



An Improved Framework for Sentiment Analysis for College Reviews

T. Sri Devi¹, R. Dhanalakshmi², S. Sankar³

^{1,2,3} KCG College of Technology, Tamil Nadu, India,

sridevi.thangapandian@yahoo.com, dhanalakshmisai@gmail.com, sankar@kcgcollege.com

ABSTRACT

This is the data era and large amount of data is processed every second for various purposes and deep learning is used to abstract the key aspects of data. Lack of core knowledge on deep learning makes this a Black Box machine which remains dark to the real world. In the recent years many researches are going in the field of Machine learning in which Deep Learning has gained more attention. The aim of the paper is to provide survey of trending deep learning algorithms. The algorithms such as Deep Neural Network (DNN), Convolutional Neural Network (CNN) and Recurrent Neural Network (RNN) are discussed in detail. Then, performance of RNN-LSTM algorithm is compared by training and predicting college reviews feedback of data set size 500 with epochs 1 to 7. It is identified that accuracy of RNN-LSTM algorithm was around 83%. This paper provides better understanding of the deep learning algorithms and their implementation in various research areas.

Key words: Deep learning, neural networks, deep neural network, convolutional neural network, recurrent neural network, DNN, CNN, RNN

1. INTRODUCTION

Deep Learning is machine learning technique that is designed similar to neurons of the neural network present in the human brain. Basic building blocks of nervous system are neurons [8][12] and each neuron take inputs, process them and produces output. It allows the system to classify and discover patterns from the raw data without being programmed explicitly. Machine Learning evolves by allowing the machine to learn and also improve from related experiences. Deep learning algorithms learn progressively from the simple to complex features from the available data. For example, in image recognition deep learning algorithm understands light and dark areas first, then lines and finally shapes to recognize face. Deep learning has gained importance in recent years and it is used in various fields like speech recognition, image/facial recognition [3], natural language processing (NLP), skin/breast/lung/brain cancer diagnosis [4], predicting natural calamities, market price forecasting etc.

This paper is organized as follows: Section II details various deep learning algorithms. Section III discusses the result achieved as part of implementation of LSTM for training and prediction of college reviews and finally conclusion is discussed in Section IV.

2. DEEP LEARNING ALGORITHM

a. McCulloch and Pitts Neuron

In 1943 McCulloch and Pitts proposed mathematical model of neuron [8]. Refer Figure 1. It consists of

- Weighted inputs w , that matches synapses
- An adder function that sums the input
- Activation function which decides the output of neuron (fire or not to fire) for the given inputs

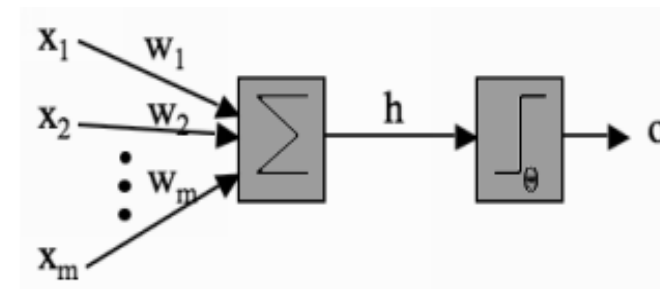


Figure 1: McCulloch and Pitts' Mathematical model of neuron

$$h = \sum_{i=1}^m w_i x_i \quad (1)$$

when h value is ≥ 0 neuron fires and produces output 1.
when h value is < 0 neuron does not fire and produces output 0.
The Activation function decides whether to fire or not and it is represented below

$$o = g(h) = \begin{cases} 1, & h > \theta \\ 0, & h < \theta \end{cases} \quad (2)$$

Where θ is Threshold value.

b. Neural Networks

Perceptron is collection of McCulloch and Pitts neurons with set of inputs and some weights. Perceptron is single layer Neural Network. Multilayer Perceptron is called Neural Networks[8]. Refer Figure2. A neural network consists of

numerous “neurons” which are technically organized as different layers. The simplest networks contain only input and output layers and no hidden layers and this network is equivalent to linear regression model. Figure 2 shows the feed-forward neural network with four input. The coefficients attached to these predictors are called “weights”. The outputs are obtained by a linear combination of the inputs. In the neural network framework weights are updated using ‘learning algorithm’.

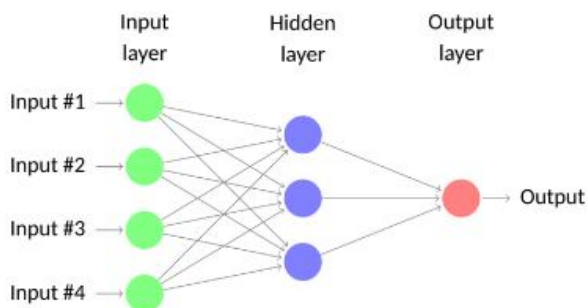


Figure 2: Neural Network

c. Deep Neural Network (DNN)

The Neural network with many hidden layers is the Deep Neural network (Fig 3). It contains multiple layers when compared to shallow network. In DNNs, data flow from input layers to output layers. DNN is designed based on Mammalian visual system that process information sequentially from edge feature, part feature, shape feature and finally abstract [6].

In deep-learning networks, nodes in each layer trains separate set of features depending on the data received from previous layers. On moving further deep into the network, the nodes can recognize even the complex features, since they consolidate features from the previous layer. Higher the layer, learning happens deeper. Various frequently used DNN models are Deep Belief Network (DBN), Sacked Autoencoder (SAE), Deep Convolutional neural networks (DCNN), Restricted Boltzmann Machines(RBMs)[1].

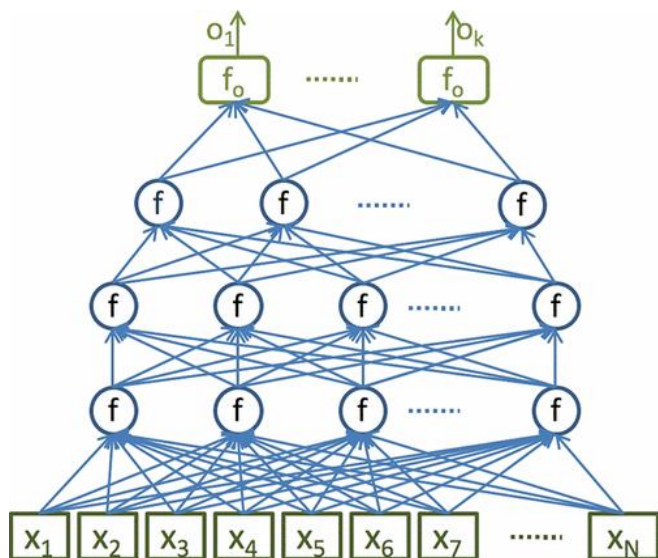


Figure 3: Deep Neural Network

Deep neural networks (DNNs) are computational models implemented successfully in many applications of artificial intelligence (AI), which includes applications like Speech recognition, Face recognition, Satellite orbit prediction [6], computer vision and NLP (Natural Language Processing)[5]. Recently, DNNs have been used in quantitative structure–activity relationship (QSAR) activities. QSAR is largely involved in drug designing which is need of the hour. This technique helps in predicting, reactivity, activity and properties of unknown molecules.

d. Convolutional Neural Network (CNN)

Convolutional neural network is similar to multilayer perceptron network with input layer, output layer and multiple hidden layers. CNN are mostly used in image analysis in medical domain. CNN is fully connected which means each neuron in one layer is connected to all the neurons in next layer. A single layer CNN has 3 phases. They are Convolution, Nonlinear Transformation and Pooling[9]. Refer Figure 4.

Convolution: when input image is passed through convolutional layer, filters in this layer extracts the important features from the image for passing further. Different convolutional kernels are used to extract important features.

Non-Linear Transformation: It is extension of convolution phase. It stripes off unwanted data to provide better feature extraction.

Pooling: Like the Convolutional Layer, the Pooling layer reduces the dimensional size of Convolved Feature. By doing so it reduces the computing power involved in processing the data via. dimensionality reduction. Also, it is helpful in retrieval of significant features. Couple of Pooling methods exist, namely Max and Average Pooling. Max Pooling will return the max value of the specific section of image which is under process by kernel while Average Pooling will return the average values of the specific section of image which is under process by kernel.

Multiple convolutional layers, pooling layers could be added before predicting the output. Adding multiple layers ensure deep learning and helps in arriving unique features of image. Output is produced through output layers which is connected to fully connect neural network. Loss function at output layer helps to calculate error gradient. Fault is propagated back to modify weight and the bias to reduce error. Thus, a training cycle includes a forward pass and a backward pass [1][6][9].

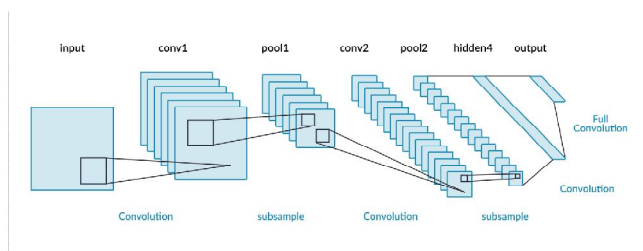


Figure 4: Design of Convolutional Neural network

CNN is applied in Image recognition, Recommender systems and Natural Language Processing, Speech processing[5][11].

e. Recurrent Neural Network (RNN)

RNN are network with loops that allows information to persist. Recurrent Neural network stores information of data during training and with that knowledge it classifies the output. Information flow in Feed forward neural network is unidirectional and it does not have memory, while information is fed back in RNN(Refer Figure 6) and it has memory, as a result Prediction are comparatively good in RNN [1]. This model has the capability to process sequence data, hence each sample is dependent on the previous data. Recurrent neural network is used with convolutional layers to identify the neighborhood effective pixels.

In deeper network, several hidden layers are added. On receiving the input, activations are applied on the first hidden layer and it is passed to the subsequent hidden layers and finally to the output layer for producing the output. Each hidden layer behaves independently and has its own weights and biases. So hidden layers cannot be combined. To combine them together, the same weights and bias needs to be maintained for the hidden layers. The inputs enter and move forward at each time step, in forward propagation. Backpropagation in the feed forward network always move backwards from final error(outputs), weights and inputs of each intermediate hidden layer. Refer Figure 5. During this process it assigns the weights responsibility for the respective error portions. Weights are assigned by calculating the partial derivatives. The learning algorithm (gradient descent) uses the derivatives for adjusting the weights in the direction which decreases error. In backward propagation weight is changed with respect to time and it is called as Back propagation through time (BPTT)[1][13].

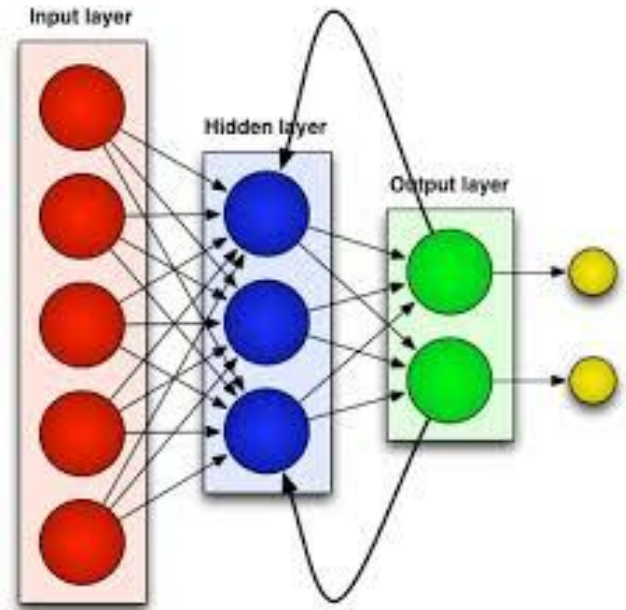


Figure 6: Recurrent Neural Network

RNN had problems of vanishing gradient and exploding gradient[1][11]. Vanishing gradient problem is difficulty identified in training artificial neural network. Normally weight of each neural network gets updated proportional to error function. In some cases, vanishing gradients will be small such that weights remain unchanged. In worst cases it stops the neural network from training. In Exploding gradient, large error gradient gets accumulated, it results in unstable model and unable to learn from data. To overcome the problem of RNN, LSTM (Long Short-Term Memory) was developed which introduced memory unit called cell in the network[10]. Refer Figure 7. LSTM consists of different memory blocks called cells and gates like Forget gate, Input gate and Output gate through which Information passes through.

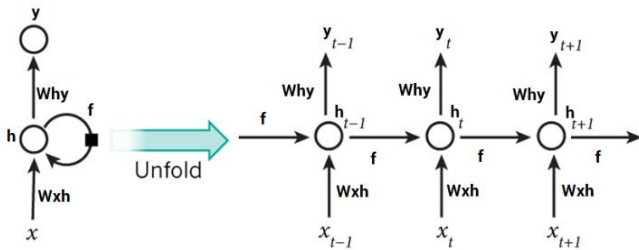


Figure 5: Backpropagation in Recurrent Neural network

In RNN, y_t is considered as predicted value, \bar{y}_t as actual value, the error (cross entropy loss) is calculated as

$$E_t(\bar{y}_t, y_t) = -\bar{y}_t \log(y_t) \tag{3}$$

$$E(\bar{y}, y) = -\sum \bar{y}_t \log(y_t) \tag{4}$$

The steps for back propagation involve computing cross entropy, unrolling network for all time steps, calculate gradient for each time step, then gradients are combined for every time steps. Accordingly, the weights are adjusted for dense layers and recurrent neurons.

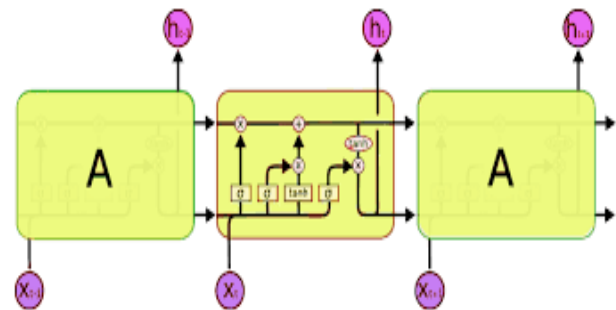


Figure 7: Architecture of LSTM Network

Input Gate retains the input value, while Forget Gate removes the unwanted data, output gate decides the output value. RNN is used in stock market prediction, video classification, sentiment analysis, sequence generation, image classification, speech or voice recognition.

3. LSTM IMPLEMENTATION AND RESULT

As part of implementation, from data collection to sentiment prediction each step was carefully evaluated and designed [2][7][13] as shown in Figure 8.

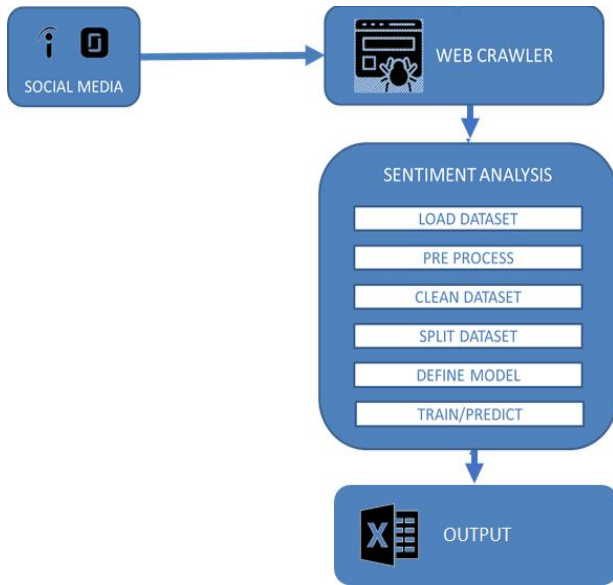


Figure 8: Functional Block diagram of Sentiment Analysis

The student’s feedback of leading engineering colleges in Chennai city is collected through open forums like Google reviews. The collected student’s feedback sentences are labeled as Positive or Negative depending on its polarity. The dataset size is 500. The labeled dataset is trained with RNN-LSTM algorithm. Then sentiment analysis is performed using RNN-LSTM algorithm. The Architecture diagram for sentiment analysis is shown in Figure 9.

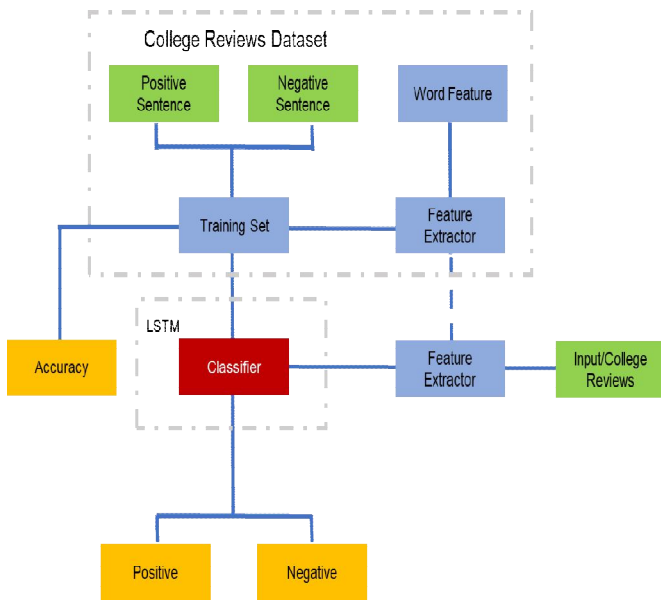


Figure 9: Architecture Diagram for sentiment analysis

In Architecture diagram Figure 9, input is represented in green color, while output is represented in orange color. Sentiment analysis is performed for each college. LSTM model is designed with 500 nodes. 80% of the data is used for training while 20% is used for testing. The Epoch (Number of times the algorithm will work through the entire dataset) is changed from 2 to 7 and performance of LSTM algorithm is captured. Validation Accuracy keeps increasing from 67% to 83% and it is shown in Table 1.

Table 1: Result: Performance of LSTM

S.No	Number of Epoch	Accuracy
1	2	71.00%
2	3	70.00%
3	4	82.00%
4	5	78.00%
5	6	80.00%
6	7	83.00%

The bar graph depicting the performance of LSTM algorithm is shown in Figure 10.

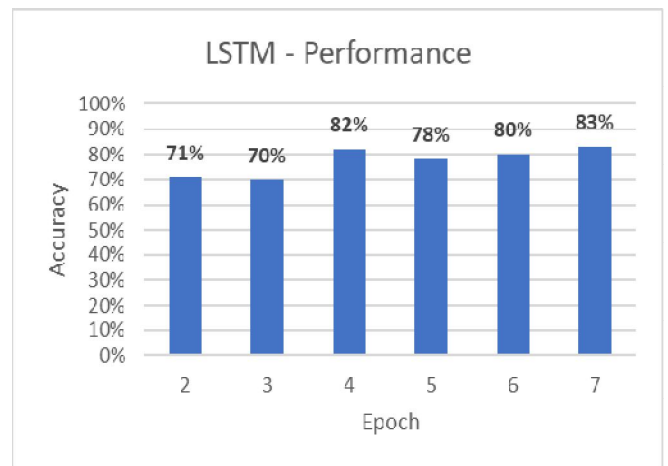


Figure 10: Result: Performance of LSTM with 500 datasets with epoch varying from 2 to 7, accuracy keeps improving from 67% to 83%

300 reviews for 5 colleges are collected. Sentiment analysis is performed for each college using LSTM algorithm. The students review comments are classified as ‘positive’ and negative and it is shown in Table 2.

Table 2: Result: Sentiment analysis of student feedback reviews using LSTM

S.No	Student Feedback	Result
1	Worst College	Negative
2	Very well developed and inspiring institution.	Positive
3	The placements were not up to the mark	Negative
4	no core company came to our campus.	Negative
5	A good college to learn and gain	Positive

	knowledge	
6	It is quite a good campus overall.	Positive
7	There are no medical facilities, no Wi-Fi and no sports.	Negative
8	No direction, guidance or training. It was literally sink or swim	Negative
9	No proper guidance	Negative
10	Good infrastructure	Positive

4. CONCLUSION

In this paper, overview of deep learning algorithms has been provided and exclusively LSTM is used for training and prediction of college reviews. Using LSTM, the college reviews were analyzed as positive and negative and various Epoch parameters were used to understand the performance of LSTM algorithm in training the dataset. As part of future work, dataset and number of Epochs would be increased this would help in improving accuracy and predictions.

REFERENCES

1. A. Shrestha and A. Mahmood, "Review of Deep Learning Algorithms and Architectures," in *IEEE Access*, vol. 7, pp. 53040-53065, 2019.
<https://doi.org/10.1109/ACCESS.2019.2912200>
2. Ahmad, S., Asghar, M.Z., Alotaibi, F.M. et al. Detection and classification of social media-based extremist affiliations using sentiment analysis techniques. *Hum. Cent. Comput. Inf. Sci.* 9, 24 (2019).
<https://doi.org/10.1186/s13673-019-0185-6>
3. Bhandary, Abhir & k b, Sudeepa & Chokkadi, Sukhada & M S, Sannidhan. (2019). "A Study on various state of the art of the Art Face Recognition System using Deep Learning Techniques. *International Journal of Advanced Trends in Computer Science and Engineering.* 8. 1590-1600. 10.30534/ijatcse/2019/84842019.
4. Bharat, Vinod & Malik, Dr. (2019). "Study of Detection of Various types of Cancers by using Deep Learning: A Survey". *International Journal of Advanced Trends in Computer Science and Engineering.* 8. 1228-1233. 10.30534/ijatcse/2019/31842019.
5. Geert Litjens, Thijs Kooi, Babak Ehteshami Bejnordi, Arnaud Arindra Adiyoso Setio, Francesco Ciompi, Mohsen Ghafoorian, Jeroen A.W.M. van der Laak, Bram van Ginneken, Clara I. Sánchez, A survey on deep learning in medical image analysis, *Medical Image Analysis*, Volume 42, 2017, Pages 60-88, ISSN 1361-8415
<https://doi.org/10.1016/j.media.2017.07.005>
6. Huang Yi, Sun Shiyu, Duan Xiusheng and Chen Zhigang, "A study on Deep Neural Networks framework," 2016 IEEE Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC), Xi'an, 2016, pp. 1519-1522.
<https://doi.org/10.1109/IMCEC.2016.7867471>
7. K. Zvarevashe and O. O. Olugbara, "A framework for sentiment analysis with opinion mining of hotel

- reviews," 2018 Conference on Information Communications Technology and Society (ICTAS), Durban, 2018, pp. 1-4.
8. Mildenerger, T. Stephen Marsland: Machine learning. An algorithmic perspective. *Stat Papers* 55, 575-576 (2014).
<https://doi.org/10.1007/s00362-012-0471-0>
9. Pamina, J. and Raja, Beschi, Survey on Deep Learning Algorithms (January 12, 2019). *International Journal of Emerging Technology and Innovative Engineering*, Volume 5, Issue 1, January 2019.
10. Pandya, Sheetal & Kalani, Nilesh. (2019). "Review on Text Sequence Processing with use of different Deep Neural Network Model". *International Journal of Advanced Trends in Computer Science and Engineering.* 8. 2224. 10.30534/ijatcse/2019/56852019.
11. Pouyanfar, Samira & Sadiq, Saad & Yan, Yilin & Tian, Haiman & Tao, Yudong & Presa Reyes, Maria & Shyu, Mei-Ling & Chen, Shu-Ching & Iyengar, S.. (2018). A Survey on Deep Learning: Algorithms, Techniques, and Applications. *ACM Computing Surveys.* 51. 1-36. 10.1145/3234150.
12. W. T. Illingworth, "Beginner's guide to neural networks," in *IEEE Aerospace and Electronic Systems Magazine*, vol. 4, no. 9, pp. 44-49, Sept. 1989
13. Zhang, Q., Yang, L.T., Chen, Z., & Li, P. (2018). A survey on deep learning for big data. *Inf. Fusion*, 42, 146-157.
<https://doi.org/10.1016/j.inffus.2017.10.006>