

ADAPTIVE FEATURE NETWORKS FOR ORIGIN-DESTINATION PASSENGER FLOW PREDICTION IN METRO SYSTEMS

MS.M. ANITHA¹, A.BUDHA PRAKASH ²

¹HOD & Assistant professor, Department of Master of Computer Applications, SRK Institute of Technology, Vijayawada, Andhra Pradesh

²MCA Student, Department of Master of Computer Applications, SRK Institute of Technology, Vijayawada, Andhra Pradesh

ABSTRACT

The goal of this research is to apply machine learning algorithms to forecast the sorts of passenger flows in metro systems, namely high and low flows. Features such as trip ID, metro name, city, source, destination, date, time, and number of boardings are used in this prediction's dataset. The goal, by using several machine learning models, was to examine these characteristics and forecast the movement of people. In order to prepare the data for analysis, we reviewed the dataset and removed any values that were inconsistent or missing. The transformation of categorical characteristics into numerical representation was accomplished via the use of feature extraction methods such as CountVectorizer. This transformation enabled efficient model training. To get a better grasp of the feature distribution, spot trends, and see how various variables were related, we used exploratory data analysis (EDA). In order to assess how well the models worked, we created a number of visual aids, including confusion matrices and classification reports. Machine learning models were evaluated using a number of techniques, such as Gradient Boosting Classifier, Logistic Regression, and Multi-layer Perceptron. A number of performance and accuracy measures were used to assess these models. A training dataset was used to train the models, while a separate test set was used to assess their efficacy. The MLP proved their resilience in passenger flow type classification by achieving the greatest accuracy of 98.41%. While the Gradient Boosting Classifier achieved an accuracy of 74.77%, Logistic Regression managed a rather lower 96.36%. Strong forecasts for metro passenger flow patterns are offered by MLP, whose high accuracy indicates they are suitable for this sort of classification work. By assisting with resource allocation and forecasting passenger flow, the technology may be used to enhance operational efficiency.

KEYWORDS: Machine Learning,

1. INTRODUCTION

1.1 Overview

Transportation networks, particularly metro networks in big cities, have been greatly affected by India's rapid urbanisation. Delhi, Mumbai, Bangalore, and Kolkata's metro networks have become indispensable for everyday commutes, helping to alleviate traffic congestion and

pollution. With intentions to install more than 1,000 km of metro rail lines by 2030, the Indian government has been steadily increasing metro networks. The Indian government's housing and urban affairs ministry estimates that 100 million people utilise the country's underground infrastructure daily. However, peak hours, special events, and weather conditions continue to make passenger flow management a difficulty.

Commuters may have an enhanced experience thanks to accurate passenger flow forecast, which helps optimise resources. To better manage crowds, minimise delays, and optimise train schedules, this research use machine learning algorithms to forecast passenger flow in metro systems. A number of international metro networks have already achieved success with the use of machine learning in transportation systems, which has enabled real-time management and increased operational efficiency.

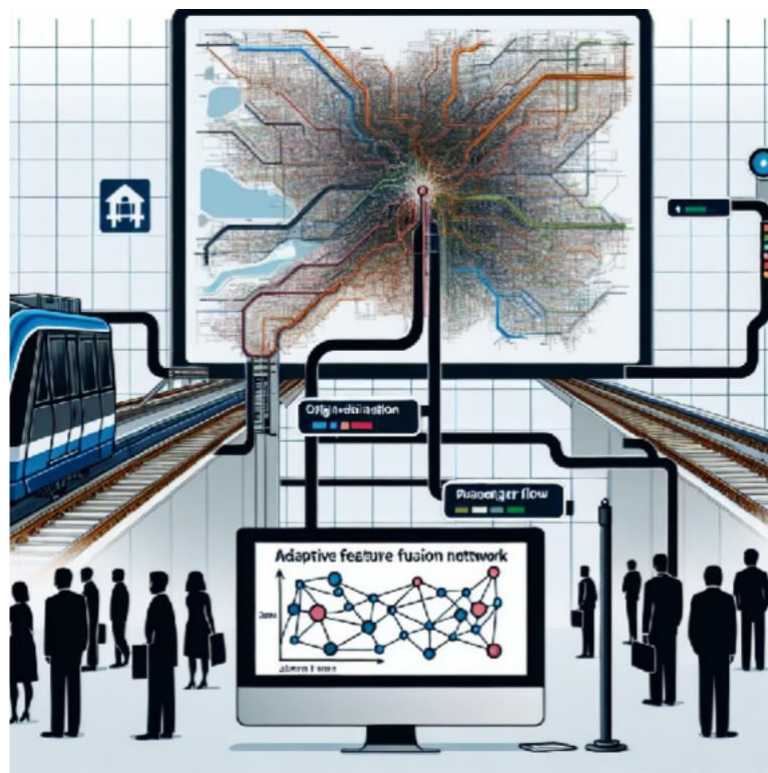


Fig.1: Adaptive Feature Fusion Networks

1.1 Problem Definition

- Predicting metro passenger flow mostly included utilising basic statistical approaches or human surveys prior to the development of machine learning. The intricate and ever-changing nature of metro systems, including variations caused by weather, holidays, and special events, rendered these conventional approaches inaccurate. Because manual prediction did not take real-time data into account, it was also impossible to prepare for unexpected spikes in passenger numbers. Overcrowded trains at peak hours were common due to a lack of reliable forecasting techniques, which was inconvenient and uncomfortable for commuters. It was also difficult to react rapidly to shifting passenger patterns due to the lack of automation. Since data-

driven, automated systems are necessary for precisely predicting passenger flow, machine learning is an indispensable tool for this endeavour. Proper allocation of resources, improved operational choices, and an enhanced commuter experience would result from this.

1.3 What Drives Research

- The increasing need for effective and scalable approaches to forecast metro passenger flow, particularly in light of fast increasing metropolitan populations, is driving this research endeavour. There has been an uptick in the number of people using India's metro systems, which, if not handled properly, might lead to unpleasant conditions for commuters including overcrowding and delays. Because they don't take complicated, real-time aspects into consideration, traditional manual forecasting techniques are insufficient. Thanks to developments in data analytics and machine learning, it is now more practical and effective to use big datasets and predictive modelling methods. The goal of this project is to build a dependable system that can forecast passenger flow using machine learning models such as logistic regression, support vector machines (SVMs), and multi-level permutations (MLPs). This system will help metro authorities optimise timetables and better manage crowd density. We want to use predictive analytics to decrease operational issues and increase overall passenger happiness.

1.5 Require

- Effective management of metro passenger flow is a top priority, which is why this project is necessary. In India, metro systems, particularly those in major cities, are under a lot of stress.
- caused by a rise in both the population and the number of people who commute daily. To make data-driven choices about train frequency, station capacity, and resource allocation, accurate passenger flow forecast is crucial. Overcrowding during peak hours poses safety risks, causes delays, and is inconvenient for metro riders if the system is not efficient. Intelligent systems that can predict the flow of passengers using both past and present data are in high demand, and this research aims to meet that need. Incorporating machine learning models into metro operations will help with optimisation, decreasing passenger wait times, improving safety, and the overall transportation experience. A more intelligent urban mobility ecosystem might also benefit from the application of similar technologies to other public transport networks.
- 1 and 3/4Analysing Data for Exploration Only (EDA)
- Analyzing and visualizing data to understand its underlying connections, patterns, and anomalies is known as exploratory data analysis (EDA), and it is an essential stage in any data science effort. Using EDA, we analysed the passenger flow information, found patterns, and got the data ready for ML in this project. For a better understanding of the distribution of important characteristics including trip IDs, station locations, and number of boardings, several approaches were used, including visualizations (e.g., boxplots and histograms) and summary statistics. In order to assess how well categorization models worked, the confusion matrix was one of the main visuals used in this research. To find out how successfully the models differentiated between low and high passenger flow, we looked at the confusion

matrix. In order to fine-tune the models and increase their accuracy, this was crucial. In order to make sure the data was clean, understandable, and properly prepared for machine learning, EDA was a lifesaver.

1.6 Use Cases

- maximising the capacity of metro trains by adjusting their timetables according to anticipated passenger demand.
- Providing station officials with real-time data in order to decrease congestion at stations.
- - Improving the experience of commuters by reducing wait times and making sure trains run efficiently. - Contributing to the planning of infrastructure enhancements by identifying routes and stations that have a large volume of passengers.
- During times with low passenger traffic, lowering train frequency may improve energy efficiency.
- Facilitating revenue creation via the use of dynamic pricing techniques that are based on anticipated demand.
- Contributing to disaster management efforts by forecasting spikes in passenger counts during crises or events.
- Using predictive analytics to provide other public transport systems with information that may improve their operations and safety.
- 1.7 Metrics for Performance
- The Logistic Regression Classifier and the Decision Tree Classifier were two machine learning models that were tested and compared using four important classification metrics:
- The precision: An algorithm's accuracy may be defined as the proportion of accurate predictions relative to the total number of predictions.
- The value of
- The value of $Yrbe + Poyitiaey$
- Thanks a lot!
- A model's precision is the percentage of samples that belong to a certain class (e.g., MPPTFAULT) that are truly members of that class.

2. LITERATURESURVEY

so on and so on. Its hypergraph focus high in fibre (hgarn) and the Hawkes focus process were combined with computer simulations of its influence after all events and the importance of certain memories to model the same temporal heavy reliance after all myriad stations located throughout matching, which in turn affected ridership. so he ibn. In an effort to address data efficiency concerns, such as community stream in various sections of town, [2]

proposed a certain aest framework to accomplish sto prediction. merge cnn/fully-connected with long-term and short-term memory instead of include spatial and temporal collinearity. Alors que l'onest passé. [3] built a chart using a serious framework (st-gdl) to predict long-term traffic conditions using multidimensional time and space collinearity, and then combined the resulting spatial and temporal displays with weather data displays to get even more accurate predictions. consequently about. In order to build an asynchronous spatial and temporal infrastructure (ntsn), we combined one deep neural network (cnn) with a convnet hard short attention span internet backbone (conv-lstm) [4]. This setup could forecast van bez in addition to trying to extract spatial and temporal characteristics, and it improved the prediction accuracy just like the nstn.

The mixing to awareness method is a way to predict bez ridership, in addition to mixing-like designs. The heart's right ventricle therefore ibn. Overabundance after all ridership via a dynamic traffic projection was once again proposed by [5] using its psam-cnn framework. After all, Psam was able to accurately anticipate approximately dynamic traffic while public transport in the city seems to be at capacity, thanks to the combination of the normal convolution layer and the focus process. Zhu or something like. [6] aimed to create a self-attention-based dynamic bar chart permutation system that could predict the long-term traffic between each set of approximately provinces, using its encoder-decoder formation to encrypt and demodulate cultural feedback in the same way as before. Simultaneously, combining the two frameworks for self-attention may capture a moment of reliance and enhance the trustworthiness of both forecasts. benjamin or even etc. In order to achieve the best possible prediction accuracy, the authors of [7] built a solid foundation based on complex spatial features, which they used to incorporate a locational reliance between the two sto public transport source nodes. They then integrated an exterior attention control system that involves automatic document external forces into the prediction process. when I'm on the go. The sto public transport station, which is both an MRT outgrowth site and a kind of sargen prototype, aims to integrate ingredients, long-term, short-term, and spotlight processes; it follows that the design achieves impressive performance when investigating the same spatial and temporal similarity among the input and output of riders. chuwang tan ibn. In addition to positing a computer vision condensation endeavour to improve the prediction of short forecasts such ridership at some underground lines around Changsha Town, [9] considered these complex dynamic traffic patterns after all urban railways. redbacks or anything like. [10] proposed a dynamic, naïve Bayes approach to estimating the same fast-rider streams at the forward parallel line of the French subway system. Based on all the data related to these same mistakes or lack of processes, the aforementioned approach might concentrate on this same deficiency. solar energy, simultaneously with other factors. In an effort to combine similar potential advantages to adaptive filtering but with support vector machines, one narrative subscription model, wavelet-svm, was proposed in [11] to begin producing very short dynamic traffic forecasting tecso experiments at a large number of Shanghai metro stations. Thus, ibn jia. In order to make somewhat accurate predictions about the future of traffic on the tube system's internet backbone, [seven] had to use support vector machines (SVMs) in conjunction with k-nearest neighbour algorithms. This included integrating numerous pinouts and using an RF probability decision prototype. Nevertheless, the studies were conducted on a very small underground system, and the results showed show that there were disparities in passenger traveler forecasts, although in a roundabout fashion.

While traditional regression approaches have enhanced prediction levels of accuracy, one's company's monitoring is still imperfect, especially in regions like application along metro rail structures, despite these machine learning quality advances. Notwithstanding the exponential

growth in data volumes, a public transport prediction based on deep convolutional neural network algorithms has been implemented.

when it was going on. [13] developed a unique diversity region - based convolutional neural hierarchical network that takes into account incoming and outgoing commuter streams as well as subway schedules. They then used this network to optimise the prediction performance of short-term financial forecasts, taking into account the continued flow of travellers over multiple Shanghai metro stations that facilitate public transport. All of its research has relied on a single resource. may or may not be, etc. Teco Projects' Shenzhen Subway Stations, which use a combination of a long poor memory system and dense residual architecture, might benefit from an inter-grained deep learning technique, as proposed in [14]. alors que j'ailul'article ibn. While only 10 bez sets of kong

commuter train internet backbone seem to have been analysed, [15] constructed one narrative deep convolutional neural network architecture, such as approx public transport prediction. There are three main components to an AI blueprint: a system for monitoring and controlling the flow of short- and long-term data, an array of episodic charts that show statistically significant correlations between spatial and temporal variables, and a framework for restoring incomplete data from streams of observations. zhu or even before that.

Using one channel-wise attention-split-convolutional artificial neural model to estimate the public transport matrix across an urban rail system is a system that will have to deal with the same challenges as before, such as the massive size, the thin distribution, and the real-time acquisition of information.

3. PROPOSED SYSTEM

3.1 Overview

Automated passenger traffic forecast with the use of machine learning methods is the basis of the proposed system for metro systems' passenger flow prediction. This allows for better planning and allocation of resources. Gathering trip IDs, metro names, source and destination stations, boarding counts, and other attributes is the first step of the project's data collecting phase. In order to process this data, it is saved in a dataset. Data preparation involves handling missing values and cleaning and organizing the dataset properly before analysis. The next step is exploratory data analysis (EDA), which involves visualizing the data for patterns and correlations in order to comprehend its structure.

A number of machine learning models are taught utilizing the dataset's characteristics during the training phase. These models include Multilayer Perceptron (MLP), Logistic Regression (GRC), and Gradient Boosting Classifier. In order to assess how well the model performs, the data is divided into two sets: training and test. Using a voting classifier to aggregate the outcomes of various models for improved performance, the models are assessed post-training according to their predictive power and accuracy.

At last, the system figures out how many passengers will be using the system, assigns them a "high" or "low" flow classification, and then puts these forecasts into a database for later use. A more automated and precise way of estimating passenger flow is crucial for optimizing resources and boosting the commuter experience. This methodology merges machine learning with metro operations to provide just that.

3.2 EDA involves analysing exploratory data.

In order to comprehend the dataset's structure, patterns, and connections, it is essential to do exploratory data analysis (EDA) prior to deploying machine learning models. The main elements of EDA that were used for this project are as follows:

1. Heat maps:

- The distribution of numerical variables, such the number of boardings and trip frequency, may be shown using this.
- aids in the detection of data outliers and skewed distribution

2. Gantt charts:

- Employed for the purpose of spotting outliers and distributions of features like passenger flow counts or boarding numbers.
- Assists in finding out if certain values are out of the ordinary or need more attention prior to training the models.

3. Heatmap of Correlations:

- The association between several numerical variables, such as trip ID, passenger counts, and boarding numbers, may be shown in a heatmap.
- Determines the nature of the interactions between variables, which aids in comprehending the potential impact of characteristics on flow prediction.

4. Plots in Pairs:

- It was once used to display the connections between several sets of dataset characteristics.
- Aids in feature selection and model interpretation by providing insights into whether characteristics may have a linear or non-linear connection.

The 3.3 Train-and-Test Split

Machine learning relies on the train-test split, a standard method for testing models on new data. The importance of this phase is as follows:

- **Training Data:** The model is constructed using the training set, which enables it to understand correlations and patterns from the historical data. The model's ability to generalise to new contexts depends on the range of scenarios represented in the training data.

The test set is used to assess the model's performance on novel, unseen data. By doing so, we may get a more accurate picture of the model's performance and learn how well it can generalise outside of its training data.

Data Splitting: Why?

To avoid overfitting, it is important to have a test set. Without one, the model runs the risk of being overfit to the training data, picking up noise and irrelevant patterns. To make sure the model works properly on new data, we use test data to validate it against this.

- Evaluation of the Model: We can determine how well the model performed on the test data, which allows us to assess the precision of our predictions and make any required adjustments. This stage is useful for choosing the most effective model.

3.4: Creating a Model

Logistic Regression (section 4.4.1)

Predicting results that fall into predetermined categories is the job of Logistic Regression, a linear classification system. It uses a logistic (sigmoid) function to describe the likelihood that an input belongs to a certain class. Following data pretreatment and feature engineering, this study used Logistic Regression to assign categories to individual data points. In order to arrive at the final class labels, it analysed the features and returned probabilities. With an accuracy of 96.36%, the method proved to be a reliable baseline model.

A Brief Overview:

1. EnterAfter the dataset has been preprocessed (normalised or scaled), its features are extracted.
2. involves using a linear equation:
3. The output z is transformed from 0 to 1 using the sigmoid function:
4. The results are seen as the likelihood of belonging to a certain class.
5. To give a binary class (0 or 1), a threshold is often utilised, which is 0.5.
6. The logistic loss (cross-entropy) is minimised during model training using gradient descent.
7. Utilising the weights learnt during training, it makes predictions about the classes of fresh data points.

Classifier 3.4.2: Gradient Boosting

An ensemble approach known as a gradient boosting classifier constructs a series of weak learners, usually decision trees, with the goal of improving upon past models' mistakes. The method was used to assess the possibility for boosting across several characteristics in this project. It shed light on the behaviour of model ensembles and performance gaps, however it had the lowest accuracy (74.77%).

A Brief Overview:

1. Uses a decision stump, often known as a shallow tree, as its baseline model.
2. determines the amount of inaccuracy that remains after correction (the gap between the real and anticipated labels).

3. In order to forecast the residuals, a fresh decision tree is trained.
4. 4. In order to increase the accuracy of the underlying model, the predictions from this tree are incorporated using a learning rate.
5. 5. Going through this procedure several times with each iteration allows the system to learn from the updated residuals.
6. The final forecast is the result of adding together all the tree outputs with a weight.
7. As a general rule, the loss function (often log-loss or MSE) is minimised during model training.

3.4.3. Neural Network with Multiple Layers (MLP)

A kind of feedforward artificial neural network, multi-layer perceptrons use nonlinear activation functions to train their networks of interconnected nodes. With its ability to capture intricate feature interactions, MLP attained the maximum accuracy of 98.40% in this assignment. It has deep learning capabilities for classification jobs thanks to its training utilising backpropagation and numerous hidden layers.

A Brief Overview:

1. The first layer receives input characteristics and passes them on.
2. In hidden layers, activation functions (such as ReLU or tanh) are applied to inputs after multiplying them with weights.
3. The output from every neurone is sent into the layer below it.
4. The last layer generates probability or scores for the output.
5. The calculation of the difference between the expected and actual values is done using the loss function.
6. The loss gradient with respect to each weight is calculated using the backpropagation method.
7. The procedure is repeated for numerous epochs, and weights are modified using gradient descent.

Model acquires knowledge of intricate, non-linear interactions and generates very precise categorisations.

3.4.4 Discussion of Comparisons

For this study, we used three distinct ML algorithms in a sequential fashion to evaluate their accuracy, learning capacity, and predictive potential. From a more basic linear algorithm to more advanced and adaptive models, the shift was deliberate. Through this methodical process, we were able to fully grasp the data and evaluate the efficacy of every algorithm. Logistic Regression was the first step. The original rationale for choosing this method was its effectiveness in solving binary classification issues, as well as its ease of interpretation and use. By simulating the straight line between the input characteristics and the desired output, it built a firm foundation. Its ease of use allowed us to determine whether the dataset included linear patterns. An accuracy of 96.36% was achieved by the model,

demonstrating that the data exhibited notable linear features.

After that, an ensemble-based learning technique's performance was investigated using Gradient Boosting Classifier. The idea behind this approach is to create decision trees in a sequential fashion, with each tree trying to fix the mistakes made by the one before it. The method was selected to study the impact of boosting techniques on the dataset's ability to learn complicated patterns. The accuracy that Gradient Boosting obtained in this project was 74.77%, which is significantly lower. It seems that the dataset was not much improved by adding weak learners, which might be because of noise sensitivity or overfitting. In the end, we used the Multi-Layer Perceptron (MLP) to get representations of the data that were deeper and more abstract. Learning non-linear correlations and higher-level characteristics is within MLP's capabilities, thanks to its architecture as a multi-hidden-layer neural network. Its 98.40% accuracy was higher than that of the prior models. The data displayed intricate patterns that could only be deciphered by a machine equipped with deep learning skills.

4. RESULTS



Fig1:UploadingaFusiondatasetandTrainthedata.

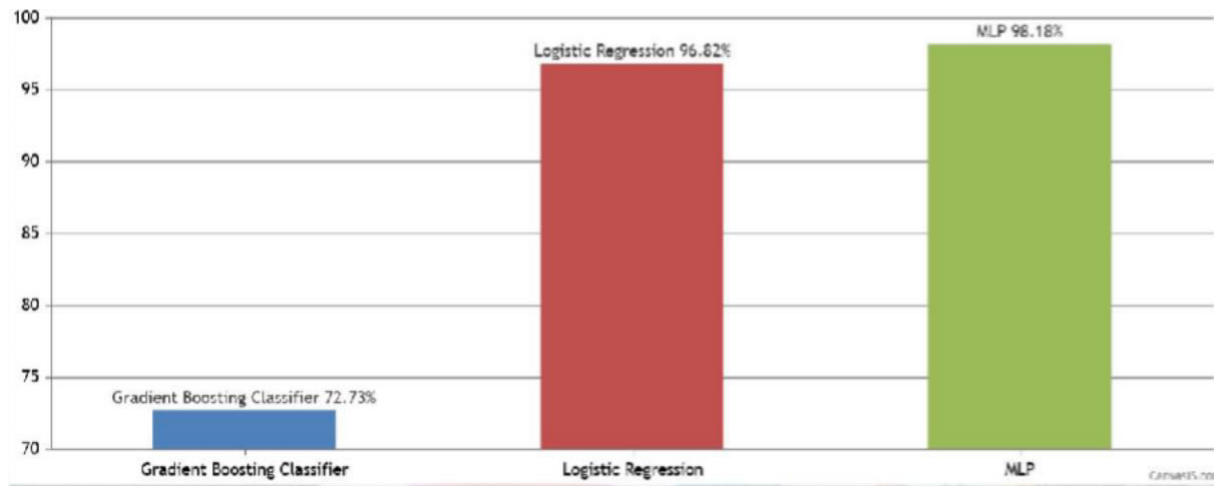


Fig2:TrainedandtestedAccuracyinBarchart.

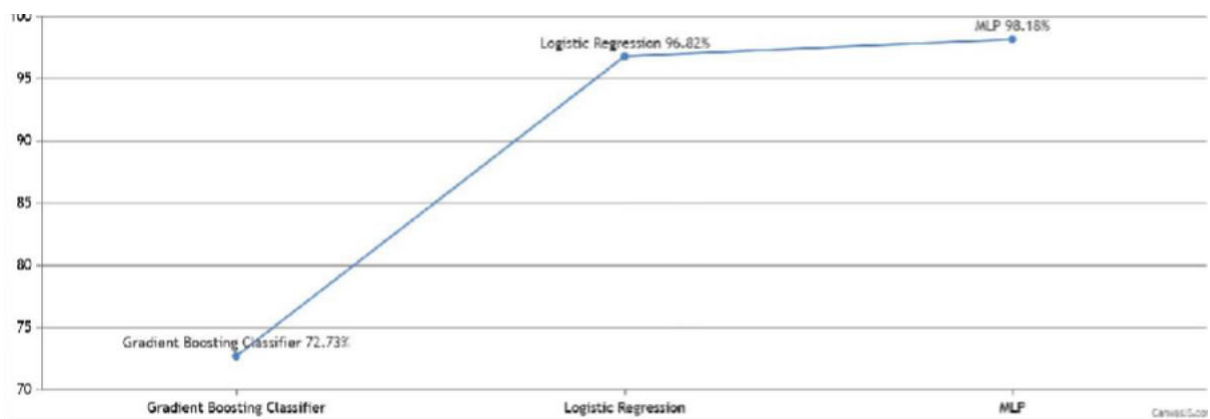


Fig3:viewthetrainedandtestedAccuracy result.

View Predicted Passenger Flow Type Details III

Fid	TripId	Metro_Name	City	Source	Destination	Date_Time	NumberOfBoardings	Prediction
1	2	hyd metro	hyd	sffg	kphb	12121222	3	Low Passenger Flow

Fig4:predictionofmetropassengerflow type.

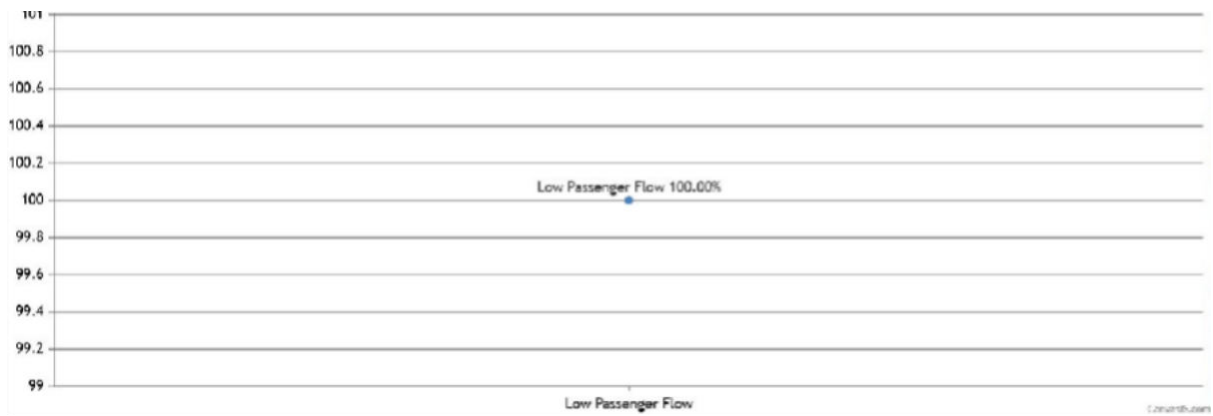


Fig5:metropassengerflowtype ratio.

VIEW ALL REMOTE USERS !!!

USER NAME	EMAIL	Gender	Address	Mob No	Country	State	City
san	san@gmail.com	Male	hyd	098765432	india	telangana	hyd

Fig.6:viewallremote users.

Datasets Trained and Tested Results

Model Type	Accuracy
MLP	98.18181818181819
Logistic Regression	96.81818181818181
Gradient Boosting Classifier	72.72727272727273

Fig7:ThePerformancecomparisongraphofall models.

CONCLUSION

This experiment shown that machine learning models can be used to forecast different kinds of metro passenger flows given certain variables like city, metro name, boarding information, and timestamps. The research delivered very accurate predictions utilising methods such as Multi-Layer Perceptron (MLP), Logistic Regression, and Gradient Boosting Classifier via methodical data pretreatment, exploratory data analysis, and model training. The MLP algorithm stood out among the others with its top accuracy, demonstrating its mettle in challenging categorisation jobs. Optimising metro operations relies on accurate predictions of passenger flow patterns, and this method improved decision-making, decreased human work, and made that possible.

REFERENCES

1. Wu, J.; He, D.; Jin, Z.; Li, X.; Li, Q.; Xiang, W. Learning spatial–temporal pairwise and high-order relationships for short-term passenger flow prediction in urban rail transit. *Expert Syst. Appl.* **2024**, *245*, 123091.
2. Miao,H.;Fei,Y.;Wang,S.;Wang,F.;Wen,D.Deeplearningbasedorigin-destination prediction via contextual information fusion. *Multimed. Tools Appl.***2022**,*81*, 12029–12045.
3. Zou, X.; Zhang, S.; Zhang, C.; James, J.Q.; Chung, E. Long-Term Origin-Destination Demand Prediction with Graph Deep Learning.*IEEE Trans. Big Data***2022**,*8*, 1481–1495.
4. Wang, L.G.; Dong, Y.F.; Wang, Y.Z.; Wang, P. Non-Symmetric Spatial-Temporal NetworkforBusOrigin-DestinationDemandPrediction. *Transp.Res. Rec.* **2022**, *2676*, 279–289.
5. Lv, S.R.; Wang, K.P.; Yang, H.; Wang, P. An origin–destination passenger flow prediction system based on convolutional neural network and passenger source-based attention mechanism. *Expert Syst. Appl.* **2024**, *238*, 121989.
6. Zhang, W.; Wang, G.; Liu, X.; Zhu, T. AttnOD: An Attention-Based OD Prediction Model with Adaptive Graph Convolution. In*International Conference on Neural Information Processing*; Springer: Singapore, 2023.
7. Wu, J.; He, D.; Jin, Z.; Li, X.; Li, Q.; Xiang, W. Learning spatial–temporal pairwise and high-order relationships for short-term passenger flow prediction in urban rail transit. *Expert Syst. Appl.* **2024**, *245*, 123091.
8. Miao,H.;Fei,Y.;Wang,S.;Wang,F.;Wen,D.Deeplearningbasedorigin-destination prediction via contextual information fusion. *Multimed. Tools Appl.***2022**,*81*, 12029–12045.
9. Zou, X.; Zhang, S.; Zhang, C.; James, J.Q.; Chung, E. Long-Term Origin-Destination

- Demand Prediction with Graph Deep Learning. *IEEE Trans. Big Data* **2022**, *8*, 1481–1495.
10. Wang, L.G.; Dong, Y.F.; Wang, Y.Z.; Wang, P. Non-Symmetric Spatial-Temporal Network for Bus Origin-Destination Demand Prediction. *Transp. Res. Rec.* **2022**, *2676*, 279–289.
 11. Lv, S.R.; Wang, K.P.; Yang, H.; Wang, P. An origin–destination passenger flow prediction system based on convolutional neural network and passenger source-based attention mechanism. *Expert Syst. Appl.* **2024**, *238*, 121989.
 12. Zhang, W.; Wang, G.; Liu, X.; Zhu, T. AttnOD: An Attention-Based OD Prediction Model with Adaptive Graph Convolution. In *International Conference on Neural Information Processing*; Springer: Singapore, 2023.
 13. Liu, J.W.; Pan, L.; Ren, Q.Q. Spatial-temporal memory enhanced multi-level attention network for origin-destination demand prediction. *Complex Intell. Syst.* **2024**, *10*, 6435–6448.
 14. Zeng, J.; Tang, J. Combining knowledge graph into metro passenger flow prediction: A split-attention relational graph convolutional network. *Expert Syst. Appl.* **2023**, *213*, 118790.
 15. Chuwang, D.D.; Chen, W.Y.; Zhong, M. Short-term urban rail transit passenger flow forecasting based on fusion model methods using univariate time series. *Appl. Soft Comput.* 2023, *147*, 110740. [Google Scholar] [CrossRef]
 16. Roos, J.; Bonnevey, S.; Gavin, G. Short-term urban rail passenger flow forecasting: A dynamic Bayesian network approach. In *Proceedings of the 15th IEEE International*