International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075, Volume-9 Issue-2S, December 2019

Deep Learning Predictive Models for Cognitive Radio System

Sandeep Bidwai, Nikhil Joshi, Saylee Bidwai, Uday Wali

Abstract: ***Cognitive Radio (CR) was introduced to improve the utilization of Radio Frequencies (RF) that remain under-utilized by the primary users (licensee). The main idea behind CR is to allow un-licensed (secondary) users to occupy vacancies in licensed bands. However, CR mandates the secondary user to vacate the frequency band within a specified time after the primary user attempts to use the frequency band. CR does not expect the primary users to share their frequency usage schedules and hence the secondary users have to scan and predict the vacancy. The advantage for the secondary users is that they do not pay for utilization of band, if they are conformal to the CR specifications. CR is the next generation of smart communication systems.***

CR requires continuous monitoring of the intended RF band in the intended geographical area. This information may be used to predict spectral vacancies (white spaces). Certain bands, e.g. Analog TV bands, will have pre declared utilization schedules but in general, spectrum utilization is a random process and hence prediction can be difficult. However, Deep Learning (DL) techniques can improve the accuracy of prediction. Deep Learning techniques require large and clean data sets to work correctly. Such data sets are also necessary to compare achievable accuracy of prediction algorithms. Towards this end, we have created data sets that can be used for simulation, training and testing of CR over GSM band (890-960MHz). A typical file with two hour of observations will have about 1.2 million samples. More than 1000 sets of data samples have been captured from urban and rural areas in India. All the data sets have been cleaned to avoid instrument errors and statistical outliers.

In this paper we have used these standardized data sets to perform a comparative analysis of three DL methods for CR, viz. Auto-encoder (AE), Long Short-Term Memory (LSTM) and Multi Layer Perceptron (MLP). Results of the comparison are discussed.

Index Terms- GSM, LSTM, Auto-encoder, MLP, Cognitive Radio.

I. INTRODUCTION

In past decade, there has been a substantial increase in utilization of Radio Frequency (RF) wireless communication systems.

Revised Manuscript Received on December 15, 2019.

***** Correspondence Author

Sandeep Bidwai, Research Scholar, Visvesvaraya Technological University, belagavi, Karnataka, India.

Nikhil Joshi, Security Researcher, Payatu Soft wares Labs Limited Liability Partnership

Mrs. Saylee Bidwai, Assistant Professor, Department of Electronics & Telecom,t ADCET, Ashta, Sangli, Maharashtra, India.

UdayWali, Professor, Dept. of EEE, KLE Dr M S Sheshgiri College of Engineering. & Technology, Belagavi, Karnataka, India.

The usage of frequency spectrum is regulated and steered by regional & international organizations like European Telecommunication Standards Institute (ETSI), International Telecommunication Union (ITU), European Conference of postal and telecommunication Administrations (CEPT), Japanese Telecommunication Technology Committee (TTC), IEEE etc.

The Radio Frequency (RF)spectrum is a limited natural source and hence it is necessary to use it efficiently. The spectrum bands are usually licensed for specific services like TV broadcasting, mobile, fixed broadcast and satellite to avoid the interference between Service bands. Modern RF electronics is controlled by digital controllers which can be programmed to suit specific requirements. Such radio equipment depends on development of software to control the RF parameters such as base band, carrier and modulation technique, power, etc. Such equipment is therefore called Software Defined Radio (SDR).

GSM traffic is typically a time series, which is a sequential data measured regularly at a uniform time interval. A time series would be a deterministic or a stochastic process. It is necessary to use mathematical models for prediction of time series that truly present the statistical characteristics of sampled data.

This paper analyzes three different prediction methods based on ANN.Features extracted from the required evaluations were compared amongMulti Layer Perceptron (MLP), Auto Encoder (AE) and Long Short Term Memory (LSTM). MLP is a type of feed-forward neural network with multiple layers used for the purpose of supervised training. AE isspeciallysuited for dimensional reductionin unsupervised training. LSTM isused to model sequential (time based) information and their long-range dependencies more precisely than conventional Recurring Neural Networks (RNN). These models are selected with an objective to compare three predictive algorithms.

II. DATA COLLECTION

In our previous work[1], a practical approach to identify the white spaces in GSM band has been carried out. we have started with design and fabrication of dual band GSM antenna (1800 MHz & 900MHz). We have used this antenna along with general purpose Spectrum analyzer to understand GSM spectrum usage. We have received the data in the form of power values of received signals. The setup is set at various places in India. The percentage of spectrum usage varies with locations.We have developed a small program to record the spectrum periodically.

This data is used by analysis software, under development. Collection of data sets for extended time periods is a slow process, which we have started afterdevelopment of spectrum recorder software. The software provides a visual feedback on the datacollection. This in turn has helped us to increase the reliability of the captured data. Figure 1 shows the software front panel.

This software captures the singals regularly in the band of 890-960 MHz. The software is scaled like spectrum analyzer to capture the data in the regular interval of 3 seconds. Total 70 MHz band is captured i.e. total 351 frequencies status report has been generated. The comparative advantages of this software is that it can record the trace data in the manually set range of channels, so it is flexible provided respective antenna type must be attached to spectrum analyzer. The trace data is recorded in the interval of 3 seconds between two traces. The measurements are recorded in rural area like Ashta, Sangli in Maharashtra, urban area like Pune, Begalavi and Sangli city.

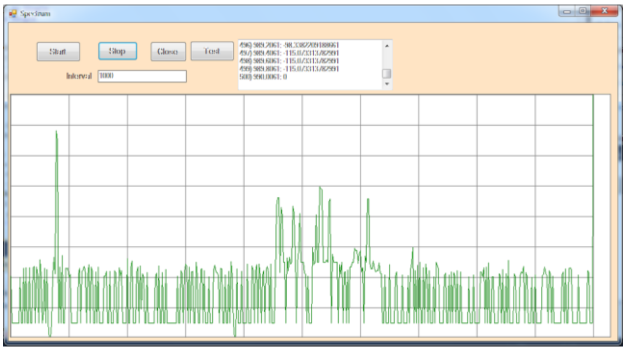


Fig. 1.Visualization of MS Windows application like real time spectrum analyzer.

III. ARTIFICIAL NEURAL NETWORKS

Artificial Neural Networksare simple processing architectureswith interconnected units named as *neurons.*Neurons are arranged into different layers. One layer can have multiple neurons. Neural network topology varies in varies among different models [10]. In this research work we have preferred following three types of neural networks i.e. LSTM, Auto encoder& MLP.

A. Long Short Term Memory

In a typical Recurrent Neural Network using back propagation through time (BPTT) learning algorithm, the error is propagated in time steps. It happens that the error either vanishes or saturates over few time steps. This limits the ability of the network to learn long time correlations in the signal. A Supervised gated version of RNN, known as Long Short Term Memory (LSTM) overcomes these problems(1997 Hochreiter). LSTM consists function units to maintain a constant error flow that can be back propagated in time and state, which allows RNN tolearn over many time steps. Each fundamental unit contains the information in gated cells indicated as linear stage in Fig.2. Cells acts like a memory where the information can be stored, written, read or cleared using gates. The cells can take the decision through the iterative process of making guesses, back propagating the errors and adjusting the weights. These gates can block or pass the received information based on strength of gating signal.

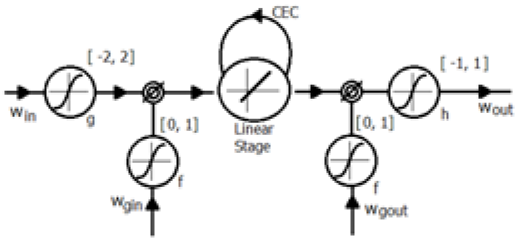


Fig. 2.structure of LSTM memory cell

Fig.2. shows the structure of LSTM memory cell. The cell has a linear unit called Constant Error Corrousel (CEC) that contains the state of the cell (short term memory) which is recalled when required. The feedback loop keeps the value intact as long as necessary. Input and output are gated scaled and activated input which is stored in the linear stage. The cell output is similarly activated, gated and scaled. Input of a cell contains current input and outputs of other cells including its own value. Gating the activated input prevents untimely erasure of a stored value. Similarly, gating the output prevents other cells from gating affected at an unexpected time. In Fig.2, the limiting values of activation functions are indicated in square brackets. The gate learn to open on arrival of correlated input presented over large interval of time to maintain the state CEC. The state of a cell can be cleared by a specific input gate. The output of the cell is also gated and presented as an activated signal to avoid overloading inputs of other cells. Each cell remembers its state infinitely, unless it is explicitly cleared. If there is a temporal correlation between two signals, two cells corresponding to the se time steps need to fire together to generate an output. The cells connections now represent a temporal connection and hence are able to respond to a time –series input. As the cells can store the state infinitely using the feedback loop, there is no need to connect the cells in sequence unlike in other recurrent neural networks. Each cell behaves like a memory cell with two distinct actions: update and emit, similar to write and read of a digital memory. The cell may also be forcefully cleared.

For ease of understanding, consider a single cell. Let the output of the cell be $y_{out}(t)$. outputs from other cells and current input be represented as $y_{net}(t)$. Gating output at the input and output be $y_{gin}(t)$and $y_{gout}(t)$ respectively. Activation functions at the gate controls are sigmoid functions with output in the range of 0-1. Activation function at the input to the cell and output of the cell is a scaled tanh(). The output of the cell is an activated value of the internal state S.output in current time corresponds to memory in previous timey step, gated by the current input. Therefore

$$y_{gout}(t) = f\,(y_{net}\,(t)\ w_{gout}) \quad (1)$$
$$y_{gin}(t) = f\,(y_{net}(t)\ w_{gin}) \quad (2)$$
$$y_{out}(t) = y_{go}(t)\ h(S(t\text{-}1)) \quad (3)$$
$$y_{in}(t) = g(y_{net}(t)\ w_{in}) \quad (4)$$

The state of the cell will have the values given by the following equations

$$S(0) = 0,\ S(t) = S(t\text{-}1) + y_{in}\,(t) \text{ for } t > 0 \quad (5)$$

Retrieval Number: B11291292S19/2019©BEIESP
DOI: 10.35940/ijitee.B1129.1292S19

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

The weights of the cell are updated using a gradient descent, keeping the error constant across the time steps. LSTM has been successfully used in handwritten character recognition, stock market prediction, cognitive radio and similar other time series predictions.

In this paper we have preferred LSTM network for predicting white spaces in GSM band.An LSTM model is developed and trained using the data set observed at various places. The large data set is transformed into the normalized subsets i.e. training and testing subsets. Training size is set 87%. The difference between actual values and predicted values is calculated as Mean square error. Input size to LSTM model is equal to the size of lookback. The main data set is transformed into a single vector to provide the input to the model. Total 128 units of LSTM have included. Adam optimizer is used with hyper parameters like learning rate (lr)=0.001, β1=0.99, β2 = 0.999. We have done execution of LSTM model which is buildusing python libraries and found that MSE is lowest for the dataset is 0.0068 for observations at Belagavi and Sangli as shown in Table 1.

B. Auto Encoder

Auto encoder is a suitable type of neural network which is used for unsupervised coding. The main objective of auto encoder is to learn a encoding for a set of data especially for dimensionality reduction by training the network to ignore the noise. Auto encoder tries to produce the output close to its input that is why the name is "Auto encoder". In the discrete manufacturing process, recognition of various defective patterns could be available that can significantly reduce diagnostics processes and increase manufacturing process stability and quality. Therefore, the effective recognizer is in demand to improve the performance of pattern recognition process [6].

Research work in [11] shows the computationally efficient method for gradient-based optimization of stochastic objective functions. This method has objective to solvemachine learning problems with large datasets and/ or high dimensional parameter space. Overall it was found that adamoptimizer is strong and suitable for wide range of optimization problems in machine learning perspective. That's why we have used adam optimizer in these models.

In this work we have used auto encoder for predicting the vacancies in GSM band. Following steps are involved in preparing the auto encoder architecture:

Step1: Declare the required libraries
Step2: Load the data set
Step3: Normalization of the data (in the range 0 to 1)
Step4: De-normalization of the data (in the range 0 to 1)
Step5: Set input dimension to look back
Step6: Construct the auto encoder by adding dense layers (input, output units=128, code units=16)
Step7: Define the accuracy function with target and predicted values.
Step8: divide the main data set into training and testing data subsets.
Step9: Get the adam optimizer and set it's hyper parameters.
Step10: Plot losses and prediction plots
Step11: Store plots in pickle.

After the execution of this model, we found the lowest MSE as 0.0075 for the dataset observed at Belagavi.

C. Multilayer perceptron (MLP)

It is the fundamental type of Artificial Neural Network. This kind of simple network has one input layer, more than one hidden layers and aoutput layer. The best practice mentioned in[7] suggests one or more hidden layers .So, in order to get the same result, it is not necessary to increase number of layers, but we can increase the number of nodes/neurons in hidden layer[8].

Figure 3 shows the one input layer, two hidden layers and output layer used in building the multilayer perceptron model.

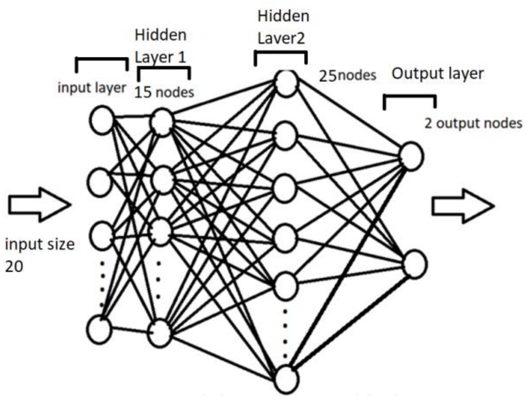


Fig.3 Architecture of Multi layer Perceptron

(MLP)showing layers, number of neurons in each layer (inforamtion transfers from Left to right)

Figure 3 above shows the architecture of Multi layer perceptron model having input size as 20, dense layers (hidden layers) of size 15 and 25 and output layer have 2 nodes.

Table 1 shows the comparative study of all the three models. It is one of the simplest architecture of the neural network. In this work we are using MLP as classifier. Input data would be classified into 0 or 1 based on the threshold value of -90 dBm

The step wise flow of MLP algorithm is as follows:

Step1: Load and generate data sets.
Setp2: Split dataset into training and testing set using Randomsampling.
Setp3: Use 33% data as test data set and remiaining data as training data set.
Setp4: Provide class weights so that MLP does not skew towards always predicting dominent class from data set. (we are calculating fraction of dataset of each class 0 & 1).
Setp5: Create layers of neural network
Step6: Apply adma optimizer with hyper parameters
Step7: Train and record loss and accuracy on each epoch
Step8: Apply prediction on traing and testing set
Step9: Print confusion matrix for prediction on training and testing sets.
Step10: Plot loss curve
Step11: Plot accuracy curve.

After the execution of this model we found the lowest MSE is 0.00 for the data set observed at pune, sangli and ashta. The training size is taken 80%.

For doing analysis of these three modules, the parameters like Means Square Error (MSE) and accuracy are chosen. All these modules are designed in python using various open source libraries like numpy, pandas, matplotlib. With number of experiments using training and testing subsets, MSEs of respective modules are observed.The hyper parameters of adam optimizer are set as, learning rate = 0.001,beta_1 = 0.99, beta_2=0.999. performance evaluation of these three modules is carried out with 20 epochs.

Mean Square Error (MSE) results after every 20 epochs of every module are shown in Table1, Table 2 shows accuracy comparison and Table 3 shows hyper parameters of optimizer and other parameters of the modules.

Table 1: Comparative study of three modules

Area	Date set	MSE after 20 epochs			Split Data Size			
		LSTM	AE	*MLP*	X_test	X_train	Y_test	Y_train
Rural	ashta_indoor_testdata-900_ 23-01-2019.csv	0.0127	0.0138	0.0574	12329,20	25030,20	12329,2	25030,2
Rural	900_ashta_outdoor_29-08-2018.csv	0.0118	0.0202	0.000	10171,20	20648,20	10171,2	20648,2
Urban	Belagavi_testdata-900_even _11-08-2018.csv	0.0068	0.0075	3.494e-04	5033,20	10216,20	5033,2	10216,2
Urban	pune_outdoor_900_19-08-2 018.csv	0.0217	0.0148	0.000	8706,20	17673,20	8706,2	17673,2
Urban	sangli_indoor-900_morning _13-08-2018.csv	0.0125	0.0100	0.000	6726,20	13653,20	6726,2	13653,2
Urban	sangli_outdoor_900_22-08-2018.csv	0.0068	0.0113	0.000	22130,20	44929,20	22130,2	44929,2

Table2. Accuracy comparison among three modules

Data set	Accuracy		
	LSTM	Auto-encoder	MLP
ashta_indoor_testdata-900_23-01-2019.csv	0.9963	0.9989	0.5051
900_ashta_outdoor_29-08-2018.csv	1	1	1
Belagavi_testdata-900_even_11-08-2018.csv	0.9994	0.9990	0.5792
pune_outdoor_900_19-08-2018.csv	1	1	0.0000
sangli_indoor-900_morning_13-08-2018.csv	1	1	0.0078
sangli_outdoor_900_22-08-2018.csv	1	1	0.000

Table3. Comparison among parameters of three modules

ADAM Optimizer Hyper parameters			
	LSTM	Autoencoder	MLP
Learning Rate	0.001	0.001	0.001
Beta_1	0.99	0.99	0.9
Beta_2	0.999	0.999	0.999
Decay	0.01	0.01	0.01
Batch size	20	100	100
General Parameters			
No. of Epochs	20	20	100
Training Size	87%	80%	80%
Look back	20	20	20
No. of units	20	Encoders=128,64,32,16 Decoders=32,64,128	Input, dense 15,25,2
Activation function	Sigmoid, tanh	Encode & Decode= Relu, decode= Sigmoid	Linear

Note:***Column number/Frequency:*** *300 (949.8MHz) to 330 (955.8MHz)*

Since the dataset is highly imbalanced it is important to provide class_weights so that MLP does not skew towards always predicting the dominant class from imbalanced dataset.In Machine learning terminology, aconfusion matrix known as error matrix when we define statistical classification problem. It us used to present the performance of classifier (or classification model) over the set of test data with true value known. It allows the visualization of the performance of an algorithm. Confusion matrix on train and test data shows true positive (TP), false positive (FP), true negative (TN) and false negative (FN) values after the successful execution of MLP algorithm.

IV. RESULTS & DISCUSSIONS

The Regerous experiments were conducted on the trace data recorded at various places (rural and urbadn) and datasets with different size are mentioned in Table 1. All together six data sets were used. Conducted experiment used open source python keras library. The main characteristics of the deep neural network is that it allows non-linear data and complexes in learninig[5]. In all the three models the errors in the neural network are calculated in term of Mean Square Error (MSE) which is given as

$$e_t = y_t - \overline{yt} \quad (7)$$

$$MSE = \sum_{1}^{n} et^{\wedge}2 \, / n \quad (8)$$

Where, n is number of predictions, y_t is actual values observed and $\overline{yt}$ is predicted values at time t, e_t is a prediction error at time t[9].

Figure 4 shows the response of MLP model in terms of accuracy and epochs for the given dataset. Figure 5 shows the variation of MSE with respect to epoches seperately on training and testing datasets.

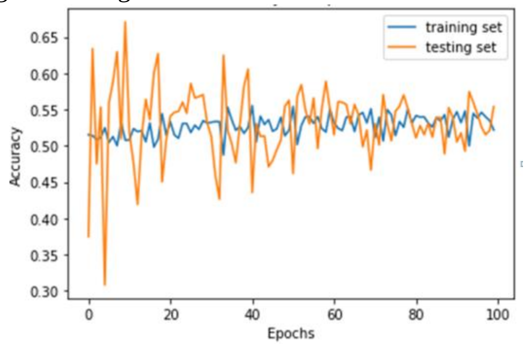


Fig.4 Accuracy Vs. Epochs plot using MLP

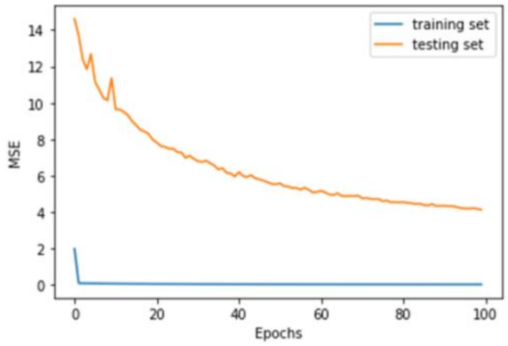


Fig. 5 MSE Vs. epochs plot for training and testing data

Figure 6 shows prediction plot of auto encoder model and Blue color indicates the ground truth, orange color Indicate training data set and green color shows testing dataset. Figure 7 shows alteration of Mean Square Error with respect to epochs.

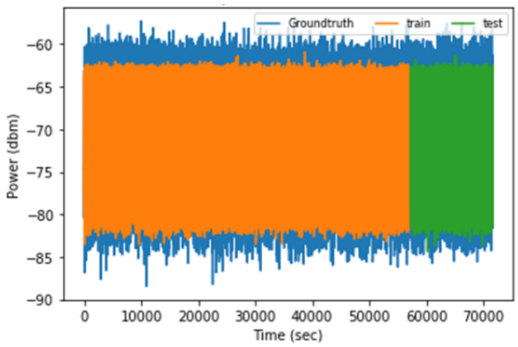


Fig.6 Prediction plot of Auto encoder on train and test sets

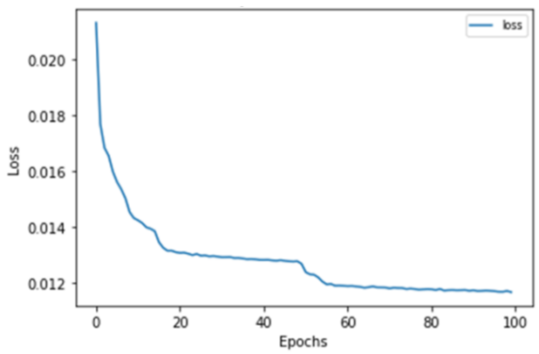
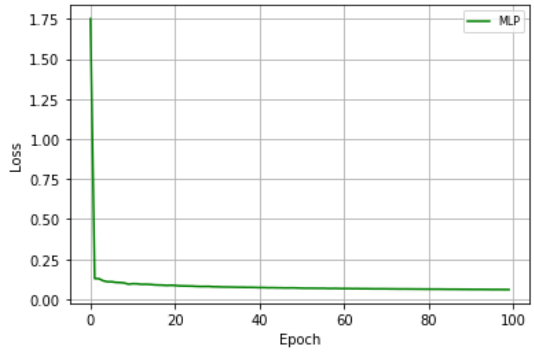
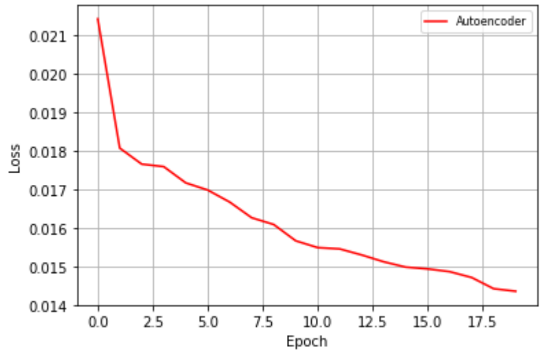


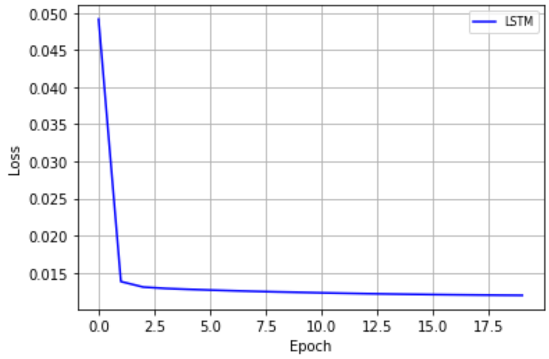
Fig.7. Mean Square Error of Auto encoder



(a) Multi Layer Perceptron model response



(b) Auto Encoder model response



(c) LSTM Model response

Fig. 8Plots for Loss Vs epochs

From the above plots, it is proved that autoencoder and LSTM have lowest minimum losses in the neural network.

V. CONCLUSION

All three ANN's have proved that they are capable of predicting the GSM network traffic based on historical data.The overall results from the observation table 1shows that MLP ishaving comparatively Less MSEthan LSTMand Auto encoder modules. LSTM is a complex type of neural network and have processing time more compare to other neural networks. These types of predictive methods can be helpful to the upcoming cognitive radio systems in spectrum sensing operation to know the status of white spaces.

As a next step, we will focus on the implementation of Convolutional Neural Network and it's performance comparison with Markov model.

Further works may be focused on another Deep Learningtechniques, like Continuous Restricted Boltzmann Machine (CRBM). The objective is to ensurestatus of spectrum availability for a fair share of bandwidth allocation and a better Quality of Servicefor a dynamic and adaptive management application in cognitive radio system.

REFERENCES

1. Sandeep Bidwai1, Uday V. Wali2, Mr. Bahubali Shirgapur3, Mrs. Saylee S. Bidwai4,"Detecting White Spaces for Cognitive Radio",International Journal of Technology and Science, ISSN (Online) 2350-1111, (Print) 2350-1103 Volume 9, Issue 2, 2016 pp. 8-11
2. S. Hochreiter and J Schmidhuber, "The Long Short Term Memory," Neural Computation, Vol 9, No.8, pp. 1735-1780,1997.
3. F.A. Gers, J. Schmidhuber& F. Cummins, "Learning to forget: Continual prediction with LSTM", Neural Computation, vol.12, no, 10, pp. 2451-2471,2000
4. F.A. Gers, N.N. Schraudolph& J.Schmidhuber, " Learning precise timing with LSTM recurrent netwroks", Journal of Machine Learning Rearch, vol.3, pp.115-143, Mar. 2003
5. Bengio, Y. 'Learning deep architectures for AI', Found. Trends Mach. Learn., Vol. 2, No. 1, pp.1–127, 2009.
6. Jianbo Yu, XiaoyunZheng, Shijin Wang, "A Deep Autoencoder feature learning method for pattern recognition", Elsevier, 2019. 0959-1524, https://doi.org/10.1016/j.jprocont.2019.05.002.
7. de Villiers, J. and Barnard, E. 'Backpropagation neural nets with one and two hidden layers', Neural Networks, IEEE Transactions on, Vol. 4, No. 1, pp.136–141,1993
8. Hornik, K., Stinchcombe, M. and White, H. 'Multilayer feedforward networks are universal approximators', Neural Networks., Vol. 2, No. 5, pp.359–366.1989.
9. Tiago Prado Oliveira, Jamil Salem Barbar and Alexsandro Santos Soares," Computer network traffic prediction: a comparison between traditional and deep learning neural networks", International Journal on Big Data Intelligence, Vol. 3, No. 1, 2016.
10. Haykin, S.,"Neural Networks: A Comprehensive Foundation", 2nd ed., Prentice Hall PTR, Upper Saddle River, NJ, USA.1998.
11. Diederik P. Kingma, Jimmy Lei Ba," adam: a method for stochastic optimization", Published as a conference paper atICLR 2015.

AUTHORS PROFILE

Sandeep Bidwai, graduated from J.N.E.C. Aurangabad. He has done his masters from Shivaji University, Kolhapur. Research Scholar at VTU, belagavi, Karnataka. His area of interest is Spectrum Sensing, Cognitive Radio, Artificial Intelligence, Machine Learning, Deep Learning, Wireless Communication. He is a member of ISTE and CSI.

Nikhil Joshi, has done his Computer Engineering from International Institute of Information Technology, Pune, Maharashtra in 2017. Presently working as security researcher in payatusoftwares labs LLP. He is currently working on implementations of ML in offensive and defensive security products. At Payatu, He has orchestrated methodologies to pen-test Machine Learning application against ML specific vulnerabilities and loves to explore new ways to hack ML powered applications. Parallelly Nikhil's research is focused on security implications in Deep Learning applications such as Adversarial Learning, Model stealing attacks, Data poisoning, etc.

Mrs. SayleeBidwai, working as Assistant professor in Department of Electronics & Telecom at ADCET, ashta, sangli, Maharashtra. She is a research Scholar in V.T.U. Belagavi, Karnataka. Her are of interest is network on chip, VLSI design, Digital Design, Programming in VHDL, Verilog, C# etc. She has been received Eminent Assistant Professor Award by CSI, Mumbai. She has contributed in encouraged many startups and successfully established as incubation center.

UdayWali is a Professor in Dept. of EEE at KLE Dr M S Sheshgiri College of Engineering. & Technology, Belagavi, Karnataka India. He has obtained Bachelor of Engineering in Electrical and Electronics Engg. from Karnataka University Dharwad (1981) and Ph.D from IIT Kharagpur (1986). He is a fellow of Institute of Engineers (India) and CEO of C-Quad Computers, Desur IT Park, Belagavi. He has 30 years of teaching and 15 years of industrial experience. His current research interests are Artificial Intelligence, Neural Networks, and Processor Design.