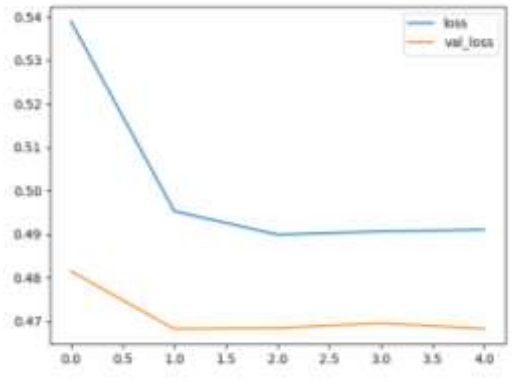

Fig 3. LSTM Model accuracy

The accuracy o f Deep Learning algorithm is shown in Fig 3. The accuracy for Long Short Term Memory is approximately 93.33%.

V CONCLUSION

Sentiment Analysis is an Natural Language Processing task to identify whether a particular product review is to be positive or negative. Neural networks and deep learning techniques are becoming popular in solving most of the machine learning classification problem. Here in this proposed model for sentiment analysis of product reviews using Long Short Term Memory model. Sentiment analysis deals with identifying and aggregating the sentiment expressed by the users. Sentiment analysis is to classify the polarity of text in document or sentence whether the opinion expressed is positive, negative, or neutral. Long Short Term Memory model is to be applied to identify the sentiment of the product reviews.

The project has achieved an accuracy of approximately 93% when Long Short Term Memory model were used. Long Short Term Memory model better when compared to other supervised machine learning algorithms for product sentiment analysis. The future work can be of analyzing the fluctuation in the performance of sentiment analysis algorithm when multiple features are considered.

REFERENCES

[1]. Manvee Chauhan , Divakar Yadav .Sentiment Analysis of Product-Based Reviews Using Machine Learning Approaches. 2017

[2]. Abinash Tripathya, Ankit Agrawalb, Santanu Kumar Rathc .Classification of Sentimental Reviews Using Machine Learning Techniques Abinash Tripathya, Ankit Agrawalb, Santanu Kumar Rathc. 2015

[3]. Jiahuan Lu, Jie Yang and Feng Mao .Sentiment analysis for e-commerce product reviews by deep learning model of Bert-BiGRU-Softmax Yi Liu.2020

[4]. P. Sasikala, L. M. I. Sheela, Sentiment analysis of online product reviews using DLMNN and future prediction of online product using IANFIS, J. Big Data, 7 (2020), 33–53.

[5]. Y. Mejova, P. Srinivasan, Exploring feature definition and selection for sentiment classifiers. In Proc. 5th Int. Conf. weblogs social media, Barcelona, Catalonia, Spain, (2011), 17–21.

[6]. D. Adnan, F. Song, Feature selection for sentiment analysis based on content and syntax models, Decis. Support Syst., 53 (2012), 704–11.

[7]. B. Liu, Sentiment analysis and opinion mining, Synth. Lect. Hum. Lang. Technol., 5 (2012), 1–167.

[8]. J. Trofimovich, Comparison of neural network architectures for sentiment analysis of Russian tweets, In Proc. Int. Conf. Dialogue 2016, RGGU, (2016).

[9]. Q. Qian, M. Huang, J. Lei, X. Zhu, Linguistically regularized LSTMs for sentiment classification. In Proc. 55th Ann. Meet. Assoc. Comput. Ling., 1 (2016), 1679–1689.

[10]. L. Nio, K. Murakami, Japanese sentiment classification using bidirectional long short-term memory recurrent neural network. In Proc. 24th Ann. Meeti. Assoc. Nat. Lang. Process., Okayama, Japan, (2018), 1119–1122.