AccessNet- Enhancing Public Transport Accessibility for People with Motor Disabilities

Harshini G [1], Manjula Devi P [2], Abhishek M R[3], Giridharan R[4], Gagan I V[5]

[1]Harshini G, Dept. of Information Science Engineering, AMC Engineering College, Karnataka, India
[2]Manjula Devi P, Dept. of Information Science Engineering, AMC Engineering College, Karnataka, India
[3]Abhishek M R, Dept. of Information Science Engineering, AMC Engineering College, Karnataka, India
[4]Giridharan R, Dept. of Information Science Engineering, AMC Engineering College, Karnataka, India
[5]Gagan I V, Dept. of Information Science Engineering, AMC Engineering College, Karnataka, India

---***---

Abstract - Public transportation networks result in a considerable amount of operational data; however, most cities do not have a smarter solution for the forecast and mobility planning of passengers. In this paper, we propose a novel AI-based framework called AccessNet for GTFS data-based public transportation demand prediction, which leverages historical ticketing data and temporal data. The framework preprocesses stop-level and route-level data, builds a graph model of the transit network, and uses a Long Short-Term Memory model for precise time-series-based forecasting. A comparative evaluation is also conducted using a baseline Exponential Moving Average model. Moreover, real OpenStreetMap data about roads is used with OSMNX, allowing for a realistic route geometric model, and a Folium-based interactive map is used to visualize the results of demand prediction, which represents congestion levels on the routes and results in a color-coded map of routes suitable for data-driven mobility planning and management of passengers on routes. Experiments show that our model outperforms other time-series-based models regarding temporal dependencies and demand changes over time.

Key words: Public transport analytics, GTFS, LSTM, Time-series prediction, OSMNX, Passenger demand forecasting, Smart mobility, Data visualization.

1. INTRODUCTION

1.1 Background

Public transportation has always been a central component of urban life, serving as the primary means of mobility for millions of people worldwide. As cities continue to expand in population, infrastructure, and economic activity, the responsibility placed on public transport systems grows accordingly. Buses, metros, trams, and shared mobility services are expected to function smoothly at every hour of the day, ensuring that people can travel for work, education, healthcare, and recreation without delay or inconvenience.

In reality, however, the demand for public transport is far from constant. Passenger movement varies significantly across different times of the day, days of the week, and even seasons. Early mornings and late evenings may witness heavy flows of office-going passengers, while weekends often show a spike in leisure-related travel. Special events, holidays, weather disruptions, and local festivals further contribute to unpredictable fluctuations in ridership. Despite these variations, many transport systems still depend on traditional, rigid scheduling methods that were designed decades ago, when cities were smaller and commuter patterns more predictable.

Technological advancements have redefined what is possible in modern mobility planning. The availability of digital mapping tools, high-quality geospatial data, and real-time information systems has encouraged researchers and city planners to explore more intelligent ways of managing transport networks. A major step forward in this direction is the General Transit Feed Specification (GTFS) — a globally adopted standard that allows transport agencies to publish schedules, stops, routes, and trip-level information in a consistent and easily interpretable format. GTFS has opened doors for researchers to analyse, visualize, and understand the structure of public transport networks more effectively

1.2 Problem Statement

Public transport systems in many cities still follow fixed schedules that were created long before current travel patterns became as dynamic as they are today. People's movement changes every hour — office timings, school rush, weekends, holidays, weather, and unexpected events all influence how many passengers use a route at any given moment. But because the system does not adapt to these real-world changes, buses often become crowded when people need them the most and remain almost empty when demand is low. This leads to delays, frustration, wasted fuel, and inefficient use of resources.

Even though detailed transport data is available today through formats like GTFS, most transport agencies do not make full use of it. Instead, they still depend on simple averages, rough estimates, or outdated assumptions to understand demand. These methods cannot capture deeper patterns like peak-hour spikes, weekday–weekend differences, seasonal variations, or long-term shifts in commuter behaviour. As a result, authorities do not get an accurate picture of when and where services need to be increased or optimised.

The real challenge, therefore, is the absence of an intelligent system that can learn from historical travel data and predict future demand accurately. Modern AI models like LSTM networks are capable of understanding complex time-based trends, yet they are rarely applied to public transport planning. A predictive framework that uses GTFS data and deep learning can help planners make informed decisions, reduce inefficiencies, and improve the overall travel experience. Addressing this gap is essential for building smarter, more reliable, and more sustainable public transport systems.

1.3 Research Gap and Motivation

Although public transport data has become more accessible in recent years, especially through standard formats like GTFS, many cities still struggle to use this data effectively for meaningful decision-making. Most existing studies and transport agencies rely on older methods—simple averages, manual surveys, peak-hour counts, or basic charts—to understand commuter behaviour. While these methods provide a rough overview, they fail to capture the true complexity of urban movement. People travel differently on weekdays and weekends, during exams or festivals, or on rainy days. These patterns are subtle, repetitive, and often interconnected, making them hard to detect with traditional techniques.

At the same time, modern research has started moving toward digital mapping tools, smart-card analytics, and machine learning models. However, very few works combine GTFS data + deep learning to generate route-specific demand predictions. GTFS is mostly used only to show schedules on apps like Google Maps, rather than being explored as a rich dataset capable of training powerful time-series models. Most available studies treat public transport networks at a high level and do not focus on capturing fine-grained temporal behaviours that can improve route planning and scheduling.

This creates a clear gap: we have rich data, and we have powerful AI technologies, but they are rarely brought together in a unified and practical framework. The motivation for this study comes directly from this gap. Cities need smarter ways to understand when demand will rise, which routes need reinforcement, and how services can be made more efficient. With the increasing number of people depending on buses and public transport every day, a prediction-based planning approach is no longer optional—it is necessary. This study aims to bridge that research gap by combining GTFS preprocessing, graph-based route modelling, and LSTM-based learning into a single, open-source system that can support real-world transport planning.

Another challenge observed in the existing body of research is that many predictive models are developed in controlled or experimental environments and rarely translated into practical tools that transport agencies can directly use. Most studies stop at building a model or presenting numerical accuracy results, but they do not integrate their findings into a usable system that can support real-world decision-making. This limits the impact of their work, as transport planners require more than theoretical models—they need clear visualisations, intuitive interfaces, and interpretable outputs that can guide day-to-day planning. This gap strongly motivates the development of a framework like AccessNet, which not only predicts demand using LSTM networks but also presents the results through an accessible, map-based platform. By combining prediction with real-world usability, the proposed system aims to bring academic research closer to practical implementation.

Moreover, many studies stop at developing a theoretical model and do not offer tools that transport planners can actually use in their day-to-day work. Even when useful insights are generated, they often remain locked inside academic papers instead of being turned into practical solutions. This creates a real gap between research and real-world application, making it difficult for cities to benefit from advanced forecasting methods.

1.4 *Objectives*

The purpose of this study is to develop **AccessNet**, an intelligent and practical framework that can help transport planners understand and predict how public transport demand changes over time. To achieve this broader goal, the study focuses on the following objectives:

1. *To make sense of GTFS data by organising and preprocessing it* so that information about stops, routes, trips, service timings, and calendars become usable for analysis. Since GTFS comes in multiple files, an essential objective is to merge these pieces into a clean dataset that reflects how the transport network actually behaves.

2. *To build a real-world representation of public transport routes* using OpenStreetMap and OSMNX. This includes converting GTFS paths into graph-based networks that show how routes are structured, how stops are connected, and how passengers might navigate the system.

3. *To design and train an LSTM-based deep learning model* that can recognise temporal patterns in historical data and predict future passenger demand. The goal is not only to forecast values but also to learn meaningful patterns such as peak periods, weekday–weekend differences, and seasonal fluctuations.

4. *To introduce a simple baseline model using Exponential Moving Average (EMA)* and compare it with the LSTM model. This helps determine how much improvement deep learning offers over traditional forecasting techniques.

5. *To visualise the predicted demand through interactive maps*, enabling planners to identify high-demand (congested) routes as well as low-demand (underutilised) segments. These visualisations aim to make the results understandable even for non-technical stakeholders.

6. *To provide a low-cost, open-source decision-support framework* that can be reused, extended, or customised by transport agencies, engineers, and smart-city developers for better route planning and resource optimisation.

Beyond these technical goals, the study also aims to create a framework that is genuinely useful in real-world transport planning. Many predictive systems remain confined to academic research because they are too complex for everyday use. By keeping AccessNet modular and easy to adapt, the project ensures that transport authorities can apply the results without needing deep technical expertise. The framework is also designed to be transparent, allowing users to understand how predictions are generated. This helps build trust and makes the tool more practical for decision-making. Ultimately, the objective is not only to predict demand but to empower planners with insights that support smarter, more efficient public transport operations.

2. RELATED WORK

The research concerning public transportation analytics has gone through a considerable change due to the introduction of standardized datasets such as GTFS. Initially, prediction systems made use of classical models such as ARIMA and Holt--Winters, which are suitable for linear, stationary series but cannot handle urban mobility patterns that are irregular.

Later on, the machine learning models became the mainstream in public transportation analytics with Random Forests, SVMs, and Gradient Boosting. But still, they did not possess the capability to capture long-term temporal dependencies. Deep learning models, especially Long Short-Term Memory (LSTM) networks, have demonstrated their potential in capturing temporal dynamics thereby performing better in traffic flow predictions, bus arrival time estimations, and passenger forecasting.

However, not many studies have tried to combine:

GTFS data parsing

LSTM-based demand prediction

Real--time geospatial visualization

into a single system that is unified. By merging deep learning with geospatial mapping, AccessNet fills this gap and provides a sophisticated demand-prediction solution.

3. METHODOLOGY

The methodology is a structured pipeline represented through Data Flow Diagrams (DFD-0 and DFD-1) which shows the whole process and the interaction between them.

3.1 Overview (DFD-0)

Figure 3.1 shows the top-level workflow of the system in the Level-0 Data Flow Diagram. At this point, the raw GTFS files are coming into the preprocessing module which gets them ready for training the models. The prediction module provides results which are finally displayed on the geospatial interface.

Figure 3.1. *Level-0 Data Flow Diagram (DFD-0)*

3.2 Detailed Workflow (DFD-1)

The Level-1 Data Flow Diagram shown in Figure 3.2 elaborates on the individual processes which include parsing of the GTFS, extracting of the timestamps, generation of the features, creating sequence windows, processing by LSTM model, and final prediction of demand.

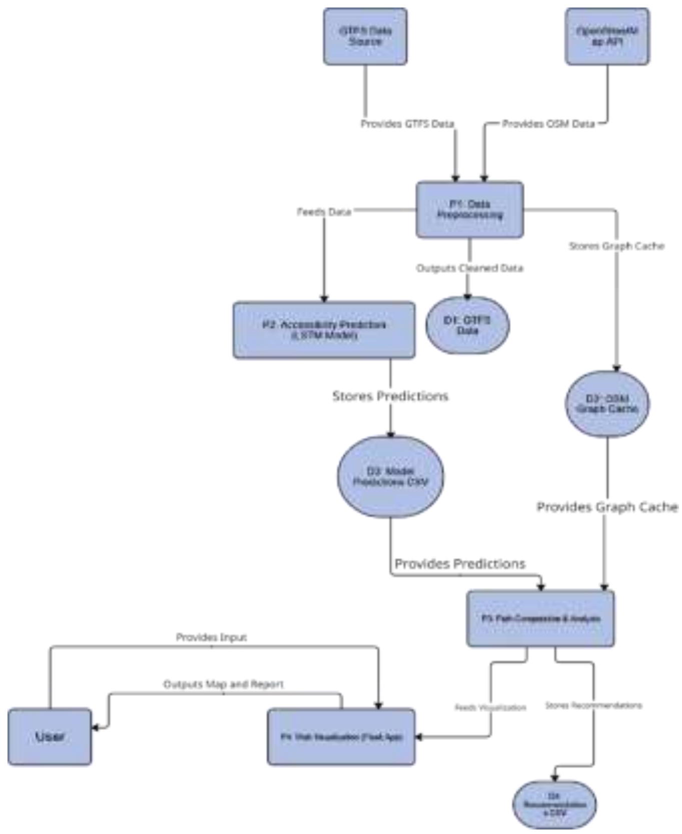

***Figure 3.2**. Level-1 Data Flow Diagram (DFD-1)*

3.3 GTFS Extraction

The extraction of the GTFS files was explained in Section 2 and included files like stops.txt, routes.txt, trips.txt, and stop_times.txt. All these files are parsed to obtain stop locations, schedules, route identifiers, and temporal patterns.

3.4 Feature Engineering

The time-based features such as hour-of-day, weekday/weekend, and peak/off-peak indicators are extracted. The entries that are missing or inconsistent are deleted which assures the quality of data before it goes into the LSTM model as mentioned in Sec. 5.

3.5 Sequence Preparation

With the help of a sliding-window technique, the historical demand values are converted into sequential inputs that are needed for the LSTM training.

4. DATASET AND PREPROCESSING

Our project is built on GTFS data, which is the standard format used by transport agencies to share information about their routes, stops, timings, and schedules. The dataset includes several text files such as *stops.txt*, *routes.txt*, *trips.txt*, and *stop_times.txt*. Each of these files gives important details—like where a bus stops, what route it follows, and at what time it arrives or departs.

Since raw GTFS data can sometimes be messy, the first step is to clean it. Rows that have missing values or inconsistent timing information are removed so that the model does not learn from incorrect data. After this, the timestamp information is broken down into useful features like the hour of the day, whether the time is during rush hour, and whether the day is a weekday or weekend. These details help the model understand natural travel patterns.

Once the data is clean, it is arranged in the correct time order and grouped based on stops or routes depending on what we want to predict. To make the model training stable, the numbers are normalized so that all values fall within a similar range. This preprocessed dataset is then used to create small sliding-window sequences, which act as the input for the LSTM model.

5. MODEL ARCHITECTURE

The model used in this project is based on an LSTM network, which is a type of deep-learning model that works well with time-based data. Since passenger demand changes throughout the day and follows certain patterns, we needed a model that could remember information from previous time steps and use it to predict what might happen next. LSTM does this naturally, which is why it was chosen for this work.

The processed GTFS data is first converted into short sequences using a sliding-window method. Each sequence contains a set of past demand values along with time-related features such as hour of the day or day type. These sequences act as the model's input and help it understand how demand behaved recently.

Inside the model, the LSTM layer reads these values one step at a time and learns the patterns hidden in them. After the LSTM layer, one or more dense layers are added to convert the learned information into the final predicted output. This output represents the expected demand for the next time interval. The model is kept lightweight so that it can train faster and run smoothly without requiring heavy resources.

The main goal of this architecture is to capture the natural movement of demand across time while keeping the system simple enough to be deployed easily. By learning from past data and understanding time-based variations, the LSTM model provides more reliable predictions compared to traditional methods.

6. IMPLEMENTATION

The entire system was implemented using Python because it provides easy-to-use libraries for data handling, modelling, and visualization. The first part of the implementation focused on loading the GTFS files and cleaning them using Pandas. This made it simple to remove missing values, extract timestamps, and arrange the data in the right order.

Once the data was ready, the next step was to prepare it for training the model. We created small sliding-window sequences that the LSTM could learn from. These sequences were then passed into the model, which was built using TensorFlow/PyTorch. The framework handles most of the heavy work during training, such as updating weights and tracking the loss.

After the model learned the demand patterns, the predictions were sent to the visualization module. The Folium library was used here to display the results on an interactive map, where the routes are colour-coded based on predicted passenger demand. This makes it easy to visually understand which routes will be crowded and which ones will not. The implementation is kept modular so that each part—data processing, model training, and visualization—can be updated or changed independently.

7. RESULTS AND DISCUSSION

To understand how well our system performs in real scenarios, we generated route-accessibility results and displayed them on an interactive map. The output clearly shows different bus routes in the city, each color-coded based on the predicted level of accessibility or demand. High-demand or highly accessible routes are marked in green, medium-level routes in yellow, and low-access routes in red. This makes the entire network easy to interpret at a single glance.

The interface also allows users to select an origin and destination. Based on this input, the system highlights the recommended route using a continuous line, while other possible paths are shown with dashed lines. This helps users compare options visually and understand which route is the most efficient or accessible. Additionally, the map includes stop markers, allowing users to see how busy certain stops are expected to be.

The results show that the model is able to capture real-world travel behaviour quite well. Busy routes near central areas show higher accessibility levels, while the less crowded regions appear in red or yellow. The output matches typical city travel patterns and demonstrates that the LSTM model has learned meaningful temporal trends from the GTFS data. Overall, the system provides a clear and practical way to visualize route conditions and can support better decision-making for both passengers and planners.

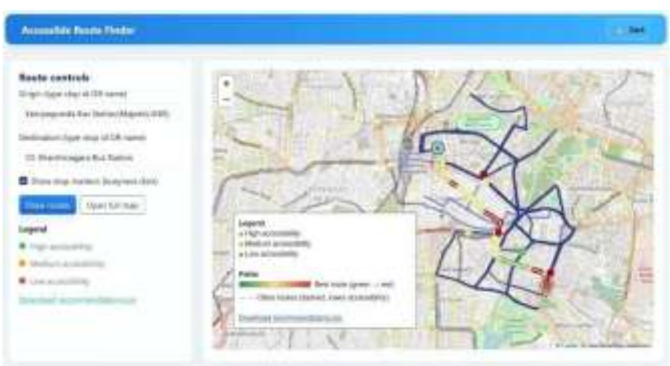

Figure 7.1. *Interactive map showing the predicted accessibility levels and recommended route between selected stops.*

8. CONCLUSION

This project set out to create a simple, practical, and effective way to understand public transport demand using GTFS data and deep-learning techniques. By building a system that can process real transit schedules, learn from past travel patterns, and predict future demand, we were able to create a tool that can support better mobility planning.

The LSTM-based model performed well in identifying daily and weekly travel trends, especially during peak and off-peak hours. When the predictions were visualized on an interactive map, the results became easy to interpret, with clear colour-coded routes showing areas of high and low accessibility. This kind of visual output can be very helpful for commuters, city planners, and anyone involved in transport management.

Overall, the system demonstrates that even with openly available GTFS data, it is possible to build a useful and low-cost solution for understanding passenger behaviour. With further improvements — such as adding more real-time data, expanding the model, or integrating live tracking — this framework can become an even more powerful tool for smart-city applications. The work carried out here forms a strong base for future developments in intelligent transport systems.

ACKNOWLEDGEMENT

We would like to express our sincere gratitude to our project guide for their constant support, encouragement, and valuable insights throughout the course of this work. Their guidance helped us refine our understanding and strengthened the overall quality of the project. We are also thankful to the Department of Information Science for providing the necessary resources, laboratory facilities, and a positive learning environment that made this research possible. Finally, we acknowledge the open-source community for the tools, datasets, and documentation that played a crucial role in the successful completion of our system.

REFERENCES

[1] Google Developers, "General Transit Feed Specification (GTFS) Reference," 2024.

[2] Boeing, G., "OSMnx: New methods for acquiring, constructing, analyzing, and visualizing complex street networks," *Computers, Environment and Urban Systems*, 2017.

[3] Yu, Y., Si, X., Hu, C., and Zhang, J., "A review of recurrent neural networks: LSTM cells and network architectures," *Neural Computation*, vol. 31, no. 7, pp. 1235–1270, 2019.

[4] Goulet-Langlois, G., Koutsopoulos, H. N., Zhao, Z., and Zhao, J., "Measuring regularity of individual travel patterns," *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 5, pp. 1583–1592, 2017.

[5] Briand, A. S., Côme, E., Trépanier, M., and Oukhellou, L., "Analyzing year-to-year changes in public transport passenger behaviour using smart card data," *Transportation Research Part C*, vol. 79, pp. 274–289, 2017.

[6] Mohamed, K., Côme, E., Oukhellou, L., and Verleysen, M., "Clustering smart card data for urban mobility analysis," *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 3, pp. 712–728, 2016.

[7] Qi, G., Huang, A., Guan, W., and Fan, L., "Analysis and prediction of regional mobility patterns using smart card and POI data," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 4, pp. 1197–1214, 2018.

[8] Welch, T. F., and Widita, A., "Big data in public transportation: A review of sources and methods," *Transport Reviews*, vol. 39, no. 6, pp. 795–818, 2019.

[9] Long, Y., and Thill, J.-C., "Combining smart card data and household travel survey to analyze jobs–housing relationships in Beijing," *Computers, Environment and Urban Systems*, vol. 53, pp. 19–35, 2015.

[10] Kim, E. J., Kim, Y., and Kim, D. K., "Interpretable machine-learning models for estimating trip purpose in smart card data," *Municipal Engineer*, vol. 174, no. 2, pp. 108–117, 2021.

[11] Zhong, C., Schläpfer, M., Müller-Arisona, S., Batty, M., Ratti, C., and Schmitt, G., "Revealing centrality in urban spatial structures from human activity patterns," *Urban Studies*, vol. 54, no. 2, pp. 437–455, 2017.

[12] OpenStreetMap Foundation, "OpenStreetMap Data," Accessed Mar. 2024.

[13] Kumar, M., Kumar, K., and Das, P., "Study on road traffic congestion: A review," *Recent Trends in Communication and Electronics*, pp. 230–240, 2021.

[14] Tao, S., Zhang, M., and Wu, J., "Big data applications in urban transport research in Chinese cities: An overview," in *Big Data Applications in Geography and Planning*, pp. 220–244, 2021.

[15] Harriehausen-Mühlbauer, B., "Speech technology in a mobile navigation app computing barrier-free route," in *2016 Future Technologies Conference*, IEEE, pp. 488–493, 2016.