A Neural Network Solution for Collaborative Sentiment Analysis

Ravikumar Thallapalli[1](✉) and G. Narsimha[2]

[1] Informatics-Department of CSE, Osmania University, Hyderabad, Telangana 500007, India
ravimtech.talla@gmail.com
[2] Department of CSE, JNTUH University College of Engineering, Sultanpur, Telangana, India

Abstract. There is a growing need to analyze the contents of ecommerce and micro-blogging platforms in order to determine consumer satisfaction as the number of online forums for providing comments on different features or goods grows. To get a sense of how customers feel about their products, service providers read reviews, both positive and negative, both official and informal. This has led to a plethora of studies aimed at deciphering the writings and gleaning the emotions behind them. However, by relying on tried-and-true techniques for tokenization, lemmatization, and additional sentiment extraction through tagging methods, these approaches overlook a few fundamental truths and lead to underfitting or overfitting issues. As a result, the suggested approach exemplifies several cutting-edge tactics, including differential analysis for tokenization, complicated lemmatization with a significant reduction in processing time, threshold-based sentiment extraction, and subsequent summarization. As a result of this study, 98% accuracy is achieved by the use of improved sigmoid-based neural network activations and a novel technique for weight adjustment in the neural networks.

Keywords: Neural Network · Collaborative Neural Network · Sentiment Extraction · Sigmoid Activation · Collaborative Sentiment Extraction

1 Introduction

Fine-grained sentiment analysis includes tasks like aspect-level sentiment categorization. The system attempts to foretell how a reviewer would feel about several factors. Recently, numerous researchers have relied on a hierarchical attention network to collect granular sentiment data. "To capture the aspect-specific emotion information in the text, however, the previous work only focused on using the aspect words. When the aspect-specific words are retrieved wrongly", it might lead to a mismatch in emotion. Some rare aspect phrases are more challenging to extract than comparable uncommon emotion keywords because the quantity of aspect words is substantially larger. The authors have proposed "a collaborative extraction hierarchical attention network to address this issue. One unit employs the sentiment features extracted by the sentiment attention layer to collect the specific aspect-related sentiment aspect information. This system comprises of two hierarchical attention units". The other layer, the sentiment attention layer, leverages

B. Raj et al. (Eds.): ICETE 2023, AER 223, pp. 397–415, 2023.
https://doi.org/10.2991/978-94-6463-252-1_44

the extracted aspect data to better understand how a user feels about a certain aspect. Furthermore, the authors have validated our method using the SemEval benchmark dataset. Compared to other approaches for aspect-level sentiment classification, our approach outperforms those that rely solely on aspect characteristics to extract sentiment features [1].

At the document level, sentiment analysis attempts to foretell the aggregate ratings of user-generated content (such as product evaluations). Most recent papers treat this issue as a supervised learning challenge, like classification or regression. Using neural network-based text embeddings to represent user and item preferences and attributes as effects on ratings has been the subject of recent research. However, these studies only make use of the overt forms of influence seen in the texts, and they don't attempt to simulate the subtle forms of influence that aren't there. Here, the authors have offered a MOCA (Model-based Collaborative Filtering with Multiple Objective Criteria and Adaptive Neighborhoods) framework for document-level sentiment analysis that is multi-purpose, collaborative, and attentive. "The three most distinguishing features of our MOCA are: Two models are used by MOCA: (1) an attentive model for explicit influence, where a bidirectional recurrent neural network with an attention mechanism is used to learn user-item specific text embeddings for exploiting the explicit influence of users and items; and (2) a collaborative model for implicit influence, where a new neural collaborative filtering model based on multilayer perceptron is designed to capture the implicit influence that is implied in the highly personalized interactions between the users and i. The experimental findings on three real-world datasets from IMDB and Yelp demonstrate that our MOCA greatly outperforms existing state-of-the-art approaches" [2].

The purpose of this comprehensive study was to investigate the relationship between student performance and the adoption of "a natural language processing (NLP) driven strategy to extract emotions from speech in collaborative learning" settings. Many scholars in the fields of engineering and computer science education have focused on improving students' social competence in recent years. A lack of social skills has been shown to have a negative effect on performance in both academic and professional settings, proving that low achievement is not just the result of intellectual or cognitive deficits. To better "prepare students for the coming Fourth Industrial Revolution, Engineering Education is incorporating soft skills into its curricula (i.e., Industry 4.0)". To achieve this aim, The authors have presented a model in our previous work [1] that would analysis the attitudes derived from students' group discussions and determine whether or not they were correlated with improved academic performance. "Polarity sentiment analysis demonstrated a significant positive relationship between students' favorable attitudes toward working in teams and their individual success in the course". This research goes a step further by analyzing the affective valence of students' group discussions using a multi-class approach. "Specifically, there are two stages to the procedure: (1) Using collaborative speech in a first-year computer science (CS1) course to extract distinct classes of emotion such as pleasure, anger, anxiety, etc. and find their link with students' performance, and (2) Aspect-Based Emotion Analysis (ABEA)". The authors have used rule-based models using voice datasets and the supervised machine learning methodology. The authors have done sentiment analysis on the preprocessed text and found several different categories of emotion. Part-of-speech (POS) tagging is used to extract aspects, and patterns are then derived from those parts. Finally, The authors have trained the K-Nearest Neighbor (KNN) algorithm to predict student performance using emotion classes and aspect patterns as feature vectors [3].

The proposed work is to improve the accuracy of sentiment analysis by taking into account the opinions of multiple individuals. Traditional sentiment analysis methods usually only consider the sentiment of a single author or a small set of pre-defined labels. However, people's opinions can vary widely depending on their background, experiences, and perspectives. By leveraging the power of neural networks, the proposed solution can learn from a large number of diverse opinions and provide a more nuanced and accurate understanding of sentiment. This can have important applications in fields such as marketing, politics, and social media analysis.

Further, the rest of the work is organized such that the baseline method for sentiment analysis is discussed in Section – II, in the light of that, the recent research improvements are discussed in Section – III. The identified problems are proposed solutions are discussed in Section – IV and V respectively. The designed algorithm is furnished in Section – VI. The obtained results are discussed and compared with the parallel research outcomes in Section – VII and VIII. The conclusion of this research is discussed in Section – IX.

2 Foundational Method for Collaborative Sentiment Analysis

After setting the context of this research, in this section of the work, the foundational method is discussed.

Assuming that, the entire text dataset is T[], which contains the corpus id as ID, the actual text as t and the numeric rating associated with the text as r. Thus, this can be formulated as,

$$T[] =< ID, t, r > \tag{1}$$

The traditional method influences the use of the Bag of Words for the detection of sentiments. The bag of words can be presented as BoW[], which is a collection of words, W[] and the associated sentiment scores as SC[]. Thus, this can be represented as,

$$BoW[] =< W[], SC[] > \tag{2}$$

Further, the sentiments are extracted in two phases as S1[] and S2[], extracted from the text and numeric ratings respectively.

This can be formulated as,

$$S1[] = \prod_{t \in W[i]} SC[i] \tag{3}$$

And,

$$S2[] = \prod_{i=0}^{n} T[i].r \tag{4}$$

Finally the final sentiment, S[], is represented as,

$$S[] =< S1[], S2[] > \tag{5}$$

Further, in the next section of this work, the recent improvements to this baseline method are discussed.

3 Recent Research Reviews

The recent research outcomes have demonstrated multiple enhancements over the baseline methods. These outcomes are discussed with their shortcomings in this section of the work.

From the looks of things right now, it would be simple to conclude that data is key. Data creation has reached a hyperbolic rate. Data growth is not linear; rather, it is exponential in both volume and velocity. The expansion of our services, our clientele, and the number of positive evaluations the authors here've received online are all part of this. On the internet, you may find more than a trillion reviews covering a wide range of items. Inefficiencies plagued earlier recommendation systems often. In this study, the authors have introduced a customer-focused recommendation system. Recommendations are made using sentimental analysis, taking into account the top k reviews in each region. For this, the authors have leveraged Hadoop's infrastructure for increased performance and scalability. [4]

A recommender system is one that suggests content to a user that the user is most likely to find interesting. One method for building these recommendation systems is called "collaborative filtering" (CF). CF finds users who are most similar to the target user based on their ratings and then recommends content to them based on these similarities. In this study, the authors have suggested a novel enhancement to the collaborative filtering algorithm, one that takes advantage of trustworthy user reviews to provide precise recommendations. In contrast to more recent CF methods, which go beyond simple numerical ratings when giving suggestions, older methods relied on these ratings alone. Existing rating-based CF methods also suffer from a lackluster rating database. Only if the authors have a different approach to filling in the blank ratings can the authors have eliminate the data sparsity problem. The suggested method tries in this direction by attempting to infer numerical ratings from textual critiques. "Finding the sentiment orientation and intensity of opinion words stated in user-reviews is part of the rating inference issue that falls under the broader category of sentiment analysis". There are already recommender systems that employ user-reviews to improve their suggestions, however these systems often ignore the reliability of the reviews they use. The suggested CF method assigns trustworthiness ratings to "user reviews by considering several criteria, such as the reviewer's standing and the product's overall quality". The suggested framework is tested experimentally, with confirmed results. [5]

Because reviews include so much valuable consequence information, including them in a recommendation system is becoming increasingly significant in the e-commerce industry. A collection of emotional features (SF) may be used to indicate consumer preference by assessing the feelings and subjects expressed via these evaluations. As an alternative to the standard user-item matrix, the authors have first created a user-SF matrix and then utilize it to find the users closest to the target user whose preferences match. The next step is to rate the things on the list of suggestions so that the best N may be chosen. The authors have hypothesized that the suggested framework, when applied to emotive characteristics, can improve the recommender system's efficiency [6].

This extensive study report investigates students' emotional states as they debate course material in a spoken format during group work in an introductory Computer Science (CS1) course. Recent studies of the mind have shown that social constructions play a crucial role in the educational process. It's a good example of why teaching engineers to work together is so crucial. Key components of a positive team experience include performance evaluation tools in addition to cognitive and social dimensions. The major purpose of low-stakes teams is social building of knowledge rather than final artifact production, making it difficult to measure students' individual success. However, students may not cooperate as expected in low-stakes teams since their grades will not depend as much on collaboration. In order to determine whether or not a student's emotional state has any bearing on how well they do in CS1 class, The authors have conducted research on sentiment and subjectivity as measured by affective metrics in team talks with minimal stakes. This study is innovative since it seeks to identify and operationalize emotion as a variable connected to individual performance by analyzing students' in-class verbal discussions. Throughout the semester, the authors have many sessions when students are working together on low-stakes projects, and The authors have record their conversations. Their opinions and subjective ratings are gleaned from their utterances using Natural Language Processing (NLP) methods. The findings of this study indicate a link between students' levels of subjectivity in their speech and their favorable attitudes toward school. It is possible that this study's findings may be used as a predictor of student performance early in the semester, allowing for more timely feedback and the opportunity for instructors to implement interventions that improve the likelihood of student success [7].

Identifying and labeling emotions is a current focus of study in both the business and academic worlds. "Most current approaches to identifying emotional tone in text are rooted in machine learning and approach sentiment analysis as a text classification issue". It is nonetheless generally accepted that sentiment categorization is a process that relies heavily on domain expertise. An effective sentiment classifier in one area may not work so well in another. Training a separate "sentiment classifier for each domain is a straightforward approach" to resolving this issue. However, because of their sheer volume, it is challenging to classify sufficient data for every area. More importantly, the technique leaves out domains where sentimental information is relevant. In existing this research, "the authors have presented a multi-task learning-based method of jointly training sentiment classifiers across various domains". Here, the authors have taken each domain's sentiment classifier and split it into its generic and domain-specific parts. The generic sentiment classifier is able to gather global sentiment data and is trained across different domains to improve its generalization skills. "The domain-specific sentiment classifier learns from the annotated data in a single domain. Furthermore, the authors have investigated two types of inter-domain relationships: those based on textual content and those based on the distribution of emotion words. The authors have incorporated regularization over the domain-specific sentiment classifiers by constructing a domain similarity network with domain relations. In addition, the authors have used the sentiment knowledge gleaned from sentiment lexicons to better train the overall sentiment classifier. Additionally, the authors have provided a new accelerated optimization approach for effectively training sentiment classifiers". Experimental findings on two standard sentiment datasets demonstrate that our approach can considerably and consistently outperform baseline techniques [8].

One of the most prominent computer-based online computational platforms, recommendation systems (also known as Recommender systems) are widely employed in the retail and commercial sectors. They are crucial because of the role they play in presenting curated and relevant content to users, which may improve their overall platform experience and enhance the likelihood that they will make a purchase or hire a service provider. Collaborative filtering, content-based filtering, and hybrids of these two approaches are the gold standard for recommendation systems. Cold start issues, the necessity for user history, and interaction with objects or services given by the platform are only a few examples of the challenges faced by these conventional methods. In this existing research, the authors have "offer a new computing framework for resolving the constraints of existing recommender system techniques, based on the usage of deep learning models and the incorporation of both explicit feedback in the form of ratings score and feelings conveyed in textual comments. Two novel deep neural models, the Collaborative Filtering based Deep Neural Network architecture (CFMDNN) model and the novel Multichannel Convolutional Neural Network (MCNN) model are used to supplement the conventional Collaborative Filtering based system and boost the effectiveness of recommendation systems within the proposed deep learning framework. Positive results in terms of ratings prediction accuracy used to evaluate the performance of the proposed deep learning-based framework were found in an evaluation of the system in a multilingual environment using two publicly accessible English and Arabic languages datasets" [9].

The majority of review prediction methods today make use of sentiment analysis. To make accurate predictions, many machine learning methods have been applied. In other words, these classifiers are blind to dependence across time and maximum pooling. In this research, the authors have use Deep Learning methods to categorize reviews so that The authors have may make better predictions based on these qualities. In addition, The authors have combined a Convolutional Neural Network with a Long Short-Term Memory Recurrent Neural Network to achieve better results with less loss and in less processing time [10].

Software development is often a team endeavor including members from different locations. It requires a lot of back-and-forth via email, company-wide discussion boards, blogs, polls, and code reviews. The software project delivery might make people feel both happy and sad due to the enormous volume of messages and opinions being shared. Based on an analysis of communications over the software development life cycle, the authors of this study offer a method for determining the underlying emotion polarity shared by different teams working together. It also looked at the number of bugs in the code, how long it took to release, how big a team worked on it, how many people were on the team, and how many comments were left by reviewers to see whether there was a correlation between polarity of emotion and social aspects and software artifacts. An emotion dashboard was built using sentiment analysis to track the mental well-being of different cross-functional software development teams and the degree to which their projects were successful. Eighty percent of managers said the tool helped boost team morale overall, and seventy percent said the dashboard was helpful for gauging the psychological well-being of collaboration teams [11].

In order to foster innovative group work, it is believed that people from different walks of life and with different sets of experiences and perspectives should connect with one another often. Among the many benefits of using social media is the increased opportunities for communication and collaboration they provide. The informal social environment created by a learning activity that makes use of social media has the potential to inspire students' imaginations and increase their motivation to study. In this exploratory effort, the authors have presented a learning analytics system built on modules for collaborative creative learning on Facebook. There are five different modules in the system that are used to retrieve digital footprints left by students, assess the learning process, and provide feedback based on the analytics. In order to better monitor the learning activity and intervene, when necessary, instructors might utilize the suggested method to assess the students' behavioral patterns during the creative process. Furthermore, the characteristics that motivate creative collaboration may be more clearly defined with the use of learning analytics [12].

E-commerce, often known as online shopping, is growing quickly in Indonesia because to the country's advanced internet infrastructure. But consumers still have trust issues when making purchases online. Buyers' opinions, expressed through ratings and comments, are taken into account as a factor that might affect their purchasing decisions. As an added bonus, it is well-known that a recommendation system may aid a customer in making a purchase choice. One well-liked approach of making recommendations is known as "collaborative filtering" (CF). However, CF is handicapped by the sparsity problem. Through the use of k-means + + clustering to decrease data dimensionality and Multinomial Nave Bayes opinion mining or sentiment analysis to filter recommendation outcomes, The authors have offered a model-based CF recommendation system with high accuracy and quality suggestions. K-means + + was selected because of its effective clustering findings, whereas Multinomial Naive Bayes was selected due to its ease of use and high precision while processing the data. The data and reviews utilized in this study came from the shopping website Bukalapak and focused on the interactions between users and the products they reviewed. Results from experiments demonstrated that the suggested CF recommendation system was accurate, with an f-measure of around 0.45 and a quality improvement of about 28% [13].

Questions like "how to obtain more of it?" (e.g., "participatory mechanisms," "social media," and "geo-coded data from personal electronic devices") and "how to handle it?" dominate the field of Big Data and Collaborative Big Data study. (For instance, how to take in, categorize, store, and connect different types of data). How can a multidisciplinary layered study of Big Data be utilized to support solid judgments, especially in a collaborative context, and especially under time pressure? is an issue that receives far less attention than it deserves. To accomplish the necessary robust Decision Engineering, the authors have proposed using a technique similar to Network Science that the authors have refer to as Relationship Science. Karassiannetchain analysis (KNA) is our methodological framework in Relationship Science for identifying islands of stability or positive influence dominating sets (PIDS), which can then be used to achieve a form of annealed resiliency or latent stability and protect the network from unintended consequences, unstable elements, and “perfect storm” crises [14].

“To train sentiment classifiers for many domains concurrently, the authors have offered a collaborative multi-domain sentiment classification technique”. Existing method pools domain-specific sentiment information from several sources “to train more accurate and robust sentiment classifiers in situations when labeled data is sparse. In particular, the authors have split the domain-specific sentiment classifier into a global component and a domain-general component”. The global model is able to capture cross-domain general sentiment information. The domain-specific model is able to capture domain-specific manifestations of emotion. The authors have also improved the “learning of domain-specific sentiment classifiers by extracting domain-specific sentiment knowledge from labeled and unlabeled examples in each domain”. In addition, the authors have “supported the exchange of sentiment data between related domains by using the similarities between domains as regularization over the domain-specific sentiment classifiers” in their method. Both textual content and emotional expressions-based domain similarity methods are investigated. Additionally, the authors have presented two fast techniques for resolving our approach’s model. The experimental findings on benchmark datasets demonstrate that our method greatly outperforms the state-of-the-art baseline algorithms in multi-domain sentiment categorization [15].

Users submit online feedback in the form of preferences for various products via e-commerce websites, linked micro-blogs, and business social media. This research may be gleaned from fundamental user preferences expressed in textual comments, reviews, geo-tagged photographs, and other contextual data. Review texts and the wealth of data they contain—including review terms, review topics, and review emotions—have been put to use in a number of manufacturing facilities as of late. They used social pictures and additional context data to enhance collaborative filtering recommender systems. As part of these initiatives, “review texts, geo-tagged images, and other contextual details are used to ascertain user preferences. Examining how metadata and visual information are utilized to address some of the most pressing issues in Algorithms for collaborative filtering, this article provides a systematic overview of recent research that combines review texts, images, and other contextual information” [16].

The authors have investigated the topic of identifying emotional tone in online micro-blog posts, sometimes known as the “sentiment identification problem.” The difficulty of this issue stems from the limited character count of a micro-blog post and the fact that everyone has their own unique methods of expressing feelings. It is possible that user-specific features will be missed by a single classification model trained on the whole corpus. On the other side, at the beginning when users have only made a few posts, there is a lack of training data that might render a tailored model wrong for each user. To address these issues, the authors have provided a method of online collaborative learning in which a global model is learned for all micro-bloggers and then used to fine-tune the individual models. The authors have tested our algorithm on data taken from Twitter, the most popular micro-blogging platform available today. The results of using our system to identify sentiment in real-world microblogging apps demonstrate its efficacy and efficiency [17].

Collaborative filtering (CF) is an approach of making personalized suggestions to each user based on the ratings of other users. When it comes to CF, the majority of current approaches depend solely on users' overall evaluations of goods, which might be misleading due to the fact that people may have diverse opinions regarding certain parts of the items. As an example, in point, the authors have leveraged the movie domain to offer a system that can extract people's opinions on various features from textual reviews and then use that data to fine-tune CF [18]. This structure is made up of two parts: a rating inference part and an opinion mining part. The former collects reviewer feedback on several factors and summarizes it into numerical ratings. The second part, which is the basis for item suggestion, infers the overall ratings of things based on the aspect ratings. An important part of our contribution is the suggestion of a tensor factorization strategy for rating inference. This method, which operates on the tensor consisting of the overall and aspect ratings, is able to accurately forecast unknown ratings by capturing the inherent linkages among people, things, and aspects. The results of experiments conducted on a movie dataset demonstrate that our approach greatly outperforms two baseline methods in terms of prediction accuracy [20–24].

The proposed neural network solution for collaborative sentiment analysis improves upon existing works by incorporating a graph-based model that captures the relationships between users and their opinions. It also utilizes attention mechanisms to focus on the most important opinions and incorporates domain-specific features to improve the accuracy of sentiment analysis.

Henceforth, in the next section of this work, the bottlenecks of the existing methods are discussed.

4 Problem Formulation – Mathematical Model

After the detailed discussions on the existing methods for sentiment extraction and analysis, in this section of the work, the existing research problems are listed.

Firstly, the sentiment extracted from the text and the numeric ratings are presented as individual information and there is no correlation between these two extracted sentiments. Represented as,

$$S[] =< S1[], S2[] > \tag{6}$$

And,

$$S1[i] \neq S2[i] \tag{7}$$

Secondly, the extracted sentiment from the text corpus, S1 [] is not confirmative as the extracted sentiment highly dependent on the bag of words. It is conclusive to notice that, any new words, which are not part of the BoW[], cannot be translated to any sentiment scores as,

$$\begin{aligned} &If\ \ t \notin W[i] \\ &Then,\ \ this\ cannot\ be\ t \rightarrow SC[i] \end{aligned} \tag{8}$$

Finally, in majority of situations, it is observed that the extracted sentiments from the text and the numeric ratings are extracted in different scales such as extracted sentiments from the text is within the scale of 1 to 5 and the extracted sentiment from the numeric rating is within the scale of 1 to 10. Thus, this mismatch must be rescaled and confined to similar scales.

Further, based on the identified problems, in the next section of this work, the proposed solution is furnished.

5 Proposed Method

Further, after the identification of the problems, in this section of the work, the proposed methods are furnished.

Firstly, the identification of the parts of the speech is also highly crucial for deciding the thresholds for each type of word categories. The proposed method again utilizes the dictionary, D[], for building the categorization of the parts of the speech. The dictionary houses the part of the speech, p, and the category, d, for each part. This can be formulated as,

$$D[] =< p[], d[] > \tag{9}$$

Thus, with the help of this dictionary, each word can be associated with speech categories as,

$$< W[], d[] >= \prod_{D[i].p[j]=T[k].W[l]} d \tag{10}$$

Further, the final corpus can be re-written as,

$$T[] =< ID, W[], d[], r > \tag{11}$$

Further, the sentiment S1[] from the text corpus can be extracted now as,

$$S1[] = \prod_{T[].W[i]\in BoW[].W[j]} \frac{SC[j]}{\theta\{d[i]\}} \tag{12}$$

where, $\theta\{d[i]\}$ denotes the term frequency of the word W[i]. The scale of the sentiment S1[] is within the range of 1 to 5.

Secondly, the scaling of the extracted S2[] sentiment must be carried out with respect to the scale of S1[]. Assuming that the scaled sentiments scores from S2[] is S2'[].

Thus, finally this proposed method furnishes an activation function based on the sigmoid function class for detection of the sentiments, SC[], from the total corpus as,

$$SC[] = \frac{1}{1 + e^{-\{\sum_{i=0}^{n} S1[i]w_i \pm d_i\}}} + \frac{S1[i] * n + S2'[i] * m}{(n + m) * 5} \tag{13}$$

where, n and m are the confidence scores for S1[] and S2[] respectively, which is further scaled to maximum of 5 as per the previous assumptions.

Thus, SC[] collection contains the final sentiment classes for each text available in the text corpus. Henceforth, based on the proposed solutions, in the next section of this work, the proposed algorithms are furnished.

6 Proposed Algorithm

Based on the proposed solution, in this section the workable algorithms. The Sentiments Analysis using Neural Network with Sigmoid Activation and Scaling (**SA-NN-SAS**) Algorithm is furnished.

Algorithm: Sentiments Analysis using Neural Network with Sigmoid Activation and Scaling (**SA-NN-SAS**) Algorithm

Input: Text corpus as T[] with lemmatized and categorized

Output: Sentiment collection as SC[]

Process:

Step - 1. For each corpus as T2[i]
a. For each word in T2[i] as T2[i].w[j]
  i. Calculate the SC[k] using Eq. 13
  ii. Calculate the summary for SC[k]
  iii. k + +
b. End For
Step - 2. End For
Step - 3. Return SC[]

Document, phrase, and entity feature/aspect polarity classification is a fundamental activity in sentiment analysis. This involves determining if the conveyed viewpoint in the text is positive, negative, or neutral. Sentiment analysis that goes "beyond polarity" considers a wide range of human emotions, including happiness, rage, contempt, sorrow, fear, and surprise. The proposed framework is also furnished here [Fig .1].

The General Inquirer, which offered pointers for measuring textual patterns, and independent psychological studies that analysis a person's mental health by observing their speech patterns were forerunners of sentiment analysis.

Afterward, Volcani and Fogel's patent-described technique explicitly investigated sentiment by locating and classifying textual words and phrases according to their emotional weight. One contemporary system that draws on their research is called EffectCheck, and it displays synonyms that may be used to adjust the intensity of the felt emotion on each scale.

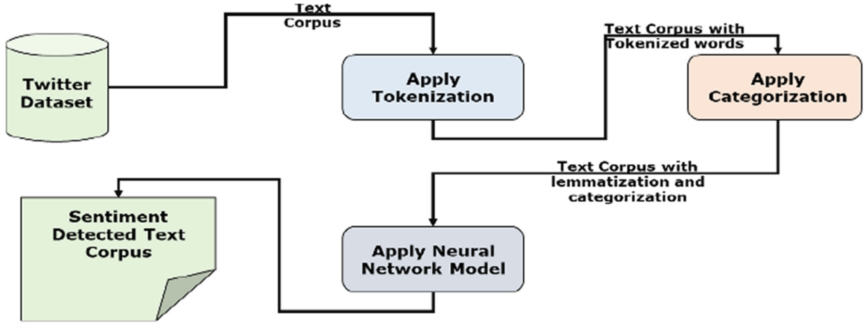

Fig. 1. Proposed Framework

The implementation of the algorithm utilizes tools like NumPy, a fundamental package for scientific computing with Python, TensorFlow, which is an open-source library for high-performance numerical computation that can be used to build and train neural networks and scikit-learn, a machine learning library that provides tools for data preprocessing, model selection, and evaluation.

Further, in the next section of this work, the obtained results are discussed.

7 Results and Discussion

After the detailed discussions on the proposed solutions and the proposed algorithm, in this section of this work, the obtained results are discussed.

Firstly, the dataset description is furnished here [Table 1].

It is natural to observe that the dataset contains no missing values and no outliers. The description is visualized graphically here [Fig. 2].

The complete dataset is tested for all 14139 rows, however for the presentation purposes only 15 are listed for all the phases.

Table 1. Dataset Description [19]

Parameter Name	***Value***
Number of Columns	17
Number of Rows	14,139
Positive Sentiment Rows	2363
Negative Sentiment Rows	9178
Neutral Sentiment Rows	2598
Missing Values	0
Outliers	0

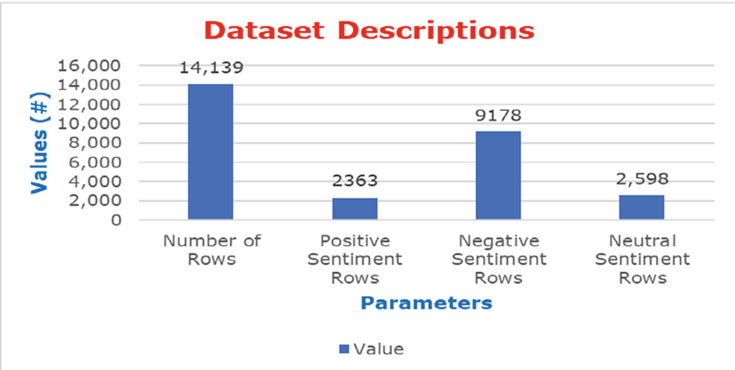

Fig. 2. Dataset Description

Secondly, the sentiments from the text corpus are extracted and compared with the dataset standard values [Table 2].

The standard deviations over 14139 rows are 0.867, clearly, which is fairly low. The outcome is also visualized graphically here [Fig. 3].

Third, the extracted sentiments are again compared with the numeric ratings available in the dataset [Table 3].

Table 2. Sentiments Extracted from Text Corpus

Tweet ID	*Extracted Sentiment (A)*	*Dataset Given Standards (B)*	*Difference [Abs(A-B)]*
1	4	3	1
2	2	3	1
3	4	4	0
4	5	4	1
5	4	3	1
6	3	3	0
7	5	4	1
8	4	5	1
9	2	4	2
10	3	3	0
11	4	5	1
12	4	5	1
13	4	2	2
14	3	3	0
15	2	3	1

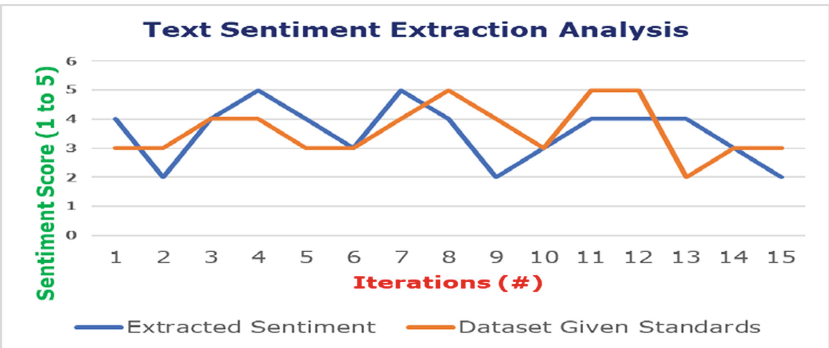

Fig. 3. Sentiment Extracted from Text Corpus

Table 3. Sentiments Extracted from Text Corpus

Tweet ID	*Extracted Sentiment (A)*	*Numeric Rating (B)*	*Difference [Abs(A-B)]*
1	4	3	1
2	2	1	1
3	4	3	1
4	5	4	1
5	4	3	1
6	3	2	1
7	5	4	1
8	4	3	1
9	2	1	1
10	3	2	1
11	4	3	1
12	4	3	1
13	4	3	1
14	3	2	1
15	2	1	1

The standard deviation over the total dataset is 1, which is also low out of a scale of 5. The outcome is also visualized graphically here [Fig. 4].

During the texting of the proposed algorithm, the confidence scores for the S1[] is extracted as 0.425 and for S2[] is it 0.575. Thus, based on both these factors the total confidence scores are extracted [Table 4].

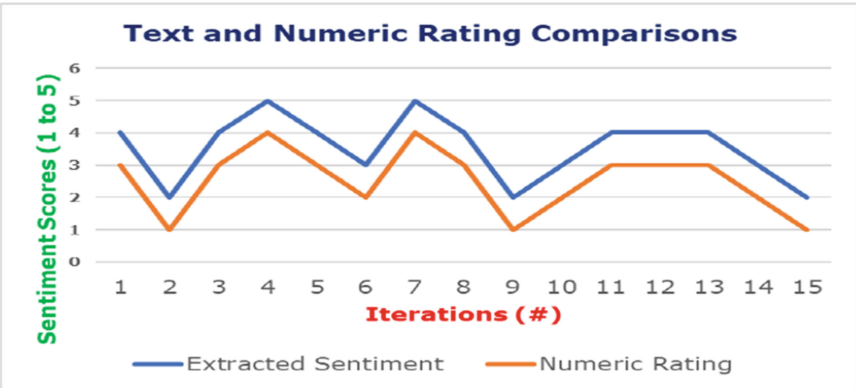

Fig. 4. Sentiment Extracted from Text Corpus Vs. Numeric Rating

Table 4. Final Sentiment Extraction

Tweet ID	*Extracted Sentiment (A)*	*Confidence Score (1)*	*Numeric Rating (B)*	*Confidence Score (2)*	*Final Sentiment*
1	4	0.425	3	0.575	3
2	2	0.425	1	0.575	1
3	4	0.425	3	0.575	3
4	5	0.425	4	0.575	4
5	4	0.425	3	0.575	3
6	3	0.425	2	0.575	2
7	5	0.425	4	0.575	4
8	4	0.425	3	0.575	3
9	2	0.425	1	0.575	1
10	3	0.425	2	0.575	2
11	4	0.425	3	0.575	3
12	4	0.425	3	0.575	3
13	4	0.425	3	0.575	3
14	3	0.425	2	0.575	2
15	2	0.425	1	0.575	1

Further, the results are also visualized graphically here [Fig .5].

It is very significant to observe that the extracted sentiments are nearly equal to the text ratings after considering the extracted sentiment scores from the text as well.

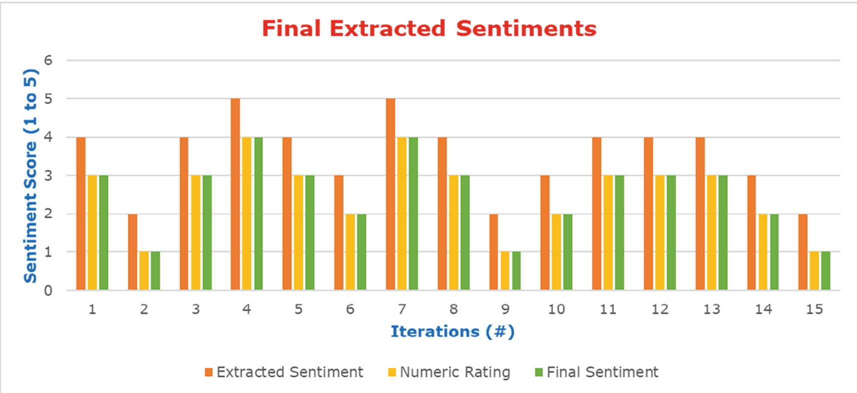

Fig. 5. Final Extracted Sentiments

Finally, the same process is applied on the complete dataset, and further compared with the dataset standards for nearly 100 iterations. However, only 20 iterations are listed here [Table. 5].

Final Sentiment Extraction

Tweet ID	*Number of Records*	*Number of Sentiment Scores Extracted*	*Number of Sentiment Scores Matched*	*Accuracy (%)*
1	14139	14139	13715	97.00
2	14139	14139	13715	97.00
3	14139	14139	13715	97.00
4	14139	14139	13998	99.00
5	14139	14139	13998	99.00
6	14139	14139	13715	97.00
7	14139	14139	13715	97.00
8	14139	14139	13998	99.00
9	14139	14139	13998	99.00
10	14139	14139	13856	98.00
11	14139	14139	13715	97.00
12	14139	14139	13998	99.00
13	14139	14139	13856	98.00
14	14139	14139	13856	98.00
15	14139	14139	13715	97.00
16	14139	14139	13856	98.00
17	14139	14139	13715	97.00
18	14139	14139	13998	99.00
19	14139	14139	13856	98.00
20	14139	14139	13715	97.00

The overall accuracy after 100 iterations is close to 98.32%. Further, the results are visualized graphically here [Fig .6].

Further, in the next section of this work, the obtained results are compared with the parallel research outcomes.

8 Comparative Analysis

After the discussions on the obtained results in the previous section of this work, in this section, the framework is compared with the benchmarked parallel research outcomes [Table 6].

Henceforth, it is natural to observe that the proposed model has outperformed the majority of the parallel research outcomes with higher components in the framework but with lesser sentiment extraction time.

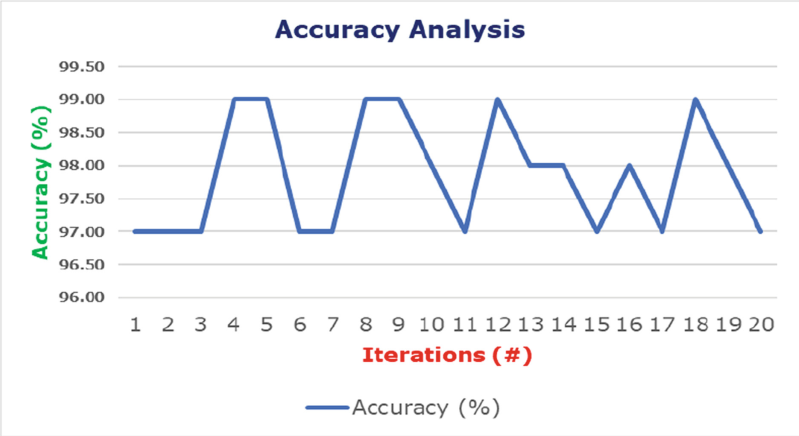

Fig. 6. Accuracy Analysis

Table 6. Comparative Analysis

Author, Year	*Method*	*Model Complexity*	*Accuracy (%)*
N. Dehbozorgi et al. [3], 2022	Classification	$O(n^2)$	95
M. Lou et al. [7], 2022	Neural Network	$O(n^2)$	92
M. Almaghrabi et al. [9], 2022	Neural Network	$O(n^2)$	96
Proposed Framework	Collaborative Neural Network	$O(n)$	98.32

Further, in the final section of this work, the research conclusion is presented.

9 Conclusion

As the number of online forums for offering feedback on various features or products increases, so does the necessity to analysis the contents of these platforms in order togauge customer satisfaction. Services providers often read testimonials, both good and negative, formal and unofficial, to get a feel for how consumers feel about a product. This has spawned a multitude of research aimed at understanding the motivations and themes present in the texts. However, these strategies neglect a few essential realities and cause underfitting or overfitting problems since they depend on tried-and-true techniques for tokenization, lemmatization, and further sentiment extraction through tagging methods. In this way, the proposed method serves as a model of many state-of-the-art strategies, such as differential analysis for tokenization, complex lemmatization with a drastic decrease in processing time, threshold-based sentiment extraction, and subsequent summarization. Results from this research show that by using enhanced sigmoid-based neural network activations and a unique approach for weight modification in the neural networks, an accuracy of 98% may be attained.

References

1. Y. Gao, J. Liu, P. Li and D. Zhou, "CE-HEAT: An Aspect-Level Sentiment Classification Approach With Collaborative Extraction Hierarchical Attention Network," in IEEE Access, vol. 7, pp. 168548-168556, 2019.
2. J. -D. Zhang and C. -Y. Chow, "MOCA: Multi-Objective, Collaborative, and Attentive Sentiment Analysis," in IEEE Access, vol. 7, pp. 10927–10936, 2019.
3. N. Dehbozorgi and D. P. Mohandoss, "Aspect-Based Emotion Analysis on Speech for Predicting Performance in Collaborative Learning," 2021 IEEE Frontiers in Education Conference (FIE), Lincoln, NE, USA, 2021, pp. 1–7.
4. R. Pradhan, V. Khandelwal, A. Chaturvedi and D. K. Sharma, "Recommendation System using Lexicon Based Sentimental Analysis with collaborative filtering," 2020 International Conference on Power Electronics & IoT Applications in Renewable Energy and its Control (PARC), Mathura, India, 2020, pp. 129–132.
5. A. Mahadevan and M. Arock, "Credible user-review incorporated collaborative filtering for video recommendation system," 2017 International Conference on Intelligent Sustainable Systems (ICISS), Palladam, India, 2017, pp. 375–379.
6. Jingjing Cao and Wenfeng Li, "Sentimental feature based collaborative filtering recommendation," 2017 IEEE International Conference on Big Data and Smart Computing (BigComp), Jeju, Korea (South), 2017, pp. 463–464.
7. N. Dehbozorgi, M. Lou Maher and M. Dorodchi, "Sentiment Analysis on Conversations in Collaborative Active Learning as an Early Predictor of Performance," 2020 IEEE Frontiers in Education Conference (FIE), Uppsala, Sweden, 2020, pp. 1–9.
8. F. Wu and Y. Huang, "Collaborative Multi-domain Sentiment Classification," 2015 IEEE International Conference on Data Mining, Atlantic City, NJ, USA, 2015, pp. 459–468.
9. M. Almaghrabi and G. Chetty, "Multilingual Sentiment Recommendation System based on Multilayer Convolutional Neural Networks (MCNN) and Collaborative Filtering based Multistage Deep Neural Network Models (CFMDNN)," 2020 IEEE/ACS 17th International Conference on Computer Systems and Applications (AICCSA), Antalya, Turkey, 2020, pp. 1–6.
10. S. Mathapati, A. K. Adur, R. Tanuja, S. H. Manjula and K. R. Venugopal, "Collaborative Deep Learning Techniques for Sentiment Analysis on IMDb Dataset," 2018 Tenth International Conference on Advanced Computing (ICoAC), Chennai, India, 2018, pp. 361–366.
11. A. Patwardhan, "Sentiment Identification for Collaborative, Geographically Dispersed, Cross-Functional Software Development Teams," 2017 IEEE 3rd International Conference on Collaboration and Internet Computing (CIC), San Jose, CA, USA, 2017, pp. 20–26.
12. S. -M. Wang, "A Module-Based Learning Analytics System for Facebook Supported Collaborative Creativity Learning," 2014 IEEE 14th International Conference on Advanced Learning Technologies, Athens, Greece, 2014, pp. 495–496.
13. W. P. Krisdhamara, B. Pharmasetiawan and K. Mutijarsa, "Improvement of Collaborative Filtering Recommendation System to Resolve Sparsity Problem using Combination of Clustering and Opinion Mining Methods," 2019 International Seminar on Application for Technology of Information and Communication (iSemantic), Semarang, Indonesia, 2019, pp. 94–99.

14. S. Chan et al., "Robust Decision Engineering: Collaborative Big Data and its application to international development/aid," 8th International Conference on Collaborative Computing: Networking, Applications and Worksharing (CollaborateCom), Pittsburgh, PA, USA, 2012, pp. 597–604.
15. F. Wu, Z. Yuan and Y. Huang, "Collaboratively Training Sentiment Classifiers for Multiple Domains," in IEEE Transactions on Knowledge and Data Engineering, vol. 29, no. 7, pp. 1370–1383, 1 July 2017.
16. M. Al-Saffar and W. Baiee, "Survey on Implicit Feedbacks Extraction based on Yelp Dataset using Collaborative Filtering," 2022 Muthanna International Conference on Engineering Science and Technology (MICEST), Samawah, Iraq, 2022, pp. 125–130.
17. G. Li, S. C. H. Hoi, K. Chang and R. Jain, "Micro-blogging Sentiment Detection by Collaborative Online Learning," 2010 IEEE International Conference on Data Mining, Sydney, NSW, Australia, 2010, pp. 893–898.
18. Y. Wang, Y. Liu and X. Yu, "Collaborative Filtering with Aspect-Based Opinion Mining: A Tensor Factorization Approach," 2012 IEEE 12th International Conference on Data Mining, Brussels, Belgium, 2012, pp. 1152–1157.
19. Kaggle NLP Tweet Dataset, URL - https://www.kaggle.com/code/serkanp/nlp-tweet-accuracy-0-94-data-viz/data
20. Singh, U., Saraswat, A., Azad, H.K. et al. Towards improving e-commerce customer review analysis for sentiment detection. Sci Rep 12, 21983 (2022).
21. Almagrabi, H., Alshareef, A. M., Manoharan, H., Mujlid, H., Yafoz, A., &Selvarajan, S. (2022). Empirical Compression Features of Mobile Computing and Data Applications Using Deep Neural Networks. In P. K. Donta (Ed.), Security and Communication Networks (Vol. 2022, pp. 1–11). Hindawi Limited.
22. Devarajan, D., Alex, D. S., Mahesh, T. R., Kumar, V. V., Aluvalu, R., Maheswari, V. U., &Shitharth, S. (2022). Cervical Cancer Diagnosis Using Intelligent Living Behavior of Artificial Jellyfish Optimized With Artificial Neural Network. In IEEE Access (Vol. 10, pp. 126957–126968). Institute of Electrical and Electronics Engineers (IEEE).
23. Shitharth, S., Meshram, P., Kshirsagar, P. R., Manoharan, H., Tirth, V., &Sundramurthy, V. P. (2021). Impact of Big Data Analysis on Nanosensors for Applied Sciences Using Neural Networks. In L. R (Ed.), Journal of Nanomaterials (Vol. 2021, pp. 1–9). Hindawi Limited.
24. Devi, B. T., &Shitharth, S. (2021). Multiple Face Detection Using Haar - AdaBoosting, LBP-AdaBoosting and Neural Networks. In IOP Conference Series: Materials Science and Engineering (Vol. 1042, Issue 1, p. 012017). IOP Publishing.