SHORT TERM PASSENGER DEMAND FORECASTING USING DEEP LEARNING TECHNIQUES

Thesis (CV 758/ Major Project/3rd Semester)

Submitted in partial fulfilment of the requirements for the degree of

MASTER OF TECHNOLOGY in

TRANSPORTATION ENGINEERING

by

DUDAM AMRUTH

(212TS012)

DEPARTMENT OF CIVIL ENGINEERING

NATIONAL INSTITUTE OF TECHNOLOGY KARNATAKA

SURATHKAL, MANGALORE -575025

November, 2022

SHORT TERM PASSENGER DEMAND FORECASTING USING DEEP LEARNING TECHNIQUES

Thesis (CV 758/ Major Project/3rd Semester)

Submitted in partial fulfilment of the requirements for the degree of

MASTER OF TECHNOLOGY in

TRANSPORTATION ENGINEERING

by

DUDAM AMRUTH

(212TS012)

Under the guidance of

Dr. RAVIRAJ H.M.

DEPARTMENT OF CIVIL ENGINEERING

NATIONAL INSTITUTE OF TECHNOLOGY KARNATAKA

SURATHKAL, MANGALORE -575025

November, 2022

D E C L A R A T I O N

I hereby declare that the Report of the P.G. Project Work (CV758/Major Project) report entitled **"Short Term Passenger Demand Forecasting Using Deep Learning Techniques"** which is being submitted to the **National Institute of Technology Karnataka Surathkal**, in partial fulfilment of the requirements for the award of the Degree of **Master of Technology in Transportation Engineering**, is a bonafide report of the work carried out by me. The material contained in this Report has not been submitted to any University or Institution for the award of any degree.

DUDAM AMRUTH

(212TS012)

Department of Civil Engineering

Place: NITK, SURATHKAL

Date: November 2022

Contents

CHAPTER 1

INTRODUCTION

1.1. General

In a modern transportation system, urban public transportation plays an important function. Comparing with other modes of travel, public transportation has the advantages of large passenger capacity, low pollution discharge, and low cost. In order to ensure the efficient and orderly operation of urban buses, not only is a good bus operation management plan required, but effective operation scheduling is also essential. Using public transportation–related information and data to make accurate public traffic passenger flow predictions can provide effective decision support for the operation and dispatch of urban buses, help transit operators control ridership inflow to avoid congestion, or adjust train. Short term passenger flow forecasting can provide real-time traffic information to help passengers make rational scheduling decisions and timetables to accommodate more passengers in peak hours.

1.2. Impact of various parameters on public bus passenger demand

Some of the factors that affect the public bus passenger flow includes travel time, ticket fares, weather conditions, route choice model and land use. Among which weather conditions and land use are the factors that are considered in this study to find out the influence on public bus passenger flow.

Weather and travel habits are intimately connected. According to a Gallup poll conducted in 2002, 40% of respondents ranked weather and road conditions as the most important piece of information in their daily lives (ITSA, 2002). Given that the severity and frequency of extreme weather conditions are expected to increase as a result of climate change and global warming, it is critical to gain a better understanding of the dynamics of the relationship between weather and travel behavior in order to inform urban planners and transportation operators on how to (re)design more weather-resilient transportation systems capable of managing and adjusting transportation services in real-time in response to variations in weather.

To capture the impact of various parameters on public bus passenger flow time series statistical analysis like Auto Regressive Integrated Moving Average Method (ARIMA), Linear Regression, etc. need to be developed.

1.3. Statistical Analysis

Statistical analysis is the collection and interpretation of data in order to uncover patterns and trends. Statistical analysis gives how the relation between the various parameters is affecting the public bus passenger flow using techniques like Linear regression (LR), Auto Regressive Integrated Moving Average Method (ARIMA), Seasonal Auto regressive Integrated moving Average Method (SARIMA) etc.

1.3.1. Linear Regression

Linear Regression is a machine learning algorithm based on supervised learning. It performs a regression task. Regression models a target prediction value based on independent variables. It is mostly used for finding out the relationship between variables and forecasting. Different regression models differ based on – the kind of relationship between dependent and independent variables they are considering, and the number of independent variables getting used.

1.3.2. Auto Regressive Integrated Moving Average Method

An autoregressive integrated moving average model is a form of regression analysis that gauges the strength of one dependent variable relative to other changing variables. The model's goal is to predict future securities or financial market moves by examining the differences between values in the series instead of through actual values.

An ARIMA model can be understood by outlining each of its components as follows:

Autoregression (AR): refers to a model that shows a changing variable that regresses on its own lagged, or prior, values.

Integrated (I): represents the differencing of raw observations to allow for the time series to become stationary (i.e., data values are replaced by the difference between the data values and the previous values).

Moving average (MA): incorporates the dependency between an observation and a residual error from a moving average model applied to lagged observations.

1.3.3. Seasonal ARIMA

Seasonal Autoregressive Integrated Moving Average, SARIMA or Seasonal ARIMA, is an extension of ARIMA that explicitly supports univariate time series data with a seasonal component. It adds three new hyperparameters to specify the autoregression (AR), differencing (I) and moving average (MA) for the seasonal component of the series, as well as an additional parameter for the period of the seasonality.

1.4. Machine Learning Algorithms

Many machine-learning approaches (MLA) have been proposed to improve the prediction accuracy of short-term traffic/passenger flow, such as linear regression and feedforward neural networks. In addition, various deep learning methods have been employed in the literature. Despite the fact that MLA has become an emerging and significant transportation technology, the challenge of efficiently using it for bus passenger flow prediction persists due to the complexity of transportation systems, and this is the focus of this study.

From the beginning of forecasting models, the passenger flow was modelled using Linear Regression. The regression works well with data that has linear relationships. However, most real-world patterns have a complex structure, making linear models insufficient. The pattern of passenger flow has nonlinear temporal dependencies. Regression models such as linear regression, support vector regression with linear kernel, and ARMA are not appropriate in this case.

Historical passenger flow data and the accompanying spatiotemporal data are the most often used input variables. Despite the fact that experimental results demonstrate that weather data can accurately represent passenger flow, only a small amount of research has been focused on the problem of forecasting passenger flow using the weather factor due to the challenges associated with data collection and fusion for weather data, spatiotemporal data, and historical passenger flow data.

To capture the temporal characteristics, Long Short-Term Memory (LSTM) model is developed and to extract spatial correlation, Graph Neural Network (GNN) is developed.

1.4.1. Long Short-Term Memory

Long Short-Term Memory networks – usually just called "LSTMs" – are a special kind of Recurrent Neural Networks (RNN), capable of learning long-term dependencies. They work tremendously well on a large variety of problems, and are now widely used. LSTMs are explicitly designed to avoid the long-term dependency problem. Remembering information for long periods of time is practically their default behavior, not something they struggle to learn.

All recurrent neural networks have the form of a chain of repeating modules of neural network. In standard RNNs, this repeating module will have a very simple structure, such as a single tanh layer.
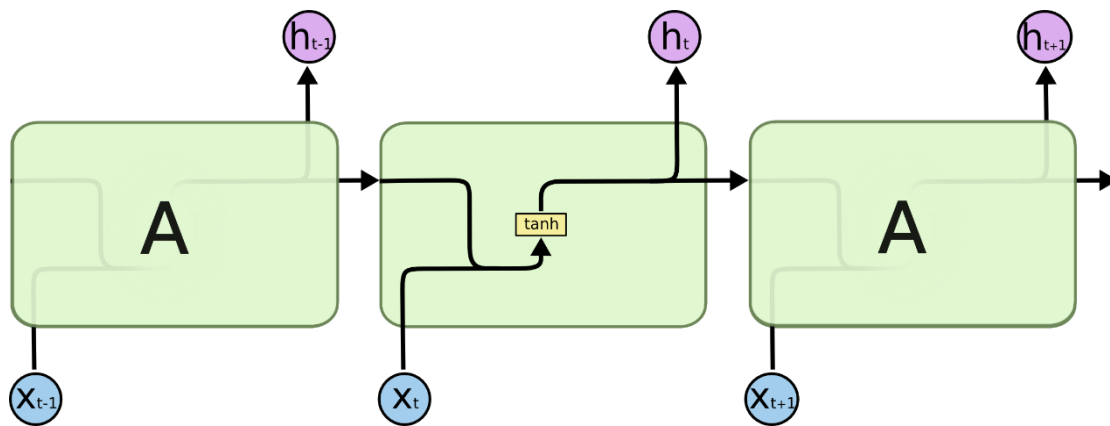

Fig 1: The repeating module in a standard RNN contains a single layer.

LSTMs also have this chain like structure, but the repeating module has a different structure. Instead of having a single neural network layer, there are four, interacting in a very special way
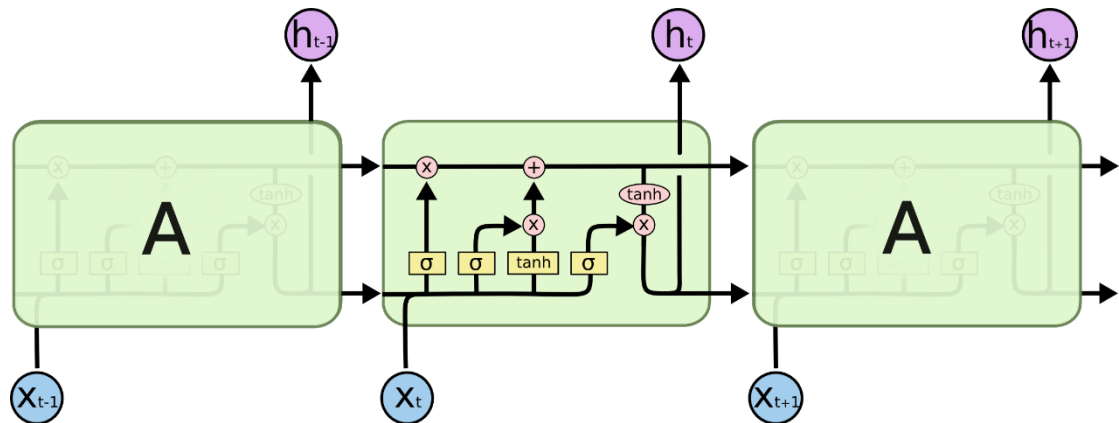
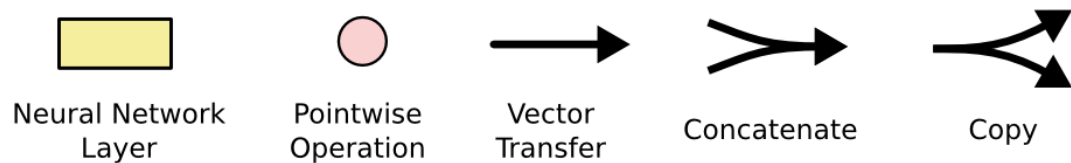**Fig 2: The repeating module in an LSTM contains four interacting layers.**

Neural Network Layer   Pointwise Operation   Vector Transfer   Concatenate   Copy

Fig 3: Operators used in LSTM networks

In the above diagram, each line carries an entire vector, from the output of one node to the inputs of others. The pink circles represent pointwise operations, like vector addition, while the yellow boxes are learned neural network layers. Lines merging denote concatenation, while a line forking denote its content being copied and the copies going to different locations.

Step by step LSTM walkthrough

1.Forget gate

The first step in our LSTM is to decide what information we're going to throw away from the cell state. This decision is made by a sigmoid layer called the "forget gate layer." It looks at h_{t-1} and x_t, and outputs a number between 0 and 1 for each number in the cell state C_{t-1}. Where 1 represents "completely keep this" while a 0 represents "completely get rid of this."
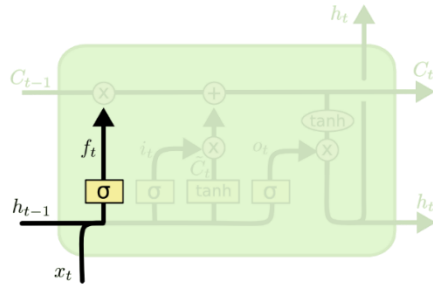
$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right)$$

Fig 4: Forget Gate

2.Input Gate

The next step is to decide what new information we're going to store in the cell state. This has two parts. First, a sigmoid layer called the "input gate layer" decides which values we'll update. Next, a tanh layer creates a vector of new candidate values, c_t, that could be added to the state. In the next step is combining these two to create an update to the state.
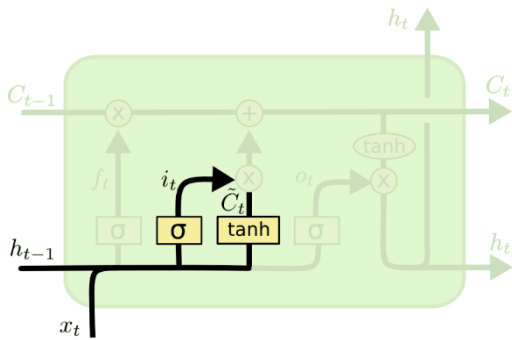
$$i_t = \sigma\left(W_i \cdot [h_{t-1}, x_t] + b_i\right)$$
$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C)$$

Fig 5: Input Gate

3.Memory Cell

It's now time to update the old cell state, C_{t-1}, into the new cell state C_t. Next, multiplying the old state by f_t, forgetting the things we decided to forget earlier. Then we add $i_{t*}\tilde{C}_t$t. This is the new candidate values, scaled by how much we decided to update each state value.
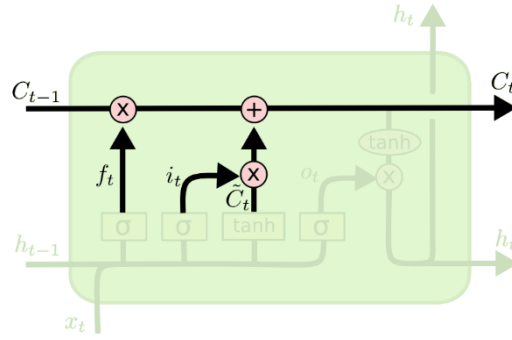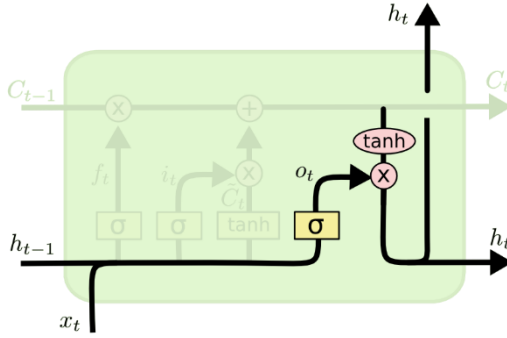
$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t$$

Fig 6: Memory Cell

4.Output Gate

In the output cell first, we run a sigmoid layer which decides what parts of the cell state we're going to output. Then, we put the cell state through tanh (to push the values to be between −1 and 1) and multiply it by the output of the sigmoid gate, so that we only output the parts we decided to.

$$o_t = \sigma \left(W_o \left[h_{t-1}, x_t \right] + b_o \right)$$
$$h_t = o_t * \tanh \left(C_t \right)$$

Fig 7: Output Gate

1.4.2. Graph Neural Network

Graph Neural Networks (GNNs) are a class of deep learning methods designed to perform inference on data described by graphs. GNNs are neural networks that can be directly applied to graphs, and provide an easy way to do node-level, edge-level, and graph-level prediction tasks.

GNN is applied in Node Classification, the task here is to determine the labeling of samples (represented as nodes) by looking at the labels of their neighbors. Usually, problems of this type are trained in a semi-supervised way, with only a part of the graph being labeled. GNN is also applied in Link prediction, here the algorithm has to understand the relationship between entities in graphs, and it also tries to predict whether there's a connection between two entities. It's essential in social networks to

infer social interactions or to suggest possible friends to the users. It has also been used in recommender system problems and in predicting criminal associations.

1.4.3. Standard Models in Deep Learning

Apart from LSTM and GNN there are other standard models in deep learning which are used in time series problems such as ANN, CNN, GCN, RNN etc.,

Artificial Neural Networks (ANN)

ANNs are nonlinear statistical models that demonstrate a complex relationship between inputs and outputs in order to uncover a new pattern. Artificial neural networks are used for a range of applications, including image recognition, speech recognition, machine translation, and medical diagnosis.

The fact that ANN learns from sample data sets is a significant advantage. The most typical application of ANN is for random function approximation. With these types of technologies, one can arrive at solutions that specify the distribution in a cost-effective manner. ANN can also offer an output result based on a sample of data rather than the complete dataset. ANNs can be used to improve existing data analysis methods due to their high prediction capabilities.

Convolution Neural Networks (CNN)

CNNs can take in an input image, assign importance (learnable weights and biases) to various aspects/objects in the image and be able to differentiate one from the other. It has various applications some of which are decoding facial recognition, analyzing documents, understanding climate etc.,

Recurrent Neural Networks (RNN)

RNN works on the principle of saving the output of a particular layer and feeding this back to the input in order to predict the output of the layer. An RNN can handle sequential data, accepting the current input data, and previously received inputs. It can be used for any time series prediction, natural language processing and machine translation.

Graph Convolution Networks (GCN)

Convolutional Networks are 3-dimensional neural networks. Most practical uses of Convolutional Neural Networks include image classification and recognition, natural

language processing and speech recognition. These models are usually more complex than the usual 2-dimensional neural network models.

1.5. Need for the study

Following are the needs for forecasting the public bus passenger flow

- ➢ To understand the various factors affecting the public bus transit passenger flow.
- ➢ To predict the bus passenger flow with good accuracy by adding the factors that affects the passenger flow.
- ➢ To develop prediction models using state of art techniques.

CHAPTER 2

LITERATURE REVIEW

2.1. General

Sui Tao et al (2018) has modelled the effect of local weather conditions on hourly bus ridership. They have derived a suite of time-series regression models (Auto Regressive Integrated Moving Average Method (ARIMAX), Seasonal Auto Regressive Moving Average Method (SARIMAX)) which are computed to capture the concurrent and lagged effects that weather conditions exert on bus ridership. Their study area was bus network in Brisbane, Australia. Brisbane is the capital of Queensland and the third most populated city in the country, with around one million population within its local government area. The data were collected from transit smart card data and weather data were acquired from the Australian Bureau of Meteorology (BOM) for the same period of time as the smart card data. The information contained in a single smart card record includes date, route, direction (i.e., inbound and outbound trips in relation to the city center), smart card ID, boarding time and stop, and alighting time and stop and journey ID for linked trips made within a one-hour transfer limit. Measurements of four weather variables, i.e., temperature, rainfall, relatively humidity and wind speed on a 30-min interval are included for 14 weather stations located across the study context. They first visually inspected hourly system-wide ridership patterns on weekdays and weekends then marked decline in ridership in late March, early and late April is in parallel with the public holidays during these days. Except for this pattern of low ridership, a strong recurring pattern of hourly ridership persists across both weekdays and weekends. They next modelled the hourly weather-ridership relationships for their four selected destinations, Following the system wide analysis, univariate ARIMA (or SARIMA) models were separately estimated for each destination. Their findings highlight that changes in particularly temperature and rainfall were found to induce significant hour-to-hour changes in bus ridership, with such effects varying markedly across both a 24-hour period and the transit network.

Md Sami Hasnine et al (2021) investigated the effects of built environment and weather on the demands for the Transportation Network Companies (TNC) in Toronto.

Their research was based on a historical dataset of Uber trips from September 2016 to September 2018 in Toronto. A wide range of built environments, sociodemographic, and weather data were generated at the dissemination area-level and fused with the monthly aggregated Uber dataset. To provide insight into the underlying factors that affect TNC demand, a series of aggregate demand models were estimated using log transformed constant elasticity demand functions, with consideration of the seasonal lag effect. To capture the weather effect, an autoregressive moving average model was estimated for the downtown core of Toronto. Their model results show that the influence of lagged ridership and seasonal lag effect have a positive correlation with TNC demand. The trip generation and attraction models revealed that TNC trips increase when the commuting trip duration is longer than 60 min. And also founded that the number of apartments in a dissemination area is positively correlated with TNC trip generation, while the number of single-detached houses has a negative correlation. Developed a time-series model which indicated that temperature and total daily precipitations are positively correlated with TNC demand.

Ming Wei (2022) also studied the influence of local weather conditions on public transit ridership in Brisbane, Australia. Based on the statistical distribution of transit ridership, this study applied a suite of geographically weighted negative binomial regression models to capture the weather–transit ridership relationship at both daily and half-hourly levels. The results revealed that weather exerts significant effects on transit ridership and its effects vary by passenger type and are not fixed across locations and temporal periods.

Yang Liu et al (2019) combined the modeling skills of deep learning and the domain knowledge in transportation into prediction of metro inbound/outbound passenger flow. They used the standard metro service/transaction data from Nanjing Metro System, the data consists of weekday records of 103 days. The information contained from a single record are user id, inbound station, outbound station, inbound time, outbound time, and type of card. Based on Long Short term Method (LSTM) they proposed an end-to-end deep learning based architecture which can reasonably address the input features of the metro passenger flow prediction problem. This architecture comprised multiple extensible components, including modeling external environmental factors such as weather and holidays, temporal dependencies, spatial characteristics,

and metro operational properties respectively. The results of this study shows that the accuracy can be improved when the daily cyclicality characteristic and the weekly trend characteristic are incorporated in the model, indicating that it has a fixed behavior patterns.

Lijuan Liu et al (2020) also developed a deep long short term memory neural network (LSTM_NN) model for predicting the metro passenger flow. The optimized traditional input variables, including the different temporal data and historical passenger flow data, were combined with weather variables for data modeling. They Constructed a Metro passenger flow forecasting(LSTM_NN) model by historical hourly passenger flow data and the corresponding temporal and whether data in a station level study with Endogenous Input variables: passenger flow direction, date, day, week, hour and Exogenous Input variables: Temp, rainfall, relative humidity, wind speed. They have compared the proposed model with weather variables and without weather variables and found that the deep LSTM_NN is a more powerful method to make the more accurate forecasts when suitable weather variables are included.

Jaison Mulerikkal et al (2022) has fed the passenger flow parameter into the layers of the deep neural network using the ST-LSTM (Spatio-Temporal Long Short Term Memory) architecture. This architecture was evaluated with passenger movement data collected from automated fare card (AFC) information from metro rail. They have used the One-Class SVM-based (Support Vector Machine) outlier detection and elimination algorithm to reduce the impact of irregular flow. The results are compared with the existing system of regression models like SVR, ANN and LSTM. The architecture has obtained a performance improvement with an error of 0.00026.

Jinlei Zhang et al (2021) proposed a deep learning architecture combining the residual network (ResNet), graph convolutional network (GCN), and long short-term memory (LSTM) (called "ResLSTM") to forecast short-term passenger flow in urban rail transit on a network scale. Firstly, they improved the methodologies of ResNet, GCN and atttention LSTM, then the model architecture was proposed wherein ResNet is used to capture deep abstract spatial correlations between subway stations, GCN is applied to extract network topology information, and attention LSTM is used to extract temporal correlations. This model architecture includes four branches for inflow, outflow, graph-network topology, as well as weather conditions and air quality. Finally,

ResLSTM is applied to the Beijing subway using three time granularities (10, 15, and 30 min) to conduct short-term passenger flow forecasting. And the results confirmed that the weather conditions and air quality were proven to have considerable influence on prediction precision.

Shengnan Guo et al (2020) proposed novel approaches in modelling the dynamics of traffic data along both the temporal and spatial dimensions, as well as consider the periodicity and spatial heterogeneity of traffic data which is Attention based Spatial-Temporal Graph Neural Network (ASTGNN) for traffic forecasting. Attention mechanism which captures local context in time series. GCN to capture the dynamics along the spatial dimension. They have considered two kinds of datasets; the first kind of datasets are about highway traffic flow in California. The second kind of datasets is about metro crowd flow of the Hangzhou metro system and then the raw data was converted to 5-minute interval. Their model outperformed all the state-of-the-art techniques.

Feifei Zhao et al (2020) proposed a model SEHNN (station-embedding-based hybrid neural network) which utilizes the VGAE (Variational Graph auto-encoder) module to embed and extract the spatial and temporal information of order interactions and geographic neighbors among the stations of a carsharing network, while the LSTM module to capture the time sequence information of the embeddings to complete the prediction of rental and return vehicles. The results from the real data of Lanzhou, China demonstrate that, compared with ElasticNet, ARIMA, LSTM, and ConvLSTM, the mean absolute error of the proposed model targeted at hourly demand forecasting is reduced by 56.5%, 47.2%, 38.7%, and 38.5%, respectively, and also found that it outperforms some widely used models among different intervals and scales including main stations and subset carsharing system

Tao Chen et al (2021) proposed a novel algorithm, namely the Spatial–Temporal Graph Sequence with Attention Network (STGSAN), to predict transit passenger flow. The algorithm mainly focused on the following three aspects: (1) a graph attention network (GAT) was used to capture the spatial correlation of various bus stops; (2) to make full use of the historical and real-time data, a bidirectional long short-term memory and attention mechanism was conducted to extract the temporal correlation of historical ridership at bus stations; and (3) external factors that affect passenger choices

were taken into account. they conducted an experiment using field data collected in Urumqi, China. They compared their model with five other models, the proposed model was proven to have excellent performance prediction.

Can Li et al (2022) provided a confidence interval-based demand forecasting, which can help transport planning and operation authorities to better accommodate demand uncertainty/variability. The proposed Origin Destination (OD) demand prediction approach well captures and utilizes the correlations among spatial and temporal information. They proposed a Probabilistic Graph Convolution Model (PGCM) which consists of two components: (i) a prediction module based on Graph Convolution Network and combined with the gated mechanism to predict OD demand by utilizing spatio-temporal relations; (ii) a Bayesian-based approximation module to measure the confidence interval of demand prediction by evaluating the graph-based model uncertainty. They used a large-scale real-world public transit dataset from the Greater Sydney area to test and evaluate the proposed approach. The experimental results demonstrated that the proposed method is capable of capturing the spatial-temporal correlations for more robust demand prediction. The proposed approach is compared with several benchmark algorithms including ARIMA, LR, HA, GRU, and LSTnet, GCRN, and STGCN. The experiments on the real-world dataset show that the proposed approach outperforms other state-of-the-art methods.

2.2. Gaps in the literature

All the research papers have been studied thoroughly. The following gaps were identified:

 - Studies related to the influence of weather conditions for the prediction of passenger demand are very less.
 - Deep learning models have not been extensively used in India for predicting the passenger flow demand, considering spatial and temporal characteristics.
 - Very few works in India have considered the impact of weather to enhance the performance of the models.

CHAPTER 3

SCOPE AND OBJECTIVES

3.1. Scope of the study

In earlier chapters various research works were studied and gaps in literature were identified. The BRTS in India is being developed at a greater rate and there are many upcoming projects of BRTS in India. As the BRTS projects are boosting the performance of these projects should be efficient, they should have minimum delay. Passengers' anxiety is reduced when they get accurate information on bus arrival and departure times at bus stops, and they may plan their trip accordingly. Estimating travel time and stopped delays at that portion of road, which includes the bus stop, can provide accurate information regarding bus arrival times. By understanding the causes of bus stop delays, suitable efforts to alleviate these delays can be taken. If the delay is caused by travel time, then adequate road network design is required, if the delay is caused by bus stop design, then proper bus stop design is required.

Public transit can be made more appealing to passengers if delays at bus stops are adequately addressed and solutions are implemented. As a result, traffic congestion on Indian highways could be decreased.

Making accurate public traffic passenger flow predictions can provide effective decision support for the operation and dispatch of urban buses and help transit operators control ridership inflow to avoid congestion, or adjust train. And also, short term passenger flow forecasting can provide real-time traffic information to help passengers make rational scheduling decisions and timetables to accommodate more passengers in peak hours.

3.2. Objectives

The objectives proposed to develop a methodology for predicting the bus passenger flow are as follows:

- ➤ To evaluate the parameters that are affecting, to examine the influence of weather and land use on bus passenger flow using ARIMA.

- To predict the future passenger demand by incorporating spatial-temporal characteristics. GNN to capture spatial relationship and LSTM to extract temporal correlations and including weather and land use parameters.
- To make use of deep learning architectures to predict the passenger flow, GNN to capture spatial relationship and LSTM to extract temporal correlations and including weather and land use parameters.
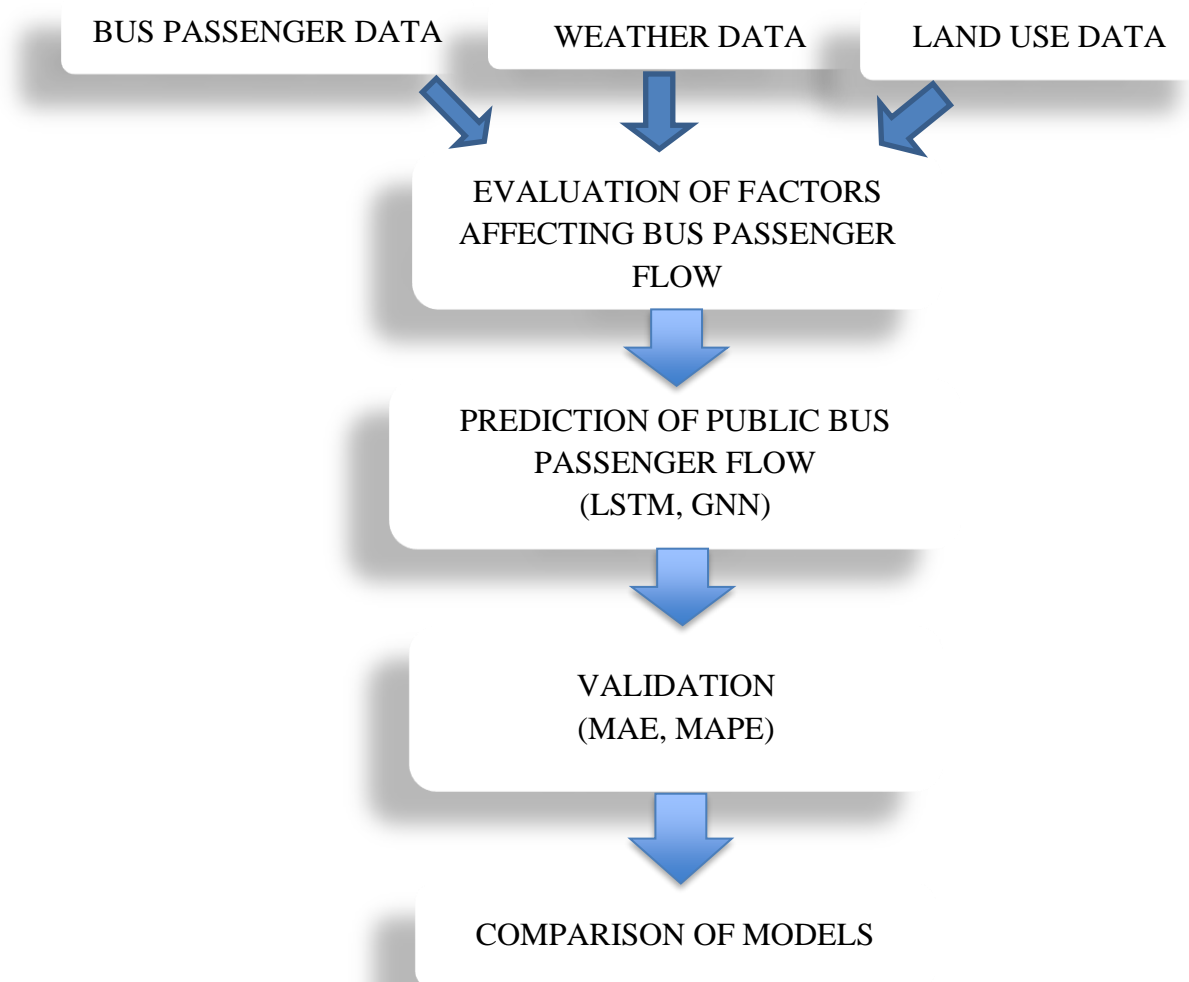- To compare the results of the developed model with the base line models.

CHAPTER 4

METHODOLOGY

4.1. General

The prediction of public bus passenger flow is done using the historical data of bus passenger data, weather data and land use data. First step is to evaluate the various parameters that are affecting the bus passenger flow. Next step is to develop the deep learning architectures to predict he passenger flow where LSTM is used to extract the temporal correlations and GNN is used to capture the spatial characteristics. Finally validating the results of proposed model with mse, mape.

4.2. Proposed Methodology

Detailed description of proposed methodology presented in the form of flow chart is given in this section.

BUS PASSENGER DATA   WEATHER DATA   LAND USE DATA

EVALUATION OF FACTORS AFFECTING BUS PASSENGER FLOW

PREDICTION OF PUBLIC BUS PASSENGER FLOW (LSTM, GNN)

VALIDATION (MAE, MAPE)

COMPARISON OF MODELS

Future Work

➢ Collecting the bus passenger data from KSRTC, Mangalore, weather data from Indian meteorological department, Bangalore and land use data.

➢ Evaluating the parameters that are affecting the bus passenger demand using ARIMA model and considering those parameters to include in the prediction of bus passenger flow.

➢ Predicting the bus passenger flow, taking into account spatial-temporal characteristics along with weather and land use data using deep learning techniques.

➢ Validating the results with MAE, MAPE and comparing the proposed model with the state of art techniques.

REFERENCES

1. Liu, Yang, Zhiyuan Liu, and Ruo Jia. "DeepPF: A deep learning based architecture for metro passenger flow prediction." *Transportation Research Part C: Emerging Technologies* 101 (2019): 18-34.

2. Liu, Lijuan, Rung-Ching Chen, and Shunzhi Zhu. "Impacts of weather on short-term metro passenger flow forecasting using a deep LSTM neural network." *Applied Sciences* 10, no. 8 (2020): 2962.

3. Mulerikkal, Jaison, Sajanraj Thandassery, Vinith Rejathalal, and Deepa Merlin Dixon Kunnamkody. "Performance improvement for metro passenger flow forecast using spatio-temporal deep neural network." *Neural Computing and Applications* 34, no. 2 (2022): 983-994.

4. Zhang, Jinlei, Feng Chen, Zhiyong Cui, Yinan Guo, and Yadi Zhu. "Deep learning architecture for short-term passenger flow forecasting in urban rail transit." *IEEE Transactions on Intelligent Transportation Systems* 22, no. 11 (2020): 7004-7014.

5. Guo, Shengnan, Youfang Lin, Huaiyu Wan, Xiucheng Li, and Gao Cong. "Learning dynamics and heterogeneity of spatial-temporal graph data for traffic forecasting." *IEEE Transactions on Knowledge and Data Engineering* (2021).

6. Zhao, Feifei, Weiping Wang, Huijun Sun, Hongming Yang, and Jianjun Wu. "Station-level short-term demand forecast of carsharing system via station-embedding-based hybrid neural network." *Transportmetrica B: Transport Dynamics* 10, no. 1 (2022): 1-19.

7. Chen, Tao, Jie Fang, Mengyun Xu, Yingfang Tong, and Wentian Chen. "Prediction of Public Bus Passenger Flow Using Spatial–Temporal Hybrid Model of Deep Learning." *Journal of Transportation Engineering, Part A: Systems* 148, no. 4 (2022): 04022007.

8. Li, Can, Lei Bai, Wei Liu, Lina Yao, and S. Travis Waller. "Graph neural network for robust public transit demand prediction." *IEEE Transactions on Intelligent Transportation Systems* (2020).

9. Hao, Siyu, Der-Horng Lee, and De Zhao. "Sequence to sequence learning with attention mechanism for short-term passenger flow prediction in large-scale metro system." *Transportation Research Part C: Emerging Technologies* 107 (2019): 287-300.

10. Tang, Liyang, Yang Zhao, Javier Cabrera, Jian Ma, and Kwok Leung Tsui. "Forecasting short-term passenger flow: An empirical study on shenzhen metro." *IEEE Transactions on Intelligent Transportation Systems* 20, no. 10 (2018): 3613-3622.

11. Wang, Peng, and Yuan Liu. "Network traffic prediction based on improved BP wavelet neural network." In *2008 4th International Conference on Wireless Communications, Networking and Mobile Computing*, pp. 1-5. IEEE, 2008.

12. Li, Yang, Xudong Wang, Shuo Sun, Xiaolei Ma, and Guangquan Lu. "Forecasting short-term subway passenger flow under special events scenarios using multiscale radial basis function networks." *Transportation Research Part C: Emerging Technologies* 77 (2017): 306-328.

13. Wang, Xuemei, Ning Zhang, Ying Chen, and Yunlong Zhang. "Short-term forecasting of urban rail transit ridership based on ARIMA and wavelet decomposition." In *AIP Conference Proceedings*, vol. 1967, no. 1, p. 040025. AIP Publishing LLC, 2018.

14. Wang, Xuemei, Ning Zhang, Yunlong Zhang, and Zhuangbin Shi. "Forecasting of short-term metro ridership with support vector machine online model." *Journal of Advanced Transportation* 2018 (2018).

15. Li, Yaguang, Rose Yu, Cyrus Shahabi, and Yan Liu. "Diffusion convolutional recurrent neural network: Data-driven traffic forecasting." *arXiv preprint arXiv:1707.01926* (2017).

16. Li, Yaguang, Rose Yu, Cyrus Shahabi, and Yan Liu. "Diffusion convolutional recurrent neural network: Data-driven traffic forecasting." *arXiv preprint arXiv:1707.01926* (2017).

17. Chan, Kit Yan, Tharam S. Dillon, Jaipal Singh, and Elizabeth Chang. "Neural-network-based models for short-term traffic flow forecasting using a hybrid exponential smoothing and Levenberg–Marquardt algorithm." *IEEE Transactions on Intelligent Transportation Systems* 13, no. 2 (2011): 644-654.

18. Williams, Billy M., Priya K. Durvasula, and Donald E. Brown. "Urban freeway traffic flow prediction: application of seasonal autoregressive integrated moving average and exponential smoothing models." *Transportation Research Record* 1644, no. 1 (1998): 132-141.

19. Danfeng, Yan, and Wang Jing. "Subway passenger flow forecasting with multi-station and external factors." *IEEE Access* 7 (2019): 57415-57423.

20. Ma, Zhenliang, Jianping Xing, Mahmoud Mesbah, and Luis Ferreira. "Predicting short-term bus passenger demand using a pattern hybrid approach." *Transportation Research Part C: Emerging Technologies* 39 (2014): 148-163.

21. Lin, Ciyun, Kang Wang, Dayong Wu, and Bowen Gong. "Passenger flow prediction based on land use around metro stations: a case study." *Sustainability* 12, no. 17 (2020): 6844.

22. Tao, Sui, Jonathan Corcoran, Francisco Rowe, and Mark Hickman. "To travel or not to travel: 'Weather'is the question. Modelling the effect of local weather conditions on bus ridership." *Transportation research part C: emerging technologies* 86 (2018): 147-167.

23. Wei, Ming. "How does the weather affect public transit ridership? A model with weather-passenger variations." *Journal of transport geography* 98 (2022): 103242.

24. Hasnine, Md Sami, Jason Hawkins, and Khandker Nurul Habib. "Effects of built environment and weather on demands for transportation network company trips." *Transportation Research Part A: Policy and Practice* 150 (2021): 171-185.