

A Comparative Study of Markov Chain and Deep Learning Predictive Models in Spectrum Sensing

Sandeep Bidwai¹, Saylee Bidwai², Uday Wali³

¹Department of Electronics and Communication Engineering, KLE Dr. M.S. Sheshgiri College of Engineering and Technology, Belagavi, Karnataka, India

²Department of Electronics and Telecommunication Engineering, Annasaheb Dange College of Engineering and Technology, Sangli, Maharashtra, India

³Department of Electronics and Communication Engineering, KLE Dr. M.S. Sheshgiri College of Engineering and Technology, Belagavi, Karnataka, India

2kl12pen02@klescet.ac.in, sayleebidwai@gmail.com, uday.wali@gmail.com

Abstract. Radio Frequency Spectrum is a limited resource. Commercial allocation of this resource may lead to inefficient use as the allottee may not use the spectrum to its full capacity. Users of Cognitive Radio (CR) need to monitor the spectrum continuously to predict channel occupancy, which is dependent on three basic parameters, viz. location, time and RF band. Several statistical methods have been used to predict channel occupancy. However, use of Machine Learning (ML) and Deep Learning techniques can help to improve accuracy of prediction. This paper attempt to compare some of these techniques to evaluate their use in real time applications. As the RF bands are inherently noisy and time variant, use of standardized data models is necessary to compare the algorithms reliably. The work done in this paper describes the channel occupation and comparing their accuracies and implementation efficiency. As a reference, a Markov model for channel prediction has been presented. Several DL techniques viz. , Multilayer perceptron (MLP), Auto encoder (AE), Convolutional Neural Network(CNN) and long short term Memory (LSTM) have been implemented and their results are compared. The results shows that an adaptation of CNN gives results that are comparable with the Markov model.

Keywords: GSM, Spectrum Sensing, Markov Chain, Deep Learning, Predictive Models.

1. Introduction

Spectrum needs to be managed carefully to avoid the spectrum scarcity issue. As the numbers of wireless users are growing, the demand of the spectrum is also increasing. CR is considered a future matured technology that can address spectrum scarcity issues. Several studies during the last few decades have shown that most of the allocated spectrum is underutilized temporally and spectrally. The predictive models working in cognitive radio hardware could be helpful to choose the vacant band in the present electromagnetic environment without interference to primary users (PUs). The spectrum prediction process is a complex and therefore advanced statistical techniques like Markov chains are indeed an important option. The state of occupancy of a frequency at a specific time can be either be free or busy. The states of the sub-bands need to be observed over successive time slots to define PU activity. In this paper, a comparative study of a Neural Network based predictive models is presented. Assume that the signal is present if the strength is above -70 dBm for clear understanding of the potential GSM signals. These models are used for predicting the channel status using historical data obtained from the spectral energy measurement. The performance parameters used in the process of performance evaluation are Mean Square Error (MSE), Root Mean Square Error (RMSE), Mean relative Error (MRE) and accuracy.

This research paper is arranged into five sections. The first section introduces the research work. Second section gives background of the work. The third section includes research methodology and data capture process. The construction of Markov & Deep learning models for spectrum prediction is presented in the fourth section. The comparative analysis of these predictive models is given in the fifth section and the sixth section includes results and discussion with the comparative analysis of Markov and Deep learning models.

2. Background

The spectrum sensing is one of the first and important step in cognitive radio engine. There are various spectrum sensing methods used to sense the electromagnetic spectrum. The major spectrum sensing methods are energy detection, matched filter detection and cyclostationary feature detection. Energy detection is found to be a simple detection technique V. I. Kostylev (2002), Adebola, E., Annamalai, A. (2014), F. F. Digham, M. S. Alouini, and M. K. Simon,(2003), E.H. Gismalla, and E. Alsusa (2011), Wang, Y., Zhang, M., Shi, Q.,(2014), S. Atapattu, C. Tellambura and H. Jiang, (2011), as it require only the information about the energy of the signal. The experimental set to capture the GSM data from surrounding electromagnetic environment is reported in Sandeep Bidwai, et. al. (2016). The data sets from four different locations have been captured. The features from these datasets are used to train the neural models. The use of neural network models have been reported in D. Han et al.,(2017), S. Pattanayak, P. Venkateswaran and R.

Nandi,(2012) . In context of Cognitive Radio, Markov or Semi-Markov process can signify the temporal relation between states of PUs S. Geirhofer, L. Tong and B. M. Sadler,(2007), Y. Saleem and M.H. Rehmani, (2014). For spectrum sensing application, the authors proposed a CNN algorithm and stated that CNN model performance is considerably well as compared to the traditional method like Estimator-correlator detector and HMM in terms of detection accuracy. The work done in T. W. Rondeau, C. J. Rieser, T. M. Gallagher and C. W. Bostian, (2004) , C. Park, S. Kim, S. Lim and M. Song,(2007), I. A. Akbar and W. H. Tranter,(2007), Q. Zhao and A. Swami, (2007), assumes that Markov chain be present in RF transmissions. Most of the models look at PUs as free or idle in the available spectrum band. However, Machine Learning methods have abilities to extract non-linear information of training data. This ability improves accuracy of prediction significantly. The author validated the existence of Markov chain in PU channel utilization with real time measurements in a paging band (928-948 MHz) C. Ghosh, C. Cordeiro, D. P. Agrawal and M. B. Rao, (2009).

In a typical Recurrent Neural Network (RNN) 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 a few time steps. This is caused by repeated multiplication by the same weight in RNN. If the weight is less than 1, the result vanishes (becomes zero) or if the weight is more than 1, the output becomes very large. This limits the ability of the network to learn long time correlations in the signal. A supervised gated version of RNN known as LSTM overcome these problems F.A. Gers, N.N. Schraudolph & J. Schmidhuber, (2003). The spectrum prediction problem is reported as binary hypothesis testing problem in D. Han et al.,(2017). A deep neural network (DNN) based data driven test statistics is proposed to overcome the problem of existing statistics which does not characterize the real environment. The detection probability was found to be 96.7% which is outperform than ideal optimal detector. Even though deep learning algorithms have wide applications in the past few years and its applications in the RF domain are still in their infant state Y. Li, Y. Dong, H. Zhang, H. Zhao, H. Shi and X. Zhao,(2010), X. Xing, T. Jing, Y. Huo, H. Li and X. Cheng, (2013) , C. Zhang, P. Patras and H. Haddadi,(2019), Y. He et al., (2017). The research work in Danilo López, Johana Hernández, Edwin Rivas, (2016) shows the computationally efficient method for gradient based optimization of stochastic objective functions. The method solves machine learning problems with high dimensional parameter space (large datasets). In O. Omotere, J. Fuller, L. Qian and Z. Han, (2018), a multi-class classification problem is formulated using deep learning techniques viz. CNN, LSTM and DNN for accurate spectrum prediction under complicated coexistence scenarios.

In our previous work, a LSTM model was used for channel prediction in the GSM band as mentioned in Sandeep Bidwai et.al. (2018), and the comparative

study of Deep learning algorithms (LSTM, MLP and Auto encoder) was conducted for cognitive radio systems as elaborated in Sandeep Bidwai et.al. (2019). The performance comparison between Markov chains and LSTM predictive models is done in Bidwai S. (2021) and found that the LSTM outperforms the Markov chains. Further, a spectrum prediction framework was designed based on neural networks as presented in KunweiLan, Hang sheng Zhao, Jianzhao Zhang, Cao Long, MenglinLuo, (2014). This framework was optimized by genetic algorithms to avoid local optimization problem.

3. Research Methodology

The research work carried out in this paper is divided into three major steps i.e. Data capture & preparation, data modelling and Model evaluation. Among these steps, data capture procedure is elaborated below.

3.1. Data capture

As mentioned earlier, we have captured the GSM trace data using an experimental setup containing a Visual Basic application. A screenshot of the windows application with a captured signal is shown in figure 1.

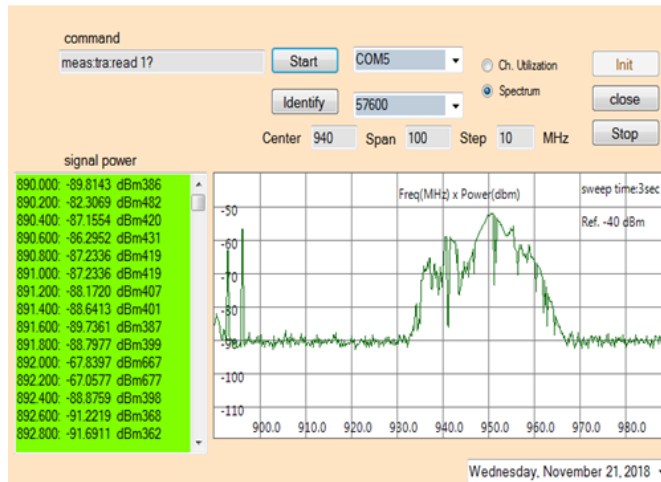


Fig. 1 Screenshot of data capture application

Data has been captured in the form of an array of frequencies and time slots. In the span of 100MHz, a total of 501 channels are recorded based on the time interval in each data set. These datasets are made available at the GitHub repository <http://github.com/profssbssb/GSM-datasets.git>. It takes minimum 3 seconds to scan the entire GSM band on the spectrum analyser.

3.2. Data Cleaning

Figure 2 (a) represents the data before cleaning. As the data includes NaN values

and some missed values, the sequences have to be cleaned for data errors. Corrected data appears like vertical lines as represented in Figure 2 (b). Each vertical line corresponds to a channel over a period of time. Blue indicates the free channel and red shows the occupied channel. White space is all intermediate values. The data captured is filtered. This filtration is done to match the frequencies which are used by a known base station. This data is used for further analysis.

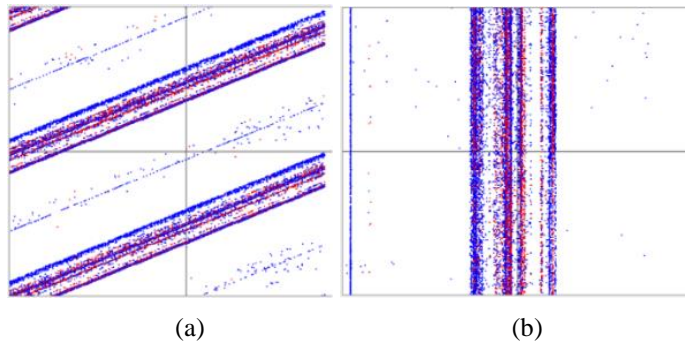


Fig. 2 Pattern of incoming data (a) pattern before data cleaning (b) pattern after data cleaning

The transition matrix is generated from training set for Markov models. Once the data is cleaned, this cleaned data is applied to the input of deep learning models used in this work. The method used to train and predict the occupancy of channels is shown in figure 3.

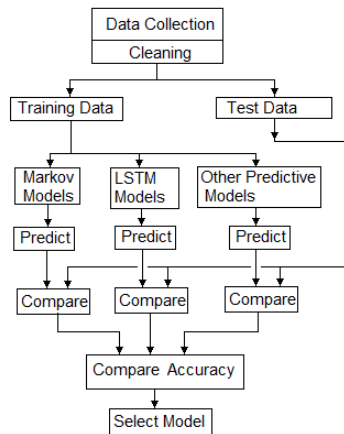


Fig. 3 Methodology used to build Predictive models

The complexity of building a usable model depends on the accuracy of requirements. The input data has been split into three categories. By choosing thresholds on the incoming data, the categories are: Signals having energy level

below -80dBm are considered as White-space (W), signals having energy level above -70dBm are considered as Occupied (O) and the signals with energy level ranging in-between -70 and -80 dBm is considered as Unknown (U). The predictions are derived based on the transition probability matrices for these three states.

4. Spectrum Predictive Models

4.1. Predictive model using Markov chain

In order to prepare the predictive model, we have developed a Markov Model to predict the probability of the next state from a set of input data. Markov Chains are a classical way to model stochastic time series. Markov chains are used in telecommunication channels for classification of time series data. However, Markov models are computationally very efficient. They require to know the status of current state. Using Hidden Markov Models, we can predict the status of the next channel based on the past channel states C. Park, S. Kim, S. Lim and M. Song, (2007). Using following equation, the process of Markov chain can be elaborated as

$$P(q_{n+1}) = S_{n+1} | (q_n = S_n) \tag{1}$$

The process of building the Markov model consists of estimating the probability of every state from the available training data. This can be easily accomplished by binning the input data into states of the system. In case of a RF transmission system, the signal strength indicates the state of the transmission system; if the signal is good, transmission exists; if the signal is below a threshold, the channel is free; otherwise the state is unknown. Therefore the model consists of three states as shown in Figure 4. As the predicted value of the system depends on the present state, a state transition matrix indicates the probabilities of transition from the current state to the next state. These probabilities are used to predict the next possible state of the system. Given any particular state, the next state can be any one of the three possible states. Therefore, each row of the transition matrix will have a cumulative probability of 1. This allows sampling of the next state based on the current state.

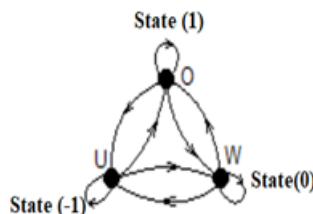


Fig. 4 Proposed Markov Model

However, it is possible to consider the transition probability as a whole and estimate the next transition, instead of the state.

All deep learning modules used in this work are constructed using python open source libraries like numpy, pandas, matplotlib. The common configuration for these deep learning models includes, hyper parameters i.e. learning rate=0.001, beta_1=0.99, beta_2=0.999, decay=0.01, batch size = 100, Number of epochs = 100, look back = 20, training set=80%, test set=20%. Deep learning techniques are well versed with the random RF data. Due to randomness of RF signals, we have applied an unsupervised learning algorithm for predicting the spectrum occupancy in GSM-900 band. The performance parameters like accuracy, Mean Square Error (MSE), Root mean square error (RMSE), Mean Relative Error (MRE) are observed to bring the characteristics of predictive models mentioned in Y. Duan, Y. L.V. and F. Wang, (2016) . As the methodology shows the path for selecting the best predictive model, it must be selected based on highest accuracy.

4.2. Predictive model using LSTM

The cleaned data is normalized using the following equations.

$$X_n = (X - X_{\min}) / (X_{\max} - X_{\min}) \quad (2)$$

$$X_scaled = X_n * (X_{\max} - X_{\min}) + X_{\min} \quad (3)$$

Where, X is input samples, Xn is the normalized value generated by applying the formula over the dataset. Xmax is maximum value Xmin is minimum value in selected column. The samples in the dataset are in the range of -40 dBm to -120 dBm which is converted into the range of 0 to 1 using equation (2) and (3). The activation functions used in DL models works in the range of 0 to 1. The normalized data set is fitted to Artificial Neural Network (ANN) models for the processing. We can choose any one channel among 501 frequencies specially from uplink (890-915 MHz) and downlink (935-960 MHz). We can select a single or multiple channels manually for occupancy predictions. The performance of Markov model and Deep learning models used in this research work has been evaluated using test data set for occupancy prediction. The Deep Learning models are implemented using methodology mentioned in figure 3. After the input dataset (total samples) is split into train and test sub datasets. The size of the data set is common to train all the Deep Learning models in this work. The results of LSTM model in terms of predictions, accuracy, and MSE are shown in figure 13.

4.3. Predictive model using Autoencoder (AE)

The Autoencoder has significant importance in the prediction process of RF transmission of primary users. The Autoencoders are the special types of ANNs used for dimensionality reduction and to match the outputs with inputs to produce the optimized results. We have followed a simple implementation steps to develop

the Autoencoder model. With the declaration of python libraries, the data set is loaded with the help of pandas library function. The data is normalized from the range -40 dBm to -120 dBm to the range 0 to 1. The data needs to be de-normalize using reverse process. The input dimension is set equal to look back. The architecture of Autoencoder is built using fully connected dense layers. The number of nodes in hidden layers are chosen to optimize the output of the model (input and output nodes =128, code units=16). Performance parameters like accuracy, MSE, RMSE, MRE are calculated. The results of prediction and MSE are plotted. The results of the Autoencoder model in terms of prediction, accuracy and MSE for single channel are shown in figure 12.

4.4. Predictive model using Multilayer perceptron (MLP)

The architecture of the Multilayer perceptron model consists of two hidden layers with 128 nodes and 64 nodes respectively and three output layers as shown in figure 5. The nodes in hidden layers or number of hidden layers can be changed according to input data size to produce the better results. The data is loaded and later split into training and test data sets. The steps used in constructing the Multilayer perceptron model are as mentioned below. The input size is set to be 20 as the most suitable one. Data set is loaded with a function from pandas library. The class weights are provided so that MLP does not skew towards always predicting dominant class from the data set. We are calculating fraction of dataset of each class 0, 1 and -1. The results in terms of accuracy, predicted values and MSE are shown in figure 14.

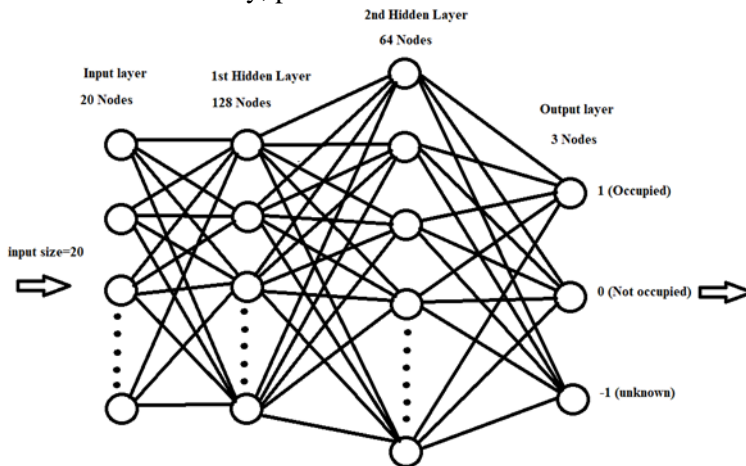


Figure 5 Feed forward Neural model using Multilayer perceptron

4.5. Predictive model using CNN

Total data samples in the input data frame are 1077. A single channel is selected for processing the input dataset. The size of input data vector would be (1077,1). Look back is 20. The shape of the dataset after applying lookback would be X=(1056,20,1). Training and test samples would be 844 and 212 respectively. X and Y

are the variables used for creating a new dataset. Whereas X has input data samples and Y has respective categorized labels for samples in X. As the architecture of CNN includes the convolutional layers, pooling layers, a flattening layer and a dense layer. Two pairs of convolutional layers and average pooling layers are used. The number of filters is set to be 5 and the filter size is chosen to be 3. The performance of the CNN model have been evaluated with parameters like accuracy, MSE, RMSE, MRE by choosing the appropriate activation function. The softmax function is chosen at the output dense layer. Dense layer_2 consists of 3 nodes to produce the probabilities of three different states. For compiling the designed model, the optimizer is used to update the weights of the neural network. Adam optimizer is found to be a popular choice among the researchers. Hyper parameters of Adam optimizer are tuned to check the performance of the model.

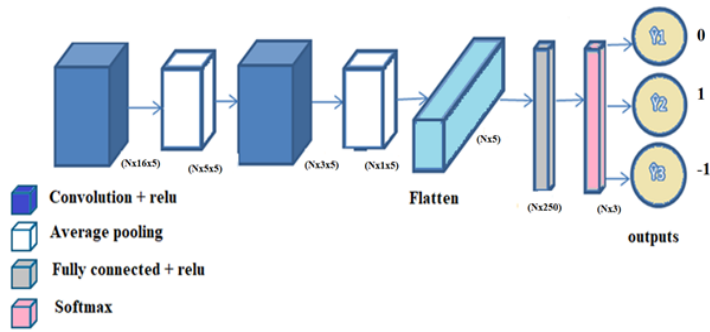


Fig. 6 Operational block schematic of CNN

The CNN operational architecture is as shown in figure 6. Major components of the model are convolution layers, average pooling layers and dense layers. The input to the model is equal to look back. The dimension of the First convolution layer is calculated

$$O1 = ((n + 2P - F)/stride) + 1 \tag{4}$$

Where, $O1$ = output of first convolution layer, n = input dimension (look back), P = number of padding, F = filter size. In this case, the stride is 1. When we set n = length of look back = 20, $P= 0$, $F=5$ then the output dimension would be $((20 + 2 * 0 - 5) / 1 + 1=16$ so the output would be $(N, 16,5)$. Here N is batch size as Keras process the data in batches. In this dimension, 5 is the number of filters. Likewise, the output of the next stage is calculated. For generating the probabilities from 3 output nodes of dense layers. These nodes produce the probabilities of occupancy, non-occupancy and unknown states. CNN model is evaluated for single channel as well as multiple channels. Figure11 (a),(b),(c) shows the prediction, accuracy and MSE responses produced by CNN predictive model.

5. Results and Discussion

The results of the performance of predictive models are compared to come up with

better predictive model for GSM band. The model parameters are listed in table 1. It shows that the prediction accuracy of CNN model is preferably greater than other models.

5.1. Comparative Studies of Markov and Deep Learning Models

Deep learning models are compared with the Markov chain model based on the performance parameters mentioned in table 1 and table 2. Table 1 shows the performance parameters of single channel and table 2 shows the performance parameters for the multiple channels. The dataset from location 1 has been chosen.

Table 1 Performance parameters of Markov & Deep learning predictive models for single channel

Sr. No.	Parameters	LSTM	CNN	MLP	AE	Markov
1	Accuracy-Training	53.89%	71.33%	47.74%	71.71%	58.07%
2	Accuracy-Test	48.69%	69.34%	44.48%	69.33%	46.76%
3	MSE-Training	0.06	0.14	0.99	0.0030	5.44
4	MSE –Test	0.11	0.14	0.90	0.0029	4.97
5	RMSE-Training	0.24	0.58	0.89	4.39	2.33
6	RMSE-Test	0.33	0.61	0.88	4.74	2.23
7	MRE-training	0.12	0.20	0.31	0.04	0.42
8	MRE-test	0.19	0.23	0.29	0.05	0.35

Table 2 Performance parameters of Markov and Deep Learning predictive models for multi-channels

Sr. No.	Parameters	LSTM	CNN	MLP	AE	Markov	LSTM	CNN	MLP	AE	Markov
		25 Uplink channels					25 Downlink channels				
1	Accuracy-Training	99.50%	99.46%	99.49%	99.38%	98.91%	95.39%	85.01%	85%	83.93%	74.35%
2	Accuracy Test	99.37	99.52	99.41	99.30	98.94	95.22	84.80	84.85	83.79	84.58
3	MSE-training	0.0020	0.0035	0.0035	0.12	4.01	0.0319	0.086	0.0852	0.135	5.10
4	MSE-test	0.0031	0.0032	0.0043	0.1269	4.01	0.0122	0.087	0.0864	0.157	3.84
5	RMSE-training	0.05	0.09	0.09	30.59	2.0	0.18	0.39	0.39	29.99	2.26
6	RMSE-test	0.06	0.10	0.12	23.75	2.0	0.11	0.39	0.39	34.61	1.96

7	MRE-training	0.00	0.00	0.00	0.02	0.99	0.07	0.14	0.06	0.03	0.128
8	MRE-test	0.01	0.00	0.00	0.02	0.99	0.04	0.14	0.06	0.03	0.045

The channel no. 300 (949.8MHz) has been chosen for the single channel evaluation of performance of the models. Uplink channels were selected from 890 MHz to 894 MHz and downlink channels were selected from 955 MHz to 960 MHz. Input matrix size for these DL models was (1077, 1). For multiple channels the size would be (1077, 25). When we apply the look back, the size of data frame would change. The new size of data frame would be a single column vector of size (26925, 1). The table 1 shows the performance parameters of Deep Learning models for single channel and multi channels. Figure 7 shows accuracy comparative response for downlink channels and figure 8 shows accuracy response of uplink channels.

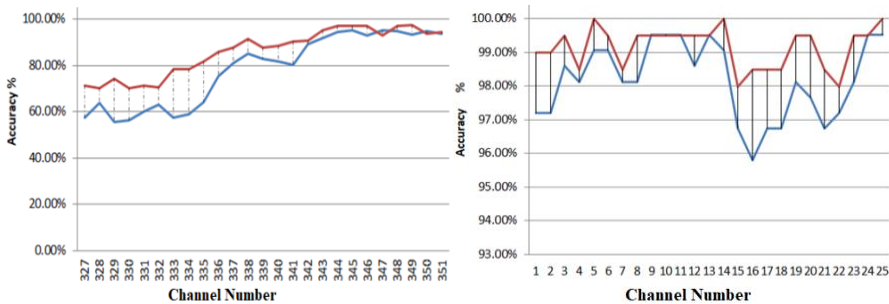


Figure 7 Comparative prediction accuracy of Markov and CNN (downlink)

Figure 8 Comparative Prediction accuracy of Markov and CNN (Uplink)

6. Comparative results of Markov and other Deep Learning models

Figure 9 shows the performance of models for a bunch of uplink channels. The average accuracy of the CNN model is 99.15% which is considerably good. Where as in Figure 10 the CNN model accuracy was found to be 86.04%.

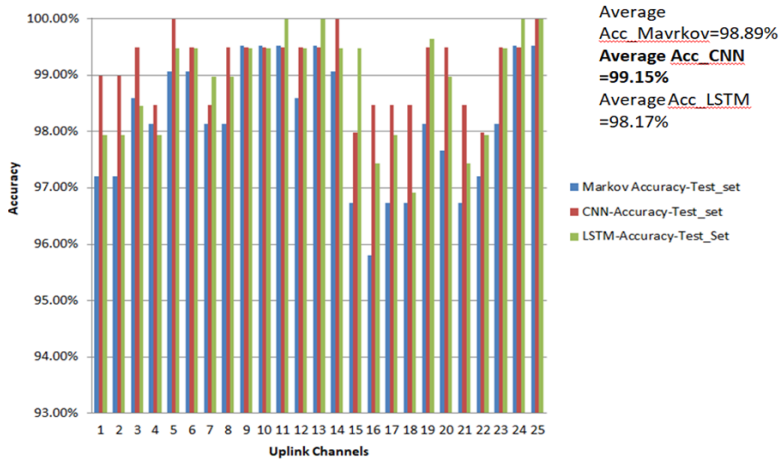


Figure 9 Accuracy comparison for uplink

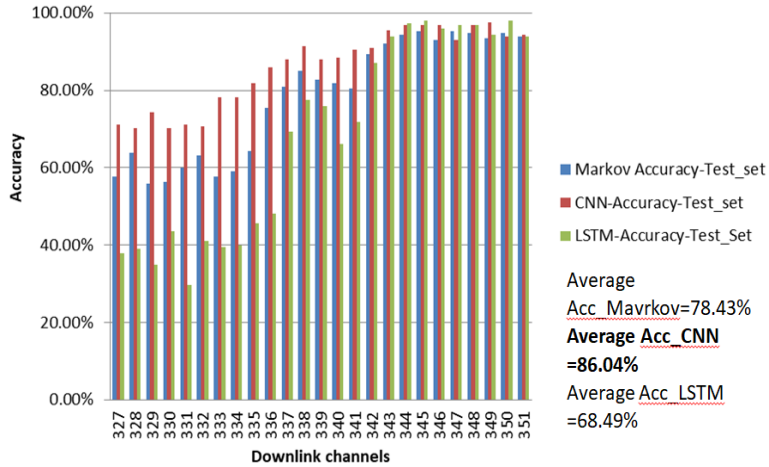


Figure 10 Accuracy Comparison for downlink

The performance of Deep Learning models are verified with single and multiple channels. In order to know the single channel behavior, the response of each model for a single channel (channel 949.8MHz) is observed.

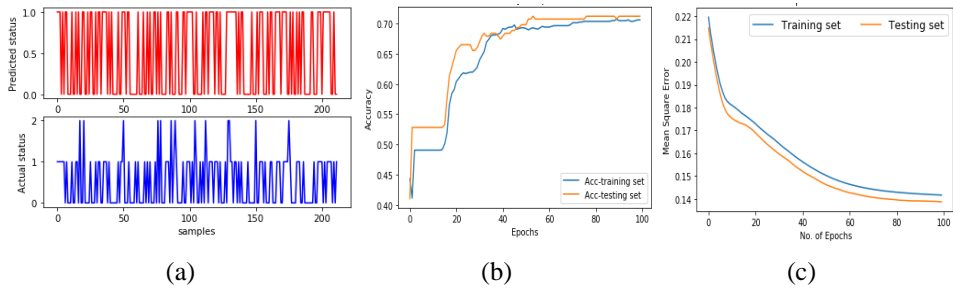


Fig. 11 Single channel CNN response (a) prediction response, (b) accuracy response, (c)

MSE response

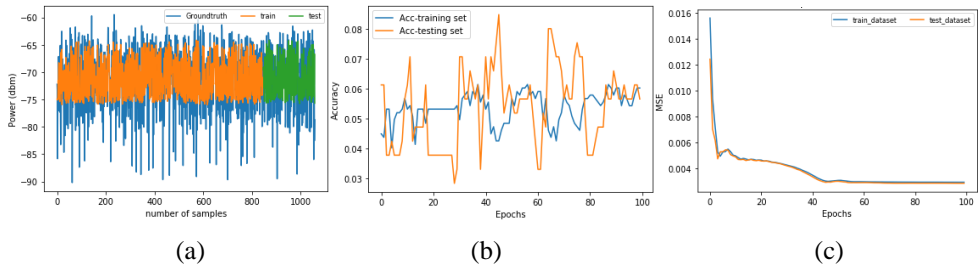


Fig. 12 Single channel Auto-encoder response (a) prediction response, (b) accuracy response, (c) MSE response

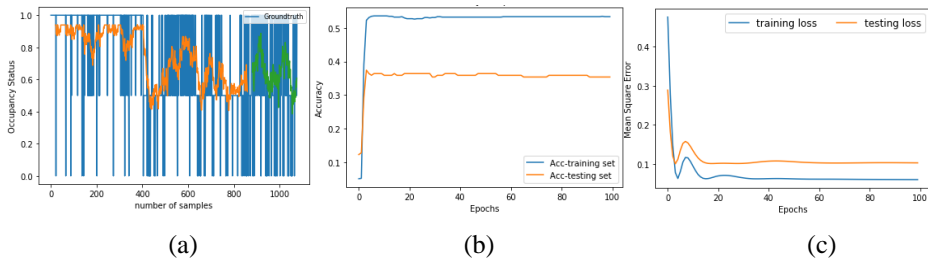


Fig. 13 Single channel LSTM response (a) prediction response, (b) Accuracy response, (c) MSE response

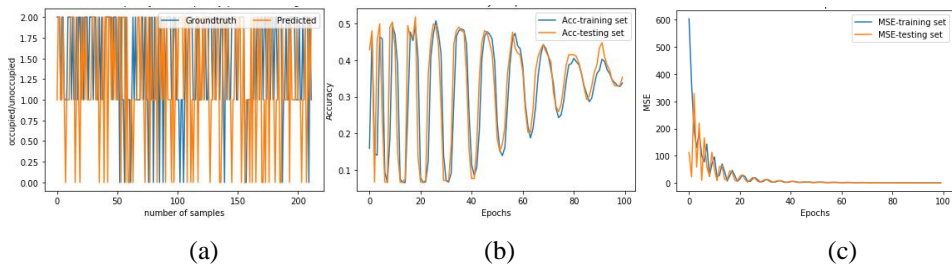


Fig. 14 Single channel MLP response (a) prediction response, (b) accuracy response, (c) MSE response

The performance of Deep Learning models are verified with single and multiple channels. In order to know the single channel behavior, the response of each model for a single channel (channel 949.8MHz) is observed. It is observed that CNN model do not show any predictions in the range 1 to 2 as data labeling is done as per the range of power levels in the data set. 100 is labeled for signals having power level below -80 dBm, 010 is labeled for the signals having power between -70 dBm to -80 dBm and 001 is labeled for signals having power level above -70 dBm. 0,1 and 2 are indexing values of these three categories. The data points are plotted according to actual power level.

7. Conclusion

After the simulations for frequency analysis to know the status of the occupancy in each channel, it was found that the Uplink of the GSM band viz. 890 to 915 MHz is mostly remaining vacant whereas downlink 935 to 960 MHz is always occupied. The approach used in this research work could be a helping hand to know the occupancy status of the GSM band based on geographical location for cognitive radio application. ANN methods proved to be effective in predicting the occupancy in the licensed band based on historical data. These types of predictive methods can be useful in the upcoming cognitive radio systems in spectrum sensing operation to know the status of white spaces. The predictive models should be implemented in appropriate hardware for better CR operation.

When we compare the performance of deep learning models, it is found that CNN and LSTM show comparatively better results for both uplink and downlink channels. The actual test data set is compared with the predicted data set statistically rather than as a time sequence. The exact prediction is nearly impossible due to the random nature of RF signals. The performance of predictive models using Deep learning techniques improves by increasing input data size. It is possible to extend the Markov model to consider the states in previous several instances, and then use a Reinforcement Learning (RL) algorithm to predict the channel occupancy.

References

- Adebola, E. and Annamalai, A., (2014). Unified analysis of energy detectors with diversity reception in generalised fading channels. *IET Commun.*, 8 (17), 3095–3104.
- Akbar, I. A. and Tranter, W. H., (2007). Dynamic spectrum allocation in cognitive radio using hidden Markov models: poisson distributed case, *Proceedings 2007 IEEE SoutheastCon*, 196-201.
- Atapattu, S., Tellambura, C. and Jiang, H., (2011). Energy detection based cooperative spectrum sensing in cognitive radio networks, in *IEEE Transactions on Wireless Communications*, 10 (4), 1232-1241.
- Bidwai Sandeep, Joshi Nikhil and Bidwai Saylee, (2018). LSTM model for channel occupation prediction in GSM band, *Publisher: IEEE*, 331-335.
- Bidwai Sandeep, Joshi Nikhil, Bidwai Saylee and Wali Uday, (2019). Deep learning predictive models for cognitive radio system, 3rd Intconf on Data Engineering & communication systems, RNSIT, Bengaluru, Karnataka, India, 19-20th dec. 2019.

International Journal of Innovative Technology and Exploring Engineering (IJITEE), 9(2S), 491-496.

Bidwai, S., Mayannavar, S. and Wali, U.V., (2021). Performance comparison of Markov Chain and LSTM models for spectrum prediction in GSM bands. In: Joshi A., Khosravy M., Gupta N. (eds) Machine Learning for Predictive Analysis. *Lecture Notes in Networks and Systems*, 141.

Bidwai Sandeep, Wali Uday V., Shirgapur Bahubali and Bidwai Saylee (2016). Detecting white spaces for cognitive radio, *International Journal of Technology and Science*, 9 (2), 8-11.

DaniiloLópez, Johana Hernández and Edwin Rivas, (2016). Algorithm and software based on multilayer perceptron neural networks for estimating channel use in the spectral decision stage in cognitive radio networks, World Academy of Science, Engineering and Technology, *International Journal of Computer and Information Engineering*, 10 (12).

Digham, F. F., Alouini, M. S. and Simon, M. K., (2003). On the energy detection of unknown signals over fading channels, *IEEE International Conference on Communications*, 55 (1), 21-24.

Duan, Y., L.V. Y. and Wang, F., (2016). Travel time prediction with LSTM neural network, *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*, 2016, 1053-1058.

Geirhofer, S., Tong, L. and Sadler, B. M., (2007). Cognitive radios for dynamic spectrum access - dynamic spectrum access in the time domain: modeling and exploiting white space, in *IEEE Communications Magazine*, 45 (5), 66-72.

Gers, F.A., Schraudolph, N.N. and Schmidhuber, J., (2003). Learning precise timing with LSTM recurrent networks, *Journal of Machine Learning Research*, 3, 115-143.

Ghosh, C., Cordeiro, C. D., Agrawal, P. and Rao, M. B., (2009). Markov chain existence and Hidden Markov models in spectrum sensing, *2009 IEEE International Conference on Pervasive Computing and Communications*, 1-6.

Gismalla, E.H. and Alsusa, E., (2011). Performance analysis of the periodogram-based energy detector in fading channels, *IEEE Transactions on Signal Processing*, 59 (8), 3712-3721.

GitHub link <https://github.com/profssbssb/GSM-datasets.git>.

Han, D. et al., (2017). Spectrum sensing for cognitive radio based on convolution neural network, *2017 10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*, 1-6,

He, Y. et al., (2017). Deep-Reinforcement-Learning-Based Optimization for Cache-Enabled Opportunistic Interference Alignment Wireless Networks, in *IEEE Transactions on Vehicular Technology*, 66 (11), 10433-10445.

Kostylev, V. I., (2002). Energy detection of a signal with random amplitude, *IEEE International Conference on Communications*, 3, 1606–1610.

Lan Kunwei, Zhao Hangsheng, Zhang Jianzhao, Long Cao and Luo Menglin, (2014). A spectrum prediction approach based on neural networks optimized by genetic algorithm in cognitive radio networks”, *10th IntConf on Wireless Communications, Networking and Mobile Computing – (WiCOM 2014) IET Digital Library*, Beijing, China.

Li, Y., Dong, Y., Zhang, H., Zhao, H., Shi, H. and Zhao, X., (2010). Spectrum usage prediction based on high-order Markov model for cognitive radio networks, *2010 10th IEEE International Conference on Computer and Information Technology, Bradford*, 2784-2788.

Omotere, O., Fuller, J., Qian, L. and Han, Z., (2018). Spectrum occupancy prediction in coexisting wireless systems using deep learning, *2018 IEEE 88th Vehicular Technology Conference (VTC-Fall)* , 1-7.

Park, C., Kim, S., Lim, S. and Song, M., (2007). HMM based channel status predictor for cognitive radio, *2007 Asia-Pacific Microwave Conference, Bangkok*, 1-4.

Pattanayak, S., Venkateswaran, P. and Nandi, R., (2012). Artificial neural networks for cognitive radio: a preliminary survey, *2012 8th International Conference on Wireless Communications, Networking and Mobile Computing*, 1-4.

Rondeau, T. W., Rieser, C. J., Gallagher, T. M., and Bostian, C. W., (2004). "Online modeling of wireless channels with hidden Markov models and channel impulse responses for cognitive radios," *2004 IEEE MTT-S International Microwave Symposium Digest (IEEE Cat. No.04CH37535)*, 2, 739-742.

Saleem, Y. and Rehmani, M.H., (2014). Primary radio user activity models for cognitive radio networks: a survey, *Journal of Network and computer applications*, 43, 1-16.

Wang, Y., Zhang, M. and Shi, Q., (2014). Unified approach to performance analysis of energy detection in generalised fading channels, *Electron. Lett.*, 50 (11), 837–839.

Xing, X., Jing, T., Huo, Y., Li, H. and Cheng, X., (2013). Channel quality prediction based on Bayesian inference in cognitive radio networks, *2013 Proceedings IEEE INFOCOM*, 1465-1473.

Zhang, C., Patras, P. and Haddadi, H., (2019). Deep learning in mobile and wireless networking: a survey, in *IEEE Communications Surveys & Tutorials*, 21 (3), 2224-2287.

Zhao, Q. and Swami, A., (2007). A decision-theoretic framework for opportunistic spectrum access, in *IEEE Wireless Communications*, 14 (4), 14-20.