Patient Text Feedback Based Optimized Deep Learned Model to Identify the Impact of Therapy

Jagriti, Vikas Khullar, Harjit Pal Singh, Manju Bala

Abstract: ***Text classification or Text mining is a very demanding field because the content created by the user in natural language is not easily understandable. It becomes very important to systematically identify and extract subjective information from user content so that it can be easily understandable. The whole process is done by assigning a particular class to text. In the field of opinion mining, most of the work has been done in common areas like restaurants, electronic goods, movie feedback, etc. and a lot of work needs to be done in the area of healthcare and medical. So, the proposed work has been carried over healthcare. The aim of this study is to classify the text feedback of patient using optimized deep learning model to identify the impact of therapy. In proposed method comparison of CNN with machine learning algorithms has been done, in which, CNN gave better results in terms of accuracy (99.98%), precision (0.981), recall (0.981), mean squared error (0.282). Further, we have implemented the CNN with N-gram technique and found that this method improved the results of CNN based on precision (0.999), recall (0.999), mean squared error (0.001), area under curve (0.998) but accuracy remained the same.***

Keywords**: *Sentiment, Opinion, Health, Therapy, Text Classification, Deep learning.***

I. INTRODUCTION

Natural Languages Processing (NLP) is the subfield of computer science and artificial intelligence that plays an intermediate role between computer and human because it provides the way to computer to analysis, understand and make meaning of human language in intelligent and easy way that can be easily understood by computers [1][2][3][4]. It is not so easy task to understand human language because computers can only understand the structured text written in spreadsheets or tables and human language has in the form of unstructured data like texts and voices, so it can't be easily understood by computer. So, to solve this problem there is a need of NLP [3][5][6].There are many application areas of NLP i.e. text classification/text mining, sentiment analysis, stemming, word processing, personal assistants like Siri, Cortana, and Alexa. We have worked in the field of text classification/ text mining [1][2][3].

Text classification/ categorization is the major task of Natural Language Processing (NLP), It plays a crucial role in most of the NLP applications since it has been gaining more interest day by day, in the research area [2][6][8]. In other words we can say that text classification is a treatment of text subjectivity with large applications i.e. opinion mining, spam detection social media monitoring and fake news detection [1][3][6].

On social media, the amount of data has been increasing day by day and the amount of public feedback has also been growing so fast, that the task of analyzing opinions has become more difficult that can't be done manually. For this problem, text classification is needed approach that subjectively retrieves text [3][6][8]. This is the method that converts unstructured data into structured manner. Text classification has been used in many business areas like goods, materials, travel, housing, etc. and there has been most of the work done in these common fields but there has not so much work done in the field of healthcare so we have chosen the field of healthcare [2][3][4].In healthcare, once any therapy introduced, trials are done on a small number of test subjects. From where general effects of therapy, are detected. Patients can also participate in these trials. Trials have been done on different patients of different age groups. It is also very important to conclude how the people use particular therapy, understand its safety, side-effects and effectiveness [5][6][8]. After taking Therapy some people share their experiences on the different social media, websites, blogs in the form of text and text classification worked on that data to retrieve feedback of patient towards particular therapy too helps others to choose the best one. There has been large number of machine learning methods used for this technique such as support vector machine, radial bases function, decision tree, naïve byes and k-nearest but now a day deep learning models Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), Long Short Term Memory (LSTM) has been gaining more attention in the field of NLP. The main aim of our work is to make use of deep learning algorithm for text classification in the healthcare area.

Most of the previous work on text classification has used many technologies. Machine learning techniques combined with support vector machine (SVM) are the most usually used technology. SVM has been widely used algorithm for text classification on online data [6][8][9][10][11][12]. Certain work has been done for text classification in the medical field using machine learning algorithms. Long-Sheng Chen et. al. [10] had used neural network-based methods to identify the emotions of the customers to allow companies to respond appropriately to what customers have said. Dataset had been taken from Reviewcentre.com. Felix Greaves et. al. [12] had analyzed the sentiments of patients about different aspects like the hospital's cleanliness, dignity, and overall recommendation.

Dataset had been taken from NHK UK's site. Wang et. al.

Revised Manuscript Received on October 05, 2019.

Jagriti, CSE, CT Institute of Engineering, Management and Technology, Jalandhar, India. Email: jagritiit186@gmail.com

Vikas Khullar, CSE, CT Institute of Engineering, Management and Technology, Jalandhar, India. Email: vikas.khullar@gmail.com

Dr. Harjit pal singh, ECE., CT Institute of Engineering, Management and Technology, Jalandhar, India. Email: drhpsingh@gmail.com

Dr. Manju Bala, CSE, Khalsa College of Engineering and Technology, Jalandhar. India, Email: drmanju571@gmail.com

[13] had been introduced hybridization of machine learning and rule-based classifiers for classification the sentiments in suicide notes. Victoria Bobicev et. al. [14] had used machine learning algorithms for analysis sentiment of users and also introduced the technique for personal health information on twitter. Tanveer Ali et. al. [16] used machine learning methods like SVM, Naïve Bayes, and Logistics regression to classify sentiments of users towards the hearing loss. Wenbo A.Akay et. al. [17] had classified the opinions of users towards the specific drug that uses for lung cancer. Data had taken from cancer-forums.net. Duyu Tang et. al. [18] had developed a deep learned system that classifies message level twitter sentiments. Anthony Rios and Rmakanth Kavuluru [19] used deep learned model for biomedical text classification with a particular use-case. Peng Wang et. al. [20] proposed a unite structure to elaborate short text. Word embedding technique used with deep learning algorithm to classification of text. Zhao Jianqiang and Gui Xiaolin [21] used deep learning method for tweets sentiment classification. Results compared with traditional machine learning methods. Arturo Lopez et. al. [22] implemented a deep learning algorithm that is Long Short Term Memory with Recurrent Neural Network (LSTM-RNN) on human and some veterinary records. Work compared with decision tree and random forest algorithms. Fusheng Wei et. al. [23] had done an empirical study on 25000 legal documents and also compared the deep learning algorithm with the machine learning algorithm. Abdalraouf Hassan and Ausif Mahmood [2] implemented deep learning algorithms to classify the review of users. Different types of dataset have been taken for this work. Seyed Mahdi Rezaeinia et. al. [24] had proposed a novel architecture for text classification using deep learning algorithm on different datasets. It also proposes an improved method to increase the accuracy.

The literature is employed that the most of the work had proved that the deep learning algorithm like Convolutional Neural Network (CNN) gives better results as compared to machine learning algorithm [26] and also CNN with N-gram language model gives much better result than CNN [27]. By this inspiration we test CNN with N-gram in binary text classification. The result acquired has been compared with some baseline algorithm. We have executed text classification to the medical domain, especially gave attention to public reviews on anxiety therapy.

II. METHODS

A. Classification Techniques

There is the number of classification methods but we use SVM, RBFN, NN as baseline methods. These methods applied as the function approximation techniques. Firstly, we have compared our proposed model with baseline methods then compared with simple CNN. The description of these methods is as follows.

1. **SVM (Support Vector Machine)**

Support vector machine (SVM) critical classifier officially described by a separating hyperplane. In other words, the SVM algorithm aims to find out hyperplane in an N-dimension (N- no. of features) classifies the data points. SVM is the combination of associated supervised learning techniques used for classification and regression [20]. To differentiate the two data points, there are no. of hyperplane can be chosen. The main objective of the SVM is to find out the maximal distance between data points that belong to two different classes. Maximal margin gives some kind of reinforcement for future data points that can be classified with more sureness [15]. After designing hyperplane, the boundaries between the input classes and input elements are constructing by SVM. From a given training set of positive and negative labeled sample, maximum margin hyperplane splits the data. The result maximize the distance between the margin and the hyperplane [24][28][30].

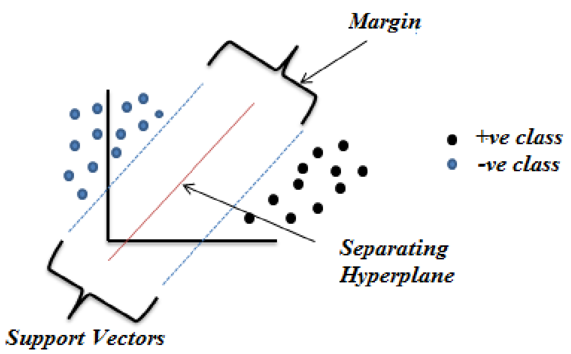

Fig. 1 support vector machine

2. **RBFN (Radial Basis Function Neural network)**

A RBFN (Radial Basis Function Neural network) is a particular type of neural network. It contains the maximum no. of simple and highly connected artificial neurons. It is a three-layered feed-forward network namely, input layer, hidden layer, and output layer [30].

Input layer: Input data enters to the input layer and performs the transfer function that indicates the output of one terminal is the input of other terminal.

Hidden layer: This layer of RBF contains the hidden neurons, and the activation. The function of these neurons is a Gaussian function. The hidden layer creates a signal according to an input vector in the input layer.

Output layer: The output layer contains output vector y with element y_j (i=1 to M). There is also a weight factor w_{kj} (k= 1 to N). This layer reacts to the input model and signal that generates by hidden layer. To get output, neurons process the input signal through an activation function of the neuron. Activation function of hidden neuron is ø(X) [30][31].

$$ø(x) = e^{-(x-\mu)2/\sigma 2}$$

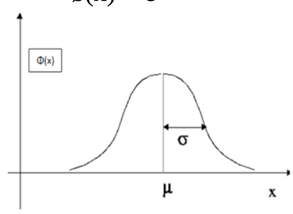

Fig. 2 Radial basis function

3. **Convolutional Neural Network (CNN)**

Convolutional Neural Network (CNN) is a class of Neural Network that has proven very efficient in the field of classification and image recognition. CNN gives successful results in some NLP task in past years; one particular task is text classification. CNN model has been classified as a key driver for NLP tasks [32]. For text classification, CNN automatically adjust parameters. CNN reduces interventions in pre-treatment and post-processing.

CNN model can automatically identify the phrases in text with max-pooling layer rather than through manual feature extraction [33]. The layered architecture of CNN as described below [32][33][34].

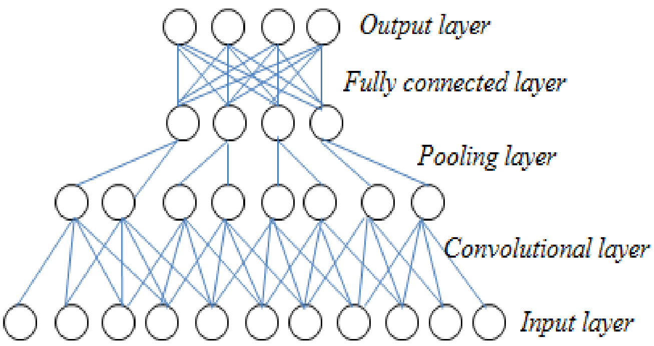

Fig. 3 CNN basic architecture

Input Layer: This layer accepts raw data as an input in the form of arrays.

Convolution Layer: This layer has no. of filters that perform convolutional operation.

ReLU: Performs element-wise operations. It set all negative pixel to 0. This Activation is applied to Convolution layer.

Pooling: This layer reduce the size of parameters. Max Pooling is the most common approach. Pooling is a downsampling operation.

Flattening: It converts all resultant 2-d array from pooled map into a single long continuous linear vector.

Fully connected layer: Each neuron in one layer connected to each of neuron in the another layer. It is same as the traditional multi-layered perceptron. Flatten data goes through this layer to classify objects.

Output layer: Finally Object is classified.

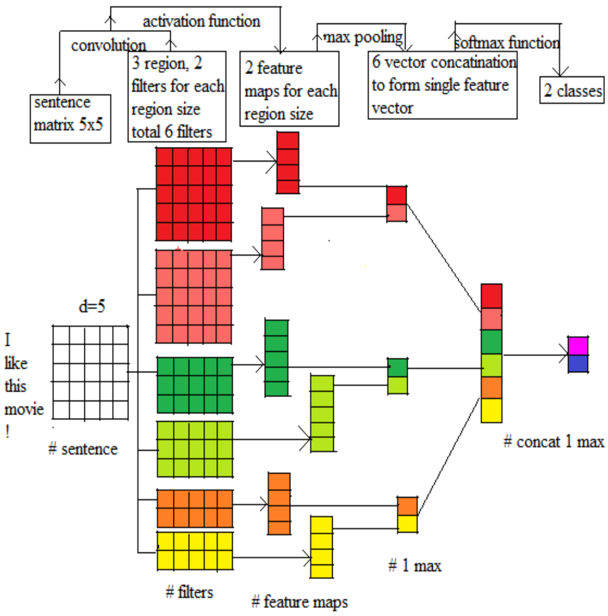

Fig. 4 CNN architecture for text classification

4. N-Gram language model

N- gram is the simplest model for assigning probabilities to sentences or sequence of words [35][36]. An N-gram is a sequence of N (no. of words) words like 1- gram (uni-gram) is one word sequence, 2- gram or bi-gram is a two word sequence and 3- gram also called tri-gram, is a three word sequence [35][36][37]. For example:

Unigram: "the", "purpose", "of", "this", "study", "was"
Bi-gram: "the purpose", "of this", "study was"
Tri-gram: "the purpose of", "this study was"

For computing N-gram, probability of word will be calculated by P(w|h), 'P' is the probability of word 'w*h', w is the word and h is the given history. For example 'h' is "this room is so dirty that" and we find the probability that the next word is 'the'. This is represented by:

P (the| this room is so dirty that)

B. Optimizers

Optimization is an essential part of machine learning algorithms. It begins by defining loss function and terminates with minimizing that function and improves accuracy. Optimizers perform the optimization task. Optimizers are the algorithms helps in either minimizing or maximizing the value of an algorithm. The optimizers improve the speed and performance for training a particular model. Keras is neural network library coded in python runs on the top of Tensor Flow [38][39][40].

There are a number of optimizers presents in Keras library that are used to bring improvement in results. Gradient Descent also known as Stochastic Gradient Descent (SGD) is the well- known algorithm used to perform optimization and is the most common way to optimize neural network. This algorithm works on all types of machine learning problems. Gradients are the small changes in weights so it calculates the small change in each individual weight then adjusts each weight or take a small step to identify direction. This process going on until loss function gets minimum as possible. Adaptive Gradient optimizer shortens in AdaGrad optimizer based on gradient optimization. This algorithm simply adapts the learning rates to parameters and work with variable learning rate. Node *weights* which have previously had large gradients are assigned large gradients and whilst node weight which have small gradients historically had assigned small gradients. AdaGrad improves SGD by assigning weight historically and it is well suited for dealing sparse data. Adadelta is an SGD technique rely on the adaptive learning rate for each dimension. It is adaptation of Adagrad and has each gradient update on each weight. It will be a weighted sum of the current gradient and average of a rolling window of past gradient updates. This algorithm is more robust and this learning rate is more stable. RMS Prop is the abbreviation of Root-Mean-Squared Propagation algorithm. It is mostly the same as Momentum algorithm. It resolves the issue of uniformity decreasing learning rate in AdaGrad. It also maintains the learning rate of each parameter that is adapted to the average of how quickly it is changing. This algorithm works well on online and noisy problems. Adam stands for Adaptive Moment algorithm. It is the combination of Momentum RMS Prop algorithms. It is a very powerful algorithm which optimizes the results very fast and also known as fast optimizer. It uses past gradients to calculates current gradients [38][39][40][41]

C. Data Acquisition

We have taken the required dataset form popular web resource www.askpatient.com. Figure 6 displays the sample format of data that we have taken. The study tells us that mostly text classification is used for a particular domain due to the kind of language used or we can say that it is domain-specific. But in our work, we focus text classification for single domain i.e. anxiety therapy review. In our dataset, we focus on different kinds of drugs used for anxiety therapy that are alprazolam, Atarax, Ativan, Buspar, diazepam, Equanil, Klonopin rapidly disintegrating, Librium, lorazepam, Miltown, Niravam, Serax, Tranxene sd, valium, Vistaril, Xanax, Xanax XR. These are the different drugs used for anxiety treatment. Each drug is made up of different salt. Dataset used for evaluation to give rigid conformation of the evaluation to be taken. We manually created a dataset for therapy chosen. We have extracted the review of every user along with their rating score. We constructed the text file in excel. We have taken only comments and rating score for each review. The comments are labeled on the basis of their rating. Each review containing text describes the rating scores of review. Comments given by different people vary in rating. Rating is changes from person to person according to their satisfaction. Rating score varies from 1 to 5. Then the text is classified into two different classes. The class is assigned positive if the rating score is >3, and is assigned negative if the rating score is <3. A total of 2346 comments are taken, from which 1246 are positive and 1100 are classified as negative.

Rating	Reasons	Side-effects	Comments	Sex	Age	Duration	Date
4	Anxiety attacks	Dry mouth	I could not function without them. Butch, they forgot to tell me there highly addicting, and you will die if you try to get off them.	F	52	20 Years	7/2/2019

Fig. 5 Sample data from www.askapatient.com

D. Procedure

The precise detail of our methodology to design and test the proposed model, as follows (Fig. 1).

i. Collect data from source (www.askspatient.com).

ii. Execute data pre-processing task (remove numbers, acronyms punctuations, URL links).

iii. Apply the N-gram model with different values of n to selected anxiety therapy reviews.

iv. Apply the following models using specific training dataset

   a. SVM.
   b. The deep learned model CNN.
   c. CNN with N-gram.

v. Calculate text in the form of positive or negative for each review.

vi. Compare the calculated results with actual values.

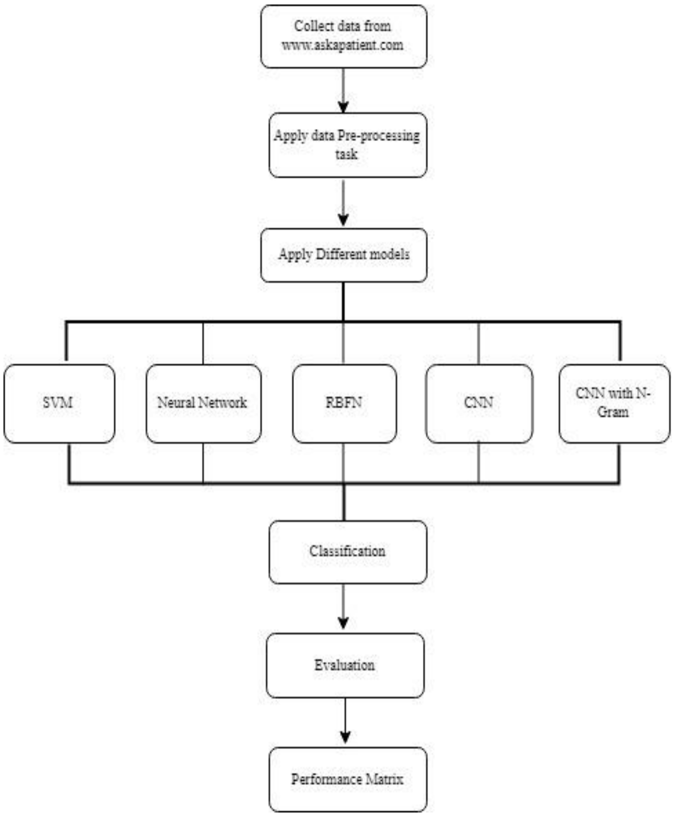

Fig. 6 Model design

III. RESULTS AND DISCUSSIONS

Different methods for evaluation of the text classification techniques have been used. In our work, the results obtained are evaluated by different models. The results obtained according to various parameters such as precision, recall, mean squared error, area under curve for each model. Results obtained for drug reviews for specific illness that is anxiety.

A. Performance Parameters

To calculate performance parameters, a confusion matrix is generated that contains the actual values and predicted values which as follows [42].

Table I Confusion matrix

Actual class/ Predicted class	Predicted class (Yes)	Predicted class (No)
Actual class (Yes)	True Positive (TP)	False Negative (FN)
Actual class (No)	False Positive (FP)	True Negative (TN)

The different measures used for evaluation different parameters, among these parameters accuracy is the most important one.

1. Accuracy

Accuracy shows the closeness towards the fact or actual result. It can be measure by:

Accuracy = (TP + TN) / (TP + FP + TN + FN)

Retrieval Number: L29061081219/2019©BEIESP
*DOI: 10.35940/ijitee.*L2906.*1081219*

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

2. Precision

Precision used to measure the correctness of classifiers. Higher precision means low false positive (FP) and lesser precision means high false-positive also it is the number of classes correctly classified divided by total number of classified classes. Correctness should be high as possible.

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

3. Recall

Recall is used to measure completeness or sensitivity. Higher the recall means lesser the false negative (FN) and lower the recall means higher the false negatives. Recall is the no. of correctly classified reviews divided by the total no. of reviews.

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

4. Area Under Curve (AUC)

AUC differentiate the classes well and AUC between two points found by calculating definite integral between them. Higher the AUC means higher the correctness of prediction.

5. F-1 Score

It defines the harmonic mean of precision and recall. If its value is 1 then it is best and if its value is 0 then it is worst.

$$\text{F-1 Score} = 2 * ((\text{precision} * \text{recall}) / (\text{precision} + \text{recall}))$$

B. Experimental Results

Experimental results show that the optimized CNN with N-gram can give better performance in the field of text classification or opinion mining as compared to the simple CNN. On the other side when our proposed work has compared with some baseline methods then the result clearly indicated that our method gave effective results based on different performance parameters.

a. *Comparative analysis between Optimizers*

Table- II: Optimizers

Optimizers	Loss	Accuracy	Precision	AUC	MSE	Recall
Adadelta	0.945	0.496	0.523	0.495	0.512	0.368
AdaGrad	0.365	0.848	0.652	0.625	0.373	0.628
Gradient Descent	0.682	0.647	0.567	0.538	0.461	0.543
Proximal AdaGrad	0.326	0.852	0.662	0.639	0.359	0.656
Proximal Gradient Descent	0.559	0.723	0.585	0.559	0.439	0.582
RMS Prop	0.022	0.998	0.742	0.726	0.272	0.743
Adam	2.645	1	0.930	0.926	0.730	0.930

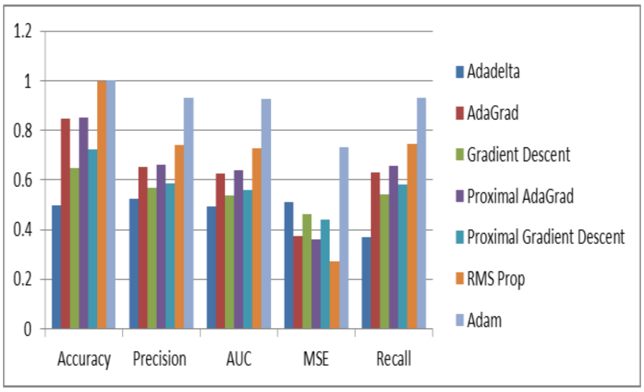

Fig. 7 Comparison analysis among different parameters

In the above result we have found that among various optimizers, Adam gives the best results in all parameters except loss. Then we have further used this optimized CNN and compared according to different parameters.

b. *Comparison between CNN and CNN with N-gram*

In this section, we have compared the performance of models CNN and CNN + N-grams based on different parameters Mean Squared Error (MSE), Precision, Recall, and Area Under Curve (AUC). In a previous study, we found that the N-gram works well with the value n=3 [43][44]. So, we have calculated results according to n=3 and obtained results show that CNN + N-grams gave better results than simple CNN in all parameters. These results are given below.

Table- III: Mean Squared Error (MSE)

Steps	Simple CNN	CNN + N-gram
100	.7575	0.01634
300	.261	0.00552
500	.1582	0.0033
700	.0930	0.00238
900	.1336	0.00182
1100	.00725	0.00148
1300	.0061	0.00126
1500	.0053	0.00108

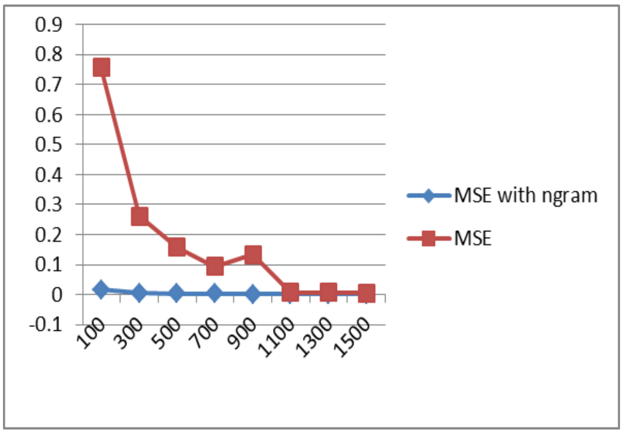

Fig. 8 Shows difference between MSE

Table- IV: Precision

Steps	Simple CNN	CNN + N-gram
100	.929	0.9887
300	.9749	0.9961
500	.9847	0.9976
700	.9890	0.9982
900	.9914	0.9970
1100	.9929	0.9989
1300	.9940	0.9990
1500	.9945	0.9991

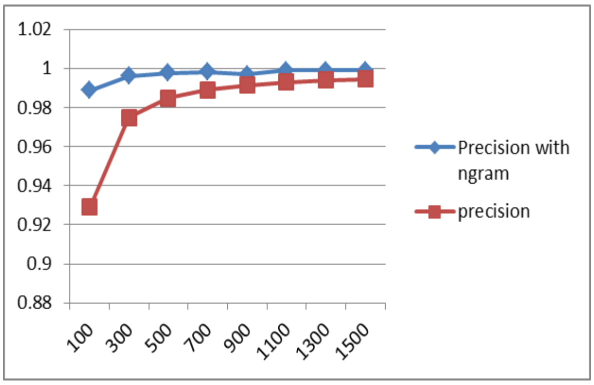

Fig. 9 Comparison between Precision

Table- V: Recall

Steps	Simple CNN	CNN + N-gram
100	.9305	0.991
300	.9726	0.99666
500	.9852	0.99792
700	.9892	0.99654
900	.9916	0.99884
1100	.9931	0.99904
1300	.9941	0.99918
1500	.9949	0.99928

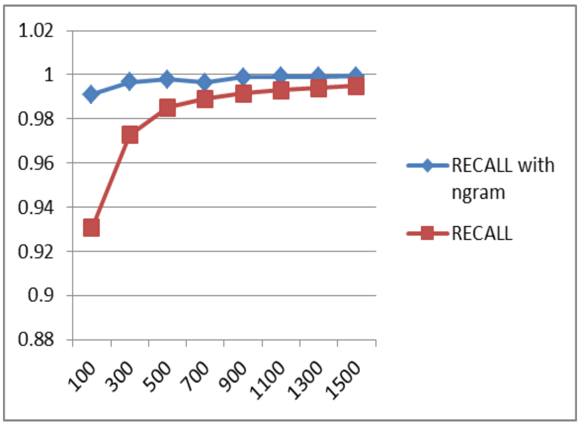

Fig. 10 Comparison between Recall

Table- VI: AUC

Steps	Simple CNN	CNN + N-gram
100	.9249	0.97488
300	.9736	0.9914
500	.9840	0.99478
700	.9885	0.99626
900	.991	0.99706
1100	.9926	0.99758
1300	.9937	0.99794
1500	.9945	0.99822

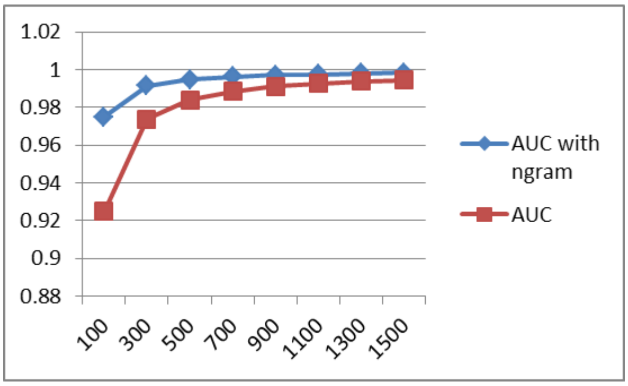

Fig. 11 Define difference between AUC parameter

Table 3-7 are the summarized results of all classification methods. Results are calculated with a comparison of two methods, first is with simple CNN and other with CNN combines with N-grams. In Table 1 Mean squared error is compared among both models and CNN with N-gram gives better results than simple CNN. In all parameters CNN combines with N-gram gives better result than simple CNN.

c. *Comparison between baseline methods*

In this section we have compared some machine learning methods with our proposed method and found that among machine learning technique Support Vector Machine (SVM), Radial Basis Function Neural Network (RBFN) and Neural Network (NN), RBFN performs better. Further, we have compared RBFN with CNN and our proposed model and our proposed model performs better than all gives 99.9% Accuracy and 0.004 MSE, 99.6% Precision, and 99.7% Recall. These results are shown below.

Table- VII: Baseline Methods

models	Precision	Recall	Mean squared error	Accuracy
SVM	0.581	0.569	0.535	56.94%
NN	0.521	0.514	0.408	51.36%
RBFN	0.66	0.629	0.4975	62.87%
CNN	0.981	0.981	0.282	99.98%
CNN+N-Gram	0.996	0.997	0.0041	99.99%

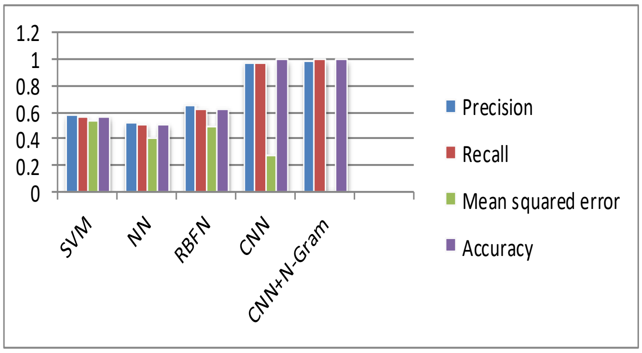

Fig. 12 Comparison with baseline methods

Results has obtained show that among different methods, the combination of CNN and N-gram technique gives better results on basis of different parameters; Mean Squared Error, Area Under Curve, Precision, Recall, Accuracy and loss.

Firstly, supervised machine learning methods such as Support Vector Machine, Neural Network, and Radial Bases Function were implemented. Among all, Radial Bias Function gave better results. From a previous study, we found that Deep learning gives better results than Machine learning (Wei, F., Qin, H., Ye, S., & Zhao, H., 2019) so we implemented CNN on our dataset. We have also compared the different optimizer's results and chose the best one (Khullar, V., Bala, M., & Singh, H. P., 2019; Thirunarayanan, R., Ruffieux, D., Scolari, N., & Enz, C., 2016).

In this research Adamax, Adam and Adadelta perform better with low losses and high accuracy. But in our work RMS Prop gives least losses and Adam performs better in other parameters.

So we have used Adam optimizer for better results. Further, we enhanced the pre-processing by implementing N-gram Technique and comparison was done with simple CNN and found that the combination of CNN and N-gram gives better results than simple CNN.

IV. CONCLUSION

In this work, we have proposed a deep learning technique for text classification by combining the advantages of CNN and N-gram. This work gives 99.98% accuracy in comparison of SVM (56.9%), NN (51.36%), RBFN (62.87%). Also the CNN is superior as it had achieved precision (0.981), recall (0.981), mean squared error (0.282). Further the N-gram technique has been implemented along with CNN which improves the results of simple CNN on the basis of precision (0.999), recall (0.999), mean squared error (0.001), area under curve (0.998) and accuracy remained the same. The proposed system summarizes that CNN with N-gram is the best choice for text classification to identify the impact of therapy from patient's text feedback.

In future the work can be done over other deep learning techniques and models can be shifted over Big data platforms.

REFERENCES

1. S. M. Rezaeinia, A. Ghodsi, and R. Rahmani, "Text Classification based on Multiple Block Convolutional Highways," 2018.
2. A. Hassan and A. Mahmood, "Convolutional Recurrent Deep Learning Model for Sentence Classification," IEEE Access, vol. 6, pp. 13949–13957, 2018.
3. G. Vinodhini and R. M. Chandrasekaran, "Opinion mining using principal component analysis based ensemble model for e-commerce application," CSI Trans. ICT, vol. 2, no. 3, pp. 169–179, 2014.
4. K. Xu, S. S. Liao, J. Li, and Y. Song, "Mining comparative opinions from customer reviews for Competitive Intelligence," Decis. Support Syst., vol. 50, no. 4, pp. 743–754, 2011.
5. Jagriti, M. Bala, H. P. Singh, and V. Khullar, "Experimental Review on Text Classification of Patient Feedback Using Supervised Learning Methods."
6. A. S. H. Basari, B. Hussin, I. G. P. Ananta, and J. Zeniarja, "Opinion mining of movie review using hybrid method of support vector machine and particle swarm optimization," Procedia Eng., vol. 53, pp. 453–462, 2013.
7. S. Baccianella, A. Esuli, and F. Sebastiani, "Multi-facet rating of product reviews," Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 5478 LNCS, pp. 485–496, 2009.
8. J. Zhu, C. Xu, and H. S. Wang, "Sentiment classification using the theory of ANNs," J. China Univ. Posts Telecommun., vol. 17, no. SUPPL. 1, pp. 58–62, 2010.
9. B. Pang, L. Lee, and S. Vaithyanathan, "Sentiment Classification using Machine Learning Techniques," Proc. - Symp. Log. Comput. Sci., no. July, pp. 97–106, 2011.
10. L. S. Chen, C. H. Liu, and H. J. Chiu, "A neural network based approach for sentiment classification in the blogosphere," J. Informetr., vol. 5, no. 2, pp. 313–322, 2011.
11. Q. Ye, Z. Zhang, and R. Law, "Sentiment classification of online reviews to travel destinations by supervised machine learning approaches," Expert Syst. Appl., vol. 36, no. 3 PART 2, pp. 6527–6535, 2009.
12. F. Greaves, D. Ramirez-Cano, C. Millett, A. Darzi, and L. Donaldson, "Machine learning and sentiment analysis of unstructured free-text information about patient experience online," Lancet, vol. 380, p. S10, 2012.
13. W. Wang, L. Chen, M. Tan, S. Wang, and A. P. Sheth, "Discovering Fine-grained Sentiment in Suicide Notes," Biomed. Inform. Insights, vol. 5s1, p. BII.S8963, 2012.
14. V. Bobicev, M. Sokolova, Y. Jafer, and D. Schramm, "Learning sentiments from tweets with personal health information," Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 7310 LNAI, pp. 37–48, 2012.
15. P. Yin, H. Wang, and L. Zheng, "Sentiment classification of Chinese online reviews: Analysing and improving supervised machine learning," Int. J. Web Eng. Technol., vol. 7, no. 4, pp. 381–398, 2012.
16. T. Ali, M. Sokolova, D. Schramm, and D. Inkpen, "Opinion learning from medical forums," Int. Conf. Recent Adv. Nat. Lang. Process. RANLP, no. September, pp. 18–24, 2013.
17. A. Akay, A. Dragomir, and B. E. Erlandsson, "Network-based modeling and intelligent data mining of social media for improving care," IEEE J. Biomed. Heal. Informatics, vol. 19, no. 1, pp. 210–218, 2015.
18. D. Tang, F. Wei, B. Qin, T. Liu, and M. Zhou, "Coooolll: A Deep Learning System for Twitter Sentiment Classification," no. SemEval, pp. 208–212, 2015.
19. J. Nayak, B. Naik, and H. S. Behera, "A Comprehensive Survey on Support Vector Machine in Data Mining Tasks: Applications & Challenges," Int. J. Database Theory Appl., vol. 8, no. 1, pp. 169–186, 2015.
20. P. Wang, B. Xu, J. Xu, G. Tian, C. L. Liu, and H. Hao, "Semantic expansion using word embedding clustering and convolutional neural network for improving short text classification," Neurocomputing, vol. 174, pp. 806–814, 2016.
21. M. Ashley, "Deep learning facilitates rapid classification of human and veterinary clinical narratives," vol. 983, 2018.
22. Z. Jianqiang, G. Xiaolin, and Z. Xuejun, "Deep Convolution Neural Networks for Twitter Sentiment Analysis," IEEE Access, vol. 6, no. c, pp. 23253–23260, 2018.
23. F. Wei, H. Qin, S. Ye, and H. Zhao, "Empirical Study of Deep Learning for Text Classification in Legal Document Review," Proc. - 2018 IEEE Int. Conf. Big Data, Big Data 2018, pp. 3317–3320, 2019.
24. S. M. Rezaeinia, A. Ghodsi, and R. Rahmani, "Text Classification based on Multiple Block Convolutional Highways," 2018.
25. S. Tan and J. Zhang, "An empirical study of sentiment analysis for chinese documents," Expert Syst. Appl., vol. 34, no. 4, pp. 2622–2629, 2008.
26. A. S. M. Alharbi and E. de Doncker, "Twitter sentiment analysis with a deep neural network: An enhanced approach using user behavioral information," Cogn. Syst. Res., vol. 54, pp. 50–61, 2019.
27. Z. Luo, L. Liu, J. Yin, Y. Li, and Z. Wu, "Deep learning of graphs with ngram convolutional neural networks," IEEE Trans. Knowl. Data Eng., vol. 29, no. 10, pp. 2125–2139, 2017.
28. F. Colas and P. Brazdil, "On the behavior of SVM and some older algorithms in binary text classification tasks," Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 4188 LNCS, pp. 45–52, 2006.
29. A. Sun, E. P. Lim, and Y. Liu, "On strategies for imbalanced text classification using SVM: A comparative study," Decis. Support Syst., vol. 48, no. 1, pp. 191–201, 2009.
30. W. Shen, X. Guo, C. Wu, and D. Wu, "Forecasting stock indices using radial basis function neural networks optimized by artificial fish swarm algorithm," Knowledge-Based Syst., vol. 24, no. 3, pp. 378–385, 2011.
31. D. K. Ghose, S. S. Panda, and P. C. Swain, "Prediction of water table depth in western region, Orissa using BPNN and RBFN neural networks," J. Hydrol., vol. 394, no. 3–4, pp. 296–304, 2010.
32. P. Llewellyn Jones, "NLP Advances in natural Language processing," Unpubl. Lect. handouts, pp. 2–3, 2001.
33. A. Rios and R. Kavuluru, "Convolutional neural networks for biomedical text classification," pp. 258–267, 2015.

Retrieval Number: L29061081219/2019©BEIESP
*DOI: 10.35940/ijitee.*L2906.*1081219*

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

34. H. Ide and T. Kurita, "Improvement of learning for CNN with ReLU activation by sparse regularization," Proc. Int. Jt. Conf. Neural Networks, vol. 2017-May, pp. 2684–2691, 2017.
35. "N-grams for Text Classification Using Supervised Machine Learning Algorithms KENNEDY ODHIAMBO OGADA A thesis submitted to the School of Computing and Information Technology in partial fulfillment for the degree of Doctor of Philosophy In Information Techn," 2016.
36. W. B. Cavnar, W. B. Cavnar, and J. M. Trenkle, "N-Gram-Based Text Categorization - Cavnar, Trenkle, Mi - Unknown.pdf," Proc. Sdair-94, 3Rd Annu. Symp. Doc. Anal. Inf. Retr., pp. 161--175, 1994.
37. W. B. Cavnar and J. M. Trenkle, "N-Gram-Based Text Categorization N-Gram-Based Text Categorization," no. December 2012, 2001.
38. N. Ketkar, "Deep Learning with Python," Deep Learn. with Python, pp. 95–109, 2017.
39. P. Vidnerova and R. Neruda, "Evolving keras architectures for sensor data analysis," Proc. 2017 Fed. Conf. Comput. Sci. Inf. Syst. FedCSIS 2017, no. December, pp. 109–112, 2017.
40. V. Khullar, M. Bala, and H. P. Singh, "Performance Evaluation of Autism Diagnosis Tool at Diverse Learning Rates," vol. 6, no. 2, pp. 92–99, 2019.
41. R. Thirunarayanan, D. Ruffieux, N. Scolari, and C. Enz, "A 51.4 Mb/s FSK transmitter employing a Phase Domain Digital Synthesizer with 1.5 µs start-up for energy efficient duty cycling," Eur. Solid-State Circuits Conf., vol. 2016-October, pp. 129–132, 2016.
42. L. C. Briand, J. Wüst, J. W. Daly, and D. Victor Porter, "Exploring the relationships between design measures and software quality in object-oriented systems," J. Syst. Softw., vol. 51, no. 3, pp. 245–273, 2000.

AUTHORS PROFILE

Jagriti is the Research Fellow in the department of Computer Science and Engineering, CT institute of Engineering, Management and Technology, Jalandhar, India. Her area of interest is Opinion Mining, Cloud Computing, Data Mining, Neural Network, Deep learning.

Vikas Khullar is working as Assistant Professor in the Department of Computer Science and Engineering, at CT Institute of Engineering, Management, & Technology, India. His area of research is Autism Spectrum Disorder, Big Data, IoT, Machine Learning, and Deep Learning. His teaching experience is 10 years and his research findings have been published through more than 20 research papers in refereed/indexed journals and international conferences.

Dr Harjit Pal Singh is Associate Professor & Head, Department of Electronics and Communication at CT Institute of Engineering, Management and Technology, India. His area of interest is Wireless Sensor Network, Autism, IoT, VoIP Wireless Communication, Digital Filter Design, and Wireless Sensor Networks. His research and teaching experience is 10 years. His research findings have been published through more than 50 research papers in refereed/indexed journals and international conferences.

Dr Manju Bala is Professor and Director at Khalsa College of Engineering and Technology, Amritsar, India. Her area of interest is Wireless Sensor Network, Digital Signal Processing, Autism Spectrum Disorder, Internet of Things, and Wireless Communication. Her research and teaching experience is 15 years. Her research findings have been published through more than 60 research papers in referred / indexed journals and international conferences