Journal of Computer and Electrical Sciences (JCES)- Volume-1 Issue-4, April 2026
ISSN: 3049-2602

SUBWAY PASSENGER FLOW FORECASTING

[1] M. Havila, [2] K. Naveen

[1] Assistant Professor, Department of Computer Science and Engineering

[2] Student, Department of Computer Science and Engineering

Andhra Loyola Institute of Engineering and Technology

Email: havilahmurray58@gmail.com

Email: katchallanaveen9@gmail.com

Abstract: *The rapid growth of urban populations and increasing dependency on public transportation systems have created significant challenges in managing passenger flow within subway networks. Efficient prediction of passenger traffic is essential to reduce congestion, improve operational efficiency, and enhance commuter safety. This paper presents a smart subway passenger flow forecasting system that leverages machine learning techniques to predict passenger volume at metro stations. The proposed system integrates a web-based platform developed using Django, enabling seamless interaction between administrative and user modules. The model is trained using a structured dataset containing key features such as station identity, time of travel, day of the week, weather conditions, temperature, holidays, and special events. A neural network-based approach is employed to capture complex patterns and relationships within the data, improving prediction accuracy. The system also incorporates data visualization tools, including dashboards, statistical charts, and performance metrics such as accuracy, precision, recall, and F1-score for model evaluation. Experimental results demonstrate that the proposed approach achieves reliable predictive performance and supports real-time decision-making in urban transportation systems. The developed system provides a scalable and intelligent solution for smart city applications, enabling efficient passenger flow management and improved transit planning.*

Keywords: Subway Passenger Flow, Machine Learning, Neural Networks, Smart Transportation, Prediction System, Data Analytics

I. INTRODUCTION

The rapid growth of urban populations and the increasing dependence on public transportation systems have created significant challenges in managing passenger flow within metropolitan subway networks. Efficient passenger flow management is crucial to ensure commuter safety, reduce congestion, and maintain the reliability and punctuality of transit services. In highly populated cities, issues such as peak-hour overcrowding, sudden passenger surges caused by public events, and external factors like weather conditions can disrupt normal subway operations and lead to inefficiencies. Traditional passenger flow management techniques mainly rely on historical data analysis and manual estimation methods, which provide only a limited understanding of traffic

https://doi.org/10.5281/zenodo.19609934

behaviour. These approaches often fail to adapt to real-time and dynamic changes, making it difficult for transportation authorities to respond quickly and effectively. As a result, challenges arise in scheduling, resource allocation, and crowd control, ultimately affecting service quality. With advancements in machine learning and data analytics, it is now possible to process large volumes of transportation data and identify meaningful patterns for accurate forecasting. Machine learning models can analyse complex relationships between influencing factors such as time, weather conditions, holidays, and special events, enabling more precise predictions of passenger flow. This work proposes a Smart Subway Passenger Flow Forecasting System that integrates machine learning techniques with a web-based application framework to deliver real-time predictions and interactive visualizations. The system utilizes a neural network-based model to capture nonlinear patterns in the data, significantly improving prediction accuracy. By offering actionable insights to both administrators and passengers, the system supports better decision-making, enhances operational efficiency, reduces overcrowding risks, and contributes to the development of sustainable and intelligent smart city transportation infrastructure.

II. LITERATURE REVIEW

Several research studies have been conducted on passenger flow management in public transportation systems. These studies highlight the importance of analysing passenger movement patterns to reduce congestion.

From the reviewed literature, the following key observations can be made:

- Many transportation systems use automated ticketing and sensors to track passenger movement. Data collected from these systems helps identify peak hours and overcrowded stations.
- Modern smart transportation systems integrate data analytics and software solutions to improve passenger flow. Web-based systems allow administrators to monitor passenger traffic in real time.
- The literature suggests that using software systems for passenger flow analysis can significantly improve the efficiency of transportation networks.

III. PROPOSED SYSTEM

The proposed Smart Subway Passenger Flow Forecasting System is designed using a modular architecture that integrates machine learning with a web-based interface. The system is divided into two primary components: the Admin Module and the User Module, both connected to a centralized database and machine learning engine. The User Module allows end-users to interact with the system by providing features such as registration, login, and dashboard access. Users can view summarized information, explore data visualizations, and predict passenger flow by entering parameters like station, time, weather, and temperature. The Machine Learning Engine acts as the

https://doi.org/10.5281/zenodo.19609934

core component, handling data preprocessing, feature selection, model training, evaluation, and real-time prediction generation. The trained model is stored for reuse, eliminating the need for retraining and improving efficiency. The Database Layer, implemented using MySQL, stores user details, station data, processed datasets, and prediction results for easy access and management. The data flow begins with admin dataset upload, followed by preprocessing and model training, after which user inputs are processed to generate predictions displayed through dashboards.

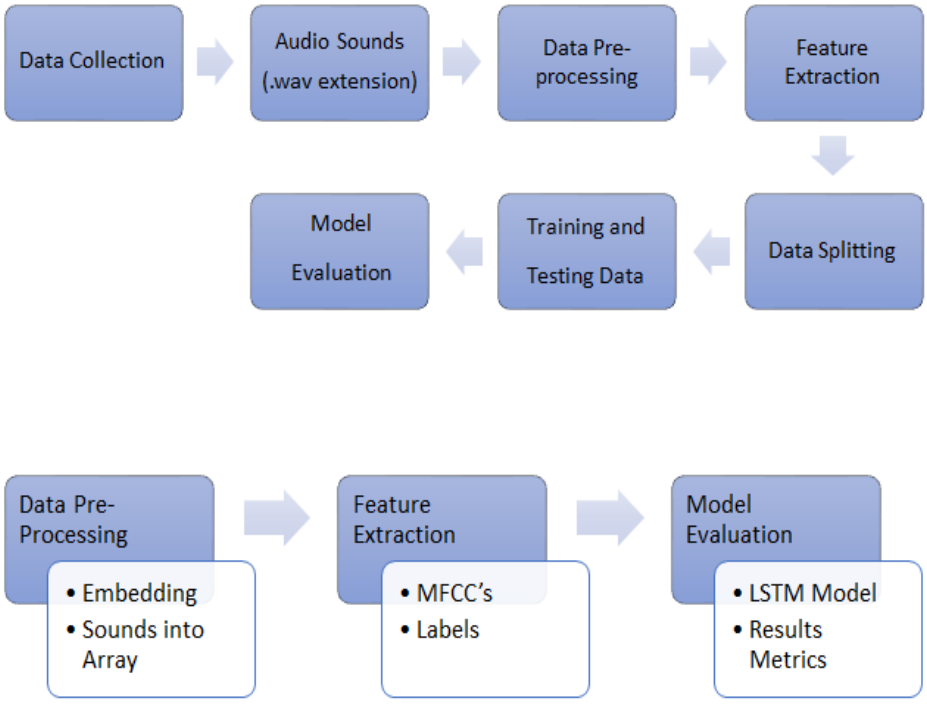

Fig. 1. System Architecture of Subway Passenger Flow Forecasting System

The overall architecture of the Smart Subway Passenger Flow Forecasting System follows a structured workflow that ensures efficient data processing and accurate prediction. The system begins with dataset collection, followed by data preprocessing to clean and prepare the data for analysis. Next, the machine learning model is trained using relevant features and evaluated to ensure accuracy and reliability. Overall, the system offers scalability, real-time predictions, seamless integration, and interactive visualization capabilities.

IV. METHODOLOGY

The proposed Smart Subway Passenger Flow Forecasting System is designed to predict passenger traffic using machine learning techniques integrated with a web-based application. The methodology consists of multiple stages, including data collection, preprocessing, feature selection, model training, evaluation, and prediction.

The methodology of the system is organised into the following steps:

https://doi.org/10.5281/zenodo.19609934

1. Data Collection:

The dataset contains structured subway passenger data including station, time, weather, and events to capture real-world travel patterns.

2. Dataset Features:

Key attributes include station ID, hour, day, weather, temperature, holidays, events, and passenger flow as the target variable.

3. User Interaction:

Data is cleaned by removing null values, duplicates, and converting categorical data into numerical form for consistency.

4. Normalization:

Numerical features like temperature are normalized to ensure uniform scaling and improve model performance.

5. Feature Selection:

Important features such as station, hour, day, weather, temperature, holiday, and events are selected based on relevance.

6. Model Development:

A neural network model is used to capture complex nonlinear relationships between input features and passenger flow.

7. Model Training:

The dataset is split into training and testing sets (80:20), and the model is trained using epochs, batch size, and optimization techniques.

8. Model Evaluation:

Performance is measured using accuracy, precision, recall, and F1-score to ensure reliable predictions.

9. Prediction Process:

User inputs are processed and converted into numerical form, and the model generates passenger flow predictions and categorize

10. Visualization Dashboard:

Graphs and charts display passenger trends, model accuracy, and confusion matrix for better analysis and decision-making.

https://doi.org/10.5281/zenodo.19609934

V. PROPOSED SYSTEM RESULTS

The performance of the proposed Smart Subway Passenger Flow Forecasting System was evaluated using a structured dataset containing multiple features related to passenger behaviour and environmental conditions. The dataset was divided into training and testing subsets using an 80:20 ratio to ensure reliable evaluation of the model.

- The model achieved 78.6% accuracy, 79.0% precision, 78.6% recall, and 78.7% F1-score, indicating reliable and balanced prediction performance.
- The results show that the model correctly predicts passenger flow for most cases while maintaining a good balance between accuracy and error reduction.
- During training, loss decreased steadily and accuracy improved across epochs, confirming effective learning of data patterns.
- The model successfully captures relationships between key features such as time, weather, and station data to improve predictions.
- Most predictions fall into correct categories, with minimal misclassification mainly occurring between closely related categories.
- Graphs like accuracy curves, confusion matrix plots, and passenger distribution charts validate consistent and stable model performance.
- The system provides real-time results including passenger flow level (Low, Medium, High) and estimated passenger count, demonstrating practical usability.

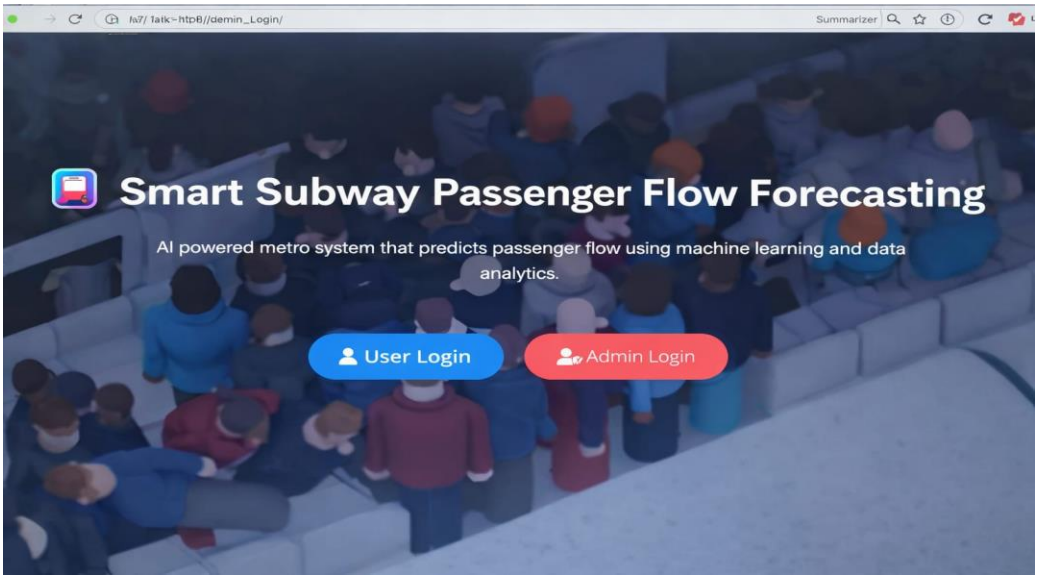

Fig. 2. User Interface

https://doi.org/10.5281/zenodo.19609934

Fig.2. represents the login interface of the Smart Subway Passenger Flow Forecasting system, where both users and administrators can securely access the platform. It highlights a clean and user-friendly design with options such as “User Login” and “Admin Login,” making it easy to navigate. The background visualization of passengers in a subway environment emphasizes the real-world application of the system, which focuses on managing and predicting metro passenger movement using advanced technologies like machine learning and data analytics.

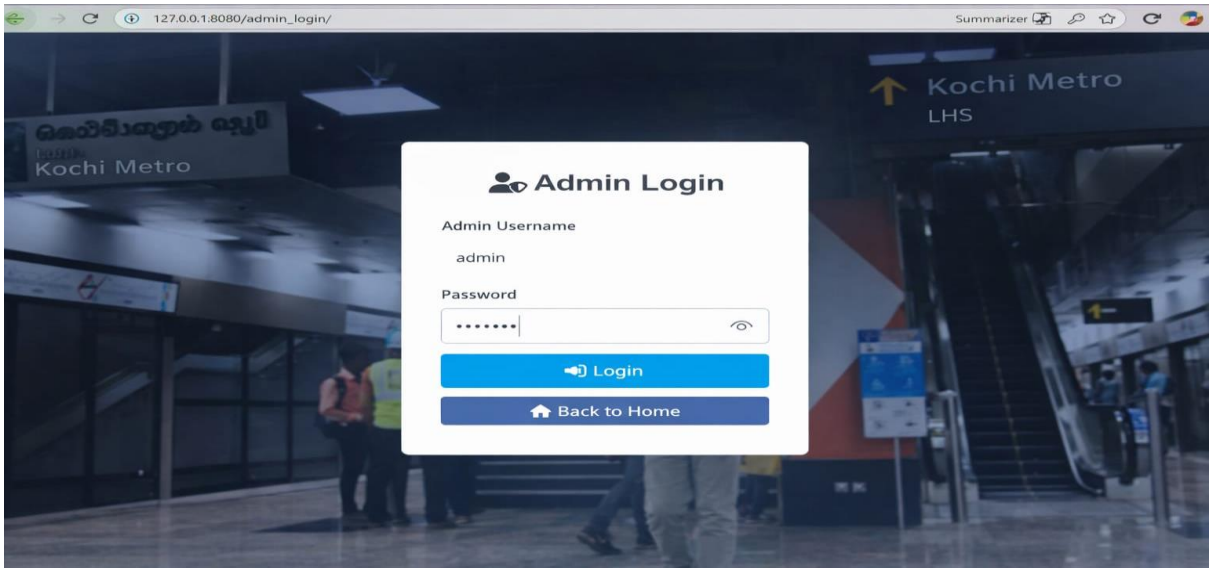

Fig. 3. Registration page

Fig.3 illustrates the admin login module, which serves as a secure gateway for administrators to manage the system. By entering valid credentials such as username and password, admins gain access to backend functionalities. This module ensures data security and controlled access, allowing only authorized personnel to monitor system activities, manage passenger data, and oversee forecasting operations efficiently. The interface design also ensures simplicity while maintaining a professional look.

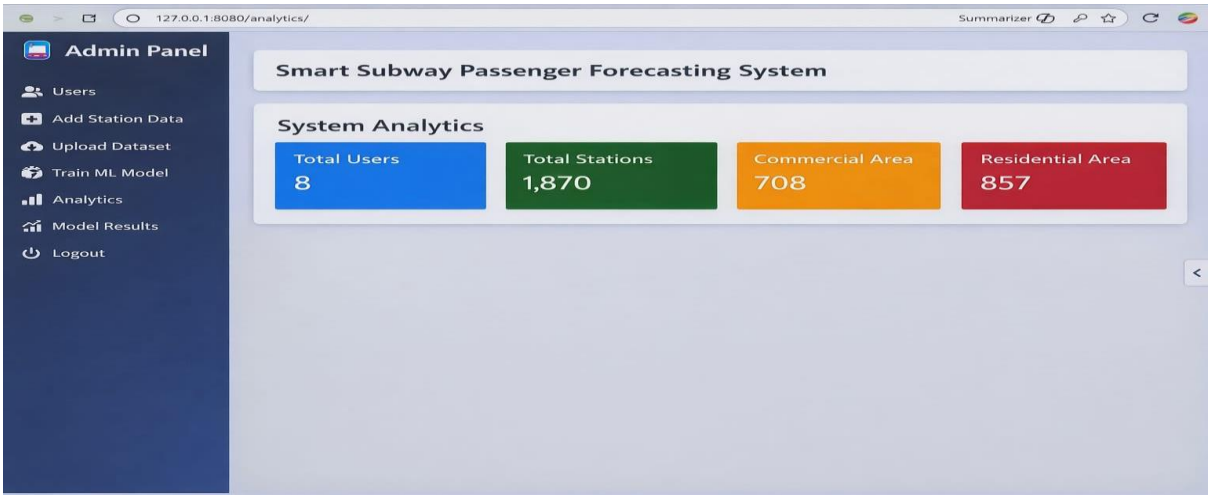

Fig. 4. Admin Page

https://doi.org/10.5281/zenodo.19609934

Fig. 4. shows the admin dashboard or analytics panel of the system, where key information is visually presented. It includes important metrics such as total users, total stations, commercial areas, and residential areas, displayed in a structured and color-coded format. The side navigation panel provides options like uploading datasets, training machine learning models, and viewing results. This dashboard enables administrators to analyze passenger flow trends, make informed decisions, and improve metro system efficiency through data-driven insights.

.

VI. CONCLUSION

The proposed Smart Subway Passenger Flow Forecasting System effectively demonstrates how machine learning techniques can be applied to predict and manage passenger traffic in modern subway networks. By integrating a neural network-based prediction model with a web-based platform developed using Django, the system delivers an efficient, scalable, and user-friendly solution for analysing and forecasting passenger flow. The model leverages multiple influencing factors such as time of day, day of the week, weather conditions, temperature, holidays, and special events to generate highly accurate predictions. The evaluation of the model using performance metrics like accuracy, precision, recall, and F1-score confirms its reliability and suitability for real-world applications. Beyond prediction capabilities, the system also incorporates interactive dashboards and visualization tools that allow users and administrators to better understand passenger patterns and trends, thereby supporting informed, data-driven decision-making. The modular design of the system ensures flexibility and scalability, enabling it to be adapted to various urban transportation scenarios and future technological advancements. Furthermore, the system enhances operational efficiency by assisting authorities in managing crowd flow, reducing congestion, optimizing train schedules, and improving overall resource allocation. By addressing key challenges in passenger flow management, the proposed system contributes significantly to the development of intelligent transportation solutions. Ultimately, it represents a progressive step toward building smarter, more efficient, and sustainable urban transit systems, aligning with the broader vision of smart city infrastructure and advanced mobility management.

https://doi.org/10.5281/zenodo.19609934

REFERENCES

[1] Y. Zheng, "Urban computing: concepts, methodologies, and applications," ACM Transactions on Intelligent Systems and Technology, vol. 5, no. 3, pp. 1–55, 2014.

[2] J. Zhang, Y. Zheng, and D. Qi, "Deep Spatio-Temporal Residual Networks for Citywide Crowd Flows Prediction," IEEE Transactions on Knowledge and Data Engineering, vol. 31, no. 1, pp. 1–14, 2019.

[3] C. Chen, D. Zhang, P. S. Castro, N. Li, L. Sun, and S. Li, "Real-time passenger flow prediction for metro systems," IEEE Transactions on Intelligent Transportation Systems, vol. 18, no. 7, pp. 1928–1939, 2017.

[4] X. Ma, Z. Tao, Y. Wang, H. Yu, and Y. Wang, "Long Short-Term Memory Neural Network for Traffic Speed Prediction Using Remote Microwave Sensor Data," Transportation Research Part C: Emerging Technologies, vol. 54, pp. 187–197, 2015.

[5] Z. Zhao, W. Chen, X. Wu, P. C. Chen, and J. Liu, "LSTM network: a deep learning approach for short-term traffic forecast," IET Intelligent Transport Systems, vol. 11, no. 2, pp. 68–75, 2017.

[6] H. Lv, Y. Duan, W. Kang, Z. Li, and F. Wang, "Traffic Flow Prediction With Big Data: A Deep Learning Approach," IEEE Transactions on Intelligent Transportation Systems, vol. 16, no. 2, pp. 865–873, 2015.

[7] Y. Liu, Z. Liu, and R. Jia, "DeepPF: A deep learning based architecture for metro passenger flow prediction," Knowledge-Based Systems, vol. 162, pp. 171–186, 2018.

[8] S. Sun, C. Zhang, and G. Yu, "A Bayesian network approach to traffic flow forecasting," IEEE Transactions on Intelligent Transportation Systems, vol. 7, no. 1, pp. 124–132, 2006.

[9] W. Min and L. Wynter, "Real-time road traffic prediction with spatio-temporal correlations," Transportation Research Part C, vol. 19, no. 4, pp. 606–616, 2011.

[10] J. Wang, Q. Gu, J. Wu, G. Liu, and Z. Xiong, "Traffic speed prediction and congestion source exploration: A deep learning method," IEEE Transactions on Intelligent Transportation Systems, vol. 21, no. 3, pp. 1167–1176, 2020.

https://doi.org/10.5281/zenodo.19609934