ISSN 2348-1218 (print)
International Journal of Interdisciplinary Research and Innovations ISSN 2348-1226 (online)
Vol. 6, Issue 3, pp: (148-156), Month: July - September 2018, Available at: www.researchpublish.com

Short Term Passenger Flow Prediction Using Deep Neural Network: A Case Study of Urban Metro Rail System

[1]Girish H., [2]Dr. K. Sunil Kumar

[1]Research Scholar, [2]Associate Professor
[1,2]APJ Abdul Kalam Technological University, College of Engineering Trivandrum, India

Abstract: **Short-term passenger flow forecasting is considered as a vital component of transportation systems which can aid in fine-tuning travel behaviors, reducing passenger congestion, supporting transportation system management and in enhancing service quality of transportation systems. For the same reason, passenger flow prediction still remains as an active research area. Many researchers have developed various travel demand prediction models based on different methodologies. Since the passenger flow prediction is complicated and non-linear in nature, the non-parametric technologies such as Neural Network, Support Vector Machine, Kalman filtering, Random Forest have been applied in passenger flow prediction. The deep-structured architecture of deep neural networks (DNN) could extract complex structure and build internal representation from the inputs, which is expected to outperform the traditional travel demand prediction models for passenger flow. This paper presents a deep neural network (DNN) based prediction model for predicting hourly passenger flow in an urban metro rail system. The input features of the model include temporal features like the day of a week, hour of a day and the holidays and spatial features include the respective metro stations and the passenger flow. These features are combined and multiple scenarios are modelled. The models are applied and evaluated with the passenger flow data from the metro system. Mean absolute error (MAE) and root mean squared error (RMSE) are used as measures of performance of these models. The experimental results showed that the DNN based prediction models effectively captures the non-linear relationship between the influential input features and the passenger flow. The model has the capability to provide an accurate and universal metro rail passenger flow prediction**.

Keywords: **Short-term passenger flow, Deep Neural Network (DNN), Metro Rail System, Root Mean Squared Error (RMSE).**

I. INTRODUCTION

Metro rail is a form of mass rapid transit public transport system employing metropolis trains. The metro rail system, unlike conventional rail-based systems is grade separated from the other traffic or provided with separate right of way (ROW) to avoid conflict with other urban transportation networks. Metro rail systems not only provide an efficient public transportation system, but also improve the urban traffic conditions while reducing air and noise pollution levels[1].

The operational issues related to urban rail transportation sector has been an active topic of research and the area covered include urban rail energy optimization, time table optimization, train scheduling problem and passenger flow prediction. Short-term forecasting of passenger flow can play an important role in addressing these operational issues. Here short-term refers to the time horizon in which the passenger flow prediction is made that is usually for an hour or less. Outputs generated from short-term forecasting models can be inputs for the operational planning, station management and real-time passenger flow monitoring. For railway operators, short-term forecasting can act as a key to the success of revenue management[2]. Short-term passenger flow forecasting can be used to fine-tune travel behaviors, reduce passenger congestion, and enhance service quality of transportation systems. The forecasting results of short-term passenger flow can be applied to support transportation system management such as operation planning, and station passenger crowd regulation planning[3].Traditional passenger flow forecasting methods fail in reflecting the gradual changes in the passenger flow distribution over time and space; therefore, advanced technologies are employed to develop prediction models for short-term and real-time passenger flow forecasts.

The short-term transportation forecasting approaches can be generally divided into two categories: parametric and non-parametric techniques. Parametric techniques and non-parametric techniques refer to the functional dependency assumed between independent variables and the dependent variable[3], [4]. In the linear prediction techniques, the forecasting approaches are based on the assumption of linearity and stationarity to infer future trends, on the basis of a changing time sequence in the past. Early research work mainly focused on the historical average model, smoothing technique, error component model and the nearest neighboring analysis model based on the assumption of linearity and stationarity to infer future transportation trends. Time-series analysis models, such as the auto regressive integrated moving average (ARIMA) and seasonal auto-regressive integrated moving average (SARIMA), have better prediction performance than other linear methods as reported by [5]. The non-linear prediction methods are able to describe non-linear characteristics in transportation systems. Compared with linear approaches, they have higher computing complexity but more accurate forecasting performance. Gaussian maximum likelihood model and nonparametric regression model fall into this category. Several algorithms, such as Bayesian networks, neural networks and support vector machines (SVM), belonging to the artificial intelligence field have also been widely applied in the intelligent transportation systems[6],[7],[8].

The prediction models can be broadly classified into traditional classical algorithms, Regressive Models, Machine learning based models and Hybrid models. Deep learning is very good at discovering intricate structures in high dimensional data and is therefore applicable to many domains of science, business, and government[9]. However, one of the drawbacks of deep learning models is low explanatory power[6]. In a recent review of short term forecasting techniques by [10], model interpretability is mentioned as one of the barriers in adapting more sophisticated machine learning models in practice. As per [11], the modelling approach and the input influential factors are both important elements to affect the performances of the passenger flow prediction model. The authors' laments that the input data, that plays a key integral part in these non-linear prediction models, do not get the needed attention and are seldom discussed in detail.

As a very important link between the modelling methodologies and the prediction performance, different types of input features are taken into consideration and needs to be systematically combined for building a more accurate and a general prediction model. The selections and combinations of different input features have a great influence on the accuracy of the prediction results [9]. Apart from the historical passenger flow, other input data features mainly include temporal features, directional features and spatial features[7], [8]. These potential features are necessary to be assembled for developing a robust and accurate passenger flow prediction model.

The present work focuses on the development of a Deep Neural Network (DNN) based prediction model that can predict hourly passenger flow for an urban metro rail network. Here, the metro rail stations are included as an additional input spatial feature to the DNN based prediction model. The work further investigates the effect of inclusion of this additional feature on the prediction performance of the developed model.

This paper is presented in four sections. Following section explains the methodology by which the proposed DNN models are developed and its execution. The models are tested and evaluated with the help of ridership data from an urban metro rail system, the details are discussed in the third section followed by the conclusions.

II. METHODOLOGY

This section describes the architecture and working of the DNN based model developed for the short term passenger flow prediction and is described below.

Deep neural network (DNN) is an artificial neural network with multiple hidden layers. Their deep-structured architecture could extract complex structure and build internal representation from the inputs, which is expected to outperform the traditional travel demand prediction models for passenger flow [11]. This paper focusses on developing a three hidden layered DNN based passenger flow prediction model using deep learning libraries in Python (version 3.6). Here, two types of DNN models are developed on the basis of inclusion or exclusion of spatial feature, 'station'. DNN Model1 considers all the aforementioned influential features excluding the spatial feature, 'station' and DNNmodel_2 considers spatial feature in addition to the temporal, directional, holiday features and historical passenger flow. The two different DNN-based models are demonstrated to evaluate their significances for the passenger flow prediction in the following case analysis section. The architecture of the developed DNN prediction model is shown in Fig. 1. As per Fig. 1, the DNN model consists of one input layer, three hidden layers and one output. The input layer captures the features of the training data with the help of labelled data. The first layer is followed by 3 hidden layers and the final layer is of the output which represents the prediction. The ridership data with the various input features like Week, Hour, Month, Stations and Passenger flow collected as input data, and a 3-hidden layer DNN is built for the prediction. Supervised learning which is based on training a data sample from a data source with correct labels already assigned is employed here for the learning. The learning procedure adapted here constitutes the following steps which are explained as follows:

(i) Parameters Initialization

The input neurons receive a feature vector $X = (x_1,\ x_2,\ x_3, \dots, x_n)$ based on the input features where $x_1,\ x_2,\ x_3, \dots, x_n$ are the individual input features and they are propagated to all the neurons in the first hidden layer.

(ii) Feature Transformation

The linear relationship between the input layer and the first hidden layer is given by the equation 1[11], where x_i refers to the individual input features from the input feature vector $X = (x_1,\ x_2,\ x_3, \dots, x_n)$ and (y_{Nj}^h) is the intermediate result of the hidden layers $Nj\ (j = 1, 2,\ 3)$ where h denotes the hidden layer. Non-linear transformation takes place between the intermediate result (y_{N1}^h) and the output of the hidden layer $(\varphi(x_i))$ according to the equation 2[11],. The output $(\varphi(x_i))$ will be considered as the inputs for the next hidden layer. This process is repeated until the last hidden layer is reached. The activation function considered for the work is Rectified Linear Unit (ReLU) and is given by the equation 3[11].

$$y_{Nj}^h = \sum_{i=1}^{n} w_{ih} x_i + b_h \quad (j = 1,2,3) \tag{1}$$

$$\varphi(x_i) = f(y_{Nj}^h) = f(\sum_{i=1}^{n} w_{ih} x_i + b_h) \quad (j = 1,2,3) \tag{2}$$

$$f(y_{Nj}^h) = f(\bullet) = \max(0, y_{Nj}^h) \quad (j = 1,2,3) \tag{3}$$

(iii) Linear Regression

The predicted passenger flow is obtained from the output layer by Linear Regression. After that, the weights w_{ih} and biases b_h of all layers are updated by minimizing the error between the generated final output (predicted passenger volume) and the input labelled data (observed passenger flow) with gradient decent in a supervised way. The fine tuning of parameters is achieved by the back propagation algorithm. The loss function also known as cost function, utilized for the present work is Mean absolute error loss. The optimizer function Adam which is an adaptive learning rate method will update the weights and biases automatically and is employed here.

All these steps are run in a loop until the error satisfies a predefined threshold. The neurons in the input layer, hidden layers and output layer are fully interconnected. The training procedure of the DNN based model is divided into two processes, namely the feed forward process and back propagation process. Firstly, the input data will be fed in the forward way and transformed layer by layer, and then the algorithm of back propagation is applied to modify the weights and bias in the back forward way. The architecture of the DNN will be optimized recursively until the error between the predicted passenger flow and the observed passenger flow to be a minimum. In order to obtain the best architecture of the DNN based prediction model, the key parameters including the number of neurons in each hidden layer, the number of hidden layers, the learning rate, the activation function and the loss function, are optimized.

III. CASE ANALYSIS

Here, the passenger flow data from one of the urban metro rail system is considered for illustration and verification of the developed passenger flow prediction model The considered metro rail is its initial phase of operation and currently it is operating between sixteen stations (S1 to S16). The case analysis is divided into two sections namely Data Collection & Analysis and Model Building. The model building procedure is further separated into three steps: training data specification, model specification and implementation.

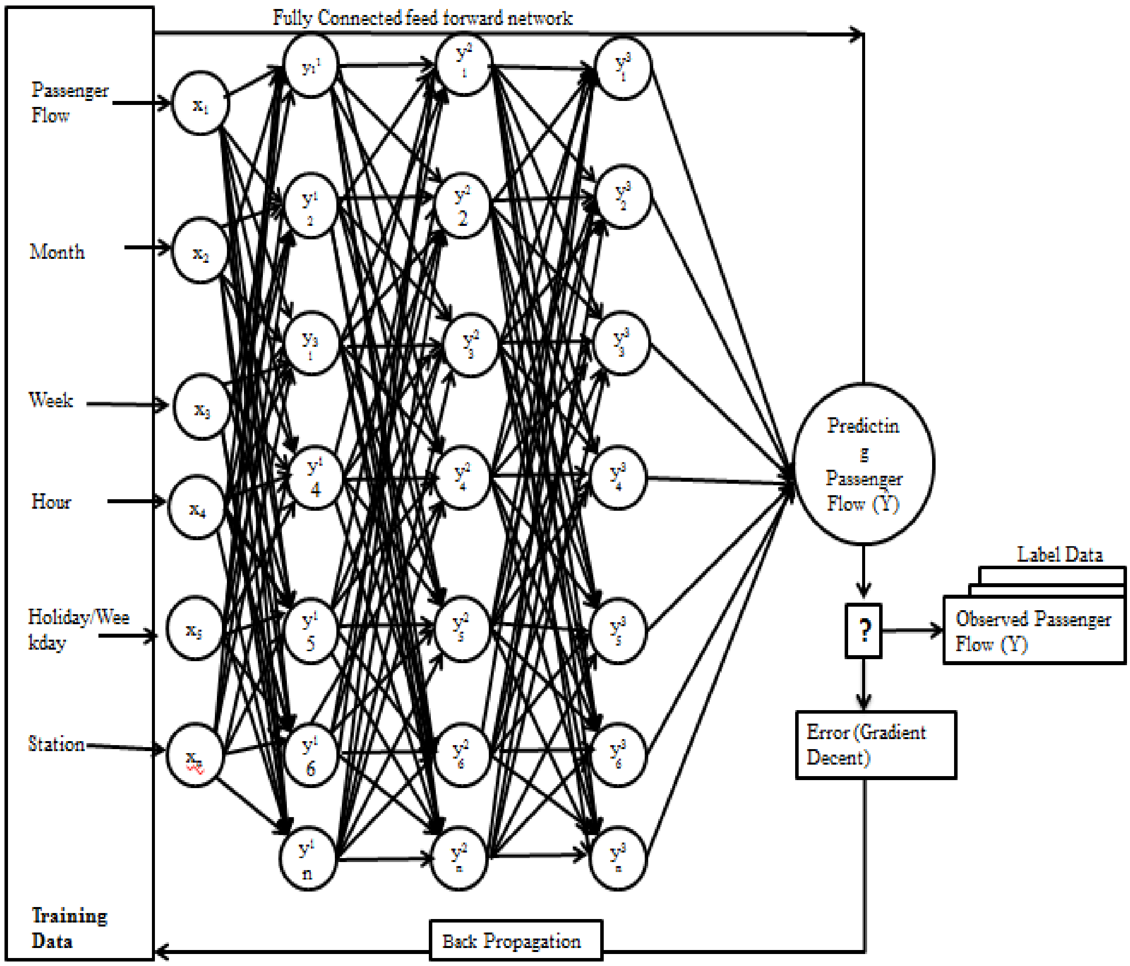

Fig 1: Architecture of the DNN based model

A. Data Collection and Analysis

Based on the ridership data collected from the Automatic Fair Collection (AFC) system of the urban metro, the total dataset in the present work is preprocessed. The total dataset for training and testing the model consists of ridership data from 4th October 2017 to 21st January 2018 comprising a total of 110 days data of passenger inflow and outflow across all the sixteen stations. Each day consists of sixteen hours of metro service starting from 06:00 in the morning to 22:00 in the night. But on Sundays, the metro service starts from 08:00 in the morning. So the passenger flow is taken as zero from 06.00 to 08:00 on Sundays. Also, passenger flow during extended service hours on days having special and social events like New Year eve are excluded from the analysis. So the ridership data comprises of the hourly inflow and outflow passenger flow for sixteen stations for 110 days. This makes a total of 56320 data points for the Deep learning model.

Initially, the ridership data was analysed to get an insight about the various influential factors that may affect the passenger flow in the network. The data is analysed separately for weekdays (Monday to Friday) and Weekends (Saturday & Sunday)/ Holidays. The monthly mean passenger inflow from October 2017 to January 2018 across sixteen stations was taken for training and prediction. It was observed that the mean passenger inflow was comparatively high in stations S1, S9 and S16 for both scenarios that is for weekdays and weekends/holidays. The average inflow was low in stations S2, S3, S4 and S5 and moderate passenger inflow was observed for other stations. A hypothesis was then constructed to check if there was any significant difference between passenger inflows for the two scenarios i.e. weekdays and weekends/holidays. Kruskal Wallis H test was performed to test the hypothesis at 5% significance level. The p value was observed to be less than 0.05 ($p<0.05$) indicating that there was a significant difference in the mean passenger inflow between weekdays and weekends/holidays. Hence, it was decided to analyze the prediction performance separately for weekdays and weekends/holidays.

ISSN 2348-1218 (print)
International Journal of Interdisciplinary Research and Innovations ISSN 2348-1226 (online)
Vol. 6, Issue 3, pp: (148-156), Month: July - September 2018, Available at: www.researchpublish.com

B. Model Building:

The model building process is illustrated using the following steps.

(i) Training Data Specification

The preprocessed training data of the ridership comprises of six input features. The specifications of the input features are as given in Table 1. Here, the temporal features of the input features such as date, month, day of the week and the hour of the day in which the considered passenger flow occurs. The spatial feature, 'station', mentions the station in which the considered passenger flow occurs. Holiday factor denoted by 'day' mentions whether the considered day is a weekday or a weekend/holiday. The total data is split into 80:20 for training and testing the model respectively. The training data and testing dataset is shuffled internally in order to take care of overfitting.

TABLE I: THE SPECIFICATIONS OF THE INPUT FEATURES

Sl No.	Input Feature	Indication	Indication Description
1	Month	1-12	1-12 represents months Jan-Dec.
2	Week	1-7	1-7 represents Monday-Sunday
3	Hour	1-16	1-16 represents each Metro Service hours i.e. from 06:00 to 22:00
4	Day	0,1	0-Weekday;1-Holiday/Weekend
5	Passenger Flow		Hourly passenger flow in a particular station
6	Station	1-16	Represents the 16 Stations starting from S1 to S16

(ii) Model Specification

In order to ascertain the influence of input features on the prediction performance, two independent DNN models were developed based on the aforementioned feature combinations. The first model is intended to make predictions for each station separately while the second model is intended to make predictions for the total metro rail system, independent of individual stations. Each of the models was trained with their respective training data. The models are described as follows.

1. **DNN Model_1:** This model comprises of the input features hour, day, week and the real passenger flow number and the model is trained for each station. During training, the day feature mentions whether the input day during the considered period is a weekday or a weekend/holiday. The hour feature mentions the hour of the day during which the passenger flow occurs (06.00 to 22.00). Week represents the day of the week in which passenger flow occurs. The model is trained using weekdays and weekends/holidays passenger flow data separately.

2. **DNN Model_2:** This model comprises of the input features hour, day, week, real passenger flow number and station. This second model adds the input spatial feature, 'station' to the first model. The model was trained considering the passenger flow data of all the stations combined.

Both the models were trained with their respective training datasets. The number of hidden layers for the model was fixed as three. The hyper parameters like number of nodes, epochs and optimizer function were determined using grid search algorithm (evaluated by Mean Squared Error (MSE)) and the optimized parameters obtained are mentioned in Table II. Thereafter the trained models were tested using the 20% test data and the performance measures were evaluated.

(iii) Implementation

Here the model performance was evaluated using the measures of Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE). They are defined as in Equations (4) and (5) gives the measurement of MAE and RMSE respectively[12].

$$MAE = \frac{1}{N}\sum_{i=1}^{N} | y_i - \hat{y}_i | \qquad (4)$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N} |y_i - \hat{y}_i|^2} \qquad (5)$$

where N is the total number of individual testing dataset points, $\hat{y}_i$ is the predicted value of the i_{th}sample and y_i is the actual value of the i_{th}sample.

For better objectivity, the training and testing processes are executed for five iterations and the prediction accuracies were evaluated by their average MAE and average RMSE over five iterations.

TABLE II: OPTIMIZED KEY PARAMETER SPECIFICATIONS OF THE TRAINING MODEL

Parameters	Training Models	
	DNN Model 1	DNN Model 2
Learning procedure	Supervised Learning	Supervised Learning
Type	DNN	DNN
Hidden Layers	3	3
Nodes	(150,150,150)	(150,150,150)
Activation Function	ReLU	ReLU
Loss Function	Mean Absolute Error	Mean Absolute Error
Epochs	50	50
Optimizer	Adam	Adam

IV. RESULTS AND DISCUSSION

DNN Model 1 was trained and tested for individual stations separately for weekdays and weekends/holidays. The performance measure obtained for the individual prediction models are summarized in Table III. It was observed that the predictions were almost near to the actual passenger flow.

TABLE III: PERFORMANCE MEASURES OF DNN MODEL 1

DNN Model1	Weekday		Weekend/Holiday	
	MAE	RMSE	MAE	RMSE
S1	54.78	88.96	78.58	107.35
S2	10.93	15.43	19.73	27.73
S3	7.68	14.14	11.5	15.7
S4	5.84	7.74	8.85	12.57
S5	8.29	11.64	13.01	18.77
S6	18.34	27.13	27.45	39.43
S7	13.8	23.21	22.24	29.8
S8	10	13.4	18.99	28.89
S9	48.64	74.52	76.88	124.9
S10	13.95	23.3	18.24	28.5
S11	17.4	33.9	24.9	33.24
S12	18.24	28.3	36.88	57.09
S13	17.68	24.73	28.84	40.32
S14	22.89	32.99	36.34	51.16
S15	25.8	33.07	33.22	50.53
S16	43.5	59.79	60.74	89.27

It is observed from Table III that the prediction errors were comparatively less for weekdays compared to weekends/holidays as signified by the MAE and RMSE values. The predictions made for most of the stations were near to the actuals especially for stations S2 to S8 & S10 to S15. Slightly higher variations were observed for Stations S1, S9 and S16 which calls for more data for training purpose. The RMSE performance measures for each station model showed higher value than its corresponding MAE value. The MAE values showed the mean absolute error of the prediction but the comparatively more RMSE values indicate that the errors are not uniformly distributed rather they are more concentrated on certain predictions.

DNN Model 2 considers the passenger flow at all stations at once. It is also trained and tested for weekday and weekend/holiday scenarios separately. Table IV summarizes the MAE and RMSE performance measures of DNN Model 2. The results show the performance measures value improved but it doesn't signify that DNN model_2 is better compared

ISSN 2348-1218 (print)
International Journal of Interdisciplinary Research and Innovations ISSN 2348-1226 (online)
Vol. 6, Issue 3, pp: (148-156), Month: July - September 2018, Available at: www.researchpublish.com

to DNN model 1. The less value of MAE and RMSE can be attributed to the Error Propagation factor. Also DNN model 2 is trained and tested with a high volume of data compared to DNN model 1 and the amount of data plays a crucial part in the performance of deep learning models.

TABLE IV: PERFORMANCE MEASURES OF DNN MODEL 2

DNN Model 2	Weekday		Weekend/Holiday	
	MAE	RMSE	MAE	RMSE
	19.98	40.64	37.38	126.31

From Tables III and IV, the results indicate that the MAE and RMSE values for the weekday scenario were different compared to weekend/holiday scenario for both models. The performance measures of weekend/holiday scenario for both models showed considerable increase in value compared to the weekday scenario. This may be due to high volume of passenger flows and variation in the passenger flow on the holidays/weekends that acts as outliers. This leads to the reduced performance of the prediction model. This signifies the feature of holiday is an important feature to be considered in training the DNN model or any machine learning based model. Overall, the model performance are satisfactory. It could be further improved by tuning the hyper parameters like nodes, activation function, epochs, etc. which is an exhaustive and time consuming task.

In order to observe the pattern of passenger inflow, an attempt is made to visualize the prediction performance of the developed models based on the testing data used for its evaluation. A curve is drawn between the predicted value and the actual observed value with the passenger volume on vertical axis and the individual data points (any given hour) on the horizontal axis. The passenger flow is predicted for next 48 data points or individual hours. Since the testing set is random, it cannot be guaranteed that the consequent data points may be consequent hours in a single day. For DNN model_1, there are sixteen stations. Here, the curve is plotted for origin station S1 having high passenger flow volume and variation. Curve is plotted for both scenarios i.e. for both weekday and weekend/holiday passenger flow model and they are shown in Fig. 2(a) and Fig. 2(b) respectively. The passenger flow volume is depicted on the y-axis and the individual data points i.e. hours at which passenger flow (both inflow and outflow) occurs are plotted on the x-axis. The blue line depicts the actual passenger flow observed and the red line depicts the predicted passenger flow by the models. Fig. 2 portrays the predicted passenger flow of Station S1 on weekday and weekend scenarios. From the figure, it can be seen that the model could predict the passenger flow to a considerable extent; and was able to capture the pattern of passenger flow effectively.

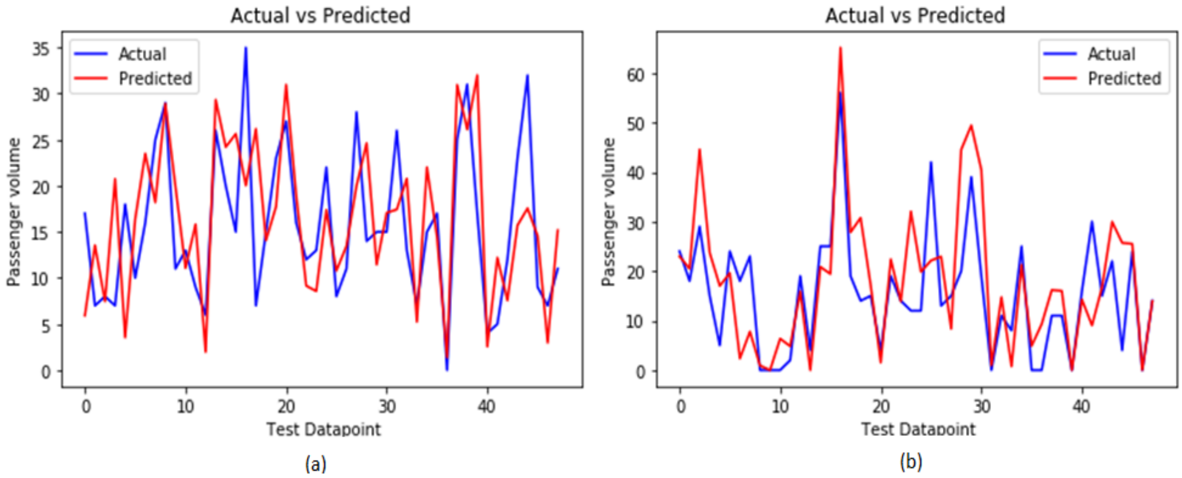

Fig 2: Actual vs Predicted passenger flow for Station S1 (a) Weekday scenario (b) Weekend/Holiday scenario

Fig. 3 portrays the comparison of predicted and actual passenger flow of DNN Model 2 on weekdays and weekend scenario respectively. From the figure, it can be observed that, compared to DNN model 1, DNN Model 2 fared much better in prediction performance as well as in capturing the passenger flow pattern effectively in both scenarios. This shows the potential of including the spatial feature 'station' with other input features in developing the deep learning based passenger flow prediction models.

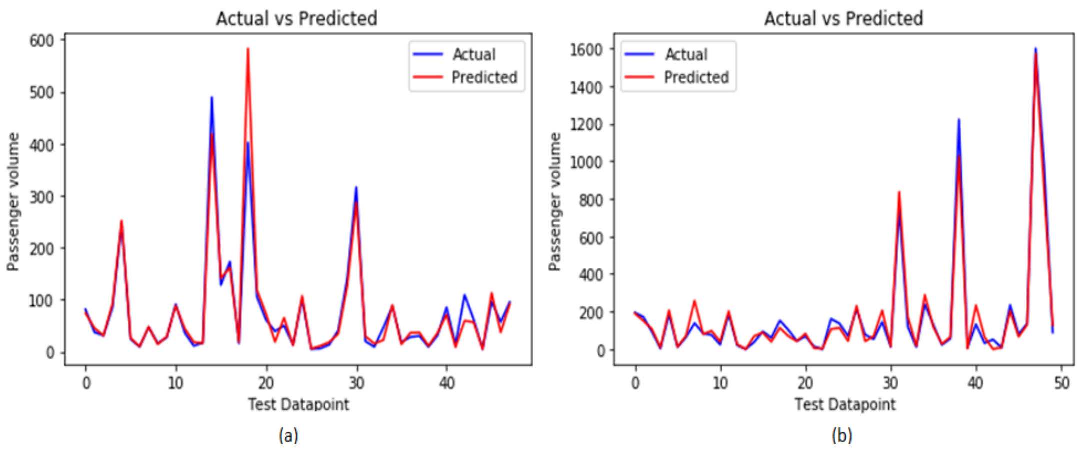

Fig 3: Actual vs Predicted passenger flow for Model DNN 2 (a) Weekday scenario (b) Weekend/Holiday scenario

V. CONCLUSIONS

In this work, a DNN- based prediction model for the hourly metro rail passenger flow has been proposed. The DNN based prediction model effectively captured the non-linear relationship between the influential input features and the passenger flow. The results of the work shows that incorporation of spatial feature 'station' along with other temporal, directional, holiday factors and the historical passenger flow improves the prediction potential. The influence between the input features and the predicted passenger flow has been analysed and verified by incorporating the feature combinations in the two models developed. The case study results showed that the holiday factor and spatial factor (station) are significant factors that can affect the performance of the prediction model besides the temporal and directional factors and the historical passenger flow. The visualization of the prediction results enabled us to understand the effectiveness of the DNN based models to capture the passenger flow pattern. The results of the DNN based short term passenger flow prediction model can aid in designing the feeder networks which also comprises of other urban transportation modes for enhancing the connectivity and further improve the passenger flow, thus enhancing the revenue for the urban metro rail corporation. The case study considered in the present work is a newly commenced metro rail and hence the passenger flow data during the period considered for the work is limited. The present work could be improved further by developing a general and robust prediction model and training the model with more passenger flow data. The effectiveness of using more advanced models like Recurrent Neural Networks (RNN) and Long Short Term Models (LSTM) for passenger flow prediction could also be investigated.

REFERENCES

[1] N. Sharma, R. Dhyani, and S. Gangopadhyay, "Critical Issues Related to Metro Rail Projects in India," J. Infrastruct. Dev., vol. 5, no. 1, pp. 67–86, 2013.

[2] M.-Q. Xie, X.-M. Li, W.-L. Zhou, and Y.-B. Fu, "Forecasting the Short-Term Passenger Flow on High-Speed Railway with Neural Networks," Comput. Intell. Neurosci., 2014.

[3] Y. Wei and M.-C. Chen, "Forecasting the short-term metro passenger flow with empirical mode decomposition and neural networks," Transp. Res. Part C, no. 21, pp. 148–162, 2012.

[4] P. Jiao, R. Li, T. Sun, Z. Hou, and A. Ibrahim, "Three Revised Kalman Filtering Models for Short-Term Rail Transit Passenger Flow Prediction," Math. Probl. Eng., 2016.

[5] Y. Jia, P. He, S. Liu, and L. Cao, "A Combined Forecasting Model for Passenger Flow Based on GM and ARMA," Int. J. Hybrid Inf. Technol., vol. 9, no. 2, pp. 215–226, 2016.

[6] N. G. Polson and V. O. Sokolov, "Deep learning for short-term traffic flow prediction," Transp. Res. Part C, no. 79, pp. 1–17, 2017.

[7] Zheng and Y. Wang, "Application of An Artificial Neural Network on Railway Passenger Flow Prediction," in International Conference on Electronic & Mechanical Engineering and Information Technology, 2011.

[8] Y. Sun, B. Leng, and W. Guan, "A novel wavelet-SVM short-time passenger flow prediction in Beijing subway system," Neurocomputing, vol. 166, pp. 109–121, Oct. 2015.

[9] L. Liu and R.-C. Chen, "A novel passenger flow prediction model using deep learning methods," Transp. Res. Part C, no. 84, pp. 74–91, 2017.

[10] I. Vlahogianni, M. G. Karlaftis, and J. C. Golias, "Short-term traffic forecasting: Where we are and where we're going," Transp. Res. Part C Emerg. Technol., vol. 43, no. 1, pp. 3–19, Jun. 2014.

[11] L. Liu and R.-C. Chen, "A MRT Daily Passenger Flow Prediction Model with Different Combinations of Influential Factors," in 31st International Conference on Advanced Information Networking and Applications Workshops, 2017.

[12] T. Chai and R. R. Draxler, "Root mean square error (RMSE) or mean absolute error (MAE)?," Geosci. Model Dev., vol. 7, 2014.