2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS) | 978-1-6654-0521-8/20/$31.00 ©2021 IEEE | DOI: 10.1109/ICACCS51430.2021.9441976

Intelligent Deep Neural Network integrated with Chaotic Particle Swarm Intelligence based Sentiment Analysis in Big Data Paradigm

T.Sreenivasulu
Department of Computer Science
VELS Institute of Science, Technology and Advanced Studies (VISTAS)
Chennai, India
sreesree82@gmail.com

R. Jayakarthik
Department of Computer Science
VELS Institute of Science, Technology and Advanced Studies (VISTAS)
Chennai, India
drrjayakarthik@gmail.com

***Abstract*— In big data paradigm usage of sentiment analysis has rapidly increased its commendable pace in any kind of environment like political, social or present affairs. As it is easy to gather the sentiments of public of entire world are stated through the assistance of social media which is more suitable for sentiment mining. This paper focuses on analyzing the service of airline industries, by applying the fuzzy induced Intelligent deep neural network (IDNN) empowered with the knowledge of chaotic particle swarm optimization, which uses the twitter sentiment analysis on their respective passengers to get their feedback or opinion. This work used the deep learning model with the fuzzy control for fine tuning the weight assignment of the DNN more precisely. The tweets are extracted and they are preprocessed to extract the essential features involved in classification of the sentiments as either positive, negative or neutral. The fuzzy deep neural network is fine-tuned by integrating chaotic particle swarm optimization. The particle swarm behaviour improves the performance of the Fuzzy DNN by avoiding random selection of population to chaotic dynamics. The performance of the developed IDNN is compared with the support vector machine and random forest classifier. The results show that the proposed model well handles with imbalance voluminous dataset in big data paradigm with higher accuracy rate.**

Keywords*— *sentiment analysis, fuzzy, deep neural network, support vector machine, random forest classifier, deep learning model

I. Introduction

An approach which analyze the piece of text as sentiment is known as sentiment analysis of opinion mining. Sentiment analysis by means of twitter has become a captivating field of research which provides more active sentiments of the public specifically in airline industries. By analyzing the sentiment of text, it allows to understand an idea of whether a piece of text is negative, positive or neutral [1]. The publics also very easily either pleased or displeased more often based on the views of twitter. While comparing with other sources like review websites or blogs are very small compared to the twitter where it is unimaginable count. Nearly 100 million people tweets daily and this number grows rapidly each day [2]. Conducting research on big data becomes easier while applying sentiment analysis. The tweets send by the mass public can be treated as opinion and their perspective of information may help the organization to understand the feelings of the publics and it may recommend them to improve their quality of service by fulfilling the needs of the customers or passengers [3].

The sentiment analysis based on the tweets can be used as a prominent instrument for organizers when they receive feedback and review from their customers especially when they release a new product. Here, sentiments are classified as the feeling of being negative or positive. The feedback depends on individual's perspective, at the same time neutral tweets shows neither negative or positive and they don't have any conclusive sentiment of it [4]. The tweets characteristics are generally language model like frequency of words, tweets lengths which is to be normally 140 characters, removing stopping words, converting the text into lowercase and removing punctuation marks are the essential preprocessing.

II. Related Work

Sindhwani and Melville [5] developed a semi supervised model of sentiment classification, which use lexical prior knowledge along with unlabeled data.

Hinton et al [6] in their work introduced a deep belief model for sentiment classification, this model effectively does the classification of sentiment with its prior knowledge of deep structure of the network and they perform well compared to other semi-supervised approaches.

Yassine et al [7] devised a sentiment analysis form product reviews using hybrid approach. The model classifies the product review as positive or negative related to the query posted in the concern product reviews offered by the amazon.

Turkey [8] in his work used point wise mutual information and retrieval of information has been analyzed on online product reviews. Semantic orientation is computed for determining mutual information among reference words and opinion-oriented words depending on the statistical information and using web search engine-based information retrieval. Mutual information of two different words phrase is used to determine whether the review is positive or negative.

ZHANG Zi et al [9] used to analyse the review of Chinese products using the concept of snippet crawling. This work

978-1-6654-0521-8/21/$31.00 ©2021 IEEE

simplifies sentiment estimation by eliminating the ratio of hit numbers of excellent and poor words.

Yuanbin Wu el at [10] discovers an opinion mining oriented product review which classifies the review as either positive or negative by applying the two-word phrases using opinion-oriented words. This helps to determine relationship among product features and the opinion expression by customers.

Won Young Kin et al [11] implement the unsupervised models and proved that it is simple to implement as supervised model. In general, supervised models have similar training dataset, but they fail to classify them in different domains, so the classifier has to be trained of the specific domain. Unsupervised approach doesn't need such requirement to annotate the trained data thus it is very simple and produce better result for all kind of domains.

Methodology of Intelligent Deep belief Network based Sentiment Analysis in Big Data Paradigm

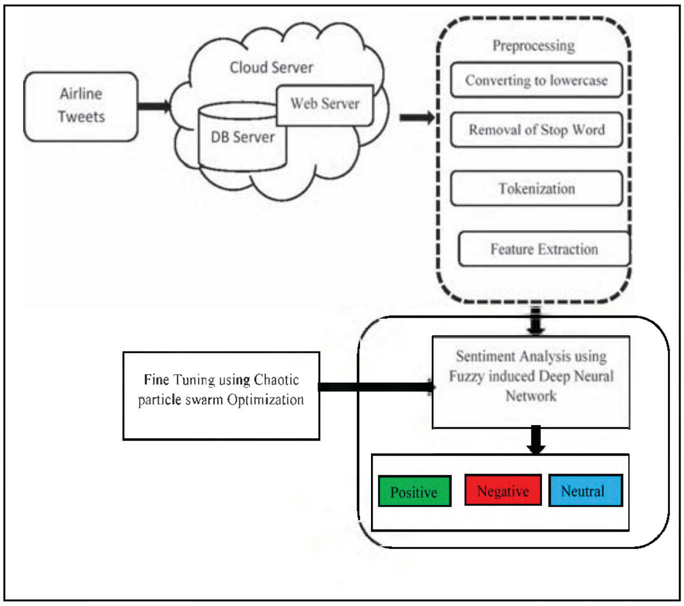

Fig. 1. Workflow of the Intelligent deep neural network for sentiment analysis on Big Data Paradigm.

This paper concentrates on sentiment analysis of airline tweets to determine the satisfaction or opinion from passengers about flight services and improve their quality based on the analysis. The airline tweet dataset which is collected from the cloud server was in raw format. The dataset undergoes preprocessing activities like converting to lowercase, removing url, tokenization of text words and feature extraction are done in this work. With the extracted features they are converted into two different subsets such as document term matrix and data frame. With the collected document matrix fed as input to the deep neural network the words association are analyzed in each layer of the DNN and the weights in the Restricted Boltzmann machine is optimized with the help of the fuzzy optimizer. The rules induced by the fuzzy DNN is scrutinized by applying chaotic particle swarm optimization. The proposed model classifies the input tweet as either positive, negative or neutral in an intelligent way.

Dataset Description

In this work U.S airlines [14] and the passenger tweets are used as the dataset for evaluating the developed model. This dataset comprised of various features like tweet id, airline sentiment, airline sentiment gold, retweet count, location of tweet, time zone of tweet, reason for negative comment, etc. This work used only the useful features like airline sentiment class label, negative reasons and the airlines details for further investigation.

Deep Neural Network:

Deep learning is a sort of model of AI that plays out the cycle of classification legitimately from text, sound or pictures. Generally, the foundation of deep learning is neural network architecture. Deep alludes to a larger number of layers for the arrangement of a deeper network. There are just a few layers in an ordinary neural network, while deep neural networks can have hundreds. This model joins numerous non-direct layers for handling and works in equal way utilizing basic components. It is created based on the motivation of functionalities of sensory nervous system, it includes a single info layer, numerous shrouded layers, and a yield layer. Each layer is interlinked by means of neurons, with each concealed layer gets its yield from the past layer.

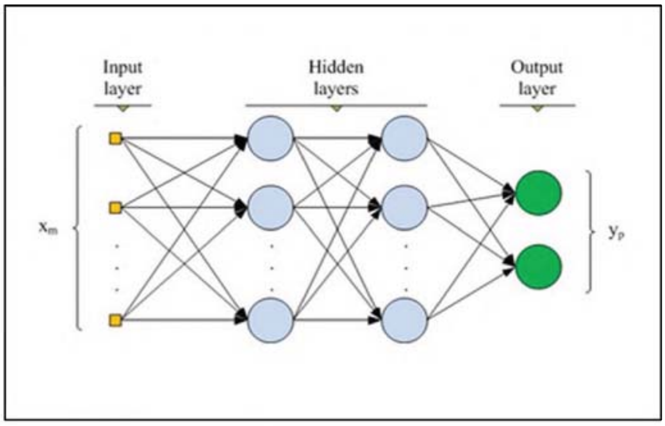

Fig. 2. Deep Neural Network

Functionalities of ordered instinctive-codifiers

An Instinctive-Codifier (IC) belongs to the type of neural network which has a single hidden layer with equal number of nodes as in input and output layer. The training set comprised is denoted as $\{ ip^{(1)}, ip^{(2)}, ip^{(3)}, \ldots \}$ where $ip^{(i)} \in TR^d$. The instinctive codifier initially translates an input $ip^{(i)}$ to a hidden node description $Z(ip^{(i)})$ calculated as shown in equation (1)

$$z(ip) = fn(CM_1 * ip + bs_1) \quad (1)$$

where CM_1 is a weight factor, bs_1 is an bias factor and fn(ip) is the activation function. Next, it decodes $Z(ip^{(i)})$ as replica $A(ip^{(i)})$ formulated in equation (2)

$$A(ip) = gf(CM_2 \cdot z(ip) + bs_2) \quad (2)$$

Where CM_2 is a decoding factor and bs_2 us a bias vector. Gf(ip) is the activation function. This model is trained using back propagation approach The back propagation assigns the weights to hidden nodes in a random fashion, the output of the network is compared with the expected output and by measuring its error rate, propagates back to the hidden node and reassign the weight values and recomputed its output accordingly and this is done in an iterative manner. The objective of this neural network is to minimized its error rate and is defined as shown in the equation (3)

$$flt(IP,O)=\frac{1}{2}\sum_{i=1}^{N}\left\|ip^{i}-o^{i}\right\|^{2} \quad (3)$$

Where IP is the set of N input records $\{ip^{(1)},ip^{(2)},ip^{(3)},...ip^{(N)}\}$, and O is the set of analogous rebuilt output $\{o^{(1)},o^{(2)},...,o^{(N)}\}$. The notation $\|\cdot\|$ signifies 2-norm of a vector.

The Ordered instinctive-codifiers are constructed with the three different layers, input layer, hidden layer and output layer. It is trained with set of instances of airline tweet dataset with class labels the weight factors CM_1^i and CM_2^i are assigned. Each time the weights are reassigned based on the input pattern and thus it completes the task of multiple instinctive codifiers' process.

Back-propagation mechanism

During the learning process of ordered instinctive-codifiers, Back Propagation (BP) is applied for reassigning the weights assigned to the hidden nodes, which involves in computation of the desired output. This back propagation is repeated for arbitrary number of times over the hidden layers, till they reach the minimum error among the expected output and the actual output. The BP assigns the weights among hidden nodes in a trial and error basis until achieves the minimum error rate or maximum iteration is reached. The gradient method is involved to determine the fault rate as formulated below by (4).

$$Flt(t)=\frac{1}{2}|err(t)|^{2} \quad (4)$$

Where err(t) is the different between the expected and actual output and using this value the weights among the neurons are modified and it is computer as shown in (5).

$$\Delta wt_j(t+1)=\eta\left(\frac{Flt(t)}{wt_j}\right)+\alpha\,\Delta wt_j(t) \quad (5)$$

where learning rate is denoted by η and the α is the constant with positive value, these parameters can either increase or decrease the degree of weight modification rate during the training period. There should be a prominent balance among learning rate because if it is higher it makes the network to be unstable, while if it's low then the training time of the neural network will be increased. To enhance the learning process of the network, this research work adapts the fuzzy control system and its working principle is explained in the following section.

Fuzzy logic induced deep neural network

A fuzzy logic is a method which is used for defining the knowledge of humans by the help of its inference system. The fuzzy system comprised of four unique models they are fuzzification, fuzzy inference process, fuzzy control system and Defuzzifier. The input of the airline tweets are converted to the fuzzy data in terms of membership value using fuzzifier module. The fuzzified data is combined using fuzzy inference module along with the control rules in order to generate the fuzzy output. Finally the Defuzzifier translates the fuzzy membership values to the real values of the airline tweet instances. The main motivation of using fuzzy logic is it has the ability of appropriate reasoning and handling the uncertain information where clearly and they permit decision making with the assessed values under incomplete situation. While adapting fuzzy control model which adjust the learning disputes of neural network based on the Mean Square Error (MSE) fault, this avoid the over fitting during learning process and avoids the local minimum. The fuzzy rules are generated with four important factors namely supportive fault (SF), varying supportive fault (VSF), mark vary Fault (MVF) and Embarrass summation of mark vary in fault (ESMVF). The four structures are described as shown in the equation (6):

$$\begin{cases} SF(t)=F(t)-F(t-1) \\ VSF=SF(t)-SF(t-1) \\ MVF(t)=1-\left\|\frac{1}{2}\left[mark\,(SF(t-1))+mark(SF(t))\right]\right\| \\ ESMVF=\sum_{m=t-4}^{t}MVF(m) \end{cases} \quad (6)$$

Assume that fuzzy system comprised of two inputs SF and VSF with two output values which varies its learning rate as Δ_η and Δ_α signifies the varying energy level.

For simplicity, we assume that the fuzzy logic system contains two inputs SF and VSF whose range lies between 0 and 1, and two output; the vary in the learning stricture Δ_η and vary in the impetus value Δ_α. The crisp value of SF, VSF and Δ_η is converted to their linguistic form of value such as negative large, negative small, positive small, positive large and zero.

Chaotic Particle Swarm Optimization for Pruning Rules

The standard Particle Swarm Optimization (PSO) keeps a population of entities which is known as particles that are guided by social interaction to reach the best prominent area of the solution search space. In general, these particles begin searching with an initial position in a random manner with to find their maximum or minimum objective function. Each particle, travel in a D problem space. Each particle x_i movement relates to its velocity Vl_i, their personal best position is denoted as $pst_i=(pst_{i1}, pst_{i2},...,pst_{id})$ and their neighbors best position $lpbst_i=(lpbst_{i1}, lpbst_{i2}, \ldots, lpbst_{id})$. The velocity of each particle is updated using the formula as follows

$$Vl_{id}=wt * Vl_{id}+C_1 * \varphi 1 * (pst_{id}-x_{id})+C_2 * \varphi_2 * (lpbst_{id}-x_{id})$$

The position of each particle is also updated with the two uniform random numbers φ_1 and φ_2 and the cognitive coefficients C_1 and C_2

$$x_{id}=x_{id}+Vl_{id}$$

To improve the searching behavior of particles the chaotic theory is integrated with particle swarm optimization [16]. chaotic variables are obtained by applying tent map which is formulated as

$$Z_{n+1}=\mu(1-2\,Z_n-0.5\| \quad 0\leq Z_n\leq 1 \;\; n=0,1,2..$$

Where µ is a bifurcation parameter which exhibits chaotic dynamic over entire particle swarm optimization to select the initial population of particles and also used to map the searching range

$x_{ij} = x_{min,j} + z_j^{(i)}(x_{max,j} - x_{min,j})$ j= 1,2,3..D

After initializing the swarm's with chaotic mapping it is defined as

$x_i = (x_{i1}, x_{i2}, \ldots x_{iD})$ i= 1,2,3…N

Fig. 3 depicts the work flow of the particle swarm optimization, it begins with initializing the population of particles involved in search space. Each particle fitness value is determined for its global best and local best positions. Update their ranking and position during each iteration and repeat the process until it reaches it maximum iteration or criteria is met.

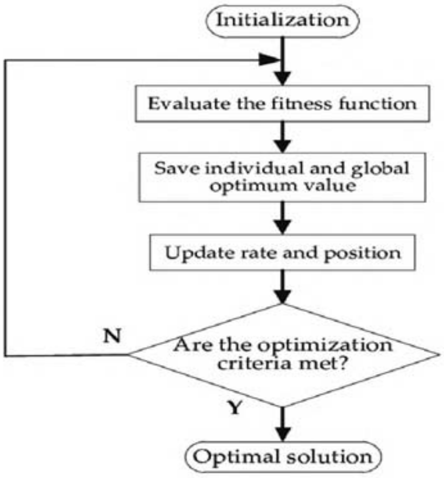

Fig. 3. Flow of Particle Swarm Optimization

The Chaotic Particle Swarm is applied on the Fuzzy Deep Neural to achieve the better classification by applying prominent rules. The figure 4 shows the personal best position and global best position of each particle's and if it reaches its local optima at their earlier stages, then instead of random movement mean best position is discover using chaotic mapping and based on it fitness value is evaluated.

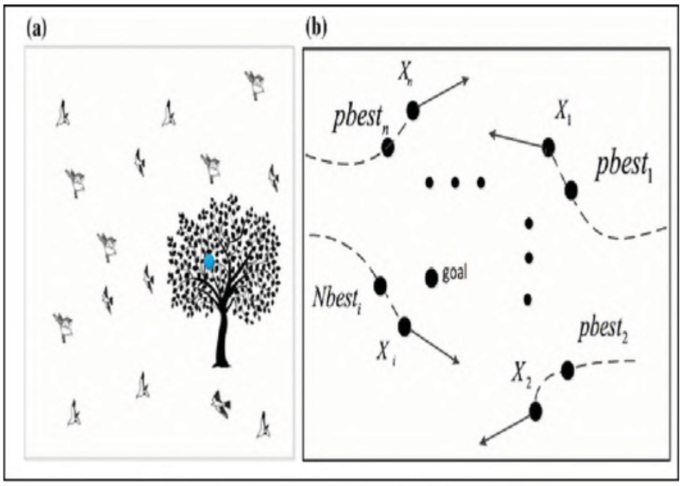

Fig.4. Chaotic Particle Swarm Optimization

Results and Discussions

This proposed work IDNN based Airline Tweet Sentiment Analysis is developed using python software. This airline tweet dataset is collected from Kaggle repositories [14] which consist of 14640 tweets. Various features are extracted from this tweet data by preprocessing them with removing stop words, conversion to lower case, removing punctuations and spaces. Tag in the tweets is removed and bag of words is used for finding the frequency of occurrence of each word. The fuzzy induced deep neural network along with the chaotic particle swarm intelligence which classifies each tweet as either negative or positive and this work eliminate the tweets which are neutral which doesn't give any useful information.

By analyzing the opinion of the passengers on tweets about the airline service it is possible to know the good and bad opinion about their flight services. The developed intelligent deep neural network [IDNN] based airline tweet sentiment analysis is compared with Random forest classifier (RFC) [6] and support vector machine (SVM) [6] with the evaluation metrics of precision, recall, f-measure and accuracy. The detailed explanation of the obtained results is discussed in the following subsections.

Metrics

The precision of the classification model is determined by its ratio of correctly predicted positive reviews of the classifier to the total number of reviews predicted as positive.

$$PRC = \frac{\text{No of Correctly Predicted Positive Reviews (TPPR)}}{\text{Total No of reviews Predicted as Postiive (TPPR+FPPR)}} \quad (7)$$

Recall of a classification model is computed by its ratio of correctly predicted positive review by the classifier to the actual number of positive reviews.

$$RCL = \frac{\text{No of Correctly Predicted Positive Reviews (TPPR)}}{\text{Acutal No of reviews Postiive (TPR+FNR)}} \quad (8)$$

F-Measure is the impact of both precision and recall measure obtained by a specific classifier which is computed as follows

$$F\text{-}Sc = \frac{2*Prcs*Rcl}{Prcs+Rcl} \quad (9)$$

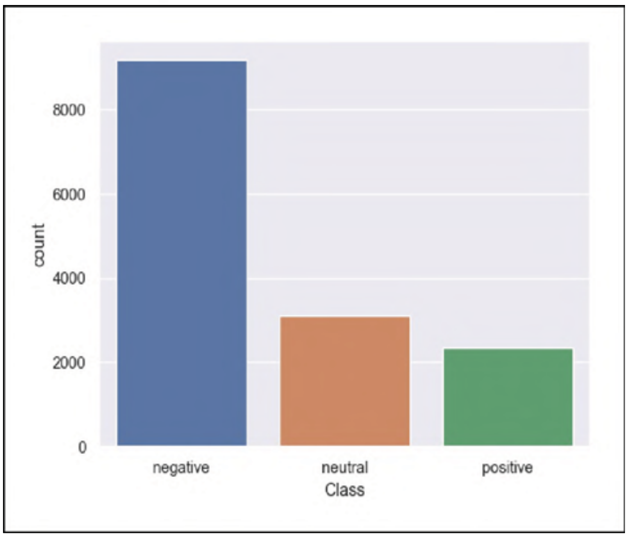

Fig. 5. Actual Count of Positive, Negative and Neutral Classes

Fig 5. portraits the actual count of positive, negative and neutral class counts of the airline tweet, the number of positive tweets is 2363, number of neutral tweets are 3088 and number of negative tweets are 9175.

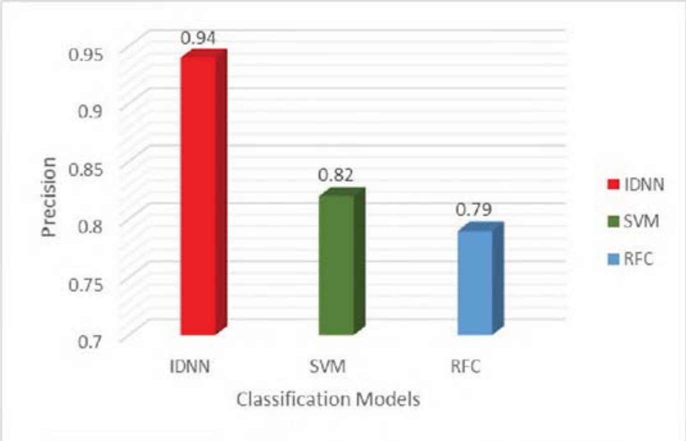

Fig. 6. Comparison Results of three sentiment analysis classification based on precision

Fig. 6. depicts the performance of three classifiers IDNN, SVM and RFC is sentiment analysis and classification of tweet about airline service as negative or positive by applying several preprocessing processes, extracting its features and analyzing their lexical patterns to classify them using three classification models. The result shows that the developed model IDNN produces high rate of accuracy in classification of opinion about the airline services.

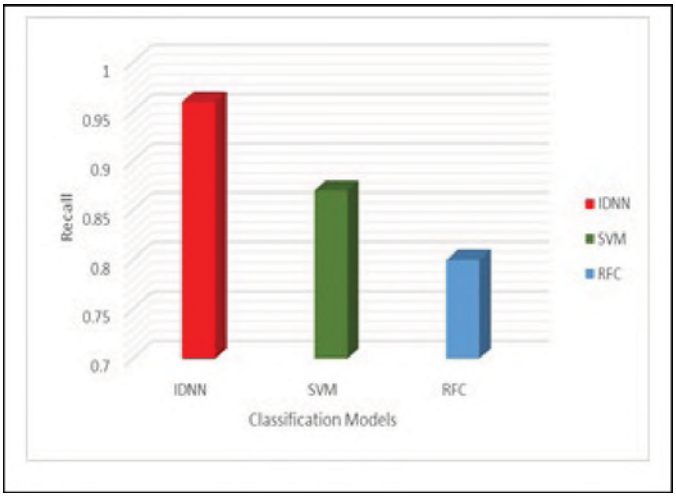

Fig. 7. Comparison Results of three sentiment analysis classification based on Recall

Fig.7. illustrates comparison results of three classifiers IDNN, SVM and RFC based on measure of recall. The opinion or sentiment analysis of tweet about airline service greatly helps in understanding the passenger's difficulties based on negative or positive reviews using three classification models. While comparing with other two models SVM and RFC, the IDNN achieves high rate of recall because it has the ability to handle imbalance dataset, where the support vector machine and random forest tree are failing to achieve it.

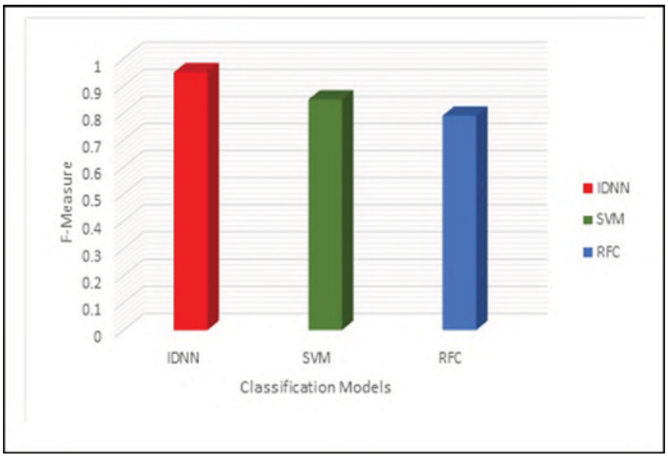

Fig. 8. Comparison Results of three sentiment analysis classification based on F-Measure

The result of Fig. 8. shows that the developed model IDNN produces high rate of F-measure in classification of opinion about the airline services. As the f-measure is produced with the harmonic measure of precision and recall they also reflect the same impact of performance comparison.

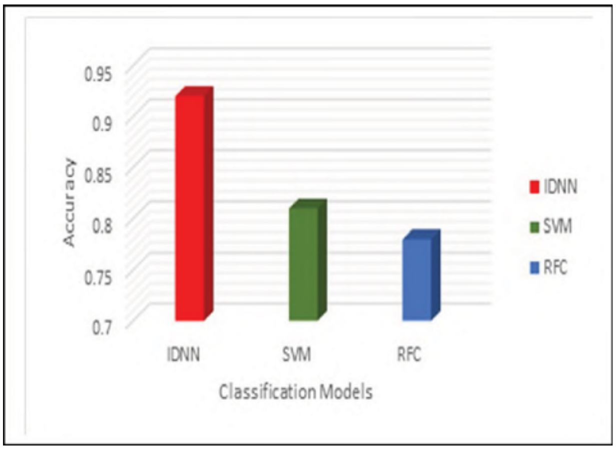

Fig. 9. Comparison Results of three sentiment analysis classification based on Accuracy

The outcome of the Fig. 9. explores that the accuracy of IDNN is higher than the other two models SVM and RBF. As the IDNN is fine-tuned using the fuzzy control measures in assignment of the weights on hidden nodes. So that airline tweet big data is handled more prominently by the IDNN, whereas SVM and RBF can able to work fine on small size datasets alone.

Conclusion

In big data computing sentimental analysis is a recent trend to recognize the needs of mass community. This paper considered the sentiment of passengers toward the service of airlines with their tweet dataset and the fuzzy induced deep neural network reveals the feeling of public about it. As the deep neural network inherit the characteristic of handling big data, this work achieves better results with certain refinement by applying fuzzy control system to emphasis the weight assignment of hidden nodes more effectively. Chaotic particle swarm optimization enhances the fuzzy control by applying chaotic nature to fine tunes the rules generated by the classifier. The analysis of performance shows that while using SVM and RFC, they can able handle the balanced data with reduced dataset size, whereas the deep neural network handles the imbalance dataset with its deep architecture and thus it produces high accuracy compared to these two models. The outcome of the results can greatly help the airlines to interpret the data and may try to improve their quality of service based on the aspect of the negative tweets in the big data paradigm.

REFERENCES

[1] T. Joachims. Making large-scale support vector machine learning practical. In B. Scholkopf, C. J. C. Burges, and A. J. Smola, editors, Advances in kernel methods: support vector learning, pages 169–184. MIT Press, Cambridge, MA, USA, 1999.

[2] D. S. Rajput, Neelu Khare, "FSS Decision Making Model for Social Networking Sites", International Journal of Social Network Mining, Inderscience, Vol 2(3) pp. 256-266, 2016.

[3] D. S. Rajput, Praveen K. Reddy, D. P. Shrivastava, "Mining Frequent term set for Web Document Data using Genetic Algorithm", Published in International Journal of Pharmacy and Technology, Vol. 8 (2), pp. 4038-4054 2016.

[4] Sindhwani, V., Melville, P.: Document-word co-regularization for semi-supervised sentiment analysis. In: IEEE International Conference on Data Mining, pp. 1025–1030 (2008).

[5] Hinton G.E, Osindero S, Teh Y W, A fast learning algorithm for deep belief nets, Neural Computing, 18, 1527–1554 (2006).

[6] Yassine Al Amrani, Mohamed Lazaar, Kamal Eddine El Kadiri, Random Forest and Support Vector Machine based Hybrid Approach to Sentiment Analysis, The First International Conference on Intelligent Computing in Data Sciences, Science Direct, Procedia Computer Science 127 (2018) 511–520.

[7] D. Devikanniga, A. Ramu, and A. Haldorai, "Efficient Diagnosis of Liver Disease using Support Vector Machine Optimized with Crows Search Algorithm," EAI Endorsed Transactions on Energy Web, p. 164177, Jul. 2018. doi:10.4108/eai.13-7-2018.164177

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

[9] X. Duan, T. He, Le Song, Research on Sentiment Classification of Blog Based on Using Deep Learning Methods 978-1-4244-6899-7/1 0 , 2010, Ieee.

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

[11] Won Young Kin, Joon Suk Ryu, Kyu Il Kim, Ung Mo Kim, A Method for Opinion Mining of Product Reviews using Association Rules ICIS 2009, November 24-26, 2009.

[12] Wu, F., Song, Y., Huang, Y.: Microblog sentiment classification with contextual knowledge regularization. In: Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence, pp. 2332–2338 (2015).

[13] Zhou, S., Chen, Q., Wang, X.: Active deep networks for semi-supervised sentiment classification. In: International Conference on Computational Linguistics, Poster, pp. 1515–1523 (2010).

[14] https://www.kaggle.com/crowdflower/twitter-airline-sentiment.

[15] Hu, M., Liu, B.: Mining and summarizing customer reviews. In: Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 168–177 (2004).

[16] B. Liu, L. Wang, Y.-H. Jin, F. Tang, and D.-X. Huang, "Improved particle swarm optimization combined with chaos," Chaos, Solitons & Fractals, vol. 25, no. 5, pp. 1261–1271, 2005.