Predicting vehicle parking space availability using multilayer perceptron neural network

M H Ismail¹, T R Razak¹, R A J M Gining¹, S S M Fauzi¹ and A Abdul-Aziz¹
¹Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA (UiTM), 02600 Arau, Perlis, Malaysia.

*Corresponding author: mohammadhafiz@uitm.edu.my

Abstract. In this study, we have investigated potential use of Multilayer Perceptron (MLP) to predict parking space availability for use within Field Programmable Gate Array (FPGA) accelerated embedded devices. While previous studies have explored the use of MLP for classification problem in FPGA, very little studies concentrated on the potential use of MLP in regression problem, especially in parking space forecasting. Therefore we formulated five Multi-Layer Perceptron (MLP) models with varying hidden units to perform single-step prediction to forecast parking space availability within the next 15 minutes based on the previous one-hour parking occupancy. The proposed models were trained on the historical data of Kuala Lumpur Convention Center dataset and evaluated against baseline ARIMA models. The results have shown that our proposed MLP model performed relatively well against baseline model with the root mean square error between (RMSE) 78.25 to 78.41 and mean absolute error (MAE) between 37.02 to 39.17.

1. Introduction

For city dwellers, finding a parking space is critical especially in a densely populated area such as the city centre. Therefore, modern commercial buildings have installed vehicle parking sensors as part of their Parking Guidance and Information System (PGIS) to detect parking space availability and utilisation. The availability of parking spaces is communicated to drivers through electronic variable message signboard (VMS) which is controlled with a series of embedded computing devices. The information provided by PGIS assists drivers in planning their route by reducing the time spent on searching and waiting for the parking space to become available. This is supported by [1], which states drivers who have prior knowledge on space availability are more successful in finding a parking spot for their vehicle compared to those who do not.

Consequently, drivers waste an average between 3.5 to 14 minutes just to find a vacant parking spot [2], which also inadvertently waste fuel and causes unnecessary pollution [3]. Reducing search time leads to less vehicle movements in the parking area, which subsequently reduces traffic congestion and noise pollution, especially near parking entrances [4][5].

PGIS detects availability of parking space through a series of embedded computer sensors placed on the parking lot. The sensors detect parking lot vacancy and send the information to a centralised computer. The parking information is stored in the database and then displayed on the variable message signboard (VMS) via computer networks at various location around the vehicle parking complex. Some PGIS posts parking information on the website or display it on mobile application to further assist drivers in finding available parking spots.
However, through our findings, most PGIS only report current parking spot availability in real-time or periodically. While this information is helpful for drivers who are already in the vicinity of parking complex, it is less useful for drivers who are still planning for a trip to the city centre. Thus, the parking space availability forecast is more useful as it could help drivers in the latter category to estimate the space availability in advance before they reached the parking complex. Moreover, parking space forecast can help drivers avoid congestion by assisting them in planning their route.

The advent of edge computing enables data to be processed directly on the edge with minimal dependency on remote computers. This is crucial as on-site forecasting can eliminate network delay problems associated with remote processing. Additionally, an implementation of PGIS which can perform on-site forecasting is preferable because it can operate independently in event of network disruption.

Recent development in electronics have enable the use of Field Programmable Gate Array (FPGA) in embedded computers. The introduction of FPGA allows embedded devices to perform complex mathematical tasks such as classifications or regression on raw data without depending on remote processing from cloud computers [6]. According to [7][8], Artificial Neural Network (ANN) is one the most used machine learning algorithms for classification tasks in FPGA. This is due to the parallel processing capabilities of FPGA which allows ANN models to be executed more efficiently [8]. In addition, Multilayer Perceptron Network (MLP) is a type of feedforward ANN which is commonly implemented on FPGA. Due to its flexibility, MLP is also used in classification problems as well as regression tasks.

While previous studies have explored the potential of implementing Multilayer Perceptron Network (MLP) in FPGA for various near real-time classification problems [9][10][11][12][13], very little research has been done in the area of regression, especially in the area related to parking space availability forecasting. Moreover, other studies on parking space forecasting implementations do not specifically target FPGA-based platforms. Thus, this study intends to focus on parking space forecasting problems which can be implemented on FPGA.

Therefore, we propose short-term parking space prediction model with MLP. Our goal is to determine the forecasting performance of MLP before implementing the network on embedded computer powered by FPGA. The model should be able to perform single-step prediction based on the previously seen one-hour parking occupancy records.

2. Literature Survey

Current research in the field has delved into predictive methods for short-term parking space availability forecasting intended for drivers. Yang et al. [15] suggested the use of neural network on historical data time series for short term parking space prediction as one of the methods to improve the efficiency of PGIS. The authors suggested a three-layer neural network model which consist of traffic flow, weather, events, and parking garage availability as multivariate input. The authors noted that this approach depends on a large historical data sample size which may be difficult to obtain and the difficulty in determining the suitable number of neurons to get accurate prediction.

Later work by Yanjie, Dounan, Blythe, Weihong and Wei [16] uses Wavelet neural network to predict short-term parking space availability. The authors trained the Wavelet Neural Network on time-series data with 1-minute interval and compared its performance to Lyapunov exponents method. The authors found out that Wavelet Neural Networks is more precise in one-step forecasting (1-minute look ahead) in regular days.

Yanxu et al. [3] compared Regression Tree, Support Vector Regression, and Neural Network Model performance in predicting parking occupancy and found the Regression Tree performed better in prediction performance than Neural Network and Support Vector Regression model. The authors also found that the computational complexity of the Neural Network model is between the Regression Tree and the Vector Regression Support Model.

The use of nonlinear auto-regressive Neural Network (NARNN) model has been proposed by [17] for predicting parking space utilization within a large time window of more than half of the day. The implementation uses a recurrent feedback connection within the network unit. However, the results are inconclusive as the dataset used in the study is limited.
Fengquan et al. [18] proposed ARIMA model to predict unoccupied parking space in real-time. This study was performed on a dataset collected from a mall parking lot. The dataset contains 2,880 datapoints representing unoccupied parking space in 15-minutes interval. The study found that the ARIMA (2,1,3) performs slightly better than the Neural Network model with a Root Mean Square Error (RMSE) of 4.47 compared to 5.45.

The use of Long Short-Term Memory (LSTM) networks and Recurrent Neural Network in parking space availability prediction has been explored by [19]. The study uses parking occupancy dataset consists of twelve million parking events recorded by various sensors in a year. The parking events data is converted into a time-series. The model can predict parking space occupancy within 1 minute, 5 minutes, 15 minutes, and 30 minutes ahead of time.

A Deep LSTM was proposed by [20] as part of smart car parking availability architecture. The Deep LSTM consist of multiple stacked LSTM, in which the output of a layer is passed to the next layer. The dataset consists of data collected from thirty parking locations within a period of 8 months. The model was trained on a time series with 1-hour interval, and 1-hour look-ahead prediction time.

Provoost et al. [21] compared the ANN, Convolutional Neural Network (CNN) and Random Forest (RF) performance in predicting the parking space occupancy from Web of Things application. The optimum ANN number of hidden layers and neurons was determined using grid search technique. The Convolutional Neural Network uses the same number of hidden layers and neurons as the ANN with a 4x4 kernel used as look-back window. Grid search is also used to determine the maximum features and optimum depth of the RF model. The study was performed on a dataset consisting of parking occupancy information, traffic flow, date, time of the day and weather conditions. However, the authors found out that time of the day and weather conditions has smaller impact on prediction accuracy.

Barreto et al. [22] has found that Field Programmable Gate Array (FPGA) enables edge computing approach of processing sensor data with Convolutional Neural Network (CNN) on the embedded computer itself. Furthermore, [23] have found that FPGA is more energy efficient compared to GPU in executing matrix multiplication and workload typical for Neural Network computing within edge computing use-case. Therefore, it is accepted that Multi-Layer Perceptron neural network (MLP) is most suitable to be implemented on edge computing device that are integrated with FPGA.

From the literature survey being done, it can be concluded that prior studies have proposed different approach for predicting parking space availability. While the previous studies’ focus is centred around forecasting performance of parking spots through various machine learning models, our proposed approach concentrates more on evaluating MLP forecasting model that are more suitable to be implemented on edge computing device.

3. Methodology

3.1. KLCC Parking Occupancy Dataset
This research uses Kuala Lumpur Convention Center (KLCC) parking availability data obtained from [24]. The original dataset comprises of 47,603 parking occupancy records from June 2016 to November 2017. The dataset attributes consist of sequence number, parking availability, and timestamp. The parking availability data is observed within 15-minutes interval.

For this study, we truncate the dataset to only include observation between June 2016 and July 2016, thus we only concentrated on the first 3,264 rows. Table 1 presents the sample data taken from the dataset.
Table 1. KLCC Parking Occupancy Dataset.

| Sequence | Space | Timestamp         |
|----------|-------|-------------------|
| 1        | 1642  | 6/1/2016 10:12    |
| 2        | 1609  | 6/1/2016 10:15    |
| 3        | 1458  | 6/1/2016 10:30    |
| ...      | ...   | ...              |
| 3197     | 1870  | 7/4/2016 17:45    |
| 3198     | 2013  | 7/4/2016 18:00    |

The “space” column denotes the number of available parking space at observation time. KLCC has 5,500 total parking space. However, the KLCC PGIS will output “FULL” if there are no more parking space available and “OPEN” if there is a problem with the sensor reading. For this study, any instance of “OPEN” will be treated as missing value.

3.2 Dataset Preprocessing and Transformation

The dataset is then preprocessed where missing values (labeled as “OPEN”) are removed and non-numeric string value of “FULL” is replaced by 0 to denote no parking space available. This results in 145 rows removed from the dataset, resulting in a timeseries with 3,119 data.

The Augmented Dickey-Fuller unit root test was performed to determine the timeseries dataset stationarity. The Null hypothesis (H₀) of the test is the timeseries dataset is non-stationary and will require differentiation. The Alternative hypothesis (H₁) is the timeseries is stationary. The p-value of ADF test is 0.001 which less than 0.05, and thus the H₀ is rejected and the timeseries is deemed to be stationary. The parking availability data is then normalized so that the value is within 0 to 1. Normalization is performed to make the time-series data suitable to be trained in MLP.

3.3 Multi-Layer Perceptron Model Design

A Multilayer Perceptron (MLP) is a type of Artificial Neural Network (ANN) which comprise of layers of neurons and their connections. Generally, the MLP layers are divided into input layers, hidden layers, and output layers.

The number of neurons in input and output layers within MLP model depends on the specific problem domain or tasks that we are trying to solve. Since the parking records observation are recorded in 15-minutes interval time, we have decided to design a single-step forecasting model which are able to predict the number of available parking space in the next 15-minutes based on the previously seen one hour (60 minutes) observation of parking occupancy data. Thus, the input unit is set to four (60 minutes = 15 minutes interval x 4). The model only has one output because we are constructing single-step prediction model.

Therefore, the input layer for the MLP model is set to four input and one output. The optimum number of hidden units is depending on the number of input units (which is four) and are constrained by the limited memory resources of embedded computing environment with FPGA [25].

Based on this consideration, we have decided to build five MLP models 8, 16, 24, 32 and 40 hidden units respectively and evaluate their performance. Figure 1 illustrates the MLP model used in our study.
The MLP models use the rectified linear unit (ReLU) activation function as it is faster than sigmoid activation function [26], [27]. ReLU also perform well on regression task in MLP [28].

### 3.4 Model Training and Evaluation

The timeseries data is split training (2,089 data points) and validation set (1030 data points) with 0.67 ratio. The timeseries data is then trained on each MLP models using mean squared error loss function and Adam optimizer. Adam optimizer is chosen because it is computationally efficient and require small memory to implement [29] which is an important factor to consider when implementing the model on resource-constrained edge computing device.

The training is facilitated using [21] and [22] as backend. Due to the stochastic nature of neural network, each training and validation run would produce a slightly different result. Thus, we use a seed value of 12 to ensure reproducibility and consistency.

We use the root square error (RMSE) and mean absolute error (MAE) as an evaluation metrics to compare the models' performance against our chosen baseline models, ARIMA (4,0,0) and ARIMA (4,0,1). ARIMA was chosen as baseline as it is a commonly used regression model in forecasting task for stationary time-series. Moreover, ARIMA conveniently includes integrated autoregressive functionality and moving average component which are useful features for creating a baseline meant for performance comparison.

### 4. Results and Discussions

The result of MLP training and model evaluation is outlined in table 2. The MLP model performance is compared against RMSE and MAE metrics, where the lowest value signifies the mean of predicted parking space availability is closest to actual data.
Table 2. Result and MLP Model Performance Comparison.

| Model                     | MSE      | RMSE      | MAE      |
|---------------------------|----------|-----------|----------|
| ARIMA (4,0,0)             | 7127.226 | 84.4229   | 42.9670  |
| ARIMA (4,0,1)             | 7131.414 | 84.4477   | 43.7597  |
| MLP (8 hidden units)      | 8046.09  | 89.7000   | 47.3162  |
| MLP (16 hidden units) **  | 6123.876 | 78.2552   | 37.0242  |
| MLP (24 hidden units)     | 6197.468 | 78.7240   | 38.0600  |
| MLP (32 hidden units)     | 6149.038 | 78.4158   | 39.1756  |
| MLP (40 hidden units)     | 6416.01  | 80.1000   | 42.1660  |

From the result, it can be seen that the MLP models performed well compared to the baseline ARIMA models. The MLP with 16 hidden units has the best forecasting performance followed by MLP (24 hidden units) and MLP (32 hidden units), the MLP performance seems to degrade with MLP (40 hidden units). We do not investigate the value beyond 40 hidden units as our aim is to implement the MLP within embedded computing environment with the least number of hidden units as possible. Increasing the number of hidden units will increase the MLP complexities and consume a lot of resources on a constrained FPGA-powered embedded computing environment.

5. Conclusion and Future Works

We have proposed and formulated several Multilayer Perceptron (MLP) models for predicting parking space availability. The proposed models are intended to be implemented in resource constrained FPGA devices. Thus, the MLP models should be simple and small without sacrificing forecasting accuracy. We have found that MLP performed comparatively well against ARIMA model with the most accurate MLP model (MLP - 16 hidden units) is able to predict the parking space availability with a mean absolute value of 37.0242 and root mean square error of 78.2552 for the KLCC parking space availability dataset. Although other models performed consistently close to each other, the MLP with 16 hidden units is preferable to be implemented in FPGA as it uses the least number of neurons and is friendly to the resource-constrained embedded computing device.

For future work, the next natural step is to implement the best performing MLP model into FPGA-based embedded computers and tests the hardware-accelerated implementation against real-time live data stream of parking space data obtained from PGIS.

References

[1] Caicedo F, Robuste F and Lopez-Pita A 2006 Parking management and modeling of car park patron behavior in underground facilities Transp. Res. Rec. 1956(1) pp. 60–67.
[2] Vlahogianni I, Kepaptsoglou K, Tsetsos V and Karlaftis M 2016 A real-time parking prediction system for smart cities J. Intell. Transp. Syst. 20(2) pp. 192–204.
[3] Zheng Y, Rajasegarar S, and Leckie C 2015 Parking availability prediction for sensor-enabled car parks in smart cities in IEEE 10th Int. Conf. on Intelligent Sensors, Sensor Networks and information Processing (Singapore) pp. 1–6.
[4] Caicedo F, Blazquez C, and Miranda P 2012 Prediction of parking space availability in real time Expert Syst. Appl. 39(8) pp. 7281–7290.
[5] Kotb A, Shen Y, and Huang Y 2017 Smart parking guidance, monitoring and reservations: a review IEEE Intell. Transp. Syst. Mag. 9(2) pp. 6–16, 2017.
[6] Gaikwad N, Tiwari V, Keskar A, and Shivaprakash NC 2019 Efficient FPGA implementation of multilayer perceptron for real-time human activity classification IEEE Access 7 pp. 26696–26706.
[7] Shawahna A, Sait S, and El-Maleh A 2018 FPGA-based accelerators of deep learning networks for learning and classification: A review IEEE Access 7 pp. 7823–7859.
[8] Tisan A and Chin J 2016 An end-user platform for FPGA-based design and rapid prototyping of feedforward artificial neural networks with on-chip backpropagation learning IEEE Trans. Ind. Informatics 12(3) pp. 1124–1133.

[9] Gschwend D, 2020 Zynqnet: An fpga-accelerated embedded convolutional neural network arXiv Prepr. arXiv2005.06892.

[10] Sarić R, Jokić D, Beganović N, Pokvić L, and Badnjević A 2020 FPGA-based real-time epileptic seizure classification using Artificial Neural Network Biomed. Signal Process. Control 62 p. 102-106.

[11] Zaiř H, Talha M, Meddah K, and Slimane S, 2019 FPGA-based system for artificial neural network arrhythmia classification Neural Comput. Appl 32 pp. 1–16.

[12] Sahin S, Becerikli Y, and Yazıcı S 2006 Neural network implementation in hardware using FPGAs Int. Conf. on Neural Information Processing (Hong Kong) pp. 1105–1112.

[13] Qiu J, Wang J, Yao S, Guo K, Li B, Zhao E, Yu J and Tang T 2016 Going deeper with embedded FPGA platform for convolutional neural network Proc. of the 2016 ACM/SIGDA Int. Symp. on Field-Programmable Gate Arrays (Monterey California) pp. 26–35.

[14] Libano F, Rech P, Tambara T Tonfaj J and Kastensmidt F 2017 On the reliability of linear regression and pattern recognition feedforward artificial neural networks in FPGAs IEEE Trans. Nucl. Sci. 65(1) pp. 288–295.

[15] Yang Y, Liu H, and Wang X 2003 The research on the key technologies for improving efficiency of parking guidance system Proc. of the 2003 IEEE Int. Conf. on Intelligent Transportation Systems (Shanghai, China) pp. 1177–1182.

[16] Ji Y, Tang D, Blythe P, Guo W, and Wang W 2014 Short-term forecasting of available parking space using wavelet neural network model IET Intell. Transp. Syst. 9(2).

[17] Zhao P, Guan H, and Zhang K 2019 A Parking Utilization Prediction Approach Based on NARNN Model Int. Conf. of Transportation CICTP (Nanjing China) pp. 4437–4448.

[18] Yu F, Guo J, Zhu X, and Shi G 2015 Real time prediction of unoccupied parking space using time series model Int. Conf. on Transportation Information and Safety (Wuhan China) pp. 370–374.

[19] Shao W, Zhang Y, Guo B, Qin K, Chan J and Salim F 2018 Parking availability prediction with long short term memory model Int. Conf. on Green, Pervasive, and Cloud Computing (Hangzhou) pp. 124–137.

[20] Ali G, Ali T, Irfān M, Draz U, Sohail M, Glowacz A, Sulowicz M and Martis C 2020 IoT Based Smart Parking System Using Deep Long Short Memory Network Electronics 9(10).

[21] Provoost J, Kamilaris A, Wismans L, Drift S and Keulen M 2020 Predicting Parking Occupancy via Machine Learning in the Web of Things Internet of Things 12.

[22] Barreto R, Lobo J, and Menezes P 2019 Edge Computing: A neural network implementation on an IoT device 5th Experiment Int. Conf. (Madeira Island Portugal) pp. 244–246.

[23] Biookaghazadeh S, Zhao M, and Ren F 2018 Are FPGAs suitable for edge computing? 1st USENIX Workshop on Hot Topics in Edge Computing (Boston USA).

[24] Ismail M H 2017 KLCC Parking Occupancy 2016 Kaggle doi:10.34740/kaggle/dsv/8000

[25] Ortigosa E M, Cañas A, Rodriguez R, Diaz J and Mota S 2006 Towards an Optimal Implementation of MLP in FPGA Int. Workshop on Applied Reconfigurable Computing Delft (Netherlands) pp. 46–51.

[26] Rynkiewicz J 2019 Asymptotic statistics for multilayer perceptron with ReLU hidden unit Neurocomputing 342 pp. 16–23.

[27] Dahl G E, Sainath T N, and Hinton G E, Improving deep neural networks for LVCSR using rectified linear units and dropout 2013 IEEE Int. Conf. on Acoustics, Speech And Signal Processing (Vancouver, Canada) pp. 8609–8613.

[28] Schmidt-Hieber J 2020 Nonparametric regression using deep neural networks with ReLU activation function Ann. Stat. 48 (4) pp. 1875–1897.

[29] Kingma D P and Ba J 2014 Adam: A method for stochastic optimization arXiv Prepr. arXiv:1412.6980 arXiv1412.6980, 2014.

[30] Chollet F 2015 Keras Github
[31] Abadi M, Barham P, Chen J, Chen Z, Davis, A, Dean J, Devin M, Ghermat S, Irving G, Isard M, Kudlur M, Levenberg J, Monga R 2016 Tensorflow: A system for large-scale machine learning 12th Symp. on Operating Systems Design and Implementation (Berkeley, California) pp. 265–283.