

CONCEPTUAL ANALYSIS OF PRODUCT EVALUATIONS USING DEEP LEARNING

R. Murugeswari* and G.Ramasakthi
Kalasalingam Academy of Research and Education, Krishnankoil.
email: murugesananth@gmail.com

Abstract— The process of classifying a small text piece into positive, negative or neutral is called as Sentiment Analysis. For example, if a message is considered, then the meaning of the sentence is identified and then analysis of the sentence has to be done. After analysis, we decide that the sentence is either positive, either negative or neutral. This is referred to as Sentiment Analysis. Nowadays, most of the people prefer the online products compared to the direct buyers. So, the online marketing is growing very high. In fact, the online marketing is fully satisfying the customer needs so that it occupies a great place in people's heart. They are also satisfying the demands of customers with high quality and low price. These are the primary reasons for the online marketing to be in a successful position. If an e-commerce company has an ability to gather information about the customer thoughts and behavior, then the online marketing will be more effective than now. In this paper, we have taken a particular e-commerce dataset called Flipkart and classify the polarity of the comments by using some of the classifiers namely Support Vector Machine Classifier, Guassian Naïve Bayes Classifier and Random Forest Classifier and Multilayer Perceptron Classifier. The existing approach is that the comments were classified based on the attitude of the customer. But, now the proposed approach has been implemented with the help of Multilayer Perceptron (MLP) Neural Network Classifier which is simulated by the SPYDER tool. The accuracy for all the four algorithms namely, Support Vector Machine, Random Forest, Naïve Bayes and Multilayer Perceptron (MLP) Neural Network algorithm has been computed and the best accuracy has been predicted by comparing them. Here, the Multilayer Perceptron (MLP) shows the best accuracy of all others and the accuracy is 99.94%. Moreover, some performance metrics has been determined here. The performance metrics namely Precision, Recall and F-measure has been evaluated for each algorithm individually and then the comparison is made among them. Then, the ROC curve is measured for the classes designated by us for revealing the recognition of parameters between two diagnostic groups. It is predicted individually for each class. At last, the Confusion Matrix is going to enumerate for each algorithm distinctly. It presents the actual and predicted values in a tabular format thereby measuring the performance of a classifier.

Keywords— Multilayer Perceptron Classifier; Random Forest Classifier; SVM Classifier; Guassian NB Classifier; Sentiment Analysis; Product Reviews

I. INTRODUCTION

The process of classifying a small text piece into positive, negative or neutral is called as Sentiment Analysis. For example, if a message is considered, then the meaning of the sentence is identified and then analysis of the sentence has to be done. After analysis, we know that the sentence is positive, negative or neutral. This is known as Sentiment Analysis.

Sentiment Analysis helps the business persons to improve their product quality and efficiency. They can identify the customer needs easily and produces the product accordingly. It helps the data analysts to analyze the data very clearly in an efficient way so that they can used this data and perform some scientific calculations.

Sentiment Analysis helps the businesses to identify the customer opinions about a product that they buy. Sentiment includes both the positive and negative sentiments. Some of the positive sentiments include like, love, share, comments etc., some of the negative comments include dislike, hatred, block, report etc., and Sentiment Analysis is classified into two types. They are i) Feature based (or) aspect-based ii) Subjectivity (or) objectivity based analysis.

The process of Sentiment Analysis is carried out by performing the step by step process. First, the dataset is collected from a particular website. Then, the dataset is loaded and then the preprocessing is done. After that, the data splitting exists. Then, the dataset is trained to the model. Finally, it designates the opinion comments as positive opinions or negative opinions or neutral comments.

Sentiment Analysis consists of three different types of classification. They are i) Polarity classification ii) Subjectivity classification iii) Opinion holder recognitions. Polarity classification means categorize opinions into Positive, Negative or Normative by identifying the sentiment orientation. Subjectivity Classification classifies an opinionated document by expressing the opinion as positive or negative opinion. Opinion Holder Recognitions recognizes the persons having same attitude and whether there are different opinions from specific persons. Sentiment Analysis consists of three different approaches. They are i) Machine Learning approach, ii) Lexicon Based approach and iii) Hybrid approach. Machine learning approach is done by using the machine learning methods and algorithms in which they automatically improve through experience only. In Lexicon Based approach, a piece of text message is named as a bag of words. If we follow this representation, then the all opinion values are assigned to all other positive, negative and neutral words that are existing within the message. Hybrid approach is the combination of the elements and the principles defined in now and the olden days. In this method,

the paper is splitted into various components such as hardware, software or mechanical components.

The residue of the paper is standardized as follows. Section 2 demonstrates the related work about the Sentiment Analysis Classification using various methods and techniques. Section 3 suspects the Multilayer Perceptron Neural Network (MLP) algorithm which is used to perform the Sentiment Classification. Section 4 manifests the proposed system architecture. The simulation result to validate our algorithm is yielded in Section 5 presents. Finally, in Section 6, we allege the conclusion of this work and explore the future possible works.

II. RELATED WORKS

The principal objective of this paper is to examine the product reviews and to carry out the Sentiment Classification process using Deep Learning techniques. Here, there are some related papers which deals with the Sentiment Analysis using various methods and techniques.

Wei Zhao et al. [1] proposed a special thing called framework here which is named as novel deep learning framework and it is used for the purpose of reviewing the products and classify them on the basis of Sentiment Classification. It hires the already presented common ratings and named them as weak supervision signals. The framework is composed of two steps: First step is to learn a high level representation which is responsible for capturing the common sentiment dissemination of sentences which were predicted through attributes named as rating information; and second step is to add a classification layer which is laid on the top of the embedding layer and it makes use of the categorized sentences for supervised fine-tuning. The reviews which are taken for this process are collected from the website called Flipkart.com. Weakly Supervised Deep Embedding trains the Deep Neural Networks, then make use of the rating information of reviews which was already presented commonly on many merchant/review websites. The training is made of two approaches: To learn the process and embedding space thereby attempts to apprehend the sentiment diffusion of sentences. This is achieved with the help of languishing the relative distances among the different kinds of sentences as per the weak labels which have been concluded from ratings; To add a softmax classifier on top of the embedding layer thereby tuning the network with the help of labeled data.

Mubashir Ali et al. [2] focused on the implementation of the latest technologies namely Deep Neural Networks, Convolution Neural Networks thereby solving numerous Sentiment Analysis problems. The challenge is that the Sentiment Analysis contains less number of adequate labeled data which is contemplated in the field of Natural Language Processing (NLP). To, solve this problem, new techniques such as the Sentiment Analysis and Deep Learning techniques have been consolidated here. In this paper, the author highlighted latest studies related to the execution of various deep learning models such as Deep Neural Networks, Convolution Neural Networks and many more models for solving the discrete problems of Sentiment Analysis. Some of the problems are Sentiment Classification, cross lingual problems, textual and visual analysis problems. Several studies have been considered in

this kind of studies which helps to amend the profound knowledge thereby leading the fortunate growth of Deep Learning applications in the province of Sentiment Analysis. Numerous problems have been decided by the two high provinces named as Sentiment Analysis and deep learning and thus achieve high accuracy also for both fields.

Akshi Kumar et al. [3] performed the Sentiment Analysis process on the tweets by considering the combination of probabilistic neural network (PNN) and a self-adaptive approach for the implementation. The proposed approach consists of two types of Probabilistic Neural Network models. First model which belongs to PNN, also mentioned as PNNS and the special feature is that it has a smoothing parameter which has been considered all over the whole network and it is a single value only. Second model, also referred as PNNC which is totally opposite to PNNS. Here, they use discrete values for each class. The train and test set was collected from the website called twitter with the help of interface named as Twitter API. Two types of Probabilistic Neural Network models are executed finally. The accuracy has been calculated for both the models, namely, PNNS and PNNC after implementation. The result shows that the PNNC has higher performance when compared to PNNS.

Sujata Rani et al. [4] described the Sentiment Analysis for Hindi movie reviews and it is mentioned here. The Hindi movie reviews are extracted from the online news papers and Web sites. The dataset consists of three native speakers of Hindi. The CNN models have 50% data on the training side and the remaining 50% percent lies in the testing side. After the accuracy of model has been predicted, the comparison is made. It has been clearly remarked that CNN model has the capability to attain more accuracy than Machine Learning algorithms and thus achieved better performance. The CNN model got the accuracy of 95%. All the experiments have been discharged using discrete values for parameters. It has been declared that CNN model having two convolution layers with filter sizes 3 and 4 represents the best among the pervasive models and has the capability to reach the maximum accuracy value of 95%.

Kazuhei Katoh et al. [5] presented the performance evaluations of for large-scale Sentiment Analysis by employing the methods of deep learning classifiers. This can be implemented by using the Rakuten Data. Most of the NLP theories and applications represents a word with the help of 1-of- K representations. Two experiments were conducted to reduce the dimensions of 1-of-K representations (1) the IMDB dataset, a small dataset (2) the Rakuten Dataset, a huge one They also tested with more hidden layers (0-6), and enacts the Neural Networks with four hidden layers and thus reached the best result.

Severyn et al. [6] proposed the Deep Learning techniques for the Sentiment Analysis of tweets presented in the twitter data. The new thing in this paper is that, he organizes a new model for the purpose of performing the initialization process of parameter weights that is done in the Convolution Neural Network. He also makes it to train the model accurately. He uses unsupervised learning method to train this model. In final stage, the previously trained parameters are used for the initialization purpose of the model. Then, he trains the latter data on the supervised learning method also. The comparison is made between the two results of the proposed approach and the system

participation in the final test set, his model ranked in the first position in both the phrase-level task. This considers being an important evidence for the value of author's solution.

Pranali Ramteke et al. [7] explained some of the data mining processes and data mining techniques. He employs these techniques to remove and encapsulate the data. The result implies that the reviews are whether positive or negative depending on the reviews given by the customer. In this paper, the required dataset is taken from the huge website called amazon.com, rediff.com and Flipkart.com. The reviews are classified into positive, negative and neutral categories which indicate the customer's opinion about a particular product. The test can also be done to separate and sieve the false reviews of the products. Most of the Opinion Mining applications are based on the process of bag-of-words. It does not seize the conditions which is prerequisite for achieving the Sentiment Analysis. The current evolution growth in Sentiment Analysis and its connected highly discharged operations are also dispensed here.

Khale Ahmed et al. [8] presented a overview about Sentiment Analysis. In this, they are labeling the variety of concepts in this area, problems and its solutions, existing APIs, descriptions of the tools used and also handover a list of open challenges related to this area. Sentiment Analysis can be performed in various firms such as social, media and industrial firms. This survey has presented many details regarding the levels of Sentiment Analysis, various techniques of opinion mining, augmentation methods for opinion mining, applications of Sentiment Analysis, list of APIs used, lexicons used, description of used tools and the existing research gaps. They have compared the state of the art techniques with all the different techniques. Moreover, they have conferred using the same data set in this paper. It is then evaluated to find the test dataset result.

Vinay James et al. [9] considers the customer feedbacks and are referred as mile stones and responsible for occupying the victorious position in many companies. The proposed work has given the accuracy as 80% in the opinion mining. The data which has been taken from the twitter page is collected with the help of the crawler Twitter4J API. Then, the attained dataset is stored to a standard database. Then they preprocessed the dataset and get rid of the stop words. Then the data presented in the dataset has been classified using POS tagging approach. By this, the tag set has been designed. The data which gets isolated from the tag set are then fed to the SVM algorithm. The frequencies of the words are calculated by using the unigram approach and then the all round product rating was being calculated.

Jhanvi et al. [10] analyzes the basic characteristics of opinion mining. He extracts the product review dataset from an particular organization called Flipkart through the Flipkart product API. By using this API, he fetches the product details such as product name, product ID, brand, product rating and also the other related information about the product. The extracted features from the dataset are subjected to clustering process. The clustering has been performed by the two most important algorithms namely, ROCK and CART algorithm. These algorithms help in classifying the reviews as positive, negative or neutral based on the comments given by the user. At last, the

author predicts the product which is having more percentage of positive reviews. First, the inputs are segregated into positive and negative and then only percentage calculation is made for both the positive and negative words here. Therefore the result survey is analyzed by the user and helps him to conclude the result based on the percentage of positive reviews of a particular product. It comprise of various methodologies such as Classification, Extraction and Clustering. By utilizing item API, there were bringing the name of the brand, audits, product rating and other essential things for item, grouping utilizing ROCK and utilizing CART calculation to order surveys as positive and negative words from the previous remarks. At last, they identify which item is having more level of positive responses. The information was classified as positive and negative words from the responses and it computes the level of positive and negative words. In this way the outcome investigation of response rate is performed and thus it encourages the client to finish up dependent on the positive survey level of the item.

III. MULTILAYER PERCEPTRON NEURAL NETWORK

A. The MLP Neural Network is considered to be one of the models of Artificial Neural Network. It belongs to class of feed forward network. It is composed of multiple layers of Perceptron. An important feature is that it utilizes the back propagation for training. Back propagation is one of the supervised learning techniques. MLP has multiple layers. It has non-linear activation function. It has a special feature that it distinguishes data in which the data is not linearly separable. The Perceptron is an algorithm in which a single straight line classifies the input data into two categories. The MLP Neural Network consists of three layers. It contains one input layer, one or more hidden layers and one output layer. These are very important for the MLP Neural Network.

IV. ARCHITECTURE DESCRIPTION

The online shoppers are increasing day-by-day. About 60 percent of the people prefer the online shopping. People in urban areas are mostly addicted to the online shopping. They get the things and other necessities in a single phone call by ordering them in online. First, we need to know about the product and its uses. Then, we need to know about the manufacturing company so that we can identify the best company. Then, only we can get the best product with high quality at low cost. So, we need a helper for this. So, we need some customer reviews. By analyzing the reviews, we can identify the best product. For this review analysis, we must need a term called Sentiment Analysis. It classifies the reviews into positive, negative or neutral, thereby we can identify the quality of a product.

In this paper, the Sentiment Analysis of product reviews of the Flipkart e-commerce website has been scrutinized here. It helps the online shoppers to choose the best product. The dataset has been collected, preprocessed, splitted into train data and test data, trained in to the proposed model and then finally accuracy has been measured.

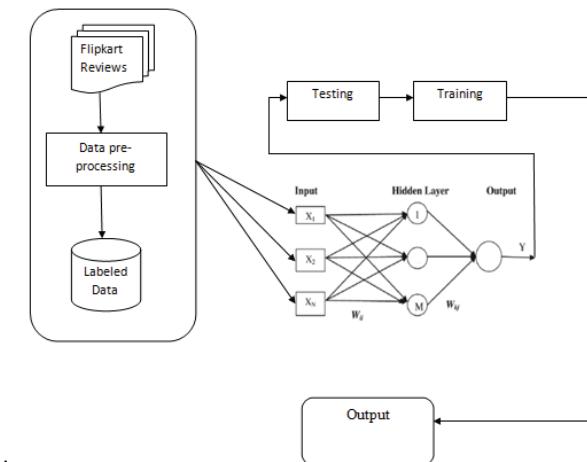


Fig. 1 System Architecture of the proposed system

Data Gathering:

The dataset required for this paper has been taken from the website named kaggle.com. In the dataset, 20,000 data were presented there. It has different types of attributes namely Product name, Unique ID for the product, brand, Retail price, Overall rating, Product rating, etc.,,

Data Preprocessing:

The dataset has the combination of numerical values and alphabetic values. We need to find the accuracy for the values. So, it is necessary that all the columns contain numeric values. So, we are going convert all the categorical alphabetic values in to the numeric labels by using the Label Encoder. It encodes the alphabetic labels into numeric labels.

Data Splitting:

After preprocessing, data splitting process is going to start. We have to split the dataset into train dataset and test dataset for implementing the dataset using one of the proposed models. Half of the data is splitted for train set and other remaining is for test set. Then, we have to train the data into the model for the purpose of finding the accuracy.

Feature Extraction:

Then, the necessary features are identified and taken into account for training the model. The identified features and their definitions are clearly explained in the table given below

TABLE 1: Extracted Features

Extracted Features	Description of the Extracted Features
Product name	It shows the name of a particular product.
Product ID	It shows the unique ID for an individual product.
Product Rating	It shows the rating which represents in the form of stars ranging from 1 to 5 given by the customers.
Brand name	It shows the brand name of a particular product.

Classification process:

Once the feature extraction process is over, then we have to classify the reviews on the basis of four attributes mentioned below.

- The four attributes are described below.
 - True Positives - The model which exactly finds the positive values only.
 - False Positives - The model which predicts the positive values wrongly as negative values.
 - True Negatives - The model which exactly finds the negative values only.
 - False Negatives - The model which predicts the negative values wrongly as positive values.

This classification predicts all the possible outcomes and it is responsible for finding the Confusion Matrix. The confusion matrix has the predicted values and true values. By having these values, we can perform some performance metrics calculation.

Sentiment Analysis:

The final process which can perform by using these values is known as Sentiment Classification. It can be executed with the help of Multilayer Perceptron Neural Network here. Then, the performance metrics, Confusion Matrix and ROC curve are estimated.

V. EXPERIMENTAL RESULTS

At first, the dataset for implementing the paper has to be collected from a website called kaggle.com. The dataset extracts from the Flipkart e-commerce sample. The implementation has been done by using the SPYDER tool. Then load the dataset. After the dataset is being loaded, then the preprocessing is to be done.

After this, we have to split the train data and test. We can split the data by giving the test size. Then, the dataset has to be trained into the various models for finding the accuracy.

The proposed model called Multilayer Perceptron Neural Network discovers the accuracy for the enrolled dataset by using the MLP classifier. Then, it is also evaluated by the other three algorithms. The accuracy of SVM algorithm is calculated using the Linear Support Vector Classifier. Then, the accuracy of Naive Bayes has been predicted using the Naïve Bayes classifier. Similarly, the accuracy of Random forest is estimated using the RFC classifier. Then, the ROC curve has been predicted individually for each class. Then, the confusion matrix for all the algorithms has been estimated. Then, the performance metrics for each algorithm is calculated. The performance metrics namely Precision, Recall and F1-Measure is calculated for each algorithm. At last, the accuracy comparison is made among all the four algorithms and the best accuracy has been shown.

Performance Metrics:

The Performance metrics are classified into four important parameters. They are

- Accuracy

- Precision
- Recall
- F-measure

The description and formula of the three important parameters are shown below.

Accuracy:

Accuracy is obtained by dividing the correctly evaluated values by the totally predicted values. In general, the accuracy denotes the overall performance.

The accuracy comparison of Multilayer Perceptron (MLP) Neural Network Classifier, Random Forest Classifier and Support Vector Machine (SVM) Classifier has been displayed in Fig 3. The names of the predicted algorithms are

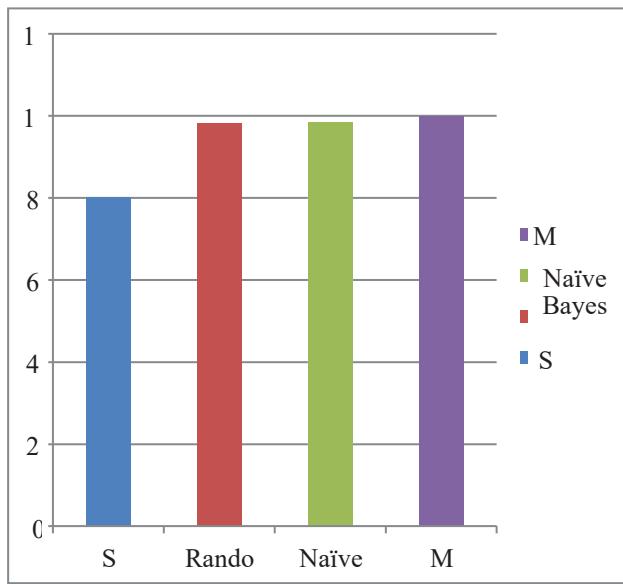


Fig., 2 Accuracy Comparison

mentioned in the X-axis where the corresponding accuracy percentage is mentioned in the Y-axis.

The graph clearly shows that the SVM contains accuracy of 80%, Random Forest has accuracy of 98.19%, Naïve Bayes has accuracy of 98.3% and the Multilayer Perceptron (MLP) consists of the best accuracy level of 99.94%.

Precision:

Precision is obtained by dividing the correctly identified positive values by the predicted positive values only. It is a measure of quality.

$$\text{Precision} = \text{TP} / (\text{TP} + \text{FP})$$

Recall:

Recall is obtained by dividing the correctly identified positive values by the actually defined positive values in the dataset. It is a measure of quantity.

$$\text{Recall} = \text{TP} / (\text{TP} + \text{FN})$$

F-measure:

F-Measure is obtained by finding the average of Recall and Precision and it also considers false positives and false negatives.

$$\text{F-measure} = 2 * (\text{Precision} * \text{Recall}) / (\text{Precision} + \text{Recall})$$

where TP = True Positive TN = True Negative

FN = False Negative FP = False Positive

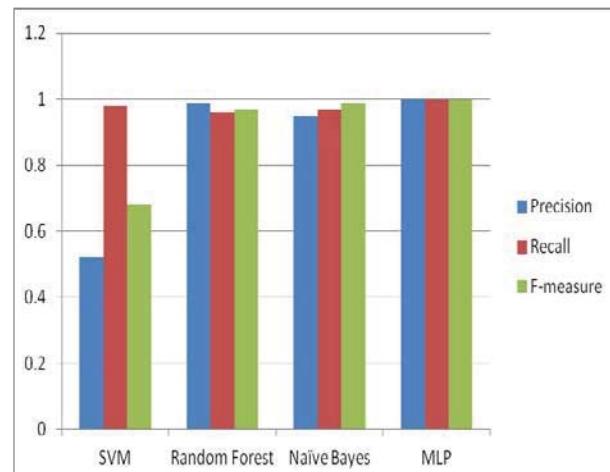


Fig. 3 Precision, Recall and F-measure comparison

ROC curve:

ROC curve defines in a graphical format in which the participants are True Positive rate and False Positive rate. It is obvious that the False Positive rate lies on the X-axis and the True Positive rate lies on the Y-axis. It is used for finding the diagnostic capability when the threshold values are varied.

The Receiver Operating Characteristic curve for the Flipkart dataset is shown in Fig 4.

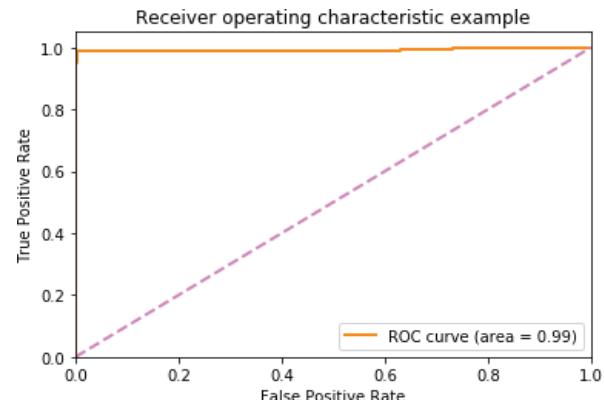


Fig 4 Receiver Operating Characteristic curve

In this graph, the area under ROC curve is 0.99 and it shows the best area. The top-left corner occupies the position of “ideal” point as it has the value one. It clearly observed that the False Positive rate remains 0 and True Positive rate occupies 1 to obtain this area.

The Receiver Operating Characteristic curve for the mentioned classes is shown in Fig 5.

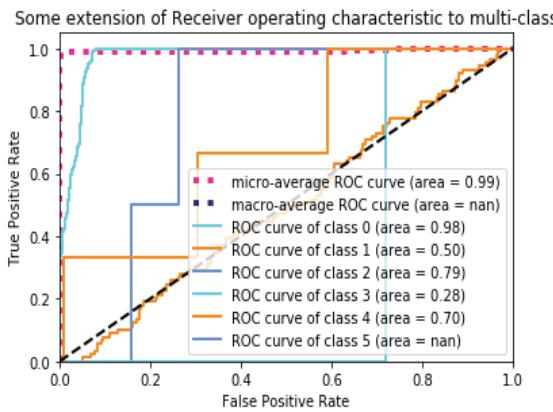


Fig. 5 ROC curve for all the mentioned classes

Confusion Matrix:

The confusion matrix contains rows and columns in which they represent the predicted and true values. It is used for performance recognition. Generally, the finding values are plotted in X axis and the actual values are plotted in Y axis. The confusion matrix of MLP Neural Network is shown in Fig. 6.

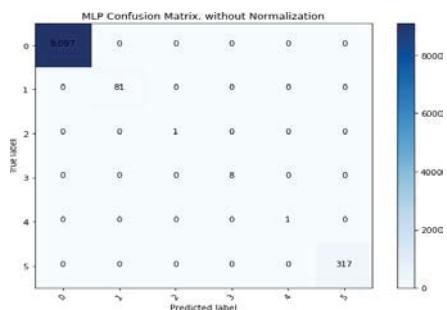


Fig. 6 MLP Neural Network Confusion Matrix

Eventually, the accuracy of Multilayer Perceptron (MLP) Neural Network Classifier is being compared with Random Forest (RF) Classifier, Support Vector Machine (SVM) Classifier and Naïve Bayes' (NB) Classifier. It is distinctly declared that the Multilayer Perceptron (MLP) Neural Network attains the best result when compared with other algorithms.

VI. CONCLUSION

Nowadays, most of the people prefer the online products compared to the direct buyers. So, the online marketing is growing very high. They are satisfying the demands of customers with high quality and low price. These are the primary reasons for the online marketing to be in a successful position. If an e-commerce company has an ability to gather information about the customer thoughts and behavior, then the online marketing will be more effective than now. In this paper, we have taken a particular e-commerce dataset called Flipkart and classify the polarity of the comments by using some of the classifiers namely Support Vector Machine Classifier,

Gaussian Naïve Bayes Classifier and Random Forest Classifier and Multilayer Perceptron Classifier. The existing approach is that the comments were classified based on the attitude of the customer. But, now the proposed approach has been implemented with the help of Multilayer Perceptron (MLP) Neural Network Classifier which is simulated by the SPYDER tool. The accuracy for all the four algorithms namely, Support Vector Machine, Random Forest, Naïve Bayes and Multilayer Perceptron (MLP) Neural Network algorithm has been computed and the best accuracy has been predicted by comparing them. Here, the Multilayer Perceptron (MLP) attains the highest accuracy when compared to all others and the accuracy is 99.94%. Moreover, some performance metrics has been determined here. The performance metrics namely Precision, Recall and F-measure has been evaluated for each algorithm individually and then the comparison is made among them. Then, the ROC curve is measured for the classes designated by us for revealing the recognition of parameters between two diagnostic groups. It is predicted individually for each class. At last, the Confusion Matrix is going to enumerate for each algorithm distinctly. It presents the actual and predicted values in a tabular format thereby measuring the performance of a classifier.

REFERENCES

- [1] Wei Zhao and Ziyu Guan, "Weakly-Supervised Deep Embedding for Product Review Sentiment Analysis", IEEE Transactions on Knowledge and Data Engineering, vol. 30, pp.185-197,2018.
- [2] Qurat Tul Ain and Mubashir Ali, "Sentiment Analysis Using Deep Learning Techniques: A Review", (IJACSA) International Journal of Advanced Computer Science and Applications, vol.8, no.6, pp. 424-434, 2017
- [3] Akshi Kumar and Ritu Rani, "Sentiment Analysis Using Neural Network", 2nd International Journal on Next Generation Computing Technologies (NGCT-2016) Dehradun, pp.262-267, 2016.
- [4] Sujata Rani and Parteek Kumar, "Deep Learning Based Sentiment Analysis Using Convolution Neural Network", Arabian Journal for Science and Engineering, pp. 1-4.,2018
- [5] Kazuhei Katoh and Takashi Ninomiya , "Deep Learning for Large-Scale Sentiment Analysis Using Distributed Representations", , The International Journal on Advances in Semantic Processing, pp. 92- 96, 2015.
- [6] Severyn and A. Moschitti," Twitter Sentiment Analysis with Deep Convolutional Neural Networks", Proc. 38th Int. ACM SIGIR Conf. Res. Dev. Inf. Retr. - SIGIR 15, pp. 959962, 2015.
- [7] Krutika Wase and Pranali Ramteke, " Sentiment Analysis of product review", International Journal of Innovations in Engineering and Science, vol. 3,no.5, pp. 8-13,2018.
- [8] Khaled Ahmed and Neamat El Tazi, " Sentiment Analysis over Social Networks: An Overview", IEEE International Journal on Systems, Man and Cybernetics, pp.: 2174-2179,2015
- [9] Minara P Anto and Nivya Johny, "Product rating using Sentiment Analysis", International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT) ,pp.:2458-2462, 2016
- [10] Janhavi N L, Santhosh Kumar K L and Jharna Majumdar, Sentiment Analysis of Customer Reviews on Laptop Products for Flipkart", International Research Journal of Engineering and Technology vol.5 Issue: 03,pp. 629-634, 2018.