International Journal of Advanced Technology & Engineering Research (IJATER)
www.ijater.com

FEATURE BASED SENTIMENT ANALYSIS OF PRODUCT REVIEWS USING DEEP LEARNING METHODS

Supriya Soni[1],Shivendra Dubey[2], Rakesh Tiwari[3],Mukesh Dixit[4]
[1]M. Tech Research Scholar, Department of Computer Science,
[2]Research Guide, Department of Computer Science,
[3]Co-Guide, Department of Computer Science,
[4]Head, Department of Computer Science,
supriyasoni2557@gmail.com[1]
shivendra.dubey5@gmail.com[2]
rits.mukesh@gmail.com[3]
Radharaman Engineering College, Bhopal

Abstract

In online product reviews users discuss about products and its features. A product may have hundreds or thousands of reviews, consumers share their experience about products and comments about products characteristics. These product reviews may have positive or negative sentiments. A positive sentiment contains good opinion about product and its features similarly a negative sentiment tells drawbacks and problems of product and its features. Features or aspects are part of the product or its characteristics. In this study we used feature/aspect based sentiment analysis and some methods for analyzing the sentiments expressed in online product reviews about the various features of products.

Keywords- Product Reviews, Sentiment Analysis, Classification, Supervised Learning, NLP

1. Introduction

Sentiment analysis is the research area where opinions expressed by humans studied, analyzed, processed and find the emotions or opinions from the given content. Contents related to any entities such as products, services, organizations, individuals, issues, events and their related attributes. Sentiment analysis is also called opinion mining mainly focus on processing of opinions and extract positive, negative and neutral sentiments. Sentiments or opinions are very important as they have big impact on human activities and are key influence on our behaviors. Purpose of study of sentiment analysis is that it provides the useful information about positivity and negativity of entity. The rapid growth of the field coincide with those of the social media on the Web, e.g., online review, blogs, micro blogs, shopping websites and social networks, because for the first time in human history, we have a huge volume of sentimental data available in digital forms. In recent years, industries use people's sentiments to analyze current trends about its products or consumer services. Sentiment analysis has found interest in commercial applications as it can use in almost every business and social domain. Sentiment analysis gives summary of online reviews, consumer services whether they have positive or negative orientation which is useful for consumers and manufacturers.

Opinion mining (often referred as Sentiment Analysis) refers to identification and classification of the viewpoint or opinion expressed in the text span; using information retrieval and computational linguistics. The opinion expressed on the topic is given significance rather than the topic itself [1]. Sentiment analysis or opinion mining extracts the subjective informationfrom the source materials such as reviews using techniques such as natural language processing, and text analytics. Opinion plays essential part in our information-gathering behavior before taking a decision. Online review sites, and personal blogs facilitate gathering of sentiments of products or object using information technologies. The main objective of opinion mining is to determine the polarity of comments (positive, negative) by extracting features and components of the object that have been commented on in each document [2, 3]. Studies related to opinion mining, on the implication of economic impact due to the reviews, issues about breach of privacy are given attention.

Generally, the opinion expressed in a review document could either be a direct opinion or comparative opinion. Direct sentiment expressions on some target objects such as products, events, topics, persons. E.g.: "The picture quality of this camera is great." Comparison opinion expresses the similarities or differences of more than one object usually stating an ordering or preference. E.g.: "car x is cheaper than car y." Different types of comparatives are Non equal Gradable (less than), Equative (same), Superlative (longest).

1.1 Architecture of Sentiment Analysis

Sentiment analysis (SA) has attracted great curiosity among the researchers both in academic and industry level. In today's world social media provide the most promising platform for sentiment analysis (SA) because of more intervention of opinions of people on internet [5]. Lots of people post their reviews; write blogs and many reviews websites are available on internet. Architecture of SA shows the steps of extracting the sentiments from large collection of reviews. Figure 1shows the basic architecture of SA.

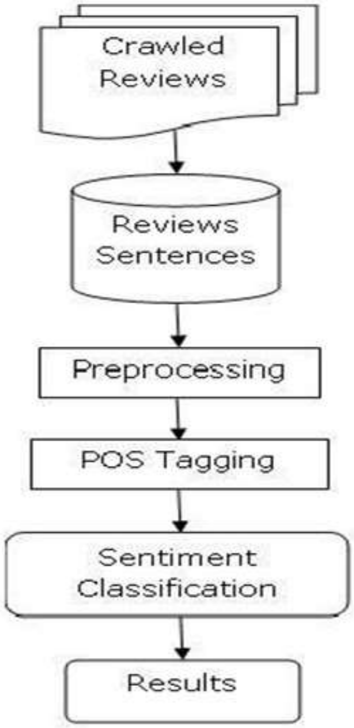

Figure 1: Architecture of Sentiment Analysis

Crawler: It is program that crawls into web pages of reviews and store the reviews inspecial format called indexing files. Indexing files are special data structures that stores text contents in TF-IDF (term frequency – inverse document frequency) format just like book index at the end of book which shows the topics with page number; therefore it is easy to search the reviews.

Preprocessing: Data processing is performed on reviews to remove the stop words andapply stemming algorithm to reduce memory usage.

POS Tagging: Parts-of-speech tagger is used in English language to assign tags to nouns,adverbs, adjective etc. These tagged words or groups (phrases) of words are extracted from the reviews using some patterns.

Semantic Orientation: Semantic orientation of words or phrases is calculated byapplying supervised or unsupervised learning methods. These different methods use statistical data to determine numerical value of semantic orientation.

1.2 Sentiment Classification

Broadly SA is classified (figure 2) as supervised learning techniques and unsupervised learning techniques [6]. There are many different methods available in both approaches that are used to classify the text content.

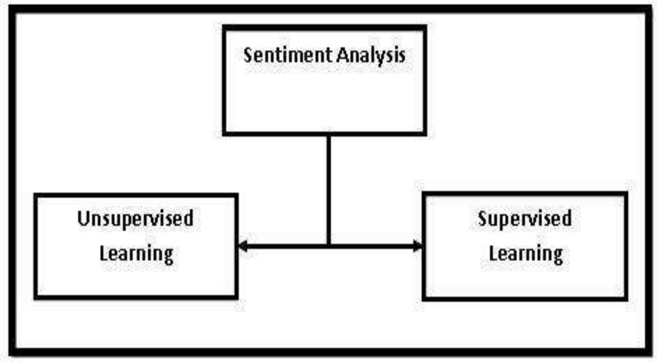

Figure 2: Basic Classifications

1.2.1 Supervised Learning

In supervised learning a labeled training data set is available which contains few training example. Each example has a pair of inputs and desired outputs. Input data has been trained and tested according to the available training examples under many different useful supervisory learning methods. It's a machine learning task of generating useful functions from training data set.

1.2.2 Unsupervised Learning

In unsupervised learning there is no labeled training data set available, there is no corpus available to train the inputs data. This is useful to discover the hidden output of the system as there are no errors or flags generate, output is not depend on training data set. In unsupervised learning outcome of inputs is unknown so it is possible to find groups of similar examples within data i.e. clustering of data, these approach most popularly used in neural and artificial neural network application.

2. Natural Language Processing

Natural language processing is an area of artificial intelligence and computational linguistics which relates computers and human in different manner. It connects computers and humans by using the natural language used by humans. It is a field of research and applications where computers learn to understand and manipulate natural languages, text contents or audio speech to perform SA. The purpose of the NLP researchers is to discover the facts that show how humans learn natural language to express the opinions so that computers can be trained to understand natural language. Use of NLP includes many disciplines like computer and information science, linguistics mathematics, electrical and electronics engineering, robotics and psychology etc. Because of its wide ability, many applications has been developed in the field of natural language text processing, lemmatization, parts-of-speech tagging, summarization, machine translation, question answering etc. Many of different tasks of NLP used in sentiment analysis methods that make SA more challenging and interesting.[4, 14]

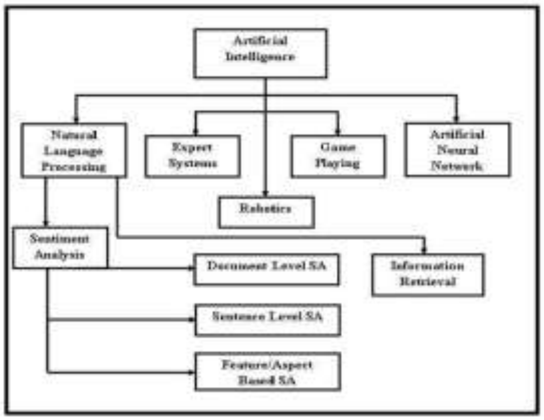

Figure 3: Overview of natural language processing and sentiment analysis

Figure 3; shows the basic classification of artificial intelligence. NLP is an important part of AI that develops learning abilities in computers similar to humans. Most recent development in NLP is statistical and corpus based methods. These are machine learning methods that include set of rules based on statistical inference. The system automatically learns these rules by analyzing real world examples.

2.1 Convolution Neural Networks (CNN)

The CNN (convolution neural network) includes pooling layers and sophistication as it gives a standard architecture to map the sentences of variable length into sentences of fixed size scattered vectors. This study has proposed a novel convolution neural network (CNN) framework for visual sentiment analysis to predict sentiments of visual content. CNN has been implemented using library Keras in Python on a Window machine. We collected data from Amazon product reviews database related to product reviews of 55740 people from various hotels. The strategy with 3 epochs has been performed for training the network. For experimental work, a dataset of hotel reviews is selected and back propagation is applied. The proposed model was evaluated on this dataset and acquired better performance than existing systems. Results shows that proposed system achieve high performance.
The main focus of this work was to initialize the weight of parameters of convolution neural network and it is critical to train the model accurately while avoiding the requirement of adding new feature. A neural language is used to initialize the word embedding and is trained by big unsupervised group of tweets. For further refining the embedding on bulky supervised corpus, a conventional neural network is used. To initialize the network, previously embedded words and parameters were used, having same architecture and training on supervised corpus. The deep learning model is applied on two tasks: message level task and phrase level task from hotel reviews to predict polarity and achieve high outcomes. By applying test-set, the proposed model lies at first rank in terms of accuracy.[16]

2.2 Recurrent Neural Network (RNN) using the Long Short Term Memory (LSTM)

As an improvement to previous machine learning methods, here I am trying to achieve better results with a Recurrent Neural Network. In a traditional recurrent neural network, during the gradient back-propagation phase, the gradients signal can end up being multiplied a large number of times (as many as the number of time steps) by the weight matrix associated with the connections between the neurons of the recurrent hidden layer. This means that, the magnitude of weights in the transition matrix can have a strong impact on the learning process. If the weights in this matrix are small (or, more formally, if the leading eigenvalue of the weight matrix is smaller than 1.0), it can lead to a situation called vanishing gradients where the gradient signal gets so small that learning either becomes very slow or stops working altogether. It can also make more difficult the task of learning long-term dependencies in the data. Conversely, if the weights in this matrix are large (or, again, more formally, if the leading eigen value of the weight matrix is larger than 1.0), it can lead to a situation where the gradient signal is so large that it can cause learning to diverge. This is often referred to as exploding gradients.
These issues are the main motivation behind the LSTM model which introduces a new structure called a memory cell. A memory cell is composed of four main elements: an input gate, a neuron with a self-recurrent connection (a connection to itself), a forget gate and an output gate. The self-recurrent connection has a weight of 1.0 and ensures that, barring any outside interference, the state of a memory cell can remain constant from one time step to another. The gates serve to modulate the interactions between the memory cell itself and its environment. The input gate can allow incoming signal to alter the state of the memory cell or block it. On the other hand, the output gate can allow the state of the memory cell to have an effect on other neurons or prevent it. Finally, the forget gate can modulate the memory cell's self-recurrent connection, allowing the cell to remember or forget its previous state, as needed.[15]

2.2.1 Word Vectors

In order to understand how deep learning can be applied, think about all the different forms of data that are used as inputs into machine learning or deep learning models. Convolution neural networks use arrays of pixel values, logistic regression uses quantifiable features, and reinforcement learning models use reward signals. Figure 4 describes the common theme is that the inputs need to be scalar values, or matrices of scalar values. When you think of NLP tasks, however, a data pipeline like this may come to mind.

Figure 4: Word vectoring using SA

This kind of pipeline is problematic. There is no way for us to do common operations like dot products or back propagation on a single string. Instead of having a string input, we will need to convert each word in the sentence to a vector. You can think of the input to the sentiment analysis module as being a 16 x D dimensional matrix. We want these vectors to be created in such a way that they somehow represent the word and its context, meaning, and semantics. For example, we'd like the vectors for the words "love" and "adore" to reside in relatively the same area in the vector space since they both have similar definitions and are both used in similar contexts. The vector representation of a word is also known as a word embedding.

2.2.2. Word to Vector

In order to create these word embeddings, we'll use a model that's commonly referred to as "Word2Vec" as shown in figure 5. Without going into too much detail, the model creates word vectors by looking at the context with which words appear in sentences. Words with similar contexts will be placed close together in the vector space. In natural language, the context of words can be very important when trying to determine their meanings. Taking our previous example of the words "adore" and "love", consider the types of sentences we'd find these words in.

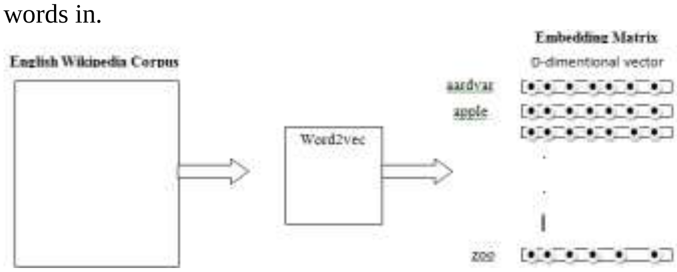

Figure 5: Word embedding using Word to Vector

This embedding matrix will contain vectors for every distinct word in the training corpus. Traditionally, embedding matrices can contain over 3 million word vectors.

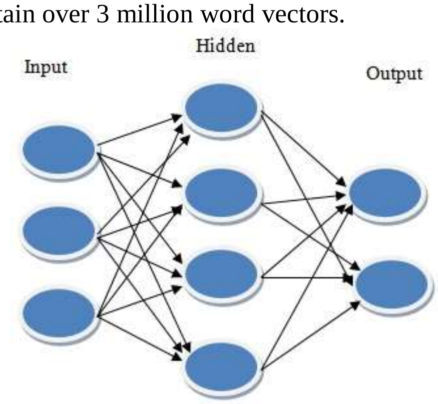

Figure 6: Feed forward network

Now that we have our word vectors as input, let's look at the actual network architecture we're going to be building. The unique aspect of NLP data is that there is a temporal aspect to it. Each word in a sentence depends greatly on what came before and comes after it. In order to account for this dependency, we use a recurrent neural network. The recurrent neural network structure is a little different from the traditional feed forward NN you may be accustomed to seeing. The feed forward network (figure 6) consists of input nodes, hidden units, and output nodes.

The main difference between feed forward neural networks and recurrent ones is the temporal aspect of the latter. In RNNs, each word in an input sequence will be associated with a specific time step. In effect, the number of time steps will be equal to the max sequence length.

3. Literature Survey

Here, we discuss various author views about that sentiment analysis. In online product reviews consumers discuss about products and its features. A product may have hundreds or thousands of reviews, consumers share their experience about products and comments about products characteristics. Feature may be part of the product or its characteristics.

[7] Peter D. Turney; in this work first time point-wise mutual information andinformation retrieval (using deep learning methods) method has been proposed which analyze the online product reviews. Calculation of semantic orientation (SO) of mutual information between opinion oriented words and reference words is depend on the statistical data and information retrieved from the web search engine. SO of mutual information of two words phrase is amount of information when we talk about positivity and negativity. Analyzing the product reviews using sentence based SA shows the polarity of number of reviews of particular products either positive or negative. Here we have the ability to further and propose a most recent algorithm modified using deep learning methods method which develops a review search engine and the web crawler for collecting the product reviews from e-commerce site.

[8] ZHANG Zi et. al.; the sentiment classification of Chinese products based on theconcepts of a crawling the snippets from web search engine. Snippets are number of hits when querying the search engine. These returned snippets are crawled from a size of ten-word window. We learn that this work is similar to original using deep learning methods method, only difference is calculation of SO is based on the returned snippets from web search engine. The distance of ten-word window is small as it may skip the many opinion oriented phrases. It simplifies the sentiment estimator by removing the ratio of number of hits of word *excellent* and *poor*.

[9] Xuinting DUAN et. al.; this work focused on extracting the sentiment from the blogcontents. Blogs writer express different emotions or feelings like "joy", "sorrow", "anger", "fear". Automatic classification of blog contents in SA is used to forsubjectivity identification. This helps search engines to report summary statistics. Using these four types of emotions ten different reference words are selected for each category.

[10] John Rothfels and Julie Tibshirani; the following thought finds unsupervisedmethods are simpler to implement as compared to supervised methods. Although supervised methods like Naïve Bayse, maximum entropy classifier and support vector machine experimented wide range of examples and acquires high accuracy on movie reviews classification but they were not able to acquire high accuracy on sentiment classification when classifying the standard topic based categorization. All supervised methods have similar trained data as per requirement while moving to another domain same classifier does not work properly and requires to annotate data to train the classifier for specific domain. Unsupervised approach is domain independent and does not require annotated trained data. These methods are easy to implement and determine the strength of opinions based on statistical ratings.

[11]. Yuanbin Wu el. at.; this work finds that in the classification of product reviewsmost of the product features and opinion oriented words are two word phrases. So introducing concept of phrase dependency parsing it is possible to extract the relation between product features and expressions of opinions. Most of the product features are nouns and opinion oriented words are combination of nouns and adjectives or nouns and verbs or adjective and adjectives but it is possible lot of noise candidates may extract which can confuse the relation extraction classifier. We point out this thing in our work and extract the combination of nouns, adjective and verbs which extract the opinion expressions from reviews.

[12]. Won Young Kin et. al.; this works uses association rules for opinion mining ofproduct reviews. SO of word is calculated from the difference between strength of its association with a set of positive words and strength of its association with a set of negative words. We find that association rule mining is based on apriori principal which is defined as if an item set is frequent then all of its subsets must also be frequent. Similarly if a feature-opinion set is frequent than all of its features must be frequent. It stores the features and opinion of products in the form of transaction T and applies the association rule mining on this transactional data and pmi method to summarize the discovered association rules.

[13]. Sanjay Kalamdhad, Shivendra Dubey et al.; The online product reviews are downloaded from shopping website using the web crawler. From these reviews, sentences are identified in which features of product are mentioned. Two-word opinion phrases are extracted from these review sentences using POS tagging which one of natural language technique by following patterns. The most important step is calculation of semantic orientation of all phrases using the reviews search engine which we developed and used in our experiment. We also extend the positive and negative reference word set to 18 words; this improves the efficiency by adding more positivity and negativity. Sentiment classification of features of products is useful for shopping websites where it is possible to give more detailed information about the product from consumer's point of view. Showing opinions about special features of products is great beneficiary to both online retailer and online buyer of products.

4. Proposed Work

The goal of the sentiment analysis issue is to analyze user's text reviews from online shopping websites. Sentiment analysis applications help buyers and manufactures to get opinions of product features for decision making. We examine the feature/aspect based SA for analyzing online reviews of products to generate opinion based summary. We present our study by implementing following tasks. Sentiment classification is not a straight forward task. Generally speaking, existing works may be camped into two major approaches, namely sentiment knowledge based approaches and machine learning based approaches. The former approaches mainly utilize the sentiment lexicon, rules and pattern matching for sentiment analysis. They build a lexicon of words with positive and negative polarities, and identify the attitude of the author by comparing words in the text with the lexicon. However, the rules and patterns of sentiment are very difficult to achieve.

The machine learning approaches employ machine learning algorithms to build classifiers from texts with manually annotated sentiment polarity. The traditional machine learning approaches are heavily based on engineered features, but it is very difficult to handcraft features to extract properties mentioned. This becomes an even harder problem especially in cases when the amount of labeled data is relatively small, e.g., thousands of examples.

For the accurate classification of sentiments, many researchers have made efforts to combine deep learning and machine learning concepts in the recent years. This section briefly describes the numerous studies, related to sentiment analysis of web contents about user's opinions, emotions, reviews toward different matters and products using deep learning techniques. Sentiment analysis tasks can be performed efficiently by implementing different models such as deep learning models, which have been extended recently. These models include CNN (Convolution Neural Networks), RNN (Recurrent Neural Networks). This section describes the efforts of different researchers toward implementing deep learning models for performing the sentiment analysis. Several researchers have used more than one model in their study, and these are mentioned under the hybrid neural network section.

4.1 Algorithm and its flow chart

Framing Sentiment Analysis as a Deep Learning Problem

As mentioned before, the task of sentiment analysis involves taking in an input sequence of words and determining whether the sentiment is positive, negative, or neutral. We can separate this specific task (and most other NLP tasks) into 5 different components. Here; figure 7 shows the flow chart for deep learning problem.

1) Training a word vector generation model (such as Word2Vec) or loading pertained word vectors
2) Creating an ID's matrix for our training set (We'll discuss this a bit later)
3) CNN and RNN (With LSTM units) graph creation
4) Training
5) Testing

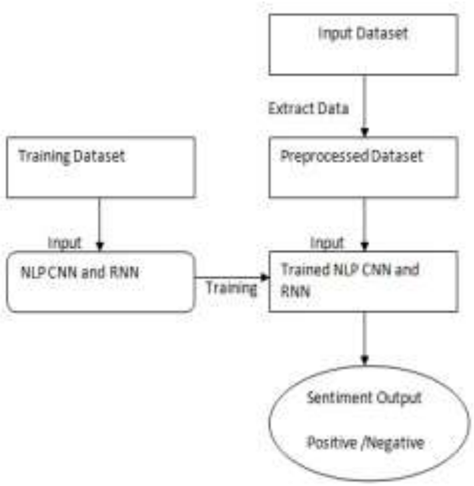

Figure 7: Flow chart for SA

Algorithm:

//Acquire annotated sentiment dataset from human-computer conversation
Inputs: human-computer conversation logs; sentiment lexicon; negation lexicon
Output: annotated sentiment dataset

1. Reserve only human's sentences and removes all questions.
2. Reserve sentences satisfied Min ≤ length ≤ Max.
3. A sentence is considered positive if it has one of the positive term, and negative if it has one of thenegative terms.
4. Remove the sentences which are both positive and negative.
5. Flip the polarity if a negation word is found immediately before a sentiment term. Remove the sentence if a negation word is found several words before a sentiment term.
6. Segment words into positive and negative.
7. Remove all one-word sentences and de-duplicate the sentences

The objective of the featured based sentiment analysis using deep learning methods is to find the sentiments about feature of products from the product reviews. The intention of the algorithm which performs sentiment analysis is provide a summary of features of products when consumers buy products from online shopping websites. This gives detailed idea about the products and its characteristics and special features. Summarization of features displays opinions of users about the products they purchase from shopping website.

4.2 Reviews Search Engine:

Our reviews search engine is different from other web search engine in perspective of database. Web search engine like Google or Yahoo has database of millions of web sites that contains all kinds of information. They have collection of all kinds of information from latest news, arts, music, religious engineering to medicines. While finding the matching documents from web search engine used in [15] [1] for phrase or combination of phrase and reference words might get all kinds web pages that are not even discussing or related with product features. So we develop a different and simple kind of search engine that has the database of product reviews only. These reviews are posted by the customers of shopping websites and discuss only about the products and its features like commenting about merits and demerits of particular feature of product.

5 Result Analysis and its parameters metrics used

Here, we use the Python version 3.6 for examination as well as its parameter which is use of this investigation. The series of steps and every one of the computations with it will be showed in this segment. In this work we use the Amazon reviews database for sentiment analysis. The applied method will be CNN and RNN. The parameter which we will use is accuracy of both the algorithms. The performance is based on their classification.
The result is based on product review taken from amazon review database. There is total no. of 55740 reviews about the mp3 player which has been sold by amazon. Then we applied deep learning methods on the dataset that is CNN and RNN. Here we compute the accuracy, precision, recall, f1 score of the methods which we discussed.

- **Accuracy;** refers toward the nearness of a calculated value to a regularor elseidentified value. For instance, if inside lab you get a weight dimension of 3.2 kg intended foraidentified substance, other than the actual or elsewell-known weight is 10 kg, then your dimension is not exact. In this case, your dimension is not close uptoward the identified value.
  Classification Accuracy is known through the relation:
  Though, there are problems through accuracy. It assumes equivalent costs for mutually type of errors. A 99% accuracy may be excellent, good, middling, poor or else awful depending leading the problem.
$$\boldsymbol{Accuracy} = \frac{\boldsymbol{TP+TN}}{\boldsymbol{TP+TN+FP+FN}}$$
- **Precision;** refers toward the nearness of two or else more dimensions to each other. Using the instance above, if you weight a certain substance five times, along withobtain 3.2 kg every time, then your dimension is extremelyaccurate. Precision is self-

governing of accuracy. For instance, but on average, your dimensionsintended for a specified substance are close toward the recognized value, but the dimensions are far from all other, then you contain accuracy with no precision.

- **Recall;**inside this perspective is also referred toward as the true positive rate or else sensitivity.
- **F1 Score;** is the weighted typical of Precision as well as Recall. Thus, this score obtaintogether false positives as well as false negatives keen on account. Naturally it is not as simpletowardunderstood as accuracy, other than F1 is generally more valuable than accuracy, particularly if you consist of an uneven class allocation. It works greatest if false positives as well as false negatives have related cost. If the costs of false positives along with false negatives areextremelyspecial, it's improved to look at both Precision along with Recall.

$$\text{F1 score} = 2 * \frac{(\text{Recall} * \text{Precision})}{(\text{Recall} + \text{Precision})}$$

- **Epoch;** An **epoch** is a calculation of the number of times every one of the training vectors are use once toward renew the weights. Intended for batch training every one of the training sample go by throughout the learning algorithm concurrently within one **epoch** earlier than weights are efficient.
- **Confusion Matrix-** A confusion matrix is an abstract of prediction outcome on a classification difficulty.

The numeral of correct along with wrong predictions are sum up with count values as well as broken down through every class. This is the key in to the confusion matrix. The confusion matrix demonstrates the method in which your classification model is confused while it makes predictions. It gives us approaching not only into the errors individual made through a classifier but further prominently the kind of errors that are being made.

	Set 1 Predicted	**Set 2 Predicted**
Set 1 Actual	TP	FN
Set 2 Actual	FP	TN

Here,

Set 1: Positive

Set 2: Negative

Explanation of the Terms:

Positive (P): Examination is positive (for case: is an apple).

Negative (N): Examination is not positive (for case: is not an apple).

True Positive (TP): Examination is positive, along with is predicted to be positive.

False Negative (FN): Examination is positive, other than is predicted negative.

True Negative (TN): Examination is negative, along with is predicted to be negative.

False Positive (FP): Examination is negative, other than is predicted positive.

Here we will show how the algorithm is performed on the dataset. First we will show the CNN based result.

- **Details of stop word in the dataset**

['a', 'about', 'above', 'after', 'again', 'against', 'all', 'am', 'an', 'and', 'any', 'are', "aren't", 'as', 'at', 'be', 'because', 'been', 'before', 'being', 'below', 'between', 'both', 'but', 'by', "can't", 'cannot', 'could', "couldn't", 'did', "didn't", 'do', 'does', "doesn't", 'doing', "don't", 'down', 'during', 'each', 'few', 'for', 'from', 'further', 'had', "hadn't", 'has', "hasn't", 'have', "haven't", 'having', 'he', "he'd", "he'll", "he's", 'her', 'here', "here's", 'hers', 'herself', 'him', 'himself', 'his', 'how', "how's", 'i', "i'd", "i'll", "i'm", "i've", 'if', 'in', 'into', 'is', "isn't", 'it', "it's", 'its', 'itself', "let's", 'me', 'more', 'most', "mustn't", 'my', 'myself', 'no', 'nor', 'not', 'of', 'off', 'on', 'once', 'only', 'or', 'other', 'ought', 'our', 'ours', 'ourselves', 'out', 'over', 'own', 'same', "shan't", 'she', "she'd", "she'll", "she's", 'should', "shouldn't", 'so', 'some', 'such', 'than', 'that', "that's", 'the', 'their', 'theirs', 'them', 'themselves', 'then', 'there', "there's", 'these', 'they', "they'd", "they'll", "they're", "they've", 'this', 'those', 'through', 'to', 'too', 'under', 'until', 'up', 'very', 'was', "wasn't", 'we', "we'd", "we'll", "we're", "we've", 'were', "weren't", 'what', "what's", 'when', "when's", 'where', "where's", 'which', 'while', 'who', "who's", 'whom', 'why', "why's", 'with', "won't", 'would', "wouldn't", 'you', "you'd", "you'll", "you're", "you've", 'your', 'yours', 'yourself', 'yourselves']

- **The words of negative reviews**

{„poor‟, "worst‟, "bad", "wrong", "defective‟, "problem‟, "terrible‟,"damage‟,"sucks‟," heavy ‟," heating ‟," ridiculous ‟," pathetic ‟,"regret‟," sad ‟," fault ‟,"annoying‟," awful ‟," useless ‟}

- **The words of positive reviews**

{"excellent‟, "fantastic‟, "good‟, "best‟, "great‟, "super‟, "amazing‟, "stunning‟, "awesome‟, "beautiful‟, "worth ‟, "nice‟, " average ‟, "brilliant‟, "decent‟, "extraordinary‟, "fine‟, "powerful‟ }

Table 1 shows that; Number of epoch compute the accuracy, Loss, time unit as well s total time for Recurrent Neural Network

Table 1: Number of epoch for RNN

No. of epoch(3)	**Total time**	**Time unit**	**Loss**	**Accuracy**
None	-	-	-	-
1 (25000/25000)	397s	16ms/step	0.4596	0.7823

2 (25000/25000)	393s	16ms/step	0.3029	0.8767
3 (25000/25000)	386s	15ms/step	0.2420	0.9056

we use 3 epoch for enhance the accuracy of the sentiments for 25000 reviews which is collected by Amazon reviews database and also describe the total times spend in each epoch and explain how much is required in each steps for iteration, Losses is also found in each epoch that is specific issue in finding accuracy.

According to Recurrent Neural Network

Accuracy: 85.45%

Confusion Matrix:

	TP	FP
TP	10866	1614
FP	2024	10476

We apply the algorithm on Recurrent Neural Network then finding out the given accuracy on various parameters as well as confusion matrix. Here; table 2 describe the result parameters using RNN.

Table 2: Accuracy outcome using RNN

Parameter	Percentage
Accuracy	85.45%
Precision Score	.8665012406947891
Recall Score	.83808
F1 Score	.85205368037869

We apply the algorithm on Convolutional Neural Network then finding out the given accuracy on various parameters as well as confusion matrix.

According to Convolution Neural Network

Confusion Matrix-

	TP	FP
TP	10588	1912
FP	1081	11419

Table 3: Number of epoch for CNN

No. of epoch	Total time	Time per step	Loss	Accuracy
None	-	-	-	-
1 (25000/25000)	198s	8ms/step	0.4538	0.7684
2 (25000/25000)	196s	8ms/step	0.2615	0.8967
3 (25000/25000)	193s	8ms/step	0.2136	0.9190

Table 3 shows that; Number of epoch compute the accuracy, Loss, time unit as well s total time for Convolution Neural Network. Here; table 4 describe the result parameters using CNN.

Table 4: Results outcome using CNN

Parameter	Percentage
Accuracy	88.03%
Precision Score	.0.8565749006076063
Recall Score	0.91352
F1 Score	0.8841314699392203

Figure 8 describes the various results parameters for CNN and RNN. In the base paper the sentiment analysis has been done through traditional machine learning classification but in this work we have shown how deep learning concepts work better than machine learning approach.

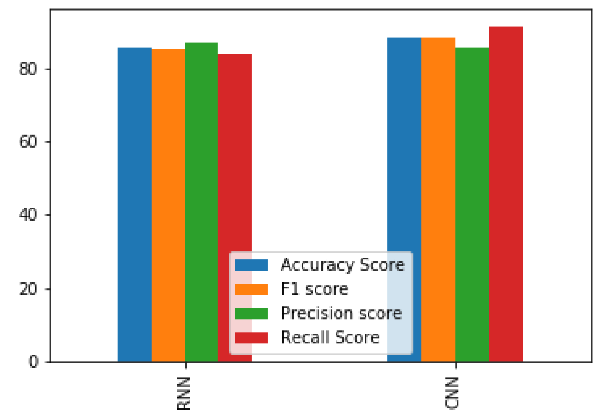

Figure 8: Representation of results parameters

In online product reviews consumers discuss about products and its features. A product may have hundreds or thousands of reviews, consumers share their experience about products and comments about products characteristics. These product reviews may have positive or negative sentiments. A positive sentiment contains good opinion about product and its features similarly a negative sentiment tells drawbacks and problems of product and its features. Feature may be part of the product or its characteristics.

6. Conclusion and Future Scope

Sentiment classification of features of products is useful for shopping websites where it is possible to give more detailed information about the product from consumer's point of view. Showing user opinions about special features of products is great beneficiary to both online retailer and buyer of products. In this paper it is observed that sentiment analysis or opinion mining plays important role while making a decision towards a particular product or a service. But it is very important to consider certain quality measures like helpfulness, usefulness and utility while analyzing each review. In the literature survey there are many sophisticated methods explained which defines the sentiment analysis with respect to different aspects.

In future, more research work is needed to improving the performance measures further. Sentiment analysis or opinion mining can be applied for any new applications which follow data mining rules. Although the techniques and algorithms used for sentiment analysis are advancing fast and giving high quality results, lot of problems in this field of study remain unresolved and also it is hard to find the fake review by reading. Sometimes fake reviews also seen as good quality review and it was modified like no one can identify their actual intension. For further improvement, we can increase the database of our reviews search engine; bigger the search database will increase the reliability of the system. Phrase extraction patterns are crucial to implement as there is possibility of useless phrases, we expect more specific opinion oriented phrases could be identified from reviews for improving performance.

Reference

[1]. Bing Liu,"Exploring User Opinions in Recommender Systems", Proceeding of the second KDD workshop on Large Scale Recommender Systems and the Netflix Prize Competition", Aug 24, 2008, Las Vegas, Nevada, USA.

[2] Dave.D, Lawrence.A, Pennock.D,"Mining the Peanut Gallery: Opinion Extraction and Semantic Classification of Product Reviews", Proceedings of International World Wide Web Conference (WWW‟03), 2003.

[3] Turney, P, "Thumbs Up or Thumbs Down? Semantic Orientation Applied to Unsupervised Classification of Reviews", ACL‟02, 2002.

[4]. G.Vinodhini and RM.Chandrasekaran, Sentiment Analysis and Opinion Mining: A Survey, Sentiment Analysis and Opinion Mining: A Survey Volume 2, Issue 6, June 2012.

[5]. V. S. Jagtap and Karishma Pawar, Analysis of different approaches to Sentence-Level Sentiment Classification, International Journal of Scientific Engineering and Technology (ISSN 2277-1581) Volume 2 Issue 3, PP : 164-170.

[6]. Gregory Grefenstette and PasiTapanainen, What is a word, what is a sentence? Problems of Tokenization, https://www.researchgate.net/publication/2360239 Rank Xerox research center April 1994.

[7]. Turney, P. 2002 Thumbs Up or Thumbs Down? Semantic Orientation Applied to Unsupervised Classification Reviews ***ACL'02***.

[8]. ZHANG Zi qiong, LI Yi-jun, YE Quang, LAW Rob Sentiment Classification for Chinese Product Reviews Using an Unsupervised Internet based Method 2008 International conference on Management Science and Engineering USA.

[9]. X. DUAN, T. HE, Le SONG, Research on Sentiment Classification of Blog Based on USING DEEP LEARNING METHODS 978-1-4244-6899-7/1 0 ©20 10 IEEE.

[10]. Gregory Grefenstette and PasiTapanainen, What is a word, what is a sentence? Problems of Tokenization, https://www.researchgate.net/publication/2360239 Rank Xerox research center April 1994.

[11].YuanbinWu, Qi Zhang, Xuanjing Huang, LideWu Phrase Dependency Parsing for Opinion Mining Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing, pages 1533–1541, Singapore, 6-7 August 2009. c 2009 ACL and AFNLP.

[12]. Won Young Kin, Joon Suk Ryu, Kyu Il Kim, Ung Mo Kim, A Method for Opinion Mining of Product Reviews using Association Rules ICIS 2009, November 24-26, 2009 Seoul, Korea Copyright © 2009 ACM 978-1-60558-710-3/09/11.

[13]. Sanjay Kalamdhad, Shivendra Dubey,"Analysis of Orientation of Product Reviews Using Sentiment Analysis" INTERNATIONAL JOURNAL OF SCIENTIFIC PROGRESS AND RESEARCH (IJSPR) ISSN: 2349-4689 Volume-21, Number - 04, 2016.

[14]. E.D. Liddy Natural Language Processing In Encyclopedia of Library and Information Science, 2nd Ed. NY. Marcel Decker, Inc.

[15]. Gers, Felix A., and E. Schmidhuber. "LSTM recurrent networks learn simple context-free and context-sensitive languages." IEEE Transactions on Neural Networks 12.6 (2001): 1333-1340.

[16]. Robinson, Piers. "The CNN effect reconsidered: Mapping a research agenda for the future." (2011): 3-11.