2018-2019 M.Sc. in Data Science and Analytics

Ryerson University
Data Science & Analytics Graduate Studies

EXPERIMENTAL RESULTS ON THE IMPACT OF MEMORY IN NEURAL NETWORKS FOR SPECTRUM PREDICTION IN LAND MOBILE RADIO BANDS

Ozan Ozyegen Supervisor: **Prof. Ayse Basar Bener**

Objective

The objective of this project is to measure the impact of memory in Neural Networks for Spectrum Prediction in Land Mobile Radio Bands

Background

In the literature, there are many proposed models and approaches to predict the occupancy status in a channel. Some of the oldest prediction models are linear prediction models where future values are predicted as linear function of past observations. Due to their simplicity they are very popular and have been widely used in various fields, including spectrum prediction. Most common models include auto-regressive (AR), moving average (MA), and auto-regressive integrated moving average (ARIMA) models [10]-[15].

Another popular family of prediction methods are ANNs, which represent a class of flexible non-linear models [15]-[20]. Yin et al. proposed using a TDNN o predict the max-min normalized power measurements of channels [19] and report an improved root-mean-squared error (RMSE) compared to using ARIMA.

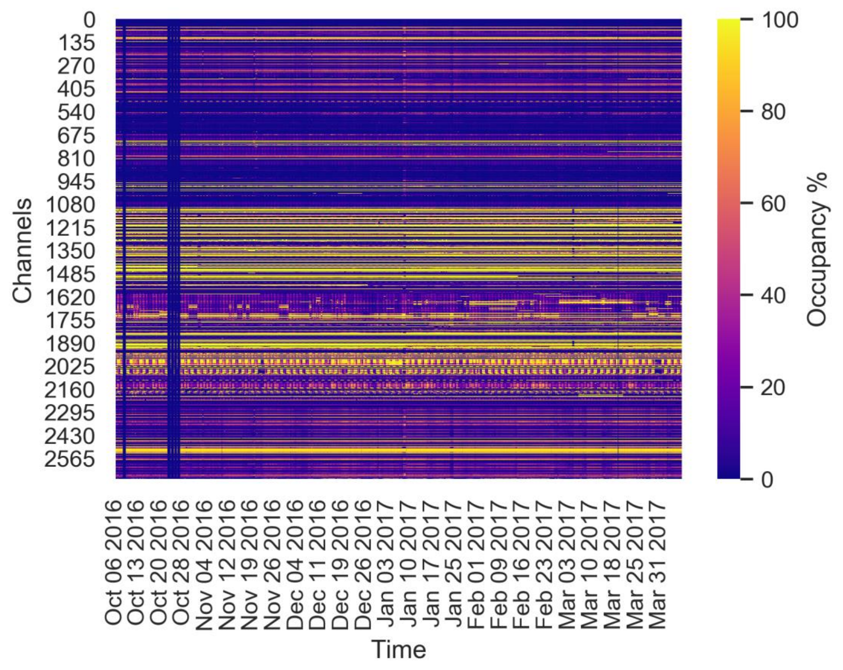
Methods such as reservoir learning [23] and long short term memory (LSTM) networks [24], [25] are used in simulated and real data to alleviate the learning problems. Dalgkitsis et al. compare LSTM performance to ARIMA and radial basis functions and found out that LSTM performs best on forecasting traffic of cellular networks [25].

The majority of the research in the literature use simulated data, assuming the traffic can be reasonably modeled through either Poisson or similar distributions [19], [27], [28]. Real life studies on the other hand concentrate on cellular traffic [24]. This might not be representative of the traffic behavior in other parts of the radio spectrum such as those of commercial dispatch services to public utilities, which are some of the users of the LMR bands. In these bands and use cases, there is a scarcity of real-life measurements and the evaluation of the performance of spectrum prediction models in the LMR spectrum.

Methodology

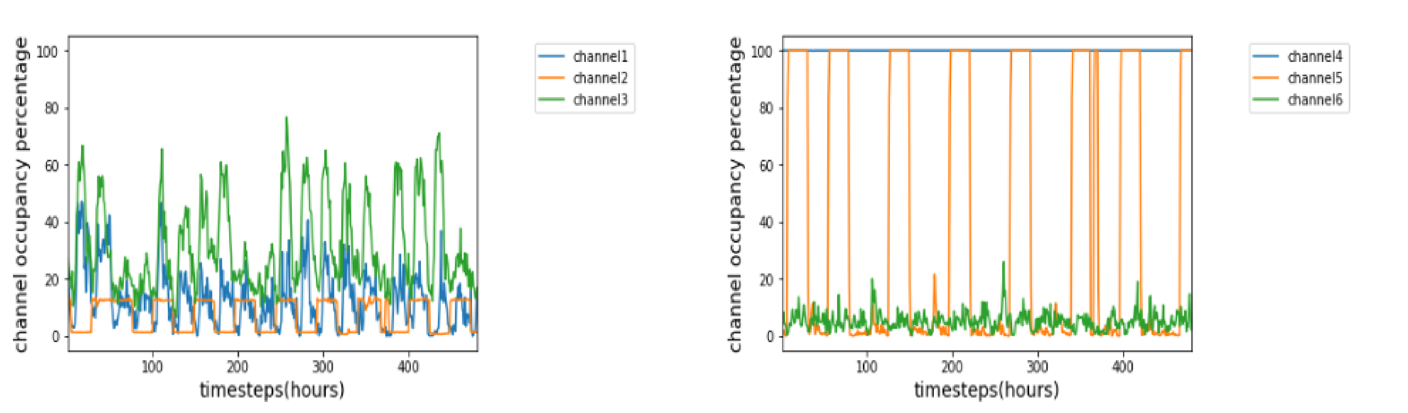
Data

The dataset used in this project consists of measurements made on the 8.663 LMR channels within LMR spectrum bands in downtown Ottawa between October 27th 2016 and April 26th 2017.



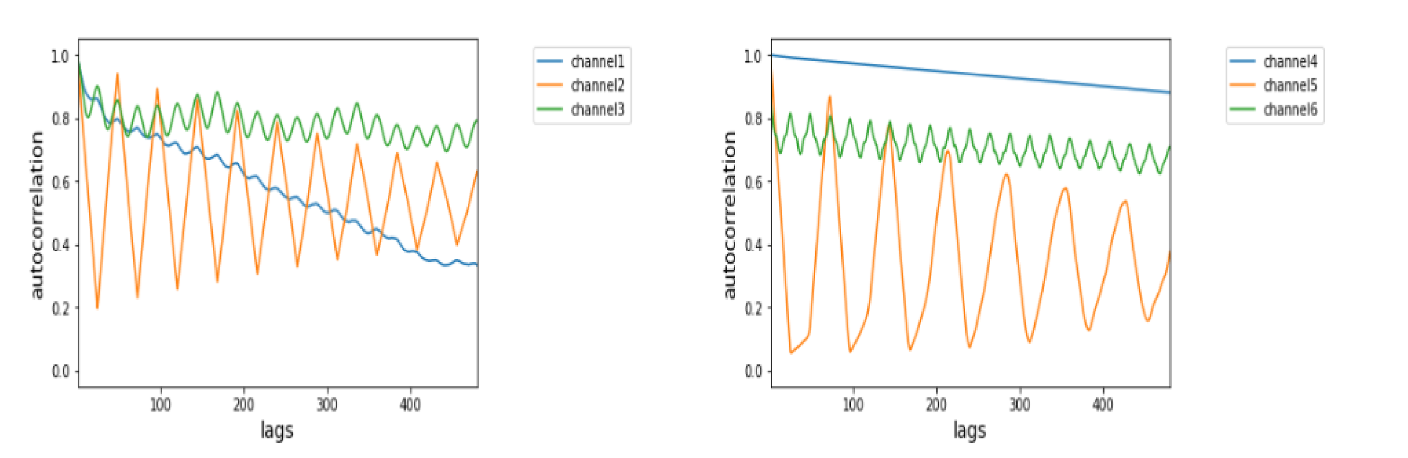
To address the missing values in our data, we have selected channels that report occupancy measurements most of the time (+95%). This leaves us with 2684 channels. This threshold was chosen to leave us a sizable channels list (approximately 31% of all channels) while removing most of the incomplete observations from the data set.

The heat-map shows the occupancy percentage values of the channels which are more than %95 complete. Brighter colours mean higher occupancy. Vertical dark lines indicate the time ranges where no data is recorded for any of the channels due to hardware issues.



(a) The occupancy percentage of three random channels (b) The occupancy percentage of three random channels

Fig. 2: The occupancy percentage of six random channels



(a) The auto correlation functions of three random channels. (b) The auto correlation functions of three random channels.

Fig. 3: The auto correlation functions of six random channels.

We have done some preprocessing on our data to understand temporal patterns, which would guide our model selection. Figures on the left illustrate the temporal occupancy behavior of 5 randomly selected channels over 240 hours. An initial look indicates possible periodic behavior in some of the channels, e.g., channel 2. These observations generalize for all channels in our dataset.

Results

Time Delay Neural Network (TDNN)

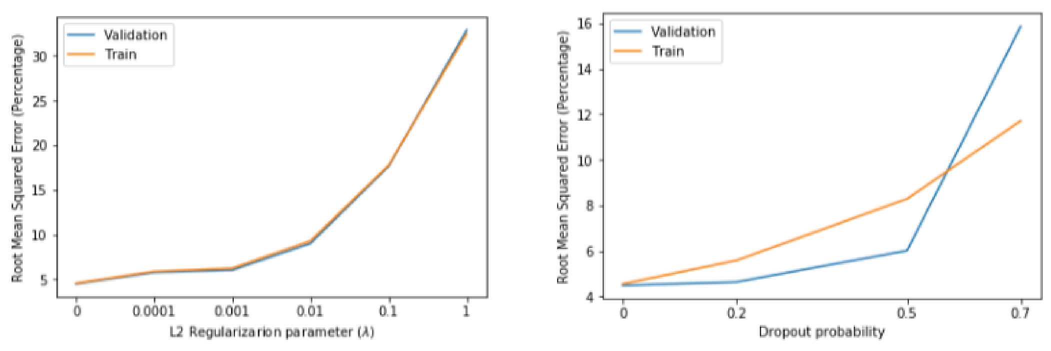
Neural networks (NN) are non-linear and non-parametric models composed of one input layer X, one or multiple hidden layers Z_j and one output layer Y. The topology or the architecture of a feed forward NN is defined by the number of units in each layer. A sliding window method is used for framing the data set. A through hyperparameter tuning is done to find out the best performing TDNN model.

Learning rate	First order momentum	Second order momentum	Epsilon
1e-4	0.9	0.95	1e-7
1e-3	0.95	0.99	0.001
1e-2	0.99	0.999	0.01

Different set of parameters for Adam Optimizer

Number of hidden units (Layer 1)	Number of hidden units (Layer 2)	Number of hidden units (Layer 3)	Activation functions
50	50	0	Sigmoid
100	100	100	'Tanh'
200	200	400	'ReLU'
400	400	-	-

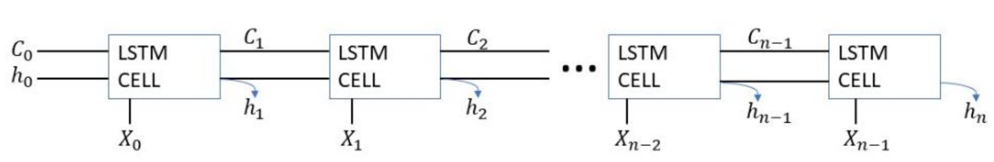
Different set of parameters for TDNN model structure



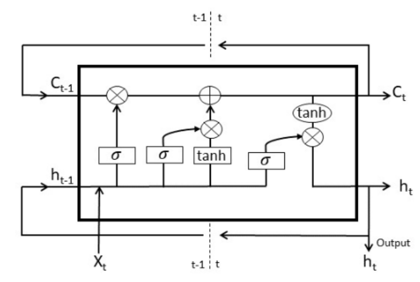
RMSE for training and validation loss for different values of L2 and Dropout probabilities

Long Short Term Memory Network (LSTM)

LSTMs are widely used in literature for time series analysis. These networks are able to capture long temporal dependencies and have a memory component which makes them superior to TDNNs and RNNs in most settings.



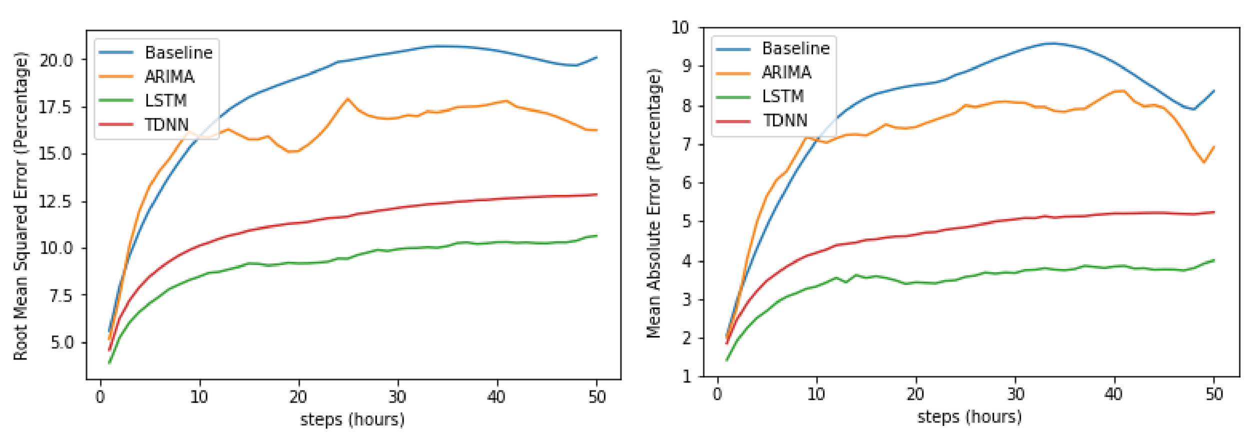
LSTM Layer Unrolled



LSTM Cell

Results

We present the performance of four models measure in MSE and numerical representation of the plots sampled every 5 steps. The LSTM showed the best prediction performance, followed by TDNN and seasonal ARIMA models. We also provide results for the baseline as it serves as an indication on how much both models are able to improve upon simple delay based guessing prediction regime.



Average performance of four predictive models

Model	$\Delta t = 1$	$\Delta t = 6$	$\Delta t = 11$	$\Delta t = 16$	$\Delta t = 21$	$\Delta t = 26$	$\Delta t = 31$	$\Delta t = 36$	$\Delta t = 41$	$\Delta t = 46$
Baseline	7.92	13.76	16.88	18.41	19.38	20.11	20.56	20.64	20.23	19.68
ARIMA	7.37	14.69	16.06	15.91	15.9	17.02	16.98	17.47	17.46	16.74
TDNN	6.24	9.25	10.47	11.1	11.46	11.85	12.23	12.47	12.64	12.74
LSTM	5.21	7.8	8.72	9.05	9.2	9.74	9.98	10.28	10.26	10.28

Results for average RMSE loss in predictions for different time horizons

Conclusions

In this report we a predictive LSTM model and evaluated its performance against a seasonal ARIMA model and a TDNN model in predicting spectrum occupancy levels in the next 50 hours with respect to each other as well as a baseline method that simply used previous values as its prediction.

The results indicated that the memory component of LSTMs, a type of recurrent neural network designed to work with sequential data, has improved the prediction performance consistently across all time horizons. While all models showed promise in predicting the next few hours, spectrum allocation in IoT environment requires longer term planning. By learning the long term dependencies in the time series data, LSTM models are able to outperform both TDNN, the simplest NN variant also designed with time series, and the seasonal ARIMA, a long established linear prediction strategy, in making predictions in medium to long term horizons.

The results discussed in this paper are still limited to predicting a little further than two days into the future, while some IoT requirements could need planning for weeks or months in advance. While this was, in part, a limitation of the hourly dataset we have used in our study, methods that are able to work with different time resolutions is still a very much active research question that should be further pursued. Finally, the results discussed in this paper are applicable to LMR bands which are narrow band channels that are used for various public safety response systems as well as some commercial dispatch services. Various spectrum bands naturally exhibit different usage patterns, which could impact the performance comparison of the prediction models discussed here and would provide natural directions for the extension of this study.