Grenze International Journal of Engineering and Technology, January Issue

Sentimental Analysis using Long Short Term Memory (LSTM) Neural Network and Inverse Document Frequency Word Embedding Technique

Abhilasha Sharma[1] and Jaya Gupta[2]
[1-2]Delhi Technological University/Software Engineering Department, New Delhi, India
Email: abhi16.sharma@gmail.com, jayagkv123@gmail.com

***Abstract*—Nowadays, there are so many platforms which allow users to write the reviews and share the comments. Reviews can be related to any product or a service. Users are free to share their feedback and experiences by writing the reviews. As the user count is increasing day by day on a very large scale which results in a drastic increase in the number of reviews. A user can write a positive or negative review according to his or her experience and to classify these reviews manually is the biggest challenge due to a very large amount of data. So, this paper proposes a model which is using deep learning techniques with inverse document frequency word embedding and LSTM (Long Short Term Memory) technique is used to train our model which helps in predicting the sentiment analysis of a review given by any user. Various studies use many machine learning and deep learning techniques to classify the reviews according to their sentiments. Our model is giving accuracy better than the present state of art model.**

***Index Terms*— Sentimental Analysis, Long Short Term Memory, Term Frequency Inverse Document Frequency.**

I. INTRODUCTION

Sentiment Analysis, as its name implies, involves discerning the sentiment or emotional stance conveyed in a given context. It essentially entails analysing and uncovering the emotional tone or purpose behind written or spoken communication, regardless of the language used. Language serves as a conduit through which we convey our thoughts and feelings to others, and every utterance carries an inherent sentiment, whether positive, negative, or neutral.

Consider a fast-food chain enterprise offering an array of culinary delights such as burgers, pizza, sandwiches, milkshakes, and more. Recently, they've launched a website enabling customers to conveniently place orders for their favourite food items. Additionally, patrons have the opportunity to share their feedback, expressing their satisfaction or dissatisfaction with the food they've ordered.

TABLE I. USER REVIEWS WITH THEIR DIFFERENT SENTIMENT

User	User Reviews	Sentiment
Review 1	I adore this cheese sandwich; it's incredibly delicious.	Positive
Review 2	The taste of this chicken burger is quite unpleasant.	Negative
Review 3	I placed an order for this pizza today.	Neutral
Review 4	I thoroughly enjoy indulging in this cake.	Positive

Grenze ID: 01.GIJET.11.1.445

So, as we can see that out of these above 4 reviews,
The initial and final reviews clearly express positive sentiments, indicating the customer's satisfaction with the sandwich. Conversely, the second review conveys a negative sentiment, suggesting the company should address issues within their burger department. The third review lacks a clear indication of satisfaction or dissatisfaction, thus can be deemed neutral. Analysing these reviews, the company can infer the necessity to emphasize the production and marketing of their sandwiches, alongside enhancing the quality of their burgers to boost overall sales.
Sentiment analysis, a form of contextual mining, assists businesses in monitoring online discussions and extracting subjective information from various sources. By assessing the social sentiment surrounding a brand, product, or service, it provides valuable insights. Recent advancements in deep learning have notably improved the capabilities of text analysis algorithms. Among the most popular tools is sentiment analysis, which categorizes incoming messages as neutral, negative, or positive based on underlying sentiment. Intent analysis takes it a step further by discerning the user's purpose behind communication, whether it involves expressing an opinion, sharing news, engaging in marketing, issuing a complaint, offering a suggestion, expressing appreciation, or posing a question.
The paper follows this structure: Section 2 conducts a thorough literature survey of the chosen area. Section 3 delineates the proposed methodology. Section 4 evaluates the experimental findings and provides an analysis thereof. Lastly, Section 5 presents the conclusion of the research paper.

II. Related Work

There has been a notable surge in the sentiment analysis task in recent times. The various research works in sentiment analysis (Ligthart et al. 2021) published an overview on Opinion mining in the earlier stage. In (Piryani et al. 2017) discusses the study topic from 2000 to 2015 and provides a framework for computationally processing unstructured data with the primary goal of extracting views and identifying their moods. Several recent surveys (Yousif et al. 2019; Birjali et al. 2021) authors has described the problem of sentiment analysis and suggested potential directions. Soleymani et al. (2017) and Yadav and Vishwakarma (2020) on sentiment classification have been published. Additionally, In the work of Yue et al. (2019) and Liu et al. (2012) conducted research on the effectiveness of internet reviews. The Authors (Jain et al. 2021b) discuss machine learning applications that incorporate online reviews in sentiment categorization, predictive decision making, and the detection of false reviews. In the work of Balaji et al. (2021) conducted a thorough examination of the several applications of social media analysis utilizing sophisticated machine learning algorithms. Authors present a brief overview of machine learning algorithms used in social media analysis (Hangya and Farkas 2017). The proliferation of community websites has given growth to various fields dedicated to analysing these webs and their contents to excerpt vital data. As a subfield of Natural Language Processing (NLP), sentiment analysis has a rich history, given its significant impact on public opinion and decision-making. Despite early works addressing sentiment analysis, development continued into the new era. Numerous everyday applications necessitate sentiment analysis for thorough examination, such as product analysis, which helps identify the constituents or merits of a product that resonate with customers in terms of quality. In work of Subhashini et al. (2021) presents the results of a comprehensive review of contemporary opinion mining literature. The scope of sentiment analysis extends to extracting script structures from ideas amidst clatter or ambiguity, representing information embedded in views, and classifying them. Mowlaei et al. (2020) proposed a method for adaptive aspect-based lexicons in sentiment classification. They outlined two approaches for creating vibrant lexicons tailored to sentiment classification based on aspects: one leveraging statistical methods and the other employing genetic algorithms. Dynamic lexicons, which can be automatically updated, offer more precise grading for context-dependent concepts (Kumar and Uma 2021). To categorize all features of journals, various lexicons from multiple dictionaries were selected. Sentiment analysis has been applied across diverse domains, from hospitality and aviation to healthcare and financial markets (Zvarevashe and Olugbara 2018). In the hospitality sector, sentiment analysis aids in understanding customer preferences and dislikes based on hotel reviews. Conversely, in financial markets, sentiment analysis, as discussed by (Valencia et al. 2019), helps in gauging trends in the stock market and cryptocurrencies based on market sentiment. Ahmad et al. (2019a) conducted sentiment analysis on tweets across various domains to analyze the sentiments expressed in those tweets. The health care domain is seeing a surge in sentiment analysis applications in recent times, customer opinion analyses (Ruffer et al. 2020; Park et al. 2020; Cor tis and Davis 2021; Arora et al. 2021), customer satisfaction analyses are few applications in the healthcare sector (Baashar et al. 2020; Miotto et al. 2018). The corporate realm has long depended on sentiment analysis to improve its operations. This analysis is utilized in managing reputation, conducting market research, analysing competitors, evaluating products, and grasping customer sentiment, among other purposes. However, sentiment analysis and natural language processing face various hurdles, like casual text styles, derision, irony, and language-specific subtleties. Many words in distinct languages hold diverse meanings and connotations depending on the context and field of application, leading to a dearth of tools and resources for all languages. Detecting sarcasm and irony has become a notable challenge that researchers are actively addressing, resulting in advancements in this domain. Despite strides forward, sentiment analysis remains fraught with numerous challenges. This study delves into these challenges, methods, applications, and processes in sentiment analysis. Comparative data analysis is presented through tables, flowcharts, and graphs to aid understanding.

III. METHODOLOGY

This section presents the proposed model namely, Sentimental Analysis using Long Short Term memory (LSTM) Neural Network and Inverse Document Frequency Word Embedding Technique which is based on the bidirectional long short term memory (LSTM) and term frequency inverse document frequency word embedding technique. Our model is also compared with bag-of-words(BOW) word embedding's, pre-trained glove word embedding's and pre-trained google vector news based on word2vec architecture word embedding's. Word embedding's play a very critical role to train the sentiment analysis model as machines can only understand numbers instead of words and sentences. Various previous studies designed many word embedding techniques which not only focus on the word to number conversion but also focuses on the contextual meaning of the sentence or text.

A. Word Embedding

Bag of Word (BOW): It is one of the oldest word embedding technique which is based on the vocabulary of the text. In bow word embedding, the dimension of embedding vector created is same as the vocabulary of the text. If we have one millions unique words in any text then we have one million dimension of each word in numbers which is not practically good moreover bag of words is not maintained any contextual meaning of the sentence. So because of that reason we restrict ourselves to use bag of words as word embedding techniques.
*Pre-trained google vector news Word embedding***:** This is the pre-trained word embedding's generated by google on the google news dataset which contains 6 million words. The dimension of each word in this word embedding is 300 dimensions and it also maintains the contextual meaning of the words in a sentence. These word embedding's based on the word2vec deep learning technique. Because of this nature, pre-trained google vector news word embedding's are widely used.
Term Frequency Inverse Document Frequency (TFID): This type of word embedding is focused on that how a word in a particular sentence is important with respect to the corpus of the text. Term TFIDF is basically two types of weights called as normalized term frequency and inverse document frequency.

B. Training Model

Various deep learning models are present for sentimental analysis which are simple recurrent NN, Long short term memory model, bidirectional long short term memory model. Our paper focuses on bidirectional long short term memory model.
Simple RNN: Simple RNN plays a very vital role in natural language processing applications as they perform in maintaining the sequences of the words in a text whereas artificial NN are failed to do so. Our paper is not using simple RNN because they are failed to maintain the long context of the text due to the vanishing gradient problem. But they perform well for smaller sentences or smaller inputs.
Long Short Term Memory (LSTM) Model: LSTM model not just contained the hidden state but also contains the cell state (for maintaining long term context) so that they can perform and maintain the longer context on a larger input. LSTM works on 3 types of gates which are forget gate, input gate, and output gate. Forget gate is designed to remove some information from the cell state by using a neural network with a sigmoid activation function. The input gate is charity to improve some data to the cell state and the output gats is used to outputs the hidden state from the LSTM cell. This complex neural network of Long Short Term Memory (LSTM) helps the model to work good on larger sentences in terms of maintaining longer context.
Bidirectional Long Short Term Memory (LSTM): Bidirectional Long Short Term memory is an advanced model of Long Short Term Memory which parses the input from both directions. Parsing the input from both directions is important because sometimes the last word contains the contextual meaning of the sentence.
Example: I like Amazon. It is a great website.

I like Amazon. It is a beautiful river.
In the above example, the contextual meaning of sentences is based on the last word. Which cannot be analysed by unidirectional Long Short Term Memory (LSTM), so that's why bit bidirectional LSTM performs exceptionally well in this type of situation and that’s why our paper proposes a model based on bidirectional LSTM.

C. Proposed Model

Initially the sentences for sentiment analysis will be tokenized into the words, then by using TF-IDF (Term frequency - Inverse Document Frequency) Word embedding technique is used for each possible word to generate the corresponding vector for n- dimension. These Word embedding’s passed through bi-LSTM Model each word at a time to maintain the sequential and contextual meaning of the words. As shown in figure below, at timestamp 0 first word embedding’s will be passed as an input in Bi-LSTM, similarly for timestamp 1 second word embedding’s will be passed as next input in Bi-LSTM Model. Finally the output from Bi-LSTM layer is passed through a dense layer where only one node of the neural network is present. Then it will calculate the loss using categorical cross entropy and the neural network will be back propagated to modify the trainable parameters using gradient descent optimizer.

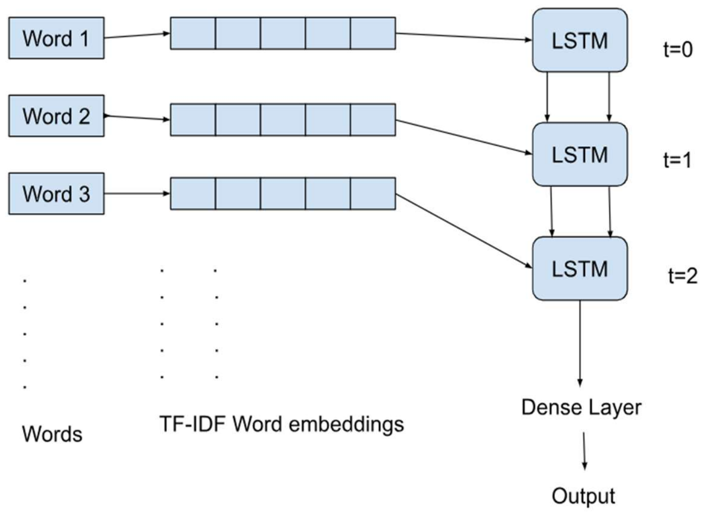

Fig.1. Block diagram of Long Short term Memory (proposed model)

IV. Experimental Results and Analysis

This segment delves into the implementation specifics of the entire configuration. The model has been applied to two datasets, FiQA and Financial PhraseBank, amalgamated into a single CSV file, and readily accessible to the public. Additionally, it discusses the process of selecting hyper parameters during training. The experiments were conducted on a machine equipped with an AMD Ryzen-5 5600H processor, running at 3.30 Gigahertz, with 8 GB of RAM, and featuring an NVIDIA GeForce RTX GPU.

D. Methods for Evaluation

In this section, precision (P), recall (R) and F1 score (F1) metrics are taken for the prediction of Sentiment Analysis based on 6 different emotions. These metrics and their mean values can be defined as:

$$Precision = TP/(TP + FP) \quad (1)$$

$$Recall = TP/(TP + FN) \quad (2)$$

$$F1\,Score = (2 * Precision * Recall)/(Precision + Recall) \quad (3)$$

V. Results

This subsection shows the comparative results of various methods on the FiQA and Financial PhraseBank dataset. Table 2 listed the result analysis for F1 score, Precision, and Recall of various models, which shows that the proposed model performed better.

TABLE II. COMAPARITIVE RESULTS OF DIFFERENT METHODS FOR FIQA AND FINANCIAL PHRASEBANK DATASET

Model	F1 (%)	Precision(%)	Recall (%)
LSTM with Bow (Bag of words)	93.6	92.71	96.23
[5]	94.31	94.21	95.06
LSTM with TF-IDF	88.27	91.83	89.65
Bi-LSTM With Bag of Words	92.21	89.88	94.85
Proposed model	95.82	92.83	97.88

VI. CONCLUSIONS

Sentiment analysis is a type of textual contextual mining that helps businesses monitor online conversations and identify and extract subjective information from source material. It does this by gauging the social sentiment surrounding a brand, product, or service. Recent developments in deep learning have significantly enhanced algorithms' capacity for text analysis. Our Proposed Model which is based on the TF-IDF word embedding technique and Bi-LSTM model performs well on the FiQA and Financial PhraseBank dataset. In future we will use word2vec and glove pre-trained word embedding techniques to capture the contextual meaning of words more expressively.

ACKNOWLEDGMENT

The authors wish to thank Mr. Kulanshu Sharma for advice on experimental design and statistical analysis.

REFERENCES

[1] Abid F, Alam M, Yasir M, Li C (2019) Sentiment analysis through recurrent variants latterly on convolu tional neural network of twitter. Futur Gener Comput Syst 95:292–308

[2] Acheampong FA, Wenyu C, Nunoo-Mensah H (2020) Text-based emotion detection: advances, challenges, and opportunities. Eng Rep 2(7):e12189

[3] Artif Intell Rev 54:5789–5829 Adomavicius G, Kwon Y (2011) Improving aggregate recommendation diversity using ranking-based tech niques. IEEE Trans Knowl Data Eng 24(5):896–911

[4] Borg A, Boldt M (2020) Using VADER sentiment and SVM for predicting customer response sentiment. Expert Syst Appl 162:113746

[5] Bose R, Dey RK, Roy S, Sarddar D (2020) Sentiment analysis on online product reviews. In: Information and communication technology for sustainable development. Springer, pp 559–569

[6] Buder J, Rabl L, Feiks M, Badermann M, Zurstiege G (2021) Does negatively toned language use on social media lead to attitude polarization? Comput Hum Behav 116:106663

[7] Calders T, Verwer S (2010) Three naive bayes approaches for discrimination-free classification. Data Min Knowl Disc 21(2):277–292

[8] Donkers T, Loepp B, Ziegler J (2017) Sequential user-based recurrent neural network recommendations. In: Proceedings of the eleventh ACM conference on recommender systems, pp 152–160

[9] Dragoni M, Petrucci G (2017) A neural word embeddings approach for multi-domain sentiment analysis. IEEE Trans Affect Comput 8(4):457–470

[10] Duric A, Song F (2012) Feature selection for sentiment analysis based on content and syntax models. Decis Support Syst 53(4):704–711

[11] Ebadi A, Xi P, Tremblay S, Spencer B, Pall R, Wong A (2021) Understanding the temporal evolution of covid-19 research through machine learning and natural language processing. Scientometrics 126(1):725–739

[12] Ebrahimi M, Yazdavar AH, Sheth A (2017) Challenges of sentiment analysis for dynamic events. IEEE Intell Syst 32(5):70–75

[13] Eke CI, Norman AA, Shuib L, Nweke HF (2020) Sarcasm identification in textual data: systematic review, research challenges and open directions. Artif Intell Rev 53(6):4215–4258

[14] El-Din DM (2016) Enhancement bag-of-words model for solving the challenges of sentiment analysis. Int J Adv Comput Sci Appl 7(1):244–252

[15] Han T, Liu C, Yang W, Jiang D (2019) A novel adversarial learning framework in deep convolutional neural network for intelligent diagnosis of mechanical faults. Knowl-Based Syst 165:474–487

[16] Kumar S, Yadava M, Roy PP (2019) Fusion of EEG response and sentiment analysis of products review to predict customer satisfaction. Inf Fusion 52:41–52

[17] Lakkaraju H, Socher R, Manning C (2014) Aspect specific sentiment analysis using hierarchical deep learn ing. In: NIPS Workshop on deep learning and representation learning, pp 1–9

[18] Mukherjee A, Venkataraman V, Liu B, Glance N (2013) What yelp fake review filter might be doing? In: Proceedings of the international AAAI conference on web and social media, vol 7

[19] Naseem U, Razzak I, Musial K, Imran M (2020) Transformer based deep intelligent contextual embedding for twitter sentiment analysis. Futur Gener Comput Syst 113:58–69
[20] Ortigosa-Hernández J, Rodríguez JD, Alzate L, Lucania M, Inza I, Lozano JA (2012) Approaching senti ment analysis by using semi-supervised learning of multi-dimensional classifiers. Neurocomputing 92:98–115
[21] Pravalika A, Oza V, Meghana N, Kamath SS (2017) Domain-specific sentiment analysis approaches for code-mixed social network data. In: 2017 8th international conference on computing, communication and networking technologies (ICCCNT). IEEE, pp 1–6
[22] Qiu G, He X, Zhang F, Shi Y, Bu J, Chen C (2010) DASA: dissatisfaction-oriented advertising based on sentiment analysis. Expert Syst Appl 37(9):6182–6191
[23] Rui H, Liu Y, Whinston A (2013) Whose and what chatter matters? The effect of tweets on movie sales. Decis Support Syst 55(4):863–870
[24] Salah Z, Al-Ghuwairi ARF, Baarah A, Aloqaily A, Qadoumi B, Alhayek M, Alhijawi B (2019) A systematic review on opinion mining and sentiment analysis in social media. Int J Bus Inf Syst 31(4):530–554
[25] Valdivia A, Luzón MV, Cambria E, Herrera F (2018) Consensus vote models for detecting and filtering neutrality in sentiment analysis. Inf Fusion 44:126–135