ISSN: 0974-5823 Vol. 7 No. 2 February, 2022

International Journal of Mechanical Engineering

Sentence Level Sentiment Analysis using Natural Language Processing, Classification with Deep Neural Network

[1]M.Inbavel, [2]Dr. P. Sivaprakasam

[1] Ph.D Research Scholar, Department of Computer Science, Sri Vasavi College, Erode, Tamilnadu, India

[2]Associate Professor, Department of Computer Science, Sri Vasavi College, Erode, Tamilnadu, India

ABSTRACT

Natural Language Processing (NLP), a field of Computer Science, Human language and Artificial Intelligence which is used by machines to understand, analyses, manipulate and interpret human's languages.It is used to perform tasks like translation, automatic summarization, speech recognition, relationship extraction, sentiment analysis etc. Among the tasks, sentiment analysis aid to monitor brand and product sentiment with customer feedback, user opinion in any type of social media.With the integration of NLP, the user text is tagged as positive, negative and neutral and classified using machine learning techniques. This research work discusses about some NLP tasks and implements improved multilayer feedforward deep neural network for the classification of sentiments. The main goal is to generate word cloud for each sentiments using text processing and to improve the accuracy of the model. The results are assessed with suitable evaluation metric sin Rapidminer tool.

Keywords –*Data Mining,Deep Learning,Dropouts, Natural Language Processing, Neural Networks.*

I. Introduction

1.1 Sentiment analysis

The world is almost fully depends in online platform and in this [1] platform reviewing products before buying is a common scenario. This analysis extracts subjective information in source material, and helps business to understand the social sentiment of their brand, product or service. Analyzing the data from these customer reviews make the data more dynamic and aid in improving business. However, reading thousands of reviews to understand about a product/service is time consuming and hence polarizing these reviews makes to understand its popularity among the buyers all over the world. People also value the experience of others based on review. Any online item with large amount of positive reviews makes trustworthy and in other hand, negative reviews make distrust and often cause sales loss. Opinions or views collected from user's experiences regarding specific products or topics straightforwardly have a huge impact on future customer purchase. In online businesses, a product may have thousands of reviews and it is hard to go through all those reviews.

1.1.1 Levels in Sentiment Analysis

There are three levels namely document level, sentence level and aspect level.

1) Document level - The document level categorizes the entire document into different sentiment such as a positive, negative or neutral. Document level classification works best when the document is written by unique user and expresses their opinion on a single entity.

Example :"I bought an iPhone a few days ago. It is such a nice phone. The touch screen is cool and the voice quality is clear too and so I simply love it!"

2) Sentence level - This type is used for reviews that contain one sentence that is written by the user.

It involves two steps:

- Subjectivity classification of a sentence into one of two classes namely objective and subjective.
- Classification of subjective sentences into one of two classes namely positive and negative.

Example: "IPhone sales are good even in this bad economy."

3) Aspect level –This level identifies aspects of given target entities and estimate the sentiment polarity for each mentioned aspect. This has two process aspect extraction and aspect sentiment classification. For aspect extraction, aspect terms present in a sentence is found by having highly frequent phrases across reviews.

Example: "The food was great, but the service was slow."

Aspects: food, service

1.1.2 Applications of Sentiment Analysis

Sentiment analysis has a number of different uses[2] mostly on social media, online business where the rise in review, ratings. Now-a days the companies are becoming increasingly interested in sentiment analysis because online opinions have become a valuable currency for businesses and companies. They filter out this valuable information (sentiments) in order to better understand consumers' view, take more effective and better-targeted action.

Some of the many different applications for sentiment analysis include:

1) Social Media Monitoring- Companies use automated sentiment analysis based on word lists in which each word being given a pre-defined sentiment value and the value of the user text is determined based on the words it contains. This kind of monitoring is an excellent way to better identify the influencers and promoters. For an instance, a restaurant might engage in social media monitoring in order to get how customer feel about their menu and what feelings people associated with their overall experience at the restaurant.

2) Public Relations- Sentiment analysis can aid companies develop and refine their public relations strategy. They use sentiment analysis to identify sales leads and spot industry trends. Sentiment analysis can also be used to identify influencers in a industry with positive sentiments.

3) Marketing: Companies are increasingly using the valuable information for their products with the online reviews instead of conducting a survey that is time consuming too. They can

analyze the content of these reviews to gain insight into the emotions of consumers toward the product or they can also analyze the comments to figure out how much knowledge that the customers have about their product.

4) Data Mining - Sentiment analysis can be used for data mining by gathering competitive intelligence about the competitors as the brand could easily track social media mentions. Hence it is an excellent way to gain a competitive edge even in today's highly competitive marketplace.

5) Political Analysis - Studies of sentient analysis on tweets and micro blogs will accurately indicate political sentiment. It is used to track political opinions. It detects consistency between the stated, actual preferences of politicians and bridge the gap between the politician and the people.

1.2 Natural Language Processing (NLP)

Natural Language Processing (NLP) is a discipline of artificial intelligence in which computers can analyze, understand, and interpret human language in a smart way. By using NLP, developers can perform tasks such as automatic summarization, relationship extraction,named entity recognition, translation, speech recognition, sentiment analysis also topic segmentation. NLP algorithms are typically based on machine learning algorithms as it use those algorithms to automatically learn rules by analyzing a set of examples and makes a statistical inference.

1.2.1 Phases in NLP

The two main phases in natural language processing are data preprocesing and algorithm development in which the former involves preparing, cleaning text data that an algorithm can work with. The preprocessing tasks includes,Tokenization, Stop word removal, Lemmatization, Stemming, Part-of-speech tagging.The two commonly usedalgorithms for NLP tasks are, 1) Rules-based systemthatuses linguistic rules and such as phonetics, phonology, morphology, syntax, and semantics, 2) Machine learning-based system that uses statistical methods to perform tasks based on training mostly has data deep learning and neural networks.

1.3 NLP in sentiment analysis

Sentiment analysis is a field of NLP used to identify the feeling, opinion, or belief such as from very negative, to neutral, to very positive. Most often, humans, communicate with each other in a variety of languages to express ourselves and whatever we say has a sentiment associated with it. It might be positive or negative or it might be neutral. For example, there is a fast-food seller in online with different food items like burgers, pizza, sandwiches, milkshakes, etc. On getting the product, the customer can provide reviews as follows,

- User Review 1: I love this sandwich, it's so delicious – Positive Tag
- User Review 2: This burger has a very bad taste – Negative Tag
- User Review 3: I ordered this pizza today but it is average – Neutral Tag

1.4 Problem Definition

Onlooking userreviews, the company can able to improve the quality of servicesfor increasing their overall sales. But, if there are hundreds and thousands of user reviews,it is impossible to go through all the reviews manually. Hence,Sentiment Analysis Model comes into play that takes huge corpus of data having user reviews and finds a pattern and comes up with a conclusion. Machine learning technique must be employed to classify the tags with high accuracy in less processing time. Also the frequency of terms used for review should be analyzed with NLP tasks.

1.5 Objective

- To generate word cloud with high occurrence of terms using NLP tasks.
- To improve the accuracy in the existing model.

▪ To lessen the processing time.

The remainder of this paper is organized as follows: In section one, Sentiment Analysis, Natural Language Processing, Problem definition and the objective is explained in detail, the related works regarding sentiment analysis was discussed in the section 2, in section 3, methodology is described and the proposed methodology is proved with the result in section 4, section 5 conclude the research work and discuss some future enhancement.

2. Literature Review

Sasikala *et al* [3] designed a sentiment analysis of online products review with an Improved Adaptive Neuro-Fuzzy Inferences System. Generally, user's opinion, belief and feelings about a product is captured through the review text which is then labeled as positive, neutral or negative sentiments. This leads to have changes on the product for better customer retaining by the companies. This work applied a Deep learning modified neural.The data values are separated into Contents-based, Grades-based along with Collaborations based and the proposed work carried out a weighting factor and classification on the product for future prediction. The results showed Collaboration based deep learning gave better results.

Pandian *et al* [4]delivered a better performing and efficient automated feature extraction technique. The traditional methodologies use the complicated manual feature extraction process but this work put forth a strong baseline to determine the predictability of the features and integrates the deep learning techniques with feature-extraction. Moreover, this research work includes three steps. The first is the development of sentiment classifiers with deep learning for comparing the performance. The second is the use of ensemble techniques and information merger to obtain the final set of sources and the third step, a combination of ensembles is introduced to categorize various models along with the proposed model. Finally experimental analysis is carried out and the to determine the best model with the deep learning baseline.

Harleen Kaur *et al* [5] analyzed Twitter data based on COVID-19 with hash tag keywords. The most dominant tool is social media and so many post their opinions on this case. This work analyzed regarding COVID-19 and classified them positive, negative and neutral sentiment scores. Supervised machine learning algorithms such as SVM with heterogeneous euclidean overlap metric, RNN and Deep learning which consists of tens or even hundreds of layers are implemented and the results are compared. From the result it is revealed, the hybrid model RNN with deep learning gave better results for sentiment classification.

Bashar *et al* [6]presented a survey related to Natural language Processing and deep learning. The deep learning being a subcategory of the machine learning produce accurate results. It can directly classify the tasks from the documents available either in the form of the text, image, or the sound by utilizing the neural network to perform the accurate classification. One of the most common deep neural networks is the convolution neural network that provides an automated way of feature extraction unlike the other techniques that extracts the features manually. This work listed the survey on the deep learning neural network architectures utilized in various applications with the technology and its features separately for deep learning and Natural language processing.

Araque *et al*7]improved the performance of deep learning techniques with traditional surface approaches based on manually extracted features. Deep learning provides automatic feature extraction but traditional surface approaches are based on complex manually extracted features. This paper applied six fold method. Initially, a deep learning based sentiment classifier using a word embeddings model and a linear machine learning algorithm was developed. Next, two ensemble techniques which aggregate baseline classifier with other surface classifiers was used . Third, two models was proposed for combining both surface and deep features to merge information. Fourth, a taxonomy for classifying the different models was introduced. Fifth, several experiments were conducted to compare the performance of the models. Seven public datasets were extracted from the micro blogging and movie reviews domain. Finally, a statistical study confirmed that the performance of these proposed models surpasses that of our original baseline.

3. Research Methodology

3.1 Natural Language Processing

The two main analyses are Syntax and semantic analysis.

Syntax techniques includes the following,

▪ Parsing- The grammatical analysis of a sentence. Parsing involves breaking the sentence into parts of speech such as noun, verb and this is useful in handling more complex processing task.
▪ Word segmentation – It takes string of text and deriving word forms from it by considering white spaces.
▪ Sentence breaking - This places sentence boundaries in large texts by splitting two sentences with the period (full stops).
▪ Morphological segmentation - Divides a single word into smaller parts called morphemes. For an instance “untestable-‘un’, ‘test’, able,” and this useful in machine translation and speech recognition.
▪ Stemming – Divide the word withits root forms.Example: Barked –bark.

Semantics techniques includes the following,

- Word sense disambiguation.-Derives the meaning of a word based on context.

Example: "The pig is in the pen." The word pen has different meanings. Here, it means a fenced area where the pig is, not a writing tool.

- Named entity recognition - This determines words that can be categorized into groups even with same spelling.

Example: "McDonald's son went to McDonald's and ordered a Happy Meal," the algorithm recognize "McDonald's" as two separate entities -- one a restaurant and one a person.

- Natural language generation - Use database to determine semantics behind words and generate new text.

Example: Automatically write a summary of findings from a business intelligence platform, maps certain words and phrases to features of the data, generate news articles or tweets based on a certain body of text based on training set.

3.2 Artificial Neural Networks (ANN)

Artificial neural networks (ANNs) are considered as a [8] computing systems that are inspired by the biological neural networks with a collection of connected nodes called artificial neurons, which resembles the neurons in a biological brain and each one transmit a signal to other neurons. At once, the neuron receives a signal, processes it and can pass to the neurons connected to it. The output of each neuron is computed by some non-linear function of the sum of its inputs. Neurons typically have a weight that increases or decreases the strength adjusts while the learning proceeds. Signals travel from the first layer (the input layer), to the last layer (the output layer), through the intermediate hidden layer.

3.2.1 Types of Artificial Neural Network

The seven most important neural networks are as follows:

3.2.1.1 Modular Neural Networks (MNN)

This neural network is characterized by a series of independent neural networks that is moderated by some intermediate network. Each independent network serves as a module, operates on separate inputs to accomplish subtask. The intermediate one takes the outputs of each module, processes them to produce the output of the network but it only accepts the modules' outputs, it does not respond also the modules do not interact with each other. Unlike a single large network which is assigned to arbitrary tasks here each module in a modular network must be assigned a specific task and connected to other modules in specific way to perform the tasks.

Applications: One of the fastest-growing areas in Artificial Intelligence.

3.2.1.2 Feedforward Neural Network – Artificial Neuron (MLFF)

The information in the network travels in one direction that is consists of input layer, hidden layer, output layer and data enter through input nodes and exit through output nodes. Activation function logistic or sigmoid is used to map the input to the output without feedback or loops. The simplest kind of neural network is a single-layer perceptron network in which the inputs are fed directly to the outputs via weights. The sum of the products of the weights and the inputs is calculated and if the value is above some threshold the neuron fires and takes the activated value otherwise it takes the deactivated value. A simple learning algorithm called delta rule that calculates the errors between calculated output and sample output trains it and the weights are adjusted with gradient descent. In contrary, Multi layer perceptron consists of multiple layers of computational units that is interconnected in a feed-forward way. Sigmoid function is used as an activation function. Back propagation is used for learning techniques. Loss function is used to compare the output values, after repeating this process for a sufficiently large number of training cycles, the network will usually converge To adjust weights properly, gradient descent is used.

Applications: Speech recognition, Computer vision.

3.2.1.3 Radial basis function network (RBF)

Typically have three layers: an input layer, a hidden layer with a non-linear RBF activation function and a linear output layer. It consider the distance of a centre with respect to the point. In the first layer, features in the inner layer are united with the Radial Basis Function and in the upcoming layer the output from the previous layer is considered for computing the same output in the next iteration.

Applications: Power Restoration Systems.

3.2.1.4 Kohonen Self Organizing Neural Network (SOM)

Vectors are input to a discrete map from an arbitrary dimension with one or two dimensions. The weight of the neurons may change depends on the value but the neuron’s location will not change. Input vector and small weight are given to every neuron in the first phase. The neuron with a closest point is considered as a winning neuron and it will have the least distance most often euclidean distance is used to calculate the distance between neurons and the point. Other neurons will also start to move along with the winning neuron in the second phase. Each neuron belongs to each kind of cluster, and through the iterations, the clustering of all the points.

Application: Medical analysis to classify diseases with higher accuracy.

3.2.1.5 Recurrent Neural Network (RNN)

The principle is to feedback the output of a layer back to the input again that helps to predict the outcome of the layer. Each neuron will act as a memory cell and the neuron will retain some information as it goes to the next time step. The prediction will improve by error correction with some changes to create the right prediction output. The learning rate will can make the correct prediction from the wrong prediction faster.

Application: Converting text to speech.

3.2.1.6 Convolutional Neural Network (CNN)

Biases and weights are given to the neurons initially. The photos are recognized by taking the input features in batch-wise and during image is converted to Grayscale from HSI or RGB. Edges are detected by finding out the pixel value. The classification of images is done into various categories. For image classification, this Neural Networks have a very high level of accuracy. Their inputs and outputs are masked by the activation function and the convolution. This includes a layer that performs a dot product (Frobenius inner product) of the convolution kernel with the layer's input matrix. Its activation function is commonly ReLU. The feature map and this map contributes to the input of the next layer, followed by pooling layers, fully connected, normalization layers.

Applications: Prediction of yield and growth in the future of a land area, Image processing and signal processing.

3.2.1.7 Long / Short Term Memory (LSTM)

This unit is composed of a cell, an input gate, an output gate and a forget gate. The cell remembers values over time intervals and the other three gates regulate the flow of information into and out of the cell. The main goal is to remember things/values for a long time in a memory cell that is explicitly defined and this values are stored in the memory cell unless told to forget by the defined "forget gate". New stuff is added through the "input gate" to the memory cell, and it is passed to the next hidden state from the cell along the vectors which is decided by the "output gate".

Applications: Composition of primitive music, Writing, Learning complex sequences.

3.3 Deep Learning or Deep Neural Network (DNN)

Deep learning is part of a machine learning based on [9] artificial neural networks with supervised, semi supervised or unsupervised. Deep learning refers to the use of multiple layers in the network with an unbounded number of layers with bounded size. In deep learning, each level transforms its input data into a slightly more abstract and composite representation and remove redundancy in representation. This kind of networks is interpreted in terms of the universal approximation theorem or probabilistic inference. This probabilistic inference led to the introduction of dropout as regularizer. There are different types of networks they always consist of the same components namely neurons, synapses, weights, biases, and functions. They are typically feedforward networks in which data flows from the input layer to the output layer without looping backwards.

Neural network architectures in DNN

The most important architectures includes:Recurrent neural networks, Long short-term memory, Convolutional deep neural networks and Multilayer feedforward neural network.

3.3 Proposed Architecture - Improved Deep Neural Network with Multilayer Feedforwardmodel (IDNN-MLFF)

Working principle

Initially, multilayer feedforward network in DNN creates a map of virtual neurons and assigns random numerical weights, to connections between them.Then, weights, bias and inputs are multiplied (equation 1) with activation function (equation 2) and return an output between zero and one with equation 3. Weight is assigned initially here it is applied as 0.3 and the bias is always one.

$$\sum(\text{weight} * \text{input}) + \text{bias} \quad \text{--- (1)}$$

$$\text{Jactivation function: } f(x) = \max(0, x) \quad \text{--- (2)}$$

$$\text{ut} = f(x) * (\text{weight} * \text{input}) + \text{bias} \quad \text{--- (3)}$$

If the network did not accurately recognize a particular pattern, which is find by loss function as in equation 4, an algorithm would adjust the weights by back propagation with gradient descent in equation 8.

$$\text{Huber Loss Function: } L(x) = \{\frac{1}{2}(y - f(x))^2; \text{ for } y - f(x)| \leq \delta$$

$$- f(x)| - \frac{1}{2}\delta^2; \ otherwise \quad \text{--- (4)}$$

The eq. 4 says, if the loss value is less than delta then Mean Squared error is used as in equation 5 otherwise Mean Absolute Error is used as in equation 6. In equation 4, delta is the hyper parameter to define the range of MAE and MSE and here it is defined as 1.

$$\text{n Squared Error (MSE)} = \frac{1}{N}\sum_{i=1}^{N}(y_i - \bar{y}_i)^2 \quad \text{--- (5)}$$

$$\text{n Absolute Error(MAE)} = \frac{1}{N}\sum_{i=1}^{N}|y_i - \bar{y}_i|^2 \quad \text{--- (6)}$$

L1 Regularization is added with loss function to avoid overfitting as in equation 7,

$$orm = \text{loss function} + \lambda \sum_{i=1}^{n} |w_i| \quad \text{--- (7)}$$

Stochastic Gradient descent is calculated as,

The magnitude and direction of the weight updateis computed by taking a step in the opposite direction of the cost gradient.

$$= -\eta \frac{\partial L}{\partial w_j} \quad \text{--- (8)}$$

Where 'ŋ' is the learning rate.

The 'Δw' has the weight updates of each weight w, which are computed as in equation 9 and followed by equation 10,

$$\text{Delta rule: } \Delta w_j = -\eta \frac{\partial L}{\partial w_j}$$

$$\sum_i (target^i - output^i)(x_j^i) \quad \text{--- (9)}$$

Weight updates in each epoch is,

$$= w_j + \Delta w_j \quad \text{--- (10)}$$

Hence, this deep learning influence more parameters such as training parameters, number of layers and number of units per layer, learning rate, initial weights until it determines the correct mathematical manipulation.

3.3.1 Existing Deep Neural Network with Multilayer Feedforward model (DNN-MLFF)

The existing work [10] use the parameters listed in TABLE1.

Table 1. Existing work (DNN-MLFF) parameters

Dataset used	Twitter sentiment dataset
Label	Binomial - Positive, Negative Tags
Training, Testing set Ratio	50:50
Number of hidden layer	3
Number of neurons	100
Dropouts	None
Activation function	ReLU
Loss function	Mean square Error (MSE)
Learning	Stochastic gradient descent
Regularization	None
Learning rate	0.1
Epochs	100

3.3.2 Proposed Improved Deep Neural Network or deep learning with Multilayer Feedforward Neural network (IDNN-MLFF)

The proposed work use the parameters listed in TABLE 2.

Table 2. Proposed work (IDNN-MLFF) parameters

Dataset used	Hotel Review sentiment dataset
Label	Polynomial - Positive, Negative, Neutral Tags
Training, Testing set Ratio	60:40
Number of hidden layer	2
Number of neurons	25
Dropouts	0.25
Activation function	ReLU
Weight	0.3

Bias	1
Loss function	Huber
Learning	Stochastic gradient descent
Regularization	L1 regularization
Learning rate	0.01
Epochs	50

Proposed Algorithm IDNN-MLFF

Step 1 Initialization of network with input layer, number of hidden layer, dropout ratio, initial weight and bias, activation function, loss function, number of epochs.

Step 2 Weight and bias multiplication with the input using equation 1.

Step 3 Calculate the output, apply activation function using equation 3.

Step 4 Calculation of loss function using equation 4-6

Step 5 Apply L1 regularization using equation 7

Step 6 Weights updation to minimize the objective function with gradient descent by backpropagation using equation 8-10.

Step 7 Repeat step 2-6 for *n* number of epoch.

Step 8 Assess the performance for each iteration.

Advantages of IDNN-MLFF

- Improved accuracy by avoiding overfitting with dropouts, L1 regularization.
- Minimized processing time due to optimal initialization of hidden layer size, neurons, dropout ratio epochs.
- Word cloud is generated with frequent terms.

4 Result and Discussion

4.1 Dataset : The dataset regarding Hotel industry is taken from Kaggle repository that includes the attributes Hotel name(categorical), Review Title(categorical), Review Text(categorical), Review percentage(numerical), Sentiments(numerical : 1-negative, 2- neutral, 3-positive).

4.2 Natural language Processing (NLP) Tasks : Filtering stopwords, Stemming, Tokenization and n-grams creation for each sentiments seperately to generate word cloud with the sentiments Negative, neutral and Positive as shown in Fig 1.

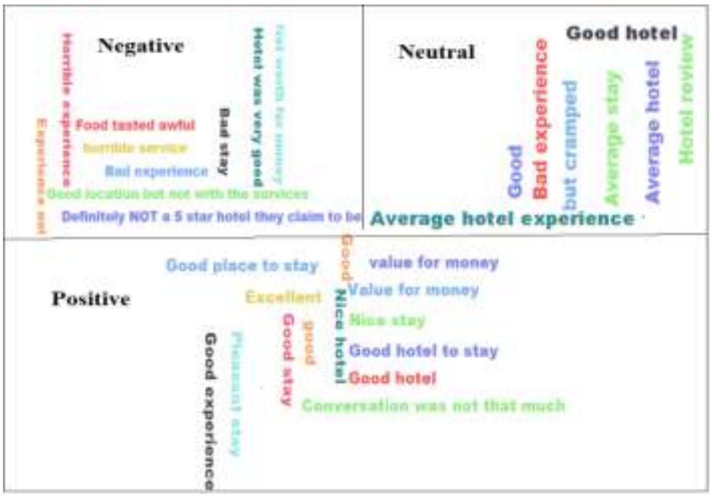

Figure 1. word cloud for three sentiments negative, neutral, positive

4.3 Confusion matrix

Confusion matrix is generated with the classified data with true positives, true negatives, false postives and false negatives. The label has three outputs (positive, negative, neutral) and so 3 X 3 matrix is formed. The diagonal elements implies correctly classified data and other entries are wrongly classified data. This matrix is used to calculate accuracy level for performance analysis in equation 11.

4.4 Performance Analysis

The performance of the existing and proposed method is assesed in Rapid miner tool with three variant evaluation metrics accuracy, processing time, kappa statistics and listed in TABLE 3. Accuracy is noted in percentage, processing time is noted in minutes and the kappa statistics is noted in range between 0 and 1 in which the valuesnearer to 1 implies a good model.

$$\text{racy} = TP + TN/\ TP + TN + FP + FN \qquad \text{--- (11)}$$

Where, TP is True positive, TN is True negative, FP is False positive, FN is False negative.

$$\text{a Statistics} = \frac{p_o - p_e}{1 - p_e} \qquad \text{--- (12)}$$

Where, p_o is observed agreement and p_e is hypothetical probability of chance agreement.

Table 3. Performance analysis

Evaluation Measure	**Existing DNN-MLFF**	**Proposed IDNN-MLFF**
Accuracy (in %)	83.67	92.45
Processing time (in min.)	2.01	1.12
Kappa statistics (in range)	0.812	0.913

In TABLE 3, the accuracy for the proposed improved deep neural multi feedforward network IDNN-MLFF is 92.45% and it higher than the existing deep neural multi feedforward network DNN-MLFF which is 83.67%. The processing time is noted in minutes and it is less for IDNN-MLFF i.e 1.12 minutes but DNN-MLFF takes 2.01 minutes. Kappa statistics generally is between the range 0 and 1 in which the values nearer to one indicates a good model and so IDNN-MLFF proves to be a good model with 0.913.

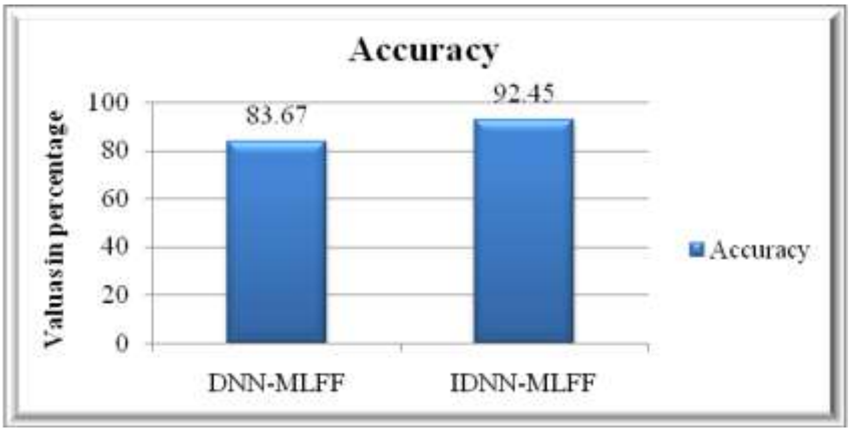

Figure 2. Accuracy of DNN-MLFF, IDNN-MLFF

In Fig 2, IDNN-MLFF has high accuracy as 92.45 % than the existing as the proposed model applies L1 regularization, dropouts that avoidoverfitting and increase the accuracy level.

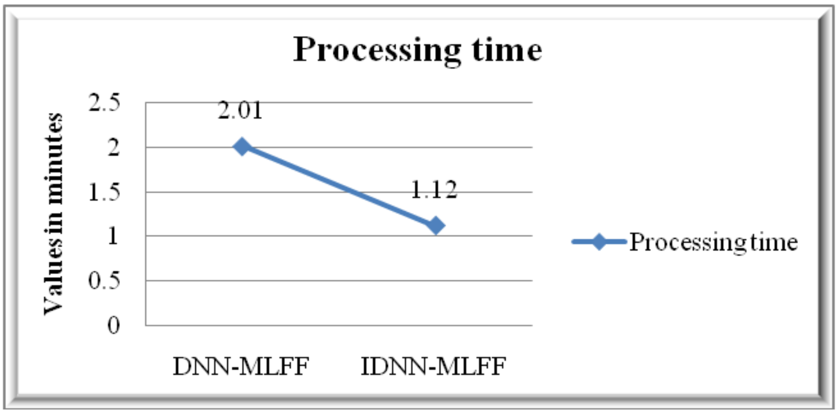

Figure 3. Processing Time of DNN-MLFF, IDNN-MLFF

In Fig 3, the processing time is less for IDNN-MLFF i.e 1.12 minutes because the learning rate 0.01 which is smaller than the existing that makes the learning faster, dropouts with the ratio 0.25, less number of epoch. Also, the hidden layer size, neurons are 2, 0.25 that are less than DNN-MLFF Hence, leads to less processing time.

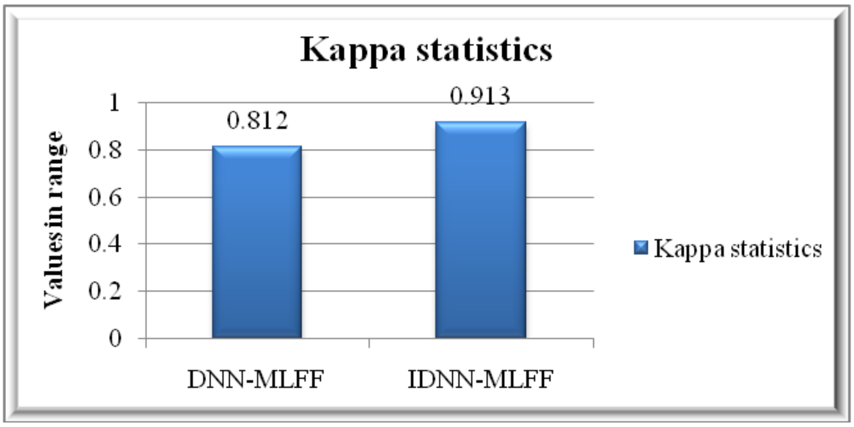

Figure 4. Kappa Statistics of DNN-MLFF, IDNN-MLFF

In Fig 4, High range 0.913 in IDNN-MLFF shows that it is a good model due to the Huber loss function which optimizes the parameter in the network and the proposed parameter values learning rate, dropout ratio influence the model to work better than the existing DNN-MLFF.

5 Conclusion and Future work

Sentiment analysis is an emerging research area at present as it helps in the improvement of sectors in online; while this analysis is automated with machine learning techniques it adds better analysis. This work analyzes deep neural network or deep learning (DNN) with multi feedforward (MLFF) architecture. The existing model DNN-MLFF is improved with the parameters namely dropouts, l1 regularization, Huber loss function, also less number of hidden layers, less number of neurons in each layer, lessen of epochs. These additional parameters in the proposed model IDNN-MLFF influence the network to work better and lead to high accuracy with 92.45%, high range of kappa statistics as 0.913 in less processing time 1.12 minutes.In future, this work can be extended in order to have high accuracy than this model. For that, loss functions such as quadratic, hinge may be tried. The number of hidden layer and neurons can be altered to assess the performance.

References

[1] Walter Kasper, Mihaela Vela, “Sentiment Analysis for Hotel Reviews”, Proceedings of the Computational Linguistics-Applications Conference, Research gate publication, 2012.

[2] SiewTheng Lai, MafasRaheem, “Sentiment Analysis of Online Customer Reviews For Hotel Industry: An Appraisal Of Hybrid Approach”, International Research Journal of Engineering and Technology (IRJET), Volume: 07 Issue: 12 | Dec 2020.

[3] Sasikala .P, Mary Immaculate Sheela .L, “Sentiment analysis of online product reviews using DLMNN and future prediction of online product using IANFIS”, Journal of Bigdata, 2020.

[4] Pandian, A. Pasumpon. “Performance Evaluation and Comparison using Deep Learning Techniques in Sentiment Analysis.” Journal of Soft Computing Paradigm 3, no. 2: 123-134.

[5] HarleenKaur, ShafqatUlAhsaan, BhavyaAlankar, Victor Chang, “A Proposed Sentiment Analysis Deep Learning algorithm for Analyzing COVID-19 Tweets”, Information Systems Frontiers, Springer, 2021.

[6] Bashar, Abul. “Survey on evolving deep learning neural network architectures.” Journal of Artificial Intelligence 1, no. 02, 2019, 73-82.

[7] Araque O., Corcuera-PlatasI.,. Sánchez-Rada J. F, Iglesias C. A., Enhancing deep learning sentiment analysis with ensemble techniques in social applications, Expert Syst. Appl. 77, 2017.

[8] ZenunKastrati, LuleAhmedi, ArianitKurti , FatbardhKadriu, DoruntinaMurtezaj, FatbardhGashi, “Deep Learning Sentiment Analyser for Social Media Comments in Low-Resource Languages”, Electronics 2021.

[9] YashIndulkar, AbhijitPatil, “Sentiment Analysis of Uber& Ola using Deep Learning”, Proceedings of the International Conference on Smart Electronics and Communication (ICOSEC 2020) IEEE Xplore, 2020.DOI: 10.1109/ICOSEC49089.2020.9215429

[10] AdyanMarendraRamadhani, Hong Soon Goo, “Twitter Sentiment Analysis using Deep Learning Methods”, 2017 7th International Annual Engineering Seminar (InAES), Yogyakarta, Indonesia, IEEE Publication, 2017, DOI: 10.1109/INAES.2017.8068556.