*International Journal on Recent and Innovation Trends in Computing and Communication*
***ISSN: 2321-8169 Volume: 11 Issue: 11s***
***DOI: https://doi.org/10.17762/ijritcc.v11i11s.8186***
***Article Received:** 29 June 2023 **Revised:** 26 August 2023 **Accepted:** 10 September 2023*

# A Deep Learning based Model using Review Associated Feature Extraction Approach for Sentiment Analysis

**Mohammad Irsad[1], Ashish Khare[2]**
[1]Centre of Computer Education and Training, Institute of Professional Studies,
University of Allahabad
Prayagraj, India
Irshad_khan209@yahoo.com
[2]Department of Electronics & Communication, JK Institute of Applied Physics and Technology
University of Allahabad
Prayagraj, India
ashishkhare@hotmail.com

**Abstract**—With the advancement of internet technologies, in the present days, online forums, social media platforms and e-commerce sites have made the product reviews process very easy. There are a lot of mobile applications, websites and forums where consumers used to share and circulate their opinions, experiences, ideas and views regarding products, brands and services. In consequence, online user reviews have become a deciding factor for many consumers prior to purchasing their selected items. The sentiment analysis is a technique to extract sentiments, feelings and insights from customer reviews and public texts. Therefore, plenty of businesses perform sentiment analysis in order to more thoroughly comprehend of their customer opinions and suggestions regarding their products and services. Furthermore, a number of scientific researchers also have a keen interest in classifying customer reviews into a set of labels employing text classification techniques. The objective of the this research work is to develop an approach to extract review associated features using Part-of-Speech (POS) tagging and design a CNN model to classify the reviews' sentiment as positive or negative. In this paper, an approach to extract review associated feature has been presented. Natural Language Processing (NLP) techniques are utilized for data preprocessing to remove uninformative data from reviews. Deep learning model CNN is used for sentiment classification and Amazon mobile reviews dataset is used for the experiment. The proposed model is experimentally evaluated and provides enhanced performance than other models also provides improved accuracy of 97.23% on Amazon mobile review dataset.

**Keywords**-Sentiment Analysis; Natural Language Processing (NLP); POS Tagging; Review Associated Feature Extraction (RAFE); Deep Learning;

## I. INTRODUCTION

With the advent of Web 2.0, web users were able to express their views on a variety of social media platforms and e-commerce portals in the form of comments, feedback, suggestions and reviews. Most of the e-commerce platforms now feature a separate area for users to write and post the valuable reviews for their products or services. By conducting a sentiment analysis on the posted reviews, valuable information can be acquired such as consumer's opinion about the product or service, aspects which are liked or disliked, main reasons for negative reviews and ratings, valuable suggestions etc. The extracted information can be used to make smart decisions for the selection of a product or service. Business sectors and organizations such as e-commerce sites, tour and travels businesses, hotels and restaurants etc. can use these reviews to enhance their performance by statistically analyzing reviews specific to their products or services. Furthermore, the reviews can also be utilized to thoroughly analyze opponents. However, due to the vast amount, variation, and speed at which reviews are continually posted online, it becomes challenging to examine them.

Opinion mining or sentiment analysis refers to the process of evaluating an individual's expression of emotions, views, and feelings towards a certain entity or subject represented in the form of text. It is the task of computationally identifying and classifying an author's sentiment represented in a piece of text. Sentiment analysis can be performed at the word level, phrase level, sentence level, and document level [1]. The sentiment analysis at the word level refers to determining sentiments about specific entities and their aspects in a text document. It comprises analysis of sentiment of an individual, service or its aspects, product or its aspect, brand, or any other entity. For example, the review, “I am fully satisfied with Samsung’s battery’s performance” expresses a positive

*International Journal on Recent and Innovation Trends in Computing and Communication*
*ISSN: 2321-8169 Volume: 11 Issue: 11s*
*DOI: https://doi.org/10.17762/ijritcc.v11i11s.8186*
***Article Received:** 29 June 2023 **Revised:** 26 August 2023 **Accepted:** 10 September 2023*

sentiment for the aspect battery of the entity Samsung. Phrase-level sentiment analysis refers to the process of detecting the sentiment (positive, negative or neutral) of a specific phrase or sentence within a larger piece of text. In sentence-level sentiment analysis, the sentiment of the complete sentence or text is detected. It is especially used for the task of subjectivity classification. In document-level sentiment analysis, the entire text document is considered to express the sentiment of single object. Document-level sentiment analysis assigns overall sentiment (positive, negative or neutral) to the entire document.

The most often used methods for sentiment analysis involves machine learning approaches, deep learning approaches, lexicon-based approaches and hybrid approaches [4,5,6,7]. The dataset is split into training and testing datasets in machine learning or deep learning methodologies. During the training phase, a training data set is employed so that the model learns to correlate the specific input text to the corresponding result to learn the documents. During the prediction phase, the testing dataset is utilized to turn concealed textual inputs into feature vectors using a feature extractor. These vectors are then entered into the model, which generates prediction labels for the relevant vectors. The lexicon-based approach employs a pre defined dictionary of words annotated with sentiments. Sentiment ratings are frequently applied in conjunction with other criteria to reduce the number of phrases containing sarcasm, dependent clauses, or negations. The rules include NLP methods such as tokenization, stemming, part-of-speech (PoS) tagging, and lexicons. The hybrid approach of sentiment analysis combines a Lexicon-based approach with a machine learning-based approach [8].

In recent years, a significant improvement in the overall performance of sentiment analysis techniques has been observed due to the advent of several types of techniques for feature extraction. These techniques play an important role for the accuracy of sentiment analysis [2,3]. Detection of unreliable and ambiguous features for classification are the major challenges encountered during the feature extraction step. In order to overcome these kinds of challenges, it needs to develop a hybrid model of feature extraction for a precise result. For concise and trustworthy feature extraction, in addition to sentence or review particular features, specific product/entity/topic, its aspect, and emoticons should be considered.

The approach proposed in this study has two distinct levels of review mining. In the first level, the reviews from each user are classified either positive or negative. In second level, sentiment bearing strings are extracted from each review and sentiment score of each sentiment bearing string is calculated. Based on the sentiment score of each sentiment bearing string in the review, the compound sentiment score of each review is calculated.

In this paper, we introduce a novel sentiment analysis approach based on extraction of review associated features, which is utilized to address the problem of generating efficient sentiment analysis results. The rest of this paper is organized as follows: section 2 describes the study of some recent research papers related to sentiment analysis and deep learning techniques. Section 3 describes the proposed methodology and deep learning model for sentiment classification. The result and discussion is described in section 4. In the last section, section 5 conclusion of the paper is explained.

## II. RELATED WORK

Since the approach presented in this work mainly focuses on an efficient way for sentiment analysis, this section includes several current research studies divided into two categories: Sentiment analysis and Deep Learning-based Sentiment analysis techniques.

### *A. Sentiment Analysis Techniques*

This section includes the many recent studies the field of sentiment analysis. Ayyub et al. [9] investigated a range of classifiers and feature sets for sentiment quantification. A quantitative performance evaluation is conducted based on the explored feature set for machine learning based techniques, ensemble based techniques and deep learning based techniques. The results indicates that differing feature sets have an effect on the performance of classifier in sentiment quantification. Furthermore, the results show that the deep learning approaches perform better in comparison to traditional machine learning based approaches.

Iqbal and Hashmi [10] provide a comprehensive framework for bridging the gap between lexicon-based approaches and machine learning-based techniques. A novel Genetic Algorithm-based feature reduction approach is developed to solve the scalability problem that arises as the feature collection expands. Using the proposed hybrid technique, the authors were able to reduce the size of the feature set by 42% while preserving accuracy.

Munuswamy et al. [11] propose a new method called sentiment-based rating prediction method to develop a recommendation system that is capable of mining valuable information from user reviews posted on social media platforms to predict exactly which products are particularly liked by users depending on their ratings. In this approach, user opinions on an item are computed using a sentiment dictionary. In the next step, in order to predict and generate suitable suggestions, item reputations are computed using the three sentiments. The n-gram approach is used as a trendy

*International Journal on Recent and Innovation Trends in Computing and Communication*
*ISSN: 2321-8169 Volume: 11 Issue: 11s*
*DOI: https://doi.org/10.17762/ijritcc.v11i11s.8186*
***Article Received:** 29 June 2023 **Revised:** 26 August 2023 **Accepted:** 10 September 2023*

feature in semantic analysis and syntax, together with SVM, to improve the accuracy of the findings for efficient classification of reviews posted on social media platforms.

Oyebode et al. [12] analyze 104 mental health applications on the App Store and Google Play using five supervised machine learning approaches to conduct sentiment classification. The experiment was performed on 88,125 user reviews. To predict the sentiment polarity of review, the classifier having best performance was used. The authors then identified the themes that reflect several aspects that impact the success of mental health applications in both good and negative ways using thematic analysis of negative and positive evaluations.

To extract the features from online posts of movie review Khan and Gul [13] utilized the Bag of Words (BoW) approach and also used this approach for vector representation of these extracted features. The authors then used the Naive Bayes machine learning technique to classify movie reviews expressed as feature vectors into positive and negative categories. In the next step, an undirected weighted network was constructed utilizing the pairwise semantic similarities between classified review sentences. The Review sentences are represent by the nodes and semantic similarity weights are represented by the edges of the graph. The weighted graph-based ranking algorithm (WGRA) was used to obtain the absolute measure for all of the reviews in the graph. Finally, the extracted summary was generated by choosing the highest-rated phrases (graph nodes) based on the highest-ranking metrics.

Zhu et al. [14] proposed a kernel optimization approach-SentiVec for sentiment word embedding. Experiment has been performed in two staged. The first stage of this research focuses on supervised learning, while the second phase focuses on unsupervised updating models such as object-word-to-surrounding-word reward models (O2SR) and context-to-object-word reward models (C2OR). The results of the experiment show that semantics and sentiment analysis features are retrieved effectively by optimum sentiment vectors and performs much better than baseline techniques on tasks including sentiment analysis, similarity detection and word analogy.

### *B. Deep Learning-Based Techniques*

Deep learning-based algorithms are currently being implemented for sentiment analysis in order to improve efficiency. With the introduction of deep learning, text representation models have been replaced by various tasks. In fact, the advancement of neural techniques overcomes the limits of hand-crafted features by utilizing machine-learning embedding models that represent a text in feature vector having low-dimensional.

Deriu et al. [15] combined two convolutional layers with two sequential pairs of pooling layers in their study. By categorizing tweet data from multilingual sentiment datasets, the suggested CNN model achieved an F1-score of 67.79%. Kim & Jeong [16] used a CNN model to perform binary sentiment classification on Amazon and MR datasets and achieved accuracy scores of 81.4% and 81.6% respectively.

Zhang et al. [17] created 3W-CNN, a three-way improved CNN model with the objective to optimize CNN networks and improve sentiment categorization. The suggested model was tested on four different datasets: customer reviews (CR), MR, MPQA dataset and the Subjectivity dataset (SUBJ). Kumar et al. [18] proposed a technique sentiment analysis which includes three methods: generating ontologies to retrieve semantic features, converting processed corpora with Word2vec, and developing convolutional neural networks (CNNs) to analyze the sentiments. For CNN parameter tuning, the technique of particle swarm optimization is used to identify non-dominant Pareto front optimum values.

To improve the capacity of categorizing the sentiment of short phrases, Bao et al. [19] proposed BERT-based text representation and a hybrid model of CNN and attention-based Bi-GRU. Attention-based Bi-GRU units and CNN were run in parallel, and their outputs were concatenated to provide the final prediction. Meena et al. [20] suggested a model for sentiment classification which employs Keras embedding to produce feature vectorization and CNN to train a final classifier on the deep information of text input. Zhang et al. [21] present a Broad Multitask Transformer Network (BMT-Net) which integrates a feature-based method with a fine-tuning technique. the objective of the proposed system is to analyze the high-level information of robust and contextual representations. Using multitask transformers, the proposed system allows global representations to be learnt across tasks. BMT-Net can firmly learn the robust contextual representation employed by the wide learning system because to its capacity to seek for acceptable features in deep and broad ways.

M Qorich, R El Ouazzani [22] used CNN model to perform their experiment on Amazon dataset and the various model approaches performed admirably to classify the reviews in positive and negative sentiment categories. Their findings show the necessity of incorporating stop-words in sentiment analysis tasks; in fact, removing stop words can lead to inaccurate sentiment prediction. In their experiment, they observed that utilizing stop words with the CNN model enhanced accuracy by 2% when compared to the CNN model that disregarded them.

*International Journal on Recent and Innovation Trends in Computing and Communication*
***ISSN: 2321-8169 Volume: 11 Issue: 11s***
***DOI: https://doi.org/10.17762/ijritcc.v11i11s.8186***
***Article Received:** 29 June 2023 **Revised:** 26 August 2023 **Accepted:** 10 September 2023*

In addition, works exploring the more advanced embedding approach, BERT, and its derivatives in enhancing sentiment analysis for reviews were discovered in the literature. To analyze the review of customers, Wu F et al. [23] proposed a model, SenBERT-CNN has been introduced. SenBERT-CNN model combines a pre-trained Bidirectional Encoder Representations from Transformers (BERT) network with Convolutional Neural Network (CNN) to capture more sentiment information in phrases. They perfoemd their experiment on customer review data of smart phone and achieved the accuracy of 95.72%.

## III. PROPOSED METHODOLOGY

It is observed from recent aforementioned recent studies that sentiment analysis has become a prominent research area due to a variety of exciting and demanding topics to investigate. Because of the broad application area, Sentiment analysis is demanding mostly in all domain of business applications. There are number of start-up firms as well as established businesses seek to provide sentiment analysis services. For longer existence in the competitive market, business organizations need to gain insight into consumer reviews about their product and services because review explains the consumer view for the product and services. Due to technical issues and operational aspects, the domain of sentiment analysis will be in high demand for long into the future.

In this section, methodologies used in this work are described. The steps used in this study are represented in Figure 1.

### *A. Data collection*

Data collection is the initial step for the task of sentiment analysis. Amazon mobile review dataset has been utilized to evaluate the proposed model and compare it to the most recent state-of-the-art approaches. The dataset has been downloaded from Kaggle website. The dataset contains 67987 data samples. In the first stage, relevant features are extracted from the dataset and eliminated all other features that are not in use. Finally, there are just two columns in the dataset which are selected for the analysis: body and rating.

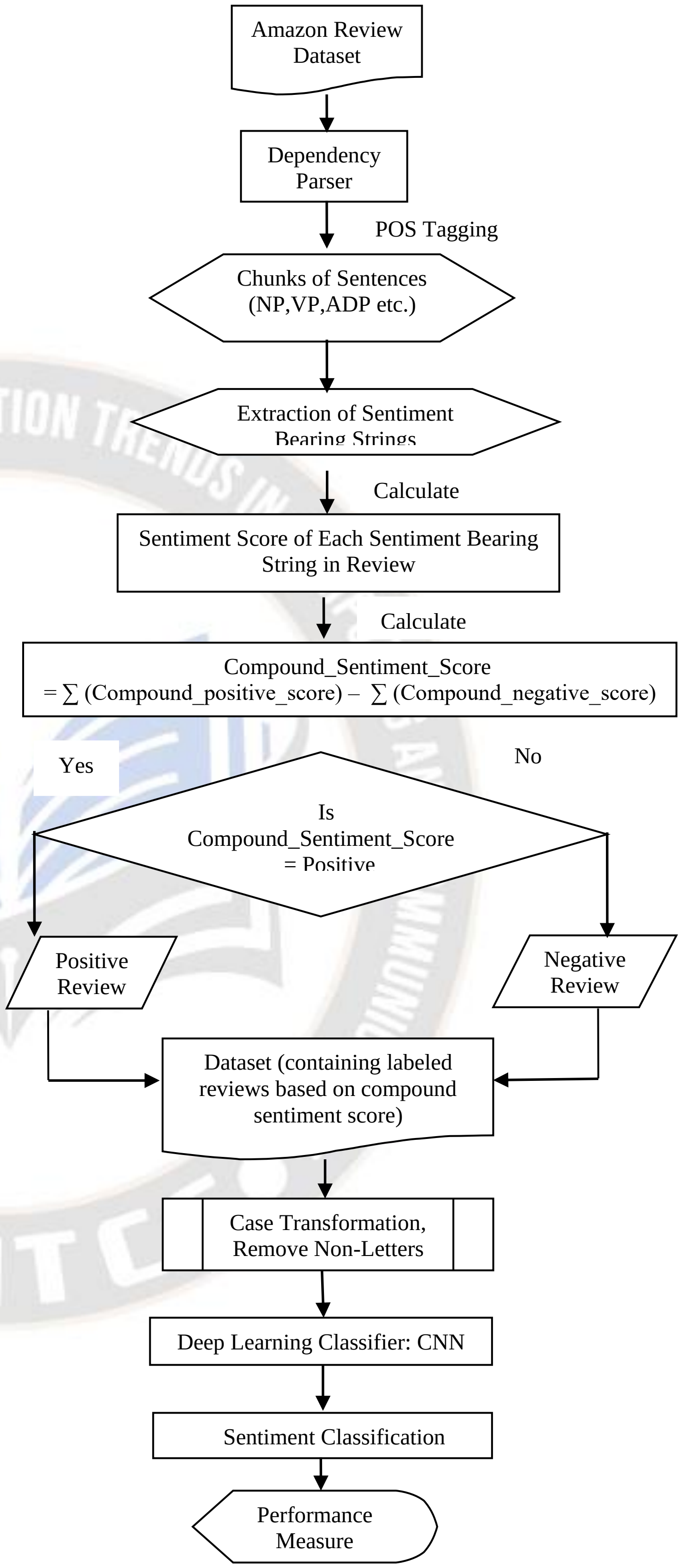


Figure 1. Process steps of proposed methodology.

### *B. Dataset description*

The dataset contains the following properties as columns.

*International Journal on Recent and Innovation Trends in Computing and Communication*
*ISSN: 2321-8169 Volume: 11 Issue: 11s*
*DOI: https://doi.org/10.17762/ijritcc.v11i11s.8186*
***Article Received:** 29 June 2023 **Revised:** 26 August 2023 **Accepted:** 10 September 2023*

**asin**: This column represents product ASIN (Amazon Standard Identification Number) which is unique for each mobile or mobile variable.
**name:** This column represents the name of the reviewer.
**rating**: This column represents the rating of the reviewer for mobile on scale of 1 to 5.
**date:** date of the review.
**verified:** This column represents valid customer as true or false. The dataset contains the reviews from 90% valid users.
**title:** Represents tile of the review.
**body:** This column represents the review contents of the reviewer.
**helpfulness:** helpfulness feedback.

*C. Part-of-Speech (POS) Tagging*

In the proposed work, sentiment analysis is performed on text data. Because It analyses text content, sentiment analysis is a Natural Language Processing (NLP) technique. The sentiments or emotions are extracted from text using sentiment analysis. The sentiment or emotion can be positive, negative or neutral. Since the sentiment analysis process is performed on text data, a number of steps are applied for preprocessing of data before the final classification. For preprocessing, Parts-of-Speech tagging (POS) technique has been applied. It includes assigning POS tags or labels to each word in each review sentence, eliminating stop words and irrelevant words, identifying and extracting commonly used words and identifying and extracting adjective from the review sentence.

Product reviews posted by customers are in text form. POS tagging is a technique used for categorizing each word present in a sentence. This technique is employed to determine the suitable part of speech to each word. It contains nouns, pronouns, verbs, adverbs, adjectives, and prepositions, among other things. Since nouns and adjectives play important role to express sentiment or emotion of an entity, therefore they are generally used for sentiment analysis. Each sentiment bearing string is extracted from the review by applying POS tagging technique.

*D. Computation of Sentiment Score*

Lexicon has been employed in previous studies to analyze sentiment. Generally, lexicons are domain specific, it is impossible to utilize same lexicon in any other domain. The proposed method does not need the use of a vocabulary or a corpus to train the model. It is based on the sentiment score and the sentiment score of a review is calculation is based on the extracted sentiment bearing strings present in the review.

A review may contain both positive as well as negative strings. Therefore the final sentiment of the review depends on all the strings presented in the review. To compute the sentiment score of all extracted sentiment bearing strings, VADER (Valence Aware Dictionary and Sentiment Reasoner) is used. It provides a list of four sentiment scores- pos (positive), neu (neutral), neg (negative) and compound. Positive, negative and neutral score are the corresponding proportions of in the text and the compound score represents overall sentiment of the text, taking polarity and intensity into consideration. The sentiment score of a review is calculated by applying following formula:

$$\text{Sentiment score} = \sum \text{compound\_positive\_score} - \sum \text{compound\_negative\_score} \quad (1)$$

Where,

Sentiment score is the final sentiment score of the review, compound positive score is the compound score of strings having positive sentiments and compound negative score is the compound score of string bearing negative sentiments.

Finally, sentiment score of a review is the difference between the entire sum of the compound scores of all positive and negative sentiment carrying strings. A review is classified as either positive or negative based on its compound sentiment score. This process is repeated for all of the reviews of dataset to produce the compound sentiment score to all reviews, and each review is assigned a class based on this score. If a review's compound sentiment score is positive, the review is classified as positive. Otherwise, it is classified as negative.

*E. Data Pre-processing*

The dataset of raw reviews is pre-processed by eliminating and correcting the complicated and ineffective text. The purpose of preprocessing is to prepare data for analysis and to organize the data so that it can be instantly comprehended and analyzed in order to obtain significant insights. In the proposed work, data preprocessing starts with data cleaning and tokenization and concludes with the elimination of digits, special characters, ineffective words.

The review data is gathered from internet sources that typically include irrelevant information such as number, scripts, URLs, HTML tags, , and other special characters. The unnecessary data is eliminated throughout from the review, in data cleaning process, and leaving only relevant text. In tokenization process, tokenization function is used to split input review into tokens. The token may contain stop-words which do not convey any sentiment. Therefore, to reduce the amount of irrelevant tokens, stop-words such as 'the', 'an', 'is', 'am', ’are’, and so on are removed. After tokenization, stemming process has been performed to reduce tokens into their base forms (e.g., “ran”, “running”, and "runner” get reduced to their base form “run”). After preprocessing, the dimension space of

*International Journal on Recent and Innovation Trends in Computing and Communication*
***ISSN: 2321-8169 Volume: 11 Issue: 11s***
***DOI: https://doi.org/10.17762/ijritcc.v11i11s.8186***
***Article Received:** 29 June 2023 **Revised:** 26 August 2023 **Accepted:** 10 September 2023*

raw review is reduced. Table I shows a few instances of reviews both before and after pre-processing.

TABLE I. SAMPLES OF REVIEWS AFTER AND BEFORE APPLYING PREPROCESSING TECHNIQUE.

| Review before preprocessing | Review after preprocessing |
|---|---|
| Due to a software issue between Nokia and Sprint this phone's text messaging capabilities don't work with Sprint's system and won't until a software patch comes out "some time in the next few months". You will have to spend at least 1 hour with Sprint's award winning customer service team to find someone who will admit this to you. The problem is that Nokia designed their phones so that incoming messages are retrieved quickly and then viewed "offline" the way most providers work. Sprint, however, likes to have people hook up to their server first and then stay connected, burning minutes while they check their inbox, compose a reply and wait for the Sprint server to respond so they can send it out. Innovation in money-making at its finest. | due software issue nokia sprint phones text messaging capabilities work sprints system software patch comes time next months spend least 1 hour sprints award winning customer service team find someone admit problem nokia designed phones incoming messages retrieved quickly viewed offline way providers work sprint however likes people hook server first stay connected burning minutes check inbox compose reply wait sprint server respond send innovation moneymaking finest |
| I'm not providing a review of the phone itself, however beware: the posting indicates very specifically that there is a SIM included in the box, but there most definitely is not. Here's hoping I don't have to pay too much more at my carrier to get one. | providing review phone however beware posting indicates specifically sim included box definitely hoping pay much carrier get one |

## *F. Review Associated Feature Extraction (RAFE)*

A review contains explanation of the entity or explanation of its aspects generally in terms of text. The sentiment of the review depends on the presence of the sentences in the review. Review associated features are a sentence-level feature representation strategy for representing sentiments, opinions and emotions from input text of reviews. The conventional techniques for feature extraction like Bag of Words (BoWs) and TF-IDF suffer from many issues such as they ignore the order of words in the text, they generate a sparse feature vector, and uninformative features can also be present in the text.

To address above issues, review associated feature extraction techniques has been used. By applying this technique sentence-level feature are extracted which contains the sentiment bearing strings. The steps used in the proposed work are as following:

1. **Input:** Reviews in raw form.
2. **Tokenization:**
   - Tokenized the input review into individual tokens by using dependency parser.
3. **Part of Speech (POS) Tagging:**
   - Assigned a POS tag to each token
   - Identified noun, pronoun, adjective, verb, adverbs other related phrases
4. **Sentiment-Bearing POS Tags identification:**
   - Identified and determined POS tags (e.g. nouns, adjectives, adverbs,) which indicates sentiment-bearing words.
5. **Sentiment-Bearing Sentence extraction:**
   - Check for each sentence, whether it contains at least one word with a sentiment-bearing POS tag.
   - If it contains, the sentence is considered as a sentiment-bearing sentence.
   - Return or store the sentiment-bearing sentences.
6. **Output:**
   - Return or save the sentiment-bearing sentences extracted from the input review text.

After applying review associated feature extraction technique on dataset, the sentences in the review contain sentiment-bearing words or phrases. Figure 2 shows the snapshot of some instances of dataset after applying RAF approach. As observed, the first column includes the only features contained in the review, the second column contains the computed sentiment score and third column sentiment either positive or negative based on the sentiment score.

| review associated features | Compound_sentiment_score | sentiment |
|---|---|---|
| ['Also damage' 'actually charges' 'remarkable feat' 'charges quickly' | -0.4489 | neg |
| however likes | 0.4215 | pos |
| ['great phone' 'nice feature'] | 1.0464 | pos |
| free accessories | 0.5106 | pos |
| clear casing | 0.3818 | pos |
| ['great deal' 'original one'] | 0.9431 | pos |
| ['amazing ringtones' 'good value' 'lucky person' 'NOW convinced'] | 2.0579 | pos |
| ['highly recommend' 'really like'] | 0.8402 | pos |
| ['cute phone' 'great phone' 'ideal phone' 'easy to use' 'useful features' | 1.8755 | pos |
| ['nice ability' 'better phone' 'useful feature'] | 1.5057 | pos |
| ['poor quality' 'GOOD reception' 'dead phone' 'really wish' | 0.9497 | pos |
| ['excellent choice' 'free option'] | 1.0825 | pos |
| ['best cell company' 'best phones'] | 1.2738 | pos |
| easy to use | 0.4404 | pos |
| definitely recommend | 0.6369 | pos |
| ['highly recommended' 'well loved' 'bad customer service' 'same mistake' | -0.6894 | neg |
| ['easily scratched' 'then perfect' 'hit accidentally'] | 0.5719 | pos |
| ['Sturdy clarity' 'easy to use'] | 0.8423 | pos |
| ['beautiful job' 'good quality' 'Very good quality'] | 1.5325 | pos |
| ['more entertainment' 'throw hard'] | 0.3727 | pos |
| ['always irritated' 'handle well' 'highly recommend'] | 0.2345 | pos |
| stick easily | 0.34 | pos |
| ['good price' 'very good price' 'highly recommend' 'better comp'] | 1.7936 | pos |
| ['great pay' 'strong signal' 'great battery life'] | 1.7074 | pos |

Figure 2. Snapshot of dataset after applying RAF approach.

*International Journal on Recent and Innovation Trends in Computing and Communication*
***ISSN: 2321-8169 Volume: 11 Issue: 11s***
***DOI: https://doi.org/10.17762/ijritcc.v11i11s.8186***
***Article Received:** 29 June 2023 **Revised:** 26 August 2023 **Accepted:** 10 September 2023*

### G. Experimental Analysis

This section describe the experiment setup applied for the proposed work. The dataset contains a column 'Rating' which stores rating value in numeric form provided by the customers corresponding to each review. The value of rating ranges from 1 to 5. The rating value 1 indicates the very negative sentiment and 5 indicates very positive sentiment about the product. Based on this rating value, the sentiments of the reviews are inferred and stored under newly appended column 'Sentiments' in dataset. The sentiment is either positive or negative and computed as following:

- Positive: if the rating value is >= 3, it is represented by 1 in 'Sentiments' column of dataset.
- Negative: if the rating value is < 3, it is represented by 0 in 'Sentiments' column of dataset.

As a consequence of this task, we created a labeled dataset for sentiment analysis. To clean the reviews, preprocessing techniques are applied on each review of dataset. The dataset is partitioned into training and testing datasets respectively in the ratio of 70:30. For sentiment analysis, Convolutional Neural Network (CNN) model is trained, tested and analyzed the result.

Further, we extracted sentiment bearing strings using review associated feature extraction approach and computed the sentiment score of each sentiment bearing string presented in the review and finally computed the compound sentiment score of review. On the basis of compound sentiment score, a review is classified either positive or negative as:

- Positive: if the value of compound sentiment score of review is positive, it labeled as 'pos' in dataset.
- Negative: if the value of compound sentiment score of review is negative, it labeled as 'neg' in dataset.

As a result, a labeled dataset is produced on the basis of sentiment score of the review. Again we portioned the dataset into training and testing parts, preprocessed the dataset, trained the model for sentiment analysis, tested the model and analyzed the result.

### H. Sentiment Classification Using CNN

We utilized CNN classifier to classify input review associated feature vector into positive or negative sentiment classes. Here, in CNN model convolutional layers, pooling layers, dense layers and dropout layers are used. In CNN model, majority of the time spent in training the neural network is utilized in convolution. Meanwhile, the full-connected layer controls the majority of the network's characteristics. The main objective of convolution is to extract the input feature and the objective of pooling is to sample the convolution matrix.

Figure 3 shows the example of CNN model, the model is simplified convolutional network of the CNN for sentence classification [24]. To begin, we must convert the phrase into a matrix, with the rows of each sentence matrix representing word vector representations.

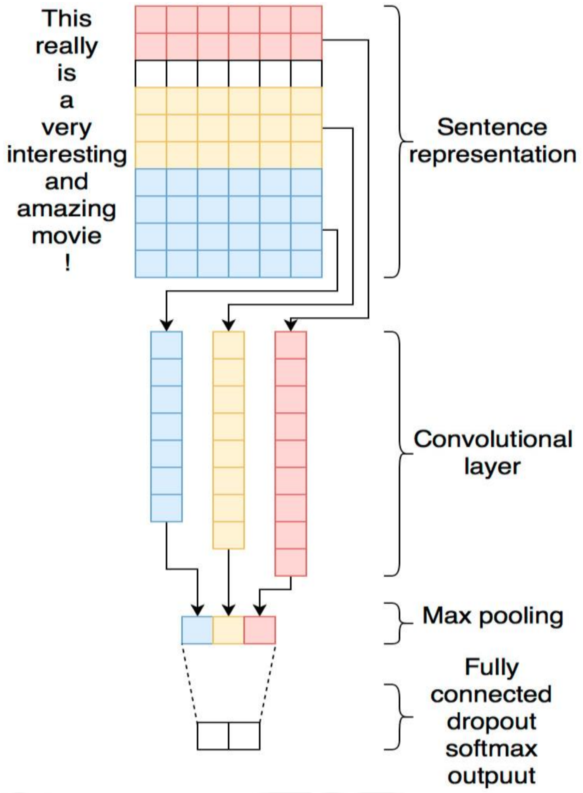


Figure 3. Architecture of CNN model [24].

In our CNN model, there are three convolution layers to process the input. The first convolution layer (Conv1D) layer has a kernel size of 4 and 128 filters (output channels). It uses 128 distinct 1D filters on the input embeddings to capture local patterns in the word sequence. Rectified Linear Unit (ReLU) is the activation function employed. After first convolution layer, a MaxPooling1D layer with a pool size of 2 is utilized to decrease the dimensionality of the feature maps while retaining the most necessary and essential information. With a kernel size of 4, the second convolution layer(Conv1D) layer includes 64 filters (output channels). It, like the first layer, applies 64 separate 1D filters on the previous layer's output, acquiring higher-level patterns and features. ReLU is employed as the activation function once more, after this another MaxPooling1D layer with a pool size of 2 is applied. The third and last convolution (Conv1D) layer includes 32 filters (output channels), with a kernel size of 4. It decreases the dimensionality even further and abstracts even higher-level characteristics. The activation function for this

*International Journal on Recent and Innovation Trends in Computing and Communication*
***ISSN: 2321-8169 Volume: 11 Issue: 11s***
***DOI: https://doi.org/10.17762/ijritcc.v11i11s.8186***
***Article Received:** 29 June 2023 **Revised:** 26 August 2023 **Accepted:** 10 September 2023*

layer is ReLU. After this convolution layer, a final MaxPooling1D layer with a pool size of 2 is used. In the next step, to convert the 2D output from the last MaxPooling1D layer into a 1D vector, flatten layer is used. This stage flattens the data in preparation for the fully connected layers. After flattening, in the nest step two fully connected Dense layers are used in the mode. The first Dense layer functions as a hidden layer and consists of 256 units. It assists in learning of higher -level representations and patterns from the features that has been extracted. In this layer, ReLU is applied as activation function. The second Dense layer consists of single unit and it represents the output layer of the model for sentiment classification. In this layer, sigmoid function is used as activation function.

## IV. RESULT AND DISCUSSION

This section describes the results of our experiment. The experiment is carried out in two steps, in the first step, the experiment is performed on dataset by applying CNN model using training dataset and performance of the model is evaluated using test dataset. The result shows the accuracy of 89.59% of CNN model. In the second step, the proposed approach, review associated feature (RAF) extraction approach has been applied on the dataset and then applied CNN model on the dataset to classify the review in its respective category. After applying review associated feature extraction approach, it is observed that the performance of the applied model is much better having accuracy of 97.23%.

In order to demonstrate the effectiveness of the review associated feature extraction approach, the performance of CNN model is compared before and after applying the proposed technique. The parameters that are used to compare the performance of the model are: precision, recall, F-1 score and accuracy.

The dataset contains both positive and negative reviews. The applied model classifies the reviews in their respective category. The performance is observed before and after applying the proposed approach. Figure 4 and figure 5 show the performance measure of the model for the classification of positive reviews and negative reviews respectively in terms of precision, recall, F1-score and accuracy. A significant improvement in the performance has been observed after applying the proposed approach for both the classes of sentiment.

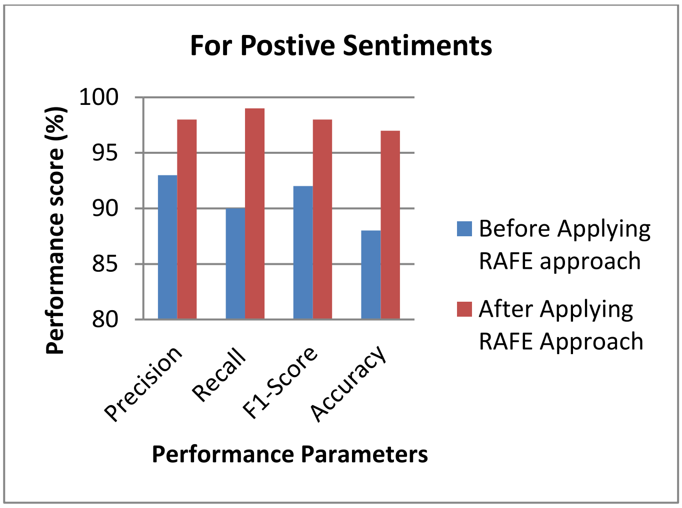


Figure 4. Classification performance of CNN model using RAFE approach to classify positive sentiments.

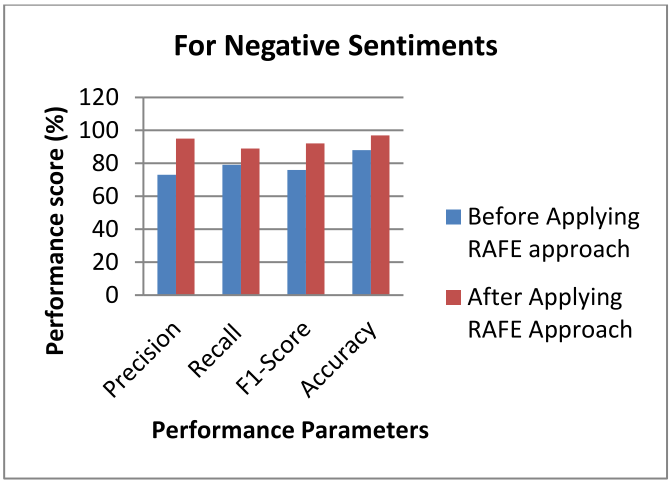


Figure 5. Classification performance of CNN model using RAFE approach to classify negative sentiments.

In addition, we have compared the classification accuracy of our CNN model with the accuracies achieve by several approaches used recently. The comparison of the accuracy of recently used models and our model is presented in Table II. As well, Figure 6 and figure 7 show the training accuracy and training loss of our proposed CNN model on Amazon Mobile Review dataset.

TABLE II. THE COMPARISON CLASSIFICATION ACCURACY OF THE BASELINE MODELS WITH PROPOSED CNN MODEL

| Reference | Model | Dataset | Accuracy (%) |
|---|---|---|---|
| [16] | CNN | Amazon Review | 81.4 |
| [17] | CNN | Amazon Review | 85.8 |
| [23] | SenBERT-CNN | Customer Review | 95.72 |
| [22] | CNN | Amazon Review | 90.00 |
| Proposed Model | RAF+CNN | Amazon Mobile Review | 97.23 |

*International Journal on Recent and Innovation Trends in Computing and Communication*
***ISSN: 2321-8169 Volume: 11 Issue: 11s***
***DOI: https://doi.org/10.17762/ijritcc.v11i11s.8186***
***Article Received:** 29 June 2023 **Revised:** 26 August 2023 **Accepted:** 10 September 2023*

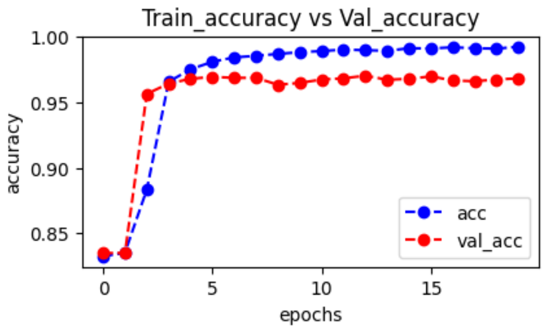


Figure 6. Training accuracy of our CNN model on Amazon Review Dataset.

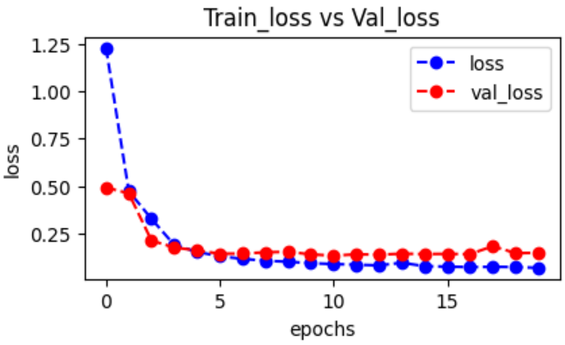


Figure 7. Training loss of our CNN model on Amazon Review Dataset.

From the obtained result, it is observed that our CNN model outperforms than other CNN model proposed recently on Amazon dataset. It also performs better than the baseline deep learning and machine learning approaches in terms of accuracy. Table 3 shows that our CNN model achieves nearly 12% better accuracy than higher baseline CNN model and also achieve nearly 2% better accuracy than SenBERT-CNN model.

## V. CONCLUSION

The main objective of this research work is to effectively perform task of sentiment analysis on Amazon mobile review dataset using review associated feature extraction approach. We addressed a number of research issues as well as alternate options to the issues encountered during the process of sentiment analysis on online user reviews. The approach review associated feature extraction, proposed in this research paper can used to transform the reviews into comparatively lower dimension feature vectors that enables effective and comprehensive sentiment analysis. The experimental outcomes show that the proposed approach significantly enhances the performance of sentiment analysis as compared to baseline deep learning approaches and other recently used approaches. The future prospects for the RAF+CNN model involves to expand the proposed model for aspect based sentiment analysis, review summarization and exploring the performance for different datasets.

## REFERENCES

[1] Singh J, Singh G, Singh R (2016) A review of sentiment analysis techniques for opinionated web text. CSIT 4:241–247
[2] Shahana PH, Omman B (2015) Evaluation of features on sentimental analysis. Procedia Comput Sci 46:1585–1592. https://doi.org/10.1016/j.procs.2015.02.088
[3] Parlar T, Özel SA, Song F (2018) QER: a new feature selection method for sentiment analysis. HCIS 8(1):1–19. https://doi.org/10.1186/s13673-018-0135-8
[4] Liu CL, Hsaio WH, Lee CH, Lu GC, Jou E. Movie Rating and Review Summarization in Mobile Environment. IEEE Trans Syst Man Cybern. 2012a;42(3):397–407. https://doi.org/10.1109/tsmcc.2011.2136334.
[5] Pannala NU, Nawarathna CP, Jayakody JTK, Rupasinghe L, Krishnadeva K. Supervised Learning Based Approach to Aspect Based Sentiment Analysis. 2016 IEEE International Conference on Computer and Informa- tion Technology (CIT). 2016. p 662–666.
[6] Schouten K, Frasincar F. Survey on aspect-level sentiment analysis. IEEE Trans Knowl Data Eng. 2016;28(3):813–30. https://doi.org/10.1109/tkde.2015.2485209.
[7] Trilla A, Alias F. Sentence-based sentiment analysis for expressive text-to-speech. IEEE Trans Audio Speech Lang Process. 2013;21(2):223–33. https://doi.org/10.1109/tasl.2012.2217129.
[8] de Kok S, Punt L, van den Puttelaar R, Ranta K, Schouten K, Frasincar F. Review-aggregated aspect-based sentiment analysis with ontology features. Progr Artif Intell. 2018;7(4):295–306. https://doi.org/10.1007/s13748-018-0163-7.
[9] Ayyub K, Iqbal S, Munir EU, Nisar MW, Abbasi M. Exploring diverse features for sentiment quantification using machine learning algorithms. IEEE Access. 2020;8:142819–31. https://doi.org/10.1109/access.2020.3011202.
[10] Nooruldeen Mohammed Kareem, & Asma Thamir Ibraheem. (2023). Developing a Frame Design for Airport Pavements Maintenance Management System. International Journal of Intelligent Systems and Applications in Engineering, 11(4s), 498–508. Retrieved from https://ijisae.org/index.php/IJISAE/article/view/2718.
[11] Flyand. A hybrid framework for sentiment analysis using genetic algorithm based feature reduction. IEEE Access. 2019;7:14637–52.
[12] Munuswamy S, Saranya MS, Ganapathy S, Muthurajkumar S, Kannan A. Sentiment analysis techniques for social media-based recommendation systems. Nat Acad Sci Lett. 2021;44(3):281–7. https://doi.org/10.1007/s40009-020-01007-w.
[13] Oyebode O, Alqahtani F, Orji R. Using machine learning and thematic analysis methods to evaluate mental health apps based on user reviews. IEEE Access. 2020;8:111141–58. https://doi.org/10.1109/access.2020.3002176.
[14] Brian Moore, Peter Thomas, Giovanni Rossi, Anna Kowalska, Manuel López. Machine Learning for Customer Segmentation and Decision Making in Marketing. Kuwait Journal of Machine Learning, 2(4). Retrieved from http://kuwaitjournals.com/index.php/kjml/article/view/218.

[15] Khan A, Gul MA, Zareei M, Biswal RR, Zeb A, Naeem M, Saeed Y, Salim N. Movie review summarization using supervised learning and graph-based ranking algorithm. Comput Intell Neurosci. 2020;2020:1–14. https://doi.org/10.1155/2020/7526580.

[16] Zhu L, Li W, Shi Y, Guo K. SentiVec: learning sentiment-context vector via kernel optimization function for sentiment analysis. IEEE Trans Neural Netw Learn Syst. 2021;32(6):2561–72. https://doi.org/10.1109/tnnls.2020.3006531.

[17] Deriu J, Lucchi A, De Luca V, et al (2017) Leveraging large amounts of weakly supervised data for multi-language sentiment classification. In: Proceedings of the 26th International Conference on World Wide Web, pp 1045–1052

[18] Dr. S. Praveen Chakkravarthy. (2020). Smart Monitoring of the Status of Driver Using the Dashboard Vehicle Camera. International Journal of New Practices in Management and Engineering, 9(01), 01 - 07. https://doi.org/10.17762/ijnpme.v9i01.81

[19] Kim H, Jeong YS (2019) Sentiment classification using convolutional neural networks. Appl Sci 9(11):2347

[20] Zhang Y, Zhang Z, Miao D et al (2019) Three-way enhanced convolutional neural networks for sentence-level sentiment classification. Inf Sci 477:55–64

[21] Kumar R, Pannu HS, Malhi AK. Aspect-based sentiment analysis using deep networks and stochastic optimization. Neural Comput Appl. 2020;32(8):3221–35. https://doi.org/10.1007/s00521-019-04105-z

[22] Bao T, Ren N, Luo R, Wang B, Shen G, Guo T. A BERT-based hybrid short text classification model incorporating CNN and attention-based BiGRU. J Organ End User Comput. 2021;33(6):1–21.

[23] Meena G, Mohbey KK, Indian A. Categorizing sentiment polarities in social networks data using convolutional neural network. SN Comput Sci. 2022;3:116.

[24] Zhang T, Gong X, Chen CLP. BMT-Net: broad multitask transformer network for sentiment analysis. IEEE Trans Cybern. 2021;52(7):6232–43. https://doi.org/10.1109/tcyb.2021.3050508.

[25] M Qorich, R El Ouazzani. Text sentiment classification of Amazon reviews using word embeddings and convolutional neural networks. The Journal of Supercomputing volume 79, pages11029–11054 (2023)

[26] Kulkarni, L. . (2022). High Resolution Palmprint Recognition System Using Multiple Features. Research Journal of Computer Systems and Engineering, 3(1), 07–13. Retrieved from https://technicaljournals.org/RJCSE/index.php/journal/article/view/35

[27] Wu F, Shi Z, Dong Z, Pang C, Zhang B (2020) Sentiment analysis of online product reviews based on SenBERT-CNN, 2020 International Conference on Machine Learning and Cybernetics (ICMLC), pp 229–234, https://doi.org/10.1109/ICMLC51923.2020.9469551.

[28] Kim Y, Convolutional Neural Networks for Sentence Classification, arXiv preprint, 2014, arXiv: 1408.5882.