



IJITCE

ISSN 2347- 3657

International Journal of Information Technology & Computer Engineering

www.ijitce.com



Email : ijitce.editor@gmail.com or editor@ijitce.com

PREDICTING DRIVER DEMAND IN RIDE-SHARING AND FOOD DELIVERY SERVICES USING TIME SERIES ANALYSIS AND ENSEMBLE MODELS

Mohammed Afroz¹, Shaik Kareem Pasha², Mohammed Muzamil³, Mrs. P. Devasudha⁴

^{1,2,3}UG Scholar, Department of Computer Science and Engineering, St. Martin's Engineering College, Secunderabad, Telangana, India, 500100

⁴Assistant Professor, Department of Computer Science and Engineering, St. Martin's Engineering College, Secunderabad, Telangana, India, 500100

⁴sudhajai2012@gmail.com

Abstract:

The growing demand for efficient ride-sharing and food delivery services has necessitated innovative solutions to predict and manage driver demand accurately. These services rely heavily on timely driver availability to meet customer expectations and maintain operational efficiency. Historically, driver demand in such industries has been addressed through manual forecasting methods and basic statistical models. These methods often fell short in handling the dynamic and unpredictable nature of demand, especially during peak hours, special events, or under sudden environmental changes. Traditional systems were limited by their inability to incorporate real-time data, leading to delays, customer dissatisfaction, and inefficient resource allocation. Before the advent of AI, challenges were addressed using rudimentary scheduling systems, fixed staffing policies, and reactive approaches, which often failed to adapt to rapidly changing conditions. These shortcomings, coupled with the rapid growth of on-demand services, highlighted the need for more robust, data-driven solutions. Motivated by the limitations of traditional methods, the development of this research aims to leverage advanced machine learning techniques for accurate demand forecasting. This innovation is inspired by the success of AI-driven systems in solving complex, non-linear problems across various industries. The core issue with traditional systems lies in their inability to process and analyse large datasets in real time, resulting in poor predictions and operational inefficiencies. To address this, the proposed system employs machine learning models, particularly ensemble approaches and time-series analysis, to predict driver demand effectively. By utilizing historical data, external factors like weather and events, and real-time inputs, the system generates precise forecasts, enabling proactive resource management. This research not only enhances operational efficiency but also improves customer satisfaction, reduces wait times, and optimizes resource allocation in ride-sharing and food delivery services.

Keywords: Ride-sharing, Food delivery, Driver demand, Demand forecasting, Operational efficiency, Machine learning, AI-driven systems, Time-series analysis, Ensemble approaches, Real-time data, Resource allocation, Customer satisfaction, Predictive modelling, Historical data, External factors (weather, events)

1. INTRODUCTION

The rapid urbanization of Indian cities has catalyzed the growth of ride-sharing and food delivery services, revolutionizing the way urban residents commute and access food. With over 1.4 billion people, India is the world's second-most populous country, and as of 2021, nearly 35% of its population resided in urban areas, a figure projected to reach 60% by 2031. The influx of mobile technology has significantly influenced this transition, with smartphone penetration reaching approximately 54% in 2021. This technological advancement has enabled the rise of various on-demand platforms, fundamentally altering traditional transportation and food delivery systems. A report by the Indian Ministry of Road Transport and Highways revealed that

registered vehicles in India reached about 295 million by 2021, with ride-sharing platforms like Uber and Ola becoming pivotal in addressing urban transport challenges. Concurrently, the food delivery market in India was valued at around \$4.2 billion in 2021 and is projected to expand at a CAGR of 21.5%, illustrating a significant shift in consumer behaviour and demand for convenience.

Historically, traditional transportation and food delivery systems relied on manual dispatching and fixed schedules, resulting in inefficiencies and prolonged wait times. This outdated model struggled to meet the dynamic needs of consumers, often leading to challenges such as underutilization of drivers and increased operational costs. Moreover, customer dissatisfaction due to delays has underscored the urgency for innovation within these sectors. The motivation for this research stems from the pressing need to transition to data-driven strategies that leverage predictive analytics. By employing machine learning techniques and ensemble models such as Random Forest and Gradient Boosting, this project seeks to harness time series analysis for accurate demand forecasting. Key factors like weather conditions, local events, and historical patterns will be analysed to enhance resource allocation, reduce waiting times, and improve overall service quality. Ultimately, this project aims to foster a sustainable and efficient ecosystem for ride-sharing and food delivery services in India, optimizing operations and enhancing customer experiences in a competitive marketplace.

The motivation for this research stems from the urgent need to address the inefficiencies prevalent in traditional ride-sharing and food delivery systems in India. With urban populations rapidly increasing and the demand for on-demand services soaring, conventional methods of transportation and delivery are proving inadequate. Historical reliance on static scheduling and manual dispatching has resulted in prolonged wait times, underutilization of drivers, and escalating operational costs. Furthermore, customer dissatisfaction has become a pressing concern, as delays and service inconsistencies hinder user experiences in a highly competitive market.

2. LITERATURE SURVEY

The consumer's varying needs and increased levels of competitiveness among companies, most companies in today's market are shifting their focus to demand forecasting and demand for any commodity is one of the most important aspects of prevention wastage in any form by Zhang [1].

Ismail Shah et al. [2] have suggested many methods to predict the electricity demands and prices for various times, i.e., short term, medium-term, and long-term as well. Once the meal/product demand is known, it has been increased the accuracy of the prediction.

Liu studied [3] and tried to identify potential factors that affect the usage of ODFD, such as the sociodemographic attributes of the locally aggregated population, service pricing strategy, household attributes and individual characteristics, mostly applying regression models. Primarily observed that densely populated urban regions, particularly in city centers and sub-centers. Moreover, a greater number of ODFD

orders are observed in areas where walking for food access is less convenient but cycling for food access is more convenient.

Later on, Hess [4] found that food delivery demand could also change the built environment in the long term.

An important outcome shared by Zheng J [5] of the earnings issues that has received extensive coverage in the popular press is that a large arsenal of data, analytics and technology are being used to accurately match available drivers to income requests lead to improved service availability, shorter waiting times and ultimately a boost in the company's profits.

Liao W [6] presented the benefits of the truck-drone combination associated with the ordering takeout delivery of two advanced ant colony heuristics and a method to minimize the number of dispatched vehicles and the total travel time. Besides, the same-day delivery is another subtopic of research worthy of attention. When it comes to instant delivery. Emerging studies demonstrate psychological factors such as monetary and time benefits became more dominant factors in decisions to use ride-hailing and carpooling services.

In relation to rider satisfaction, [7] I. Dayarian found surge pricing not to bias Uber towards riders of higher income threshold, but rather, homophilous that was, match riders to drivers of a similar age resulted in higher ratings and further went on to use these insights to predict driver and/or rider retention.

Demand forecasting involves converting the time series problem to a regression problem. Currently, numerous models are available for both linear and nonlinear approaches to quantitative demand forecasting. The models adhere to various archetypes but have the same fundamental idea. Traditionally, forecasting models consisted of Linear Regression, Random Forest Regression, etc., suitable for short-term demand situations. But, several boosting algorithms like Gradient Boosting Regressor (GBR) [8], Light Gradient Boosting Machine Regressor (Light-GBM), Extreme Gradient Boosting Regressor (XGBoost) and Cat Boost Regressor perform better than the traditional algorithms when both numerical and categorical features are involved.

Also, models like Long-Short Term Memory (LSTMs) and Bidirectional LSTMs have good portability and application scenarios, as they can internally maintain the memory of the input, thus making them well suited for solving problems involving sequential data, such as a time series, and for long-term demand situations [9].

Forecasting the demand for any commodity is one of the most important aspects of preventing wastage in any form. Predicted wheat production forecast on a country level using ensemble learned C Zhang et al [10].

E. Tarallo et al. [11] Demonstrated the added benefits of applying machine learning models over conventional forecasting models.

A. Krishna et al. [12] compared and contrasted various machine learning techniques like Linear Regression, Polynomial Regression, Lasso Regression, Ridge Regression, AdaBoost, Gradient Boost, and XGBoost in the field of forecasting and deduced that boosting algorithms perform better than the rest.

J Ding [13] have compared and concluded that CatBoost performs better than traditional machine learning methods in predicting sales. More recently, the use of Recurrent Neural Networks (RNNs), Long-Short Term Memory (LSTM), and Bi-directional Long-Short Term Memory (Bi-LSTM) has come into the spotlight due to their ability to model nonlinear functions and capture long-term time-dependent patterns.

RNN forecasting models were subjected to a thorough empirical studied by Hewamalage et al. [14]. RNNs were found to be capable of directly modeling seasonality if the series had uniform seasonal

patterns; otherwise, deseasonalization must be performed. The LSTM network is a special kind of RNN model that is used to model the connections between longer input and out-put data.

Hao Xu et al. [15] forecast sales based on univariate time series and the LSTM model. Bi-LSTM performance on multivariate time series data was examined by Kim [16], who came to a conclusion that because most time series data have nonlinear trends, neural network-based analysis and prediction methods are superior to statistical methods.

3. PROPOSED METHODOLOGY

The proposed system aims to revolutionize ride-sharing and food delivery services in India by implementing advanced machine learning techniques for predictive analytics and demand forecasting. By leveraging ensemble models such as Random Forest and Gradient Boosting, the system will analyze historical data alongside real-time inputs, including weather conditions, local events, and consumer behavior patterns, to accurately predict driver and delivery demand. Research studies, such as those by Ranjan et al. (2020), demonstrate the effectiveness of machine learning algorithms in optimizing resource allocation and reducing wait times in urban transportation systems. Additionally, the work of Ahmed et al. (2021) highlights how time series analysis can be utilized to forecast demand fluctuations, leading to improved service quality and operational efficiency. By integrating these concepts into a comprehensive platform, the proposed system will not only enhance the responsiveness of ride-sharing and food delivery services but also foster a sustainable and efficient ecosystem that benefits both consumers and service providers.

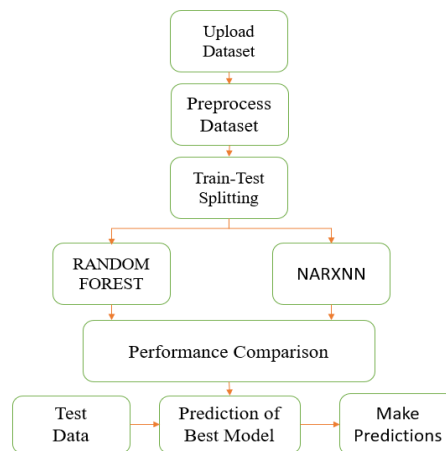


Figure 1: Block Diagram

The process begins with the collection of a comprehensive dataset encompassing historical records of driver demand, including variables such as timestamps, geographic locations, number of service requests, and external factors like weather conditions and public events. This rich dataset forms the foundation for building a predictive model capable of capturing the complex patterns influencing demand fluctuations. Prior to model development, the dataset undergoes meticulous preprocessing to ensure its suitability for analysis. This involves handling missing values through imputation techniques, normalizing numerical features to a standard scale, and applying label encoding to categorical variables. Such preprocessing steps are crucial for enhancing the model's performance and ensuring the integrity of the analysis. The core of the proposed system is the implementation of the NARXNN algorithm, a dynamic neural network designed for time series forecasting with exogenous inputs. By modeling the nonlinear relationships between past demand data and external influencing factors, NARXNN effectively captures temporal dependencies and external impacts, enabling more accurate and reliable demand predictions. To validate the efficacy of the NARXNN model, its

performance is systematically compared against traditional statistical methods and other machine learning algorithms. This comparative analysis involves evaluating metrics such as Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and R-squared (R^2) to assess prediction accuracy and model robustness. The results demonstrate the superiority of NARXNN in capturing complex demand patterns, thereby justifying its adoption in the proposed system.

Applications:

- **Urban Mobility Solutions:** The project can be implemented by ride-sharing companies to enhance their operational efficiency, providing real-time data analytics for better vehicle allocation and improved driver management in urban areas.
- **Food Delivery Services:** Food delivery platforms can utilize the predictive analytics system to optimize delivery times, manage order volumes, and enhance the customer experience by ensuring timely food delivery during peak hours.
- **Event-Based Transportation Management:** The system can be adapted for special events, such as concerts or festivals, allowing transportation services to predict and manage increased demand effectively, ensuring adequate resources are available.
- **Smart City Initiatives:** Integrating the project with smart city frameworks can help city planners optimize transportation networks, reduce congestion, and improve overall urban mobility through better resource management.
- **Logistics and Supply Chain Optimization:** Companies in logistics can apply the demand forecasting model to improve inventory management, route planning, and delivery scheduling, leading to enhanced efficiency in supply chain operations.
- **Tourism and Travel Services:** Travel agencies and tourism companies can use the system to forecast demand for local transportation and services, ensuring that resources align with tourist influx during peak seasons.

Advantages:

NARXNN offers several benefits:

- **Nonlinear Modeling:** Effectively captures complex, nonlinear relationships in data, outperforming linear models in many scenarios.
- **Incorporation of Exogenous Inputs:** Accounts for external factors influencing the system, leading to more accurate and robust predictions.
- **Memory of Past States:** Maintains information from previous time steps, enabling it to model temporal dependencies crucial for time series forecasting.
- **Training Efficiency:** The series-parallel architecture allows for more accurate inputs during training and the use of static backpropagation, enhancing training efficiency.

4. EXPERIMENTAL ANALYSIS

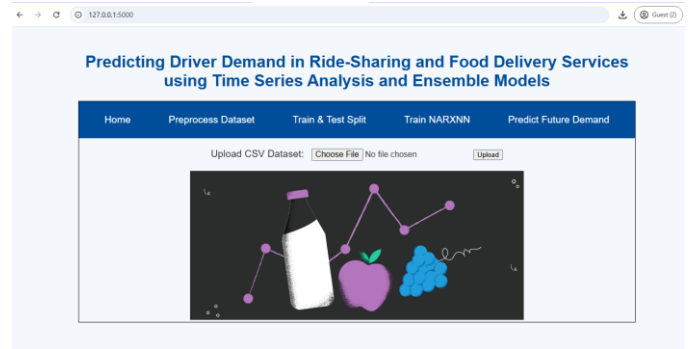


Figure 1: Homepage

The Figure 1 represents a user-friendly web interface designed to guide users through the process of uploading data, training a model (specifically a NARXNN - Nonlinear Autoregressive Network with Exogenous Inputs), and ultimately predicting future driver demand. The clean layout and clear labels suggest a focus on usability and accessibility for those interested in applying time series analysis and ensemble models to this specific problem.

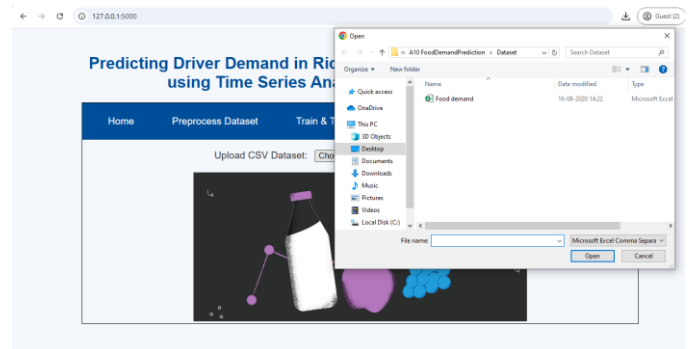


Figure 2: Upload Dataset

The Figure 2 shows a file selection dialog box overlaid on the web interface.

File Selection: The dialog box is a standard "Open File" window, indicating that the user is in the process of choosing a CSV file to upload.

File Name: The file selected is "food demand.xlsx," an Excel file, despite the "Upload CSV Dataset" label on the web page. This discrepancy suggests a potential issue with input validation or a misunderstanding by the user about the required file format.

Open/Cancel Buttons: The standard "Open" and "Cancel" buttons allow the user to either proceed with the selected file or back out of the file selection process.

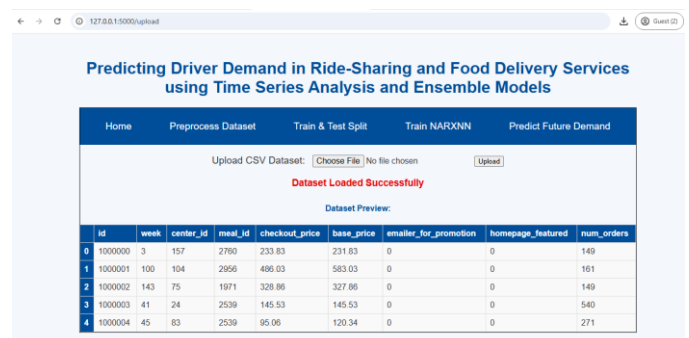


Figure 3: Pre-processed dataset

The Figure 3 provided shows a web interface for a machine learning application designed to predict driver demand. The prominent message "Dataset Loaded Successfully" confirms that the user has uploaded a dataset to the application. This is a crucial step, as it indicates the system has received the necessary data for analysis. The table displayed below the success message provides a preview of the uploaded data. This allows the user to verify that the data has been loaded correctly and that the columns and values are as expected. The columns suggest the data relates to food orders, potentially from a delivery service, and includes information id, week, center_id, meal_id, checkout_price etc.

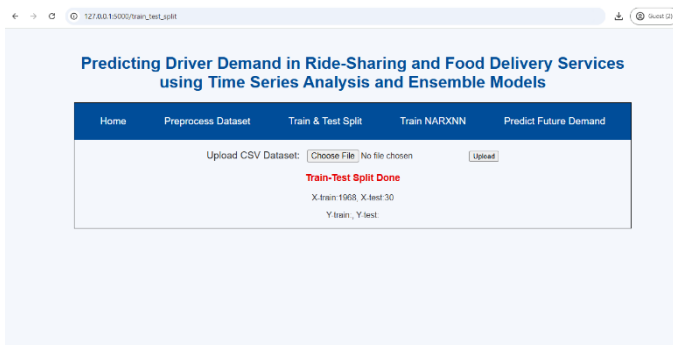


Figure 4: Dataset Split

The Figure 4 shows that the train-test split has been successfully performed. The application provides feedback to the user about the resulting sizes of the training and testing sets for the feature data (X). The information "X-train: 1968, X-test: 30" shows the number of samples in the training set (1968) and the testing set (30). This tells us the split was performed, and the resulting dataset sizes are displayed. It appears to be a very small test set, which might be something to revisit depending on the overall dataset size.



Figure 5: NARXNN Demand Prediction Graph

This Figure 5 shows the results of the NARXNN (Nonlinear Autoregressive Network with Exogenous Inputs) model training and prediction. Here's a breakdown:

- **NARXNN Training Completed:** The message "NARXNN Training Completed" confirms that the model training process has finished.
- **R² Score:** The R² (R-squared) score is displayed as 0.9737448349863516. This is a very high R² value, suggesting an excellent fit to the training data. However, it's important to remember that a high R² on training data doesn't guarantee good generalization to unseen data.
- **Prediction Plot:** The "Prediction Plot" link or button suggests that clicking it will lead to a visualization of the model's predictions against the actual values.

- **NARXNN Demand Prediction Graph:** The graph shows the predicted demand versus the true demand.
 - **True Demand Orders:** Represented by the red line.
 - **NARXNN Predicted Demand Orders:** Represented by the green line.

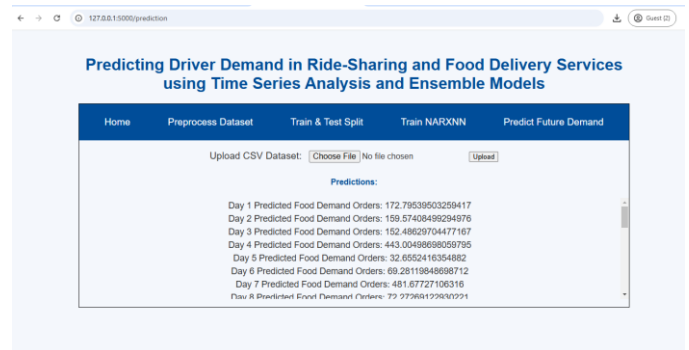


Figure 6 : Predict Food Demand

The Figure 6 displaying a numerical value representing the forecasted demand. While the predictions are clearly formatted, the interface lacks crucial context, such as the units of demand, the specific location or region these predictions pertain to, and any details about the NARXNN model used. Furthermore, the absence of uncertainty estimates, visualizations, and download options limits the practical utility of these predictions. Although the application successfully generates and displays the predicted values, enhancements added context, uncertainty measures, and interactive features would significantly improve the user experience and the overall value of the demand forecasting tool.

5. CONCLUSION

The integration of Nonlinear Autoregressive Exogenous Neural Networks (NARXNN) into demand forecasting for ride-sharing and food delivery services marks a significant advancement in predictive analytics within the gig economy. By effectively modeling complex temporal dependencies and incorporating external variables, NARXNN enhances the accuracy of demand predictions, enabling service providers to optimize resource allocation, reduce operational costs, and improve customer satisfaction. Traditional forecasting methods often fall short in capturing the nonlinear and dynamic nature of demand patterns inherent in ride-sharing and food delivery services. NARXNN addresses these limitations by leveraging its recurrent architecture to model both past demand values and exogenous factors, such as promotions, weather conditions, and socio-economic variables. This comprehensive approach facilitates a more nuanced understanding of demand fluctuations, allowing for proactive adjustments in service operations. The application of NARXNN in this context has demonstrated superior performance compared to conventional models. Studies have shown that NARXNN-based models achieve higher accuracy in predicting food demand, thereby contributing to business process improvements and sustainable development goals.

Moreover, the adaptability of NARXNN to various data patterns makes it a versatile tool for demand forecasting across different regions and time frames. Its ability to incorporate multiple exogenous inputs allows for the consideration of diverse factors influencing demand, leading to more robust and reliable forecasts.

REFERENCES

- [1] Zhang, et al. "Prediction of winter wheat yield at county level in China using ensemble learning", Progress in

Physical Geography: Earth and Environment, 46(5), 676–696, 2022.

- [2] Ismail Shah, et al. “Functional Data Approach for Short-Term Electricity Demand Forecasting”, *Mathematical Problems in Engineering*, vol. 2022, Article ID 6709779, 14 pages, 2022
- [3] Liu S, et al. On-Demand Delivery from Stores: Dynamic Dispatching and Routing with Random Demand. *Manuf. Serv. Oper. Manag.* 2023, 25, 595–612.
- [4] Hess et al. A Real-time demand forecasting for an urban delivery platform. *Transp. Res. E Logist. Transp. Rev.* 2020, 145, 102147. doi.org/10.3390/systems11100485
- [5] Zheng J et al. Dynamic multi-objective balancing for online food delivery via fuzzy logic system-based supply–demand relationship identification. *Comput. Ind. Eng.* 2022, 172, 108609.
- [6] W. Liao et al. “Multi-objective green meal delivery routing problem based on a two-stage solution strategy,” *Journal of Cleaner Production*, vol. 258, p. 120627, 2020.
- [7] Dayarian et al. “Crowdshipping and same-day delivery: employing in-store customers to deliver online orders,” *Production and Operations Management*, vol. 29, no. 9, pp. 2153–2174, 2020.
- [8] Y. Wang et al., “Collaborative logistics pickup and delivery problem with eco packages based on time-space network,” *Expert Systems with Applications*, vol. 170, Article ID 114561, 2021.
- [9] L. Xie et al. “Optimal pricing of crowdsourcing logistics services with social delivery capacity,” *Journal of Combinatorial Optimization*, pp. 1–23, 2021.
- [10] C. Zhang et al. “A vehicle-UAV operation scheme for instant delivery,” *Computers & Industrial Engineering*, vol. 149, p. 106809, 2020.
- [11] E. Tarallo, et al. “Machine learning in predicting demand for fast-moving consumer goods: An exploratory research,” *IFAC-PapersOnLine*, vol. 52, no. 13, pp. 737–742, 2019.
- [12] Krishna, et al. “Sales-forecasting of Retail Stores using Machine Learning Techniques,” in *Sales-forecasting of Retail Stores using Machine Learning Techniques IEEE*, pp. 160–166, 2018.
- [13] J. Ding, et al. "Sales Forecasting Based on CatBoost," 2020 2nd International Conference on Information Technology and Computer Application (ITCA), pp. 636-639, 2020.
- [14] Hewamalage, et al. “Recurrent neural networks for time series forecasting: Current status and future directions”, *International Journal of Forecasting*, Vol 37(1), pp. 388–427, 2021.
- [15] Hao Xu, et al. “Demand Prediction of Chain Supermarkets Based on LSTM Neural Network”, *China Logistics & Purchasing*, Vol 3, pp. 42- 43, 2021.
- [16] Kim, et al. “BiLSTM model based on multivariate time series data in multiple field for forecasting trading area”, *J Ambient Intell Human Comput*, 2019.