Sentiment Analysis using Artificial Neural Network

Niharika, Sona Malhotra

Abstract*: In the modern era, there are massive amount of web resources present such as blogs, review sites and discussion forums. These resources form the platform where users can share their opinions or reviews` about anything whether it is a product, movie or a restaurant. Analysis of public sentiments deals with the determination of the polarity of different public opinions or reviews into either the category of positive, negative or neutral. Thus, there comes the need of sentiment analysis which not only helps other individual to make a decision regarding buying a product, visiting a restaurant or watching a movie but also helps the producers of various products and owners of different restaurants to gain the knowledge of preferences of customers, so that it could be possible to increase the profit and economic value. The paper presents a survey with main focus on performance of different artificial neural networks used for opinion mining or sentiment analysis while it also includes various machine learning approaches such as Naïve Bayes, Support Vector Machine, lexicon-based approach and Maximum Entropy.

***Keywords* : Sentiment analysis, supervised learning, machine learning, opinions, artificial neural network, linear classifier, Naïve Bayes, support vector machine**

I. INTRODUCTION

Sentiment analysis is a kind of Natural Language Processing which includes the examining and analyzing of opinions regarding anything whether it is a product or a movie [1].

Thus, it is also referred as opinion mining. The opinions and reviews are posted online by the customers on different web resources that are available such as blogs or social media platforms [2]. From such reviews, a person can easily determine whether the overall review and opinion of people concerning a particular thing is positive, neutral or negative. But it would be a strenuous as well as a time consuming task for an individual to go through all the reviews regarding a particular restaurant or a movie prior to settling on a choice and making a decision. So, there comes the need and importance of classification of sentiments which could be performed at three different levels namely, document level sentiment analysis, sentence level sentiment analysis and sub-sentence level sentiment analysis which refers to examining the sentiments of a complete document, analyzing the reviews of a single sentence and classifying the opinion of the sub-expressions present in the sentence respectively [3].

There are two main approaches for analyzing the sentiments namely, Lexicon-based approach and the approach of machine learning. The techniques that come under the latter approach are Support Vector Machine (SVM), Naïve Bayes, Maximum Entropy and Artificial Neural Networks (ANN).

These techniques can be divided into two broad categories where Naïve Bayes and Maximum Entropy techniques come under probabilistic classifiers while Support Vector Machine and Artificial Neural Network form a part of linear classifiers. This paper presents a survey on these different techniques with the main focus on the utilization of Artificial Neural Networks for the task of sentiment classification.

The remainder of this survey paper is organized as follows: Section II provides a literature review on sentiment analysis and Section III elucidates common methodology adopted in research papers. Further, Section IV presents the result of analysis of several research papers surveyed in both the sub-sections. Finally, Section V concludes the study and discusses the future scope for further research in the topic of sentiment analysis.

II. RELATED WORK

The first part of this survey paper, which is sub-section 1, comprises the papers that include the comparison of the techniques that come under supervised learning such as Support Vector Machine, Naïve Bayes and Maximum entropy.Furthermore, the second sub-section includes the survey of the research papers that incorporate the artificial neural network models for opinion mining and sentiment classification and the comparative analysis of artificial neural network with other techniques.

A) Sub-Section 1 :

In the year 2006, Hang *et al.* [4] presented an idea of performing the process of sentiment analysis through the experimentation with different sets of algorithms that come under the category of machine learning, where feature selection was also conducted in order to achieve dimensionality reduction. Three algorithms were employed on online product reviews namely, discriminative classifier (Passive Aggressive (PA) algorithm), Winnow and generative model. For discriminative classifier, PA algorithm was chosen because it followed an online learning pattern and it had a predictable performance. In a like manner, Winnow also adopted an online learning algorithm which predicted the polarity of the reviews. While the generative model was based on language model which was used to state the probability with which a given sequence of words can be generated and it used the ratio between positive and negative probabilities to give the polarity. It was depicted that the performance of the classifier

Revised Manuscript Received on January 15, 2020.

***** Correspondence Author

Niharika*, Computer Science and Engineering, University Institute of Engineering and Technology, Kurukshetra University (Kurukshetra), Kurukshetra, India. Email: niharikadhawan15@gmail.com

Sona Malhotra, Computer Science and Engineering, University Institute of Engineering and Technology, Kurukshetra University University (Kurukshetra), Kurukshetra, India. Email: sonamalhotrakuk@gmail.com

was improved by high order n-grams. Moreover, for the mixed reviews, discriminative model was proved to be efficient as it gave better result than other two models. The only drawback was that classification was only limited to just positive and negative scales and did not include the classification of reviews at different scales. Later, Qiang *et al.* [5] proffered an experiment to classify the reviews that were associated with the blogs related to travelling for the seven destinations in 2009. The sentiments were classified either into positive reviews or negative reviews, so that, it could be helpful to the both travelers, and the managers before taking any decision related to choosing the destination. Three approaches of machine learning were compared namely, Support Vector Machine (SVM), Naïve Bayes and a character based N-gram model which was a dynamic language model classifier. For the analysis by Naïve Bayes, a tool was developed by VC 6.0 and for the purpose of feature selection then the method of Information Gain was applied. The SVM approach used the concept of word frequency whereas the N-gram based character language model preferred characters to be its basic unit. The dataset for the classification was the retrieved corpus from Yahoo.com and a Lingpipe DynamicLMClassifier was used for the experiment to be performed by the three algorithms. The k-fold cross validation was performed where three was chosen as the value of k. For the experiments, all the characters were converted into lowercase while the punctuation in itself was considered as a lexical item. From the experiment, it was concluded that the performance of Support Vector Machine and N-gram was better in comparison to the performance of Naïve Bayes. Also, the performances of all the three approaches were significantly equal with large training dataset and had reached the 80% level of accuracy. The limitation of this approach was that only reviews regarding western countries were taken into consideration. Further, in 2011, Shi *et al.* [6] came out with a technique of classifying the sentiments of hotel reviews on the basis of supervised learning. Two types of features were considered for the approach of supervised learning using unigram feature which were, frequency and TF-IDF which means Term Frequency- Inverse Document Frequency. The feature described latter was a weight which basically measured the significance of a word in the whole document. The dataset, which was included, contained 4000 reviews which were divided into four folds of equal size that maintained a balanced distribution of the class. The document was segmented using ICTCLAS and precision, recall and F-score were the three parameters for analyzing the performance. The experiment resulted in the realization that TF-IFD outperformed frequency in sentiment classification. Subsequently in 2011, the two approaches for sentiment classification namely, machine learning approach and the approach of lexicon-based look-up table were incorporated into a combined approach by Fang *et al.* [7] The dataset of camera reviews was chosen for sentiment classification and the classification was done on domain specific reviews and different expressions where sentiments were related to the specific domain. In order to implement the proposed system, three steps were followed where the first step was corpus filtering in which the phrases were extracted and an initial lexicon list was prepared. Secondly, the step of web search with linguistic pattern was performed with two patterns (camera aspect and seed words) which was followed by the expansion of dictionary with synonyms and antonyms. Thus, two classifiers were combined which were, Aspect classifier and Polarity classifier. Four experiments were conducted out of which the experiment performed with incorporation of domain specific lexicon and knowledge encoded in MPQA in SVM approach gave the best accuracy of 47.4%. Likewise, Dhande and Patnaik [8] in 2014 also introduced a classifier which incorporated a Naïve Bayes classifier so that it could analyze the reviews of several movies and classify them in accordance to their polarity as positive or negative. The inputs that were selected were the movie review dataset and a keyword dictionary called the WordStat dictionary. Then, the reprocessing of the data was done using the model called Bag of Words (BOW) which used occurrence of each word and unigram feature in order to conduct the process of classification. There were two inputs of sentiment analysis namely the testing file and trained Naïve Bayes classifier following which the composition matrix was evaluated that contained the data about predicted and actual classification. The experiment included simulation environment (MATLAB R2012a environment) and performance metrics (accuracy parameter). From the outcomes, it was inferred that the focused model came out to be the best with accuracy of 80.65%. After that, in 2015, Moh *et al.* [9] proposed a multi-tier architecture for the purpose of classification where the sentiment analysis was performed on a movie review dataset which was extracted from rotten tomatoes database. The introduced architecture focused on four different classifiers of supervised machine learning namely, SVM, Naïve Bayes, Random Forest and SGD (Stochastic Gradient Descent). The classification was done so as to categorize the sentiments into five different classes from very negative to very positive. The data was collected and cleaned for the purpose of pre-processing where the splitting of the train set was done according to 80-20 rule which was followed by feature selection. Then, prediction was done using three models. Model-1 used the whole review dataset with three labels while Model-2 and Model-3 considered negative-very negative and positive-very positive labels respectively. The experiments were performed which resulted that multi-tier model outperformed single-tier model. Moreover, the SGD classifiers with Scikit learn performed significantly well using the dictionaries with an accuracy of 87.23%.

B) Sub-Section 2:

The approach of neural network was implemented to classify the reviews or sentiments of the people from blogs: Live Journal and Review Centre along with movie dataset by Chen *et al.* [10] in 2011 where the concept of semantic orientation was applied and incorporated with machine learning for effective sentiment classification. The back propagation neural network (BPN) was used along with four different indexes for sentiment orientation. The first index was SO-A index that infer the semantic orientation from the association where a paradigm was prepared and strength was calculated. Second and third indexes were SO-PMI (AND) and SO-PMI (NEAR) which used Point wise Mutual Information and

Information Retrieval (PMI-IR) while the last index was SO-LSA. The proposed methodology was to prepare a dataset to form Term Document Matrix (TDM) and calculate its' SO-indexes. Then the network was trained and the performance was evaluated which concluded that the method performed better than several traditional techniques and was also succeeded in reducing the time for the classification significantly. The potential limitation of the proposed method was that it incorporated only blogs while twitter, face-book could also have been used. Later, the analysis of the sentiment for stock prediction was done using the data from Twitter by Mittal and Goel in 2012 [11] for correlating the market and public sentiments. The data contained values of DJIA (Dow Jones Industrial Average) which was fed into the preprocessor. Simultaneously, the data from twitter was also accumulated and fed into sentiment analysis algorithm which was used to classify the values of modes into four classes such as calm, happy, alert and kind. Then, both the data were fed into proposed model of Self Organizing Fuzzy Neural Network (SOFNN). The sentiment analysis algorithm comprised of steps such as generation of word list, filtering of tweets, computing and mapping of daily score. Four different learning algorithms were compared: Linear regression, logistic regression, SVN using LIBSVM library and SOFNN of five layers with Ellipsoidal Basis Function (EBF). The observations from the experiment were made that only calm and happy moods were Grauger causative of DJIA. Also, SOFNN yielded significant performance in prediction as compared to others with performance of 75.56%. Later, a portfolio management strategy was implemented to make effective buy or sell decision after pre-computation which was proved to be profitable for a period of 40 days. The potential limitation lied in the dataset as it did not map the real public sentiments. After that, Ghiassi *et al.* [12], in 2013, came out with a hybrid model which was comprised of combination of Dynamic Artificial Neural Network (DAN2) and n-gram analysis where the extraction of data was done from Twitter API v1.0 . the subject that was selected for the analysis was Justin Bieber as it was the largest twitter account of that time. The tools for gathering the information related to subject mention and references were developed. Another tool called scoring tool was developed so that, two large datasets could be made where, the first set is training dataset and the second set is called testing dataset. The task of dimensionality reduction and feature selection was performed using a twitter specific approach using linguistic analysis was used to determine n-gram which were of interest. Then, it was validated using the frequency metrics following which ten groups which were named as candidate groups were selected and then evaluated. It was depicted through the outcomes that around 10% of the tweets were neutral. Furthermore, the reduced set of lexicon had a great coverage as well as it was less complex. Later, two approaches were selected for the task of sentiment classification which were DAN2 and SVM. For such approaches, two classes were defined where one class included the emoticons in the sentiments while the other class did not and it was concluded that the presence of emoticons in the sentiments really helped in the task of sentiment analysis. Finally, the comparison made on the two supervised approach demonstrated that DAN2 performed better than SVM where the basis of evaluation of the performance were the metrics called recall and accuracy. Then, the targeted approach of twitter sentiment analysis was further enhanced by Ghiassi *et al.* [13] in 2016 itself that used supervised feature engineering and artificial neural network. The enhanced model analyzed the twitter sentiments of two brands namely Starbucks and Governor Christie. The proposed model for targeted sentiment analysis included the collection of data or tweets, cleaning of collected tweets which were followed by feature engineering and feature vectorization. Then, Dynamic Artificial Neural Network was applied for the classification of the sentiment of the tweets where classification was again done in the two sets comprising three class and five class classification. Moreover, the two system models were again considered for comparison namely Sentiment140 and Repustate that were proved to be the worst in performance. The performance was assessed by using the metrics recall, F1-measure and precision. The outcomes of the experiments concluded that the enhanced model gained F1-measure of 88% and outperformed all other systems. A Semantic Treebank was used by Socher *et al.* [14] in 2013 which was used along with the proposed model of Recursive Neural Tensor Network (RNTN) for understanding the compositionality in the sentiments as for all the nodes, it used tensor based computation. The proposed network could take the input phrases of nay length and the data that was selected for compositionality was of movie reviews taken from rottentomatoes.com where labels were extracted from the phrases that were longer in length and a standard parser was used for the purpose of parsing. Then, labeling was performed using Amazon Mechanical Turk and then labels were classified into categories of five classes from negative to somewhat positive. After that, the evaluation was performed in a bottom-up manner. The results of many approaches were compared namely, SVM, Naïve Bayes RNN, MV-RNN, RNTN and biNB (Naïve Bayes along with the bag containing bigram features). The comparisons concluded that RNTN outperformed all other approaches in predicting fine grained labels of sentiment with

performance accuracy of 80.7% which was followed by MV-RNN and RNN while for the classification of binary labels as positive or negative, again RNTN was proved to be better than other techniques with an accuracy of 85.4%. After one year, Santos and Gatti [15] used the approach of Convolutional Neural Network (CNN) for the classification of sentiments in 2014. The proffered network was named as Character Sentence Convolutional Neural Network (CharSCNN) as it was used for the purpose of extraction of character to sentence information. The datasets that were selected for the analysis of sentiments were of twitter messages and movie reviews where the approaches that were used for the two datasets were Stanford Twitter Sentiment (STS) and Stanford Sentiment Treebank (SSTb) respectively. The first step that was included in the proposed

method was to compute the score of the labels and the proposed neural network was comprised of two layers and used back propagation algorithm. Then, the word level embedding and character level

embedding of fixed size was done

where local features were produced and then combined with max operation. The next step in the model consisted of sentence level representation. The experiment for SSTb was conducted for two approaches that were with phrases and complete sentence. The approached model of CharSCNN achieved the accuracy of 85.7% for classification of movie reviews into binary classes of positive or negative while accuracy of 48.3% for fine grained. And, the accuracy of STS was 86.4%. The classification process was limited as it had a dataset of limited size. Furthermore, Liu *et al.* [16] focused on Recurrent Neural Network (RNN) for fined grained opinion mining along with word embeddings in 2015. Two datasets related to reviews were chosen namely Laptop dataset and Restaurant dataset which were provided by SemEval. Three pre-trained word embeddings which represented words in a distributed manner were used namely SENNA embedding (each word was represented as feature with an associated embedding vector in a look-up table), Google Embeddings (trained by skip-gram model) and Amazon Embeddings (containing corpus on Amazon products). The embeddings were used in three different architectures including Jordan-type RNN, Elman-type Recurrent Neural Network and Long Short- Term Memory Recurrent Neural Network. The datasets were preprocessed where the data was converted into lowercase and then vocabulary was built which was followed by addition on padding in case of boundary words. The model was assessed using and F1-score, recall and precision while it was compared with CRF baseline using paired t-test on F1-score. The results demonstrated that word embeddings improved the performance of RNN and CRF while its fine tuning significantly improved the performance of RNN. Moreover, integration of heuristic feature in RNN also improved the performance. Simultaneously, Ouyang *et al.* [17] used the Convolutional Neural Network to perform an experiment of sentiment analysis in 2015 where a framework was proposed called word2vec + Convolutional Neural Network (CNN). The dataset that was an extended movie reviews from the rottrntomatoes.com where the data was defined in five different labels from positive to negative. Word2vec was used for the transformation of the words into vectors to build sentences' vectors. Before training the network, the Google News dataset was considered for training the pre-trained vectors. Later on, the sentences' vector was supplied as input to the CNN which was built using Caffe which was an open source framework. The CNN had three Convolutional layers along with three pooling layers which extracted the features and sampled the Convolutional matrix respectively. Moreover, dropping layer was also present which was there to remove the problem of over-fitting. Finally, the performance was evaluated on the basis of fine grain which concluded that the introduced model worked better than other models like RNN, MV-RNN. Later, in 2016, an approach was proffered for sentiment analysis by Borele and Borikar [18] by using an artificial neural network (ANN) along with the comparative analysis of various methods. The dataset that was chosen for the experiment contained the reviews of several movies along with the labels of binary polarity. Several approaches were compared such as machine learning approaches that included Naïve Bayes (a simple approach for text classification but less accurate), SVM which was based on discriminative classifier and multiclass SVM variant was chosen for sentiment analysis, k-Nearest Neighbor (used similarity score as the basis for classification), Maximum Entropy and Artificial Neural Network. Other approaches that were included in the comparison were lexicon based approach where the semantic orientation of words was used to evaluate polarity, rule based approach (used semantic dictionary) and statistical model. The outcome of the comparison concluded that SVM outperformed other approaches for sentiment classification. After that, an approach was put forward that included pre-processing followed by feature extraction. Then, training of the dataset was done where weights were assigned and fuzzy logic was applied. The proposed neural network used back propagation and then testing phase was conducted. Concomitantly, Wang *et al.* [19] focused on dimensional sentiment analysis for the recognition of continuous numerical values using a regional model of CNN-LSTM in 2016 which comprised of two parts namely, LSTM and regional CNN where ratings of Valence Arousal (VA) were predicted. The prediction was done on two datasets which were Chinese Valence Arousal Text (CVAT) and Stanford Sentiment Treebank (SST). In the model, the word vectors were trained using a toolkit called word2vec. In the regional CNN model, the sentence was considered as a region and the text was divided into regions for the extraction of features. The features were then integrated in a sequential manner for the generation of text vector for VA prediction using LSTM model. The model comprised of Convolutional layer for feature extraction, max pooling layer, sequential layer and linear decoder. Two methods of lexicon based approach were chosen for the purpose of assessment which were weighted Geometric Mean (wGM) and weighted Arithmetic mean (wAM) while the regression methods that were considered were Maximum Values Regression (MVR) and Average Values Regression (AVR). The evaluation of the outcomes was done by Mean Absolute Error (MAE), Root Mean Square Error (RMSE) and Pearson correlation coefficient. The model that was put forward outperformed other conventional NN methods such as LSTM, CNN and RNN., Chen *et al.* [20] came up with a hierarchical neural network to perform the task of sentiment classification in 2016. The researchers focused on classification at document level that incorporated global user and product information. The analysis was performed on three datasets which were IMDb, Yelp2013 and Yelp2014. The datasets were splitted into three parts which included the sets called, train, development and test in the ratio of 8:1:1 followed by tokenization and sentence splitting with the use of Stanford CoreNLP. The introduction of hierarchical LSTM model was done for the purpose of generation of a joint sentence level representation and document level representation. While the traditional models only took local text information, the focused model used global user preference and product characteristics as the motive for user product attention. The model for Neural Sentiment Classification (NSC) was contrasted with other methods for document level sentiment classification such as Trigram, AvgWordvec, SSWE, Majority, RNTN+RNN, JMARS, UPNN and Paragraph Vector. The evaluation was done using two metrics that were Root Mean Square Error (RMSE) and accuracy which concluded that the introduced model performed significantly well when compared with other baseline techniques. After that, a model for performing language-agnostic analysis of the sentiments was proposed by Wehrmann *et al.* [21] with the use of a character based Convolutional Neural Network CNN) in 2017 along with character level embeddings which was capable to directly learn from the document, polarity prediction of source language as well as documents containing words from multiple languages. The sentiment analysis was done with the use of twitter corpora that was originally contained data in thirteen languages of Europe and was labeled into three classes from which only four languages were considered for sentiment analysis: English, Spanish, Portuguese and German while the classification was reduced to binary classification. The model that was put forward was called Conv-Char-S that incorporated the benefits of Conv-Emo and Conv-Char. The model processed the character level input for understanding the sentences and its architecture comprised of a Convolutional layer, one max-pooling-over-time layer and fully connected layer. The experiment was conducted and the outcomes were compared with the baselines such as SVM based approaches using performance metrics, namely accuracy and F-measure which concluded that the proposed architecture performed significantly well and was proved to be far more better than the second best baseline. Later, in 2018, Mohammad *et al.* [22] came up with a contemporary model for sentiment analysis at the aspect level was performed on hotels' reviews written in Arabic language where two models of LSTM were implemented. The first model was a Bi-LSTM-CRF which incorporated a character level bidirectional Long short Term Memory (LSTM) along with a classifier known as conditional random field for the task of extraction of opinion target expressions while the second model was an aspect based LSTM which was used for the task of sentiment polarity classification at aspect level which identified the sentiment polarity. An enhanced model LSTM was trained and the results were evaluated for the experimentation of both the tasks. For the first task, a single layer model was developed followed by its training along with word embeddings and the features of character embedding while for the second task, back propagation algorithm was used to train AB-LSTM-PC. The outcomes concluded that the introduced approaches were evaluated and demonstrated that the approaches outperformed other baselines where the performance of task of aspect OTE was enhanced with 39% while the second task had shown the enhancement of 6% in the performance. Then, in 2019, stack of deep learning algorithms was proffered by Feizollah *et al.* [23] for analyzing the sentiments. The author performed analysis on the tweets on Halal products which were, Halal cosmetics and Halal tourism. The focused system comprised of data collection, data pretty-processing which was followed by data processing. The accumulation of data was done through twitter API while the task of data preprocessing was done via removal of retweets from the dataset by recognizing the RT at the starting of tweets and the keywords were searched in two different languages which were Malay and English. After this, the dataset was fed into stacked deep learning algorithms where the result of one algorithm was propagated to the next algorithm. The experiment included the use of python library called keras for the implementation of neural network architectures. The sentiment analysis algorithms first broke the sentences into words and pretty-set sentiment value was associated with it before calculating the final sentiment score. The results concluded that an accuracy of 93.78% was achieved when the method of word2vec feature extraction was incorporated with a stack of LSTM and CNN and the polarity of the reviews were found to be positive for both the datasets.

III. METHODOLOGY

The general methodology adopted by several authors that used the concept of artificial neural network as well as other supervised learning approaches is the process that included extraction of data-set, pre-processing of data followed by splitting the data into two sets for training and testing. The partitioned datasets are then passed into the classification model and later on the result is evaluated on the basis of the prediction made as elucidated in Figure 1 [24].

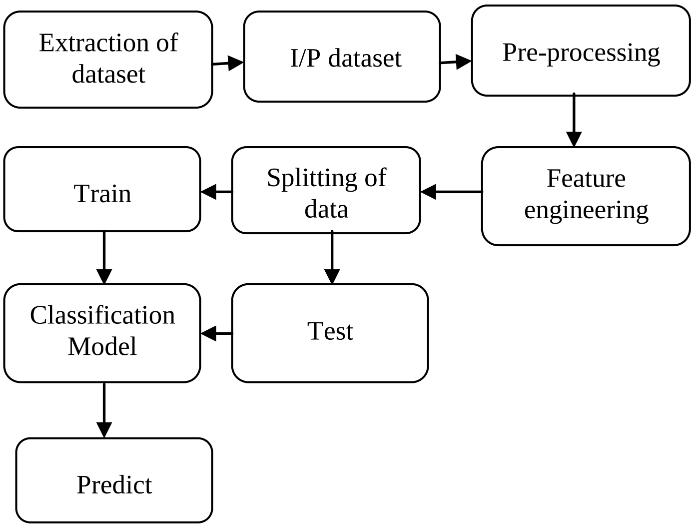

Figure 1: General methodology for sentiment analysis

Following are the steps that are included in the general methodology for sentiment analysis:

1. The first step includes the extraction of the dataset for input.
2. After that, the dataset is pre-processed so that it could be transformed into a more suitable dataset.
3. Then, the process of feature engineering is done to define features using domain knowledge.
4. Later on, the dataset is partitioned into two sets of training and testing.
5. The next step is to validate the model by the use of partitioned datasets.
6. Finally, prediction of the sentiments is done.

IV. RESULT

Following two tables give the result of the analysis of the literature survey where, the first table demonstrates the analysis of various methods used for performing the task of sentiment analysis in the research papers that has been surveyed in sub-section 1 while the second table comprises the analysis of several techniques used in the papers that have been reviewed in sub-section 2. The analysis is depicted in form of comparison which is represented in tabular form.

Table- I: Analysis of the research papers surveyed in sub-section 1.

Sr. No.	Methods used	Author	Year	Dataset	Result
1.	1. Discriminative classifier 2. Winnow 3. Generative model	Hang et al. [4]	2006	Online product review	Discriminative model outperformed other two models
2	1. Support Vector Machine (SVM) 2. Character based n-gram model 3. Naïve Bayes	Ye et al. [5]	2008	Yahoo.com corpus	SVM performed better with accuracy of 80%
3	1. TF-IFD 2. Frequency	Shi et al. [6]	2011	Hotel review dataset	TF-IFD performed better than frequency
4	Support Vector Machine	Fang et al. [7]	2011	Camera review	Accuracy of 47.4% was achieved
5	Naïve Bayes Classifier	Dhande and Patnaik [8]	2014	Movie review	Accuracy of 80.65% was achieved
6	Lexicon based approach	Moh et al. [9]	2015	Movie review	The approach achieved an accuracy of 87.23%

Table- II: Analysis of the research papers surveyed in sub-section 2.

Sr. No.	Methods used	Author	Year	Dataset	Result
1.	1. Back Propagation Neural Network 2. Navie Bayes 3. Support Vector Machine	Chen et al. [10]	2011	Movie Review	Stanford Sentiment Treebank – 85.7% Stanford Twitter Sentiment – 86.4%
2	1. Self Organizing Fuzzy Neural Network (SOFNN) 2. Linear Regression 3. Support Vector Machine (SVM) 4. Logistic regression	Mittal and Goel [11]	2012	Dow Jones Industrial Average	Accuracy of 75.56% was achieved by SOFNN.
3	1. Support Vector Machine 2. Hybrid Neural Network	Ghiassi et al. [12]	2013	Twitter subject	Hybrid neural network performed better than SVM. Strongly positive - 69.7% Mildly positive- 66.7% Mildly negative – 89.9% Strongly negative – 95.1%
4	1. Recursive Neural Tensor Network (RNTN) 2. SVM 3. Recursive Neural Network (RNN) 4. Navie Bayes	Socher et al. [14]	2013	Movie reviews	RNTN outperformed all other methods with an accuracy of 85.4%.

V. CONCLUSION

The survey concluded that artificial neural network technique performed better when compared to different traditional approaches for sentiment classification. Moreover, when used with indexes for sentiment orientation, Back Propagation Neural Network achieved better performance than several traditional techniques. Another inference was made that tuning with hyper-parameters could be significantly important for performance in case of Convolutional Neural Network. Also, hybrid model of Convolutional Neural Network (CNN) and Long-Short Term Memory (LSTM), CNN-LSTM model, performed better than conventional neural network RNTN and other techniques such as Support Vector Machine (SVM) and Naïve Bayes. Further, the limitations of several baseline approaches of sentiment analysis could be overcome with the use of different Artificial Neural Networks (ANN). The future scope includes the improvement in the accuracy which is considered as an important parameter of performance of artificial neural networks along with the removal of noise from the proposed methodology to analyze the sentiments by using an appropriate architecture of artificial neural network.

REFERENCES

1. Prediction., S. (2020). Sentiment Analysis or Opinion Mining for Stock Market Prediction.. [online] Profitaddaweb.com. Available at: https://www.profitaddaweb.com/2017/04/sentiment-analysis-or-opinion-mining.html [Accessed 12 Jan. 2020].
2. KDnuggets. (2020). Sentiment Analysis 101 – Kdnuggets . Available at: https://www.kdnuggets.com/2015/12/sentiment-analysis-101.html [Accessed 12 Jan. 2020].
3. MonkeyLearn. (2020). Sentiment Analysis. [online] Available at: https://monkeylearn.com/sentiment-analysis/ [Accessed 12 Jan. 2020].
4. C. Hang, V. Mittal and M. Datar, "Comparative Experiments on Sentiment Classification for Online Product Reviews," in AAAI'06 proceedings of the 21st national conference on Artificial intelligence, Mountain View, 2006, pp. 1265-1270.
5. Q. Ye, Z. Zhang, and R. Law, "Sentiment classification of online reviews to travel destinations by supervised," Expert Systems with Applications, vol. 36, pp. 6527-6535, 2009.
6. H. X. Shi and X. J. Li, "A Sentiment Analysis Model For Hotel Reviews Based On Supervised Learning," in Proceedings of the 2011 International Conference on Machine Learning and Cybernetics, Guilin, 10-13 July, 2011, Guilin, 2011, pp. 950-954.
7. J. Fang and B. Chen, "Incorporating Lexicon Knowledge into SVM Learning to Improve Sentiment Classification," in Proceedings of the Workshop on Sentiment Analysis where AI meets Psychology (SAAIP), IJCNLP 2011, Thailand, 2011, pp. 94-100.
8. L. L. Dhande , D. K. Patnaik, "Analyzing Sentiment of Movie Review Data using Naive Bayes Neural Classifer," International Journal of Emerging Trends and Technology in Computer Science (IJETICS), vol. III, no. 4, pp. 313-320, August 2014.
9. M. Moh , A. Gajjala , S. Gangireddy , T. Moh, "On Multi-Tier Sentiment Analysis using Supervised Machine Learning," in 2015 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology, USA, 2015, pp. 341-344.
10. C. Liu , H. Chiu L. Chen, "A neural network based approach for sentiment classification in the blogosphere," Journal of Infometrics, vol. V, no. 2, pp. 313-322, April 2011.
11. A. Mittal , A. Goel, "Stock Prediction Using Twitter Sentiment Analysis," Stanford University, CS229, vol. 15, pp. 1-5, 2012.
12. J. Skinner , D. Zimbra M. Ghiassi, "Twitter brand sentiment analysis: A hybrid system using n-gram analysis and dynamic artificial neural network," Expert Systems with Applications, vol. 40, no. 15, pp. 6266-6282, November 2013.
13. M. Ghiassi , D. Zimbra , S. Lee, "Targeted Twitter Sentiment Analysis for Brands Using Supervised Feature Engineering and the Dynamic Architecture for Artificial Neural Networks," Journal of Management Information Systems, vol. 33, no. 4, pp. 1034-1058, October 2016.
14. R. Socher , A. Perelygin , J. Y. Wu , J. Chuang, C. D. Manning , A. Y. Ng , C. potts, "Recursive Deep Models for Semantic Compositionality Over a Sentiment Treebank," in Empirical Methods in Natural Language Processings, Washington, USA, 2013, pp. 1631-1642.
15. C. N. Santos , M. Gatti, "Deep Convolutional Neural Networks for Sentiment Analysis of Short Texts," in The 25th International Conference on Computational Liguistics: Technical Papers, Dublin, Ireland, 2014, pp. 69-77.
16. P. Liu , S. Joty , H. Meng, "Fine-grained Opinion Mining with Recurrent Neural Networks and Word Embeddings," in 2015 Conference on Empirical Methods in Natural Language Processing, Lisbon, Portugal, 2015, pp. 1433-1443.
17. X. Ouyang , P. Zhou , Cheng , H. Li , L. Liu, "Sentiment Analysis Using Convolutional Neural Network," in 2015 IEEE International Conference on Computer and Information Technology; Ubiquitous Computing and Communications; Dependable, Automated and Secure Computing; Pervasive Intelligence and Computing, Liverpool, UK, 2015, pp. 2359-2364.
18. P. Borele , Dilipkumar , A. Borikar, "An Approach to Sentiment Analysis using Artificial Neural Network with Comparative Analysis of Different Techniques," IOSR Journal of Computer Engineering (IOSR-JCE), vol. 18, no. 2, pp. 64-69, March-April 2016.
19. X. Wang , W. Jiang , Z. Luo, "Combination of Convolutional and Recurent Neural network for Sentiment Analysis of Short Texts," in 26TH International Conference on Computational Linguistics: Technical Papers, Coling, 2016, pp. 2428-2437.
20. J. Wang , L.C. Yu , K.R. Lai , X. Zhang, "Dimensional Sentiment Analysis Using a Regional CNN-LSTM Model," in 54th Annual Meeting of the Association for Computational Linguistics, Berlin, Germany, 2016, pp. 225-230.
21. L. Arras , G. Montavon , K.R. Muller, W. Samek, "Explaining Recurrent Neural Network Predictions in Sentiment Analysis," in 8th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis, Copenhagen , Denmark, 2017, pp. 159-168.
22. Mohammad , B. Talafha1 , M. Ayyoub1 , Y. Jararweh, "Using long short-term memory deep neural networks for aspect-based Sentiment Analyis of Arabic Reviews," International Journal of Machine Learning and Cybernetics, pp. 1-13, March 2018.
23. A. Feizollah, S. Ainin, N. B. Anuar , N. A. B. Abdullah , M. Hazim, "Halal Products on Twitter: Data Extraction and Sentiment Analysis Using Stack of Deep Learning Algorithms," IEEE Access, vol. 7, pp. 83354-83362, June 2019.
24. Zarmeen Nasim, Quratulain Rajput, Sajdar Haider, "Sentiment analysis of student feedback using machine learning and lexicon based approaches," in 2017 5th International Conference on Research and Innovation in Information Systems (ICRIIS), 2017.

AUTHORS PROFILE

Niharika is a second year post graduation M.Tech. student at University Institute of Engineering and Technology, Kurukshetra University, Kurukshetra (India) and holds Bachelor of Technology degree in Computer Science and Engineering from Kurukshetra University, Kurukshetra (India) . Her research interests include machine learning methods, classification methods and supervised learning.

Dr. Sona Malhotra is faculty in Computer Sc. and Engineering branch at University Institute of Engineering and Technology, Kurukshetra University, Kurukshetra (India).