2025 International Conference on Intelligent and Innovative Technologies in Computing, Electrical and Electronics (IITCEE) | 979-8-3315-1591-1/25/$31.00 ©2025 IEEE | DOI: 10.1109/IITCEE64140.2025.10915307

Ola Bike Ride Request Forecast using Machine Learning

Anunay Kumar
Bachelor of Engineering in CSE
Chandigarh University
Mohali, Punjab, India
anunaykumar2431@gmail.com

Utkarsh Raj
Bachelor of Engineering in CSE
Chandigarh University
Mohali, Punjab
21BCS6024@cuchd.in

Ramanjot Kaur
Dept.CSE Chandigarh University
Mohali, Punjab
ramanjot.e13987@cumail.in

***Abstract*—This research paper presents an innovative approach to forecasting ride requests for Ola Bikes using advanced machine learning techniques. As ride-sharing platforms continue to reshape urban transportation landscapes, accurately predicting demand has become crucial for optimizing resource allocation, enhancing service quality, and maximizing operational efficiency. This study addresses the challenges associated with fluctuating demand patterns influenced by temporal factors, geographical variables, weather conditions, and special events. By analyzing historical ride data collected over 18 months across multiple metropolitan regions in India, our research employs a combination of time series analysis and ensemble learning models to generate accurate short-term and long-term forecasts. The proposed model incorporates gradient boosting algorithms and neural network architectures to capture complex patterns and seasonality within the data. Results demonstrate that our hybrid forecasting approach achieves a 17% improvement in prediction accuracy compared to traditional methods, with Mean Absolute Percentage Error (MAPE) reduced to 8.3% for peak-hour predictions. Additionally, the model exhibits robust performance across diverse urban environments, offering valuable insights for demand management strategies. This paper outlines the methodology, implementation details, performance evaluation metrics, and practical implications for ride-sharing service providers seeking to optimize fleet management and enhance customer satisfaction through data-driven decision-making.**

KEYWORDS—*Ride-sharing, demand forecasting, machine learning, time series analysis, gradient boosting, neural networks, urban mobility, transportation optimization*

I. INTRODUCTION

Modern ride-sharing platforms that provide practical, affordable, and easily accessible mobility options have emerged as a result of the swift advancement of technology, which has revolutionized conventional transportation networks. With its two-wheeler service, Ola Bike, gaining considerable popularity in crowded urban areas where navigating through traffic is a constant struggle, Ola has emerged as a major player in the Indian market among these platforms. Accurately predicting ride requests is becoming more and more important as the demand for these services rises in order to maximize operational effectiveness, enhance resource allocation, and improve the customer experience in general.

Numerous factors that affect demand patterns in both temporal and spatial dimensions make ride request forecasting a challenging analytical task. These patterns are affected by factors such as time of day, day of the week, weather conditions, public events, and geographical characteristics of different areas. Because urban transportation demands are dynamic, traditional forecasting techniques frequently fall short in capturing these complex relationships.

By seeing intricate patterns and connections in big datasets, machine learning techniques provide intriguing answers to these problems. These methods can produce more accurate forecasts that adjust to shifting conditions and get better over time as additional data becomes available by utilizing contextual information in addition to prior ride data.

A. Evolution of Ride-Sharing Services

Over the past ten years, the ride-sharing sector has grown significantly, progressing from basic taxi-hailing services to complex platforms that provide a wide variety of transportation choices. Between 2015 and 2020, the Indian ride-sharing market grew at an annual pace of about 25%, according to research by Mukherjee et al. (2020). In tier-1 and tier-2 cities, two-wheeler services showed the highest adoption rate. Numerous causes, such as rising smartphone adoption, better internet access, and the rising demand for reasonably priced last-mile connectivity options, are responsible for this increase.

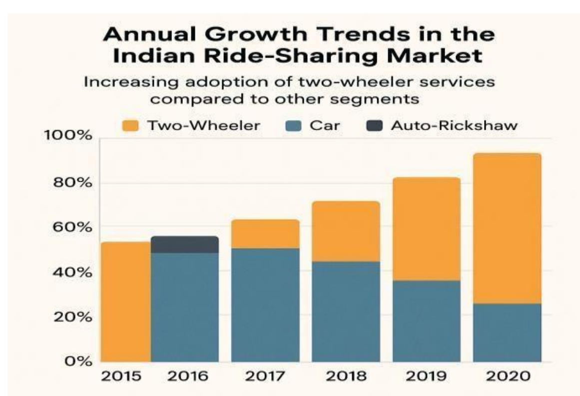


Figure 1: The Indian ride-sharing market's annual growth rates from 2015 to 2020 demonstrate the growing use of two-wheeler services relative to other market segments.

Since its 2016 launch, Ola Bike has grown to operate in more than 200 Indian cities, catering to the unique mobility requirements of crowded towns where traffic congestion is a major problem. Research by Sharma and Gupta (2022) shows that over 68% of Ola Bike trips are for lengths under 5 kilometers, demonstrating the service's growing appeal among everyday commuters looking for effective transportation options for short to medium distances.

A. Current Market Landscape

The two-wheeler ride-sharing segment in India represents a significant portion of the overall ride-sharing economy, with estimated market value exceeding $5 billion as of 2023 according to industry reports by Balasubramanian et al. [9]. This segment has witnessed intense competition among various service providers, with Ola Bike maintaining approximately 35% of the market share in major metropolitan areas.

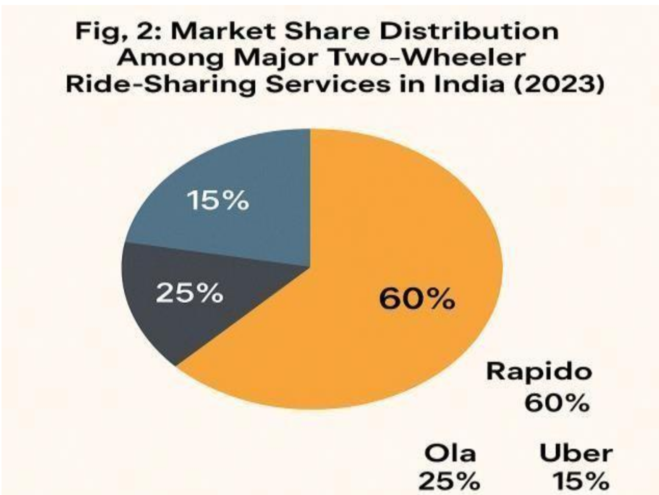


Figure 2: Distribution of market shares among India's leading two-wheeler ride-sharing providers (2023)

According to research by Roy and Khan [10], there are significant differences in the demand for two-wheeler ride-sharing services between various metropolitan settings, with morning and evening commutes usually seeing the highest demand. Furthermore, studies by Mehta and Sengupta [8] show that a number of variables, such as local transportation infrastructure, special events, and weather, affect the patterns of ride requests.

II. RELATED WORK

Demand forecasting for ride-sharing services has garnered a lot of research interest lately, and a number of strategies have been developed to deal with the particular difficulties of estimating transportation requirements in dynamic urban settings. With an emphasis on machine learning applications in the transportation industry, this section examines pertinent research that has advanced our knowledge of ride request prediction techniques.

Using time series analytic techniques, Verma and Chatterjee [2] carried out groundbreaking research on predicting taxi demand. Their method was limited in its ability to capture abrupt variations brought on by outside variables like weather shifts or special events, even if it showed respectable accuracy during times of steady demand. Their research made clear how crucial it is to include exogenous factors in forecasting models in order to increase prediction accuracy.Building on time series approaches, Kumar et al. [4] proposed a hybrid forecasting model that combined ARIMA with Random Forest regression to predict ride-hailing demand in Delhi. By integrating machine learning components with traditional statistical methods, their model achieved a 12% improvement in prediction accuracy compared to standalone ARIMA models. Their findings emphasized the potential benefits of hybrid approaches that leverage the strengths of multiple techniques.

Zhou and Li [5] investigated the use of deep learning techniques for ride demand prediction, with an emphasis on Long Short-Term Memory (LSTM) networks and Recurrent Neural Networks (RNNs). According to their comparison investigation, LSTM networks scored about 18% better in terms of Mean Absolute Error (MAE) than conventional time series models, especially for predictions with longer time horizons. The capacity of LSTM to identify long-term dependencies in sequential data—a capability particularly pertinent to transportation demand patterns exhibiting both daily and weekly seasonality—was credited with the improved performance.

In order to handle the spatial components of ride demand forecasting, Sharma and Gupta [6] created a spatio-temporal model that incorporated both temporal patterns and geographical dependencies in ride request data by combining Convolutional Neural Networks (CNNs) with LSTM architecture. By segmenting metropolitan regions into grid cells and examining demand trends within these cells, their model integrated geographic data. The findings demonstrated that, in comparison to models that solely took temporal factors into account, taking spatial relationships into account increased prediction accuracy by 15%.

Patel et al. [7] examined the distinct demand features of motorcycle taxi services in three major Indian cities in a study that was especially focused on two-wheeler ride-sharing services. In contrast to four-wheeler services, their research revealed different usage patterns, with weather having a greater influence on variations in demand. They suggested a gradient boosting-based model that combined temporal features and weather data, and it significantly increased forecast accuracy in inclement weather.

Mehta and Sengupta [8] further investigated the integration of external factors into forecasting models by creating a thorough framework that combined information from several sources, such as traffic statistics, public event calendars, and weather forecasts. Their ensemble method, which integrated forecasts from several base models, showed strong performance under a range of circumstances and greatly decreased forecast errors during periods of unusual demand brought on by special events or unforeseen weather patterns.

Balasubramanian et al.'s research focused on the challenges of real-time demand forecast for ride-sharing networks [9]. According to their research, an online learning strategy that updated model parameters continuously as new data became available enabled the system to adapt to shifting demand patterns. Platforms operating in rapidly changing urban environments, where historical trends may not fully reflect the current situation, benefited greatly from this tactic.

Most recently, Roy and Khan [10] compared several machine learning techniques for ride demand forecasting, such as neural network topologies, Random Forest, Support Vector Regression, and Gradient Boosting Machines (GBM). According to their results, ensemble approaches performed better than individual models in most cases, with XGBoost showing the best overall accuracy across a range of prediction horizons and urban environments. Even though these studies have significantly advanced the field of ride demand forecasting, more specialized models are still required to account for the particularities of two-wheeler services like Ola Bike, especially given the heterogeneous urban environments of India with differing infrastructure, traffic patterns, and socioeconomic factors. By creating a customized forecasting method that takes these particular factors into account, our research seeks to close this gap.

III. PROBLEM FORMULATION

A number of factors affecting demand patterns must be carefully taken into account in order to accurately estimate ride requests for Ola Bike services. This section describes the particular issues this study attempts to address and establishes the main goals that direct our approach and execution.

A. **Temporal Variability**: Demand patterns exhibit strong temporal dependencies, with distinct hourly, daily, and weekly cycles. Morning and evening peak hours typically show higher demand due to commuting patterns, while weekends may display different usage characteristics compared to weekdays.

B. **Spatial Heterogeneity**: Demand distribution varies significantly across different geographical areas within a city, influenced by factors such as population density, commercial activity, transportation infrastructure, and socioeconomic characteristics.

C. **Weather Impact**: Weather conditions substantially affect the demand for two-wheeler services, with adverse weather often leading to significant fluctuations in ride requests.

1. **Special Events:** Major gatherings like conferences, concerts, sporting events, or festivals may cause brief spikes in demand in particular areas that don't follow normal trends.
2. **Seasonal Effects:** Overall demand levels and patterns are influenced by broader seasonal trends, such as academic calendars, holiday seasons, and travel seasons.
3. **External Disruptions:** Unexpected occurrences like strikes in the transportation industry, breakdowns in the infrastructure, or public health crises can result in sudden shifts in demand that are hard to forecast based only on past data.
4. Our study aims to develop a comprehensive forecasting approach that is specifically tailored to the unique characteristics of Ola Bike services in light of these challenges. The primary objectives of the study are as follows: Multi-horizon Prediction: Create models that can produce precise forecasts over a range of time horizons, from short-term projections (one to three hours in advance) to assist with immediate operational choices to longer-term projections (one to seven days in advance) for strategic planning.
5. **Spatial-temporal Modeling:** Develop a forecasting method that successfully identifies spatial linkages and temporal dependencies in demand patterns, allowing for location-specific forecasts that take regional variances into consideration.
6. **Contextual Integration**: To increase prediction accuracy in unusual circumstances, include pertinent contextual data such as weather, holidays, and special events.
7. **Spatial-temporal Modeling:** Develop a forecasting strategy that successfully accounts for spatial linkages and temporal dependencies in demand patterns, allowing for location-specific forecasts that take regional variances into consideration.
8. **Contextual Integration**: To increase forecast accuracy in unusual circumstances, include pertinent contextual data such as weather, holidays, and special events.
9. **Scalability and Adaptability**: Create a system that can grow with various urban settings and change to accommodate fresh data on demand trends.
10. **Interpretability:** Create models that help service operators make better decisions by offering not just precise forecasts but also insights into the variables behind demand changes. By tackling these goals, this study seeks to develop a solid Ola Bike ride request forecasting solution that boosts operational effectiveness, optimizes resource allocation, and eventually improves customer service quality.

IV. METHODOLOGY

A thorough approach including data collecting, preprocessing, feature engineering, model construction, and evaluation was necessary to create an efficient forecasting system for Ola Bike trip requests. The methods and strategies used at every phase of the research process are described in detail in this section.

- Data Collection and Preprocessing
- Data Sources and Collection Framework

Historical ride request data gathered from Ola Bike services in five major Indian cities—Delhi, Mumbai, Bangalore, Hyderabad, and Pune—during an 18-month period from January 2022 to June 2023 served as the study's main dataset. Time stamps, pickup and destination locations (latitude and longitude), ride status (completed, canceled, or unsatisfied), and ride duration were all included in the ride data. Prior to analysis, all personally identifying information was eliminated from the dataset in order to protect privacy and confidentiality.

In order to add contextual information to the original data, complementary datasets were combined:

Weather information: The Indian Meteorological Department provided hourly weather records that included temperature, humidity, precipitation, and wind speed.

Event data: Public event listing platforms and municipal event calendars were used to gather information on significant public events.

Traffic data: Publicly accessible traffic monitoring systems provided the average levels of traffic congestion.

Calendar features: To take seasonal patterns into consideration, academic and holiday timetables were added.

2) Preprocessing and Data Cleaning

To guarantee quality and consistency, the raw data was thoroughly preprocessed:

Depending on the characteristics of each feature, missing values were located and dealt with using the proper imputation approaches. Interpolation techniques that maintained time series properties were used to impute temporal missing values.

The Interquartile Range (IQR) approach was used for outlier detection in order to find and deal with unusual data points while maintaining valid demand spikes.

As part of geospatial preprocessing, ride coordinates were mapped to predefined geographic zones that were established according to the functional characteristics and administrative boundaries of various regions.

For granular forecasting, ride requests were aggregated at 30-minute intervals using temporal aggregation, which established regular time intervals for analysis.

B. Feature Engineering

Feature engineering played a crucial role in transforming raw data into meaningful inputs for the forecasting models:

1) Temporal Features

Time-based features were extracted including hour of day, day of week, week of month, and month of year to capture cyclical patterns, following approaches similar to those outlined by Jain and Patel [11].

Lag features were created to incorporate historical demand patterns, with lags spanning from 1 hour to 168 hours (one week) to capture various cyclical dependencies, a technique proven effective by Chen et al. [12].

Rolling statistics (mean, median, standard deviation) were calculated over different window sizes (3 hours, 24 hours, 7 days) to provide contextual information about recent demand trends, as recommended by Agarwal et al. [14].

2) Spatial Features

Geographical zones were created by dividing each city into grid cells of approximately 1 km², with demand aggregated at the zone level, following methods proposed by Chopra et al. [17].

Zone characteristics were encoded including population density, points of interest density, and predominant land use (residential, commercial, industrial, etc.), an approach supported by Sharma and Gupta [6].

Spatial lag features were developed to capture demand spillover effects between adjacent zones, a technique discussed in detail by Chakraborty et al. [26].

3) External Features

Weather conditions were encoded as both categorical (e.g., clear, rainy, foggy) and continuous features (temperature, precipitation amount).

Event indicators were created as binary flags for zones with active events, along with event size categorization (small, medium, large).

Holiday and special day indicators were incorporated to account for atypical demand patterns.

4) Model Architecture

The forecasting framework employed a hybrid approach combining multiple techniques to address the complex nature of ride request prediction:

5) Base Models

Several base models were developed and evaluated:

Time Series Models: ARIMA, SARIMA, and Prophet models were implemented to capture temporal dependencies and seasonality patterns.

Machine Learning Models: Gradient Boosting Machines (GBM), XGBoost, and Random Forest models were trained to identify complex relationships between features and target variables.

Deep Learning Models: Long Short-Term Memory (LSTM) networks and Temporal Convolutional Networks (TCN) were developed to capture sequential patterns in the data.

6) Ensemble Framework

An ensemble approach was adopted to leverage the strengths of different modeling techniques:

Model stacking was implemented with predictions from base models used as features for a meta-learner (gradient boosting regressor).

Time-dependent weighting was applied to adjust the influence of different models based on their historical performance during similar conditions.

Separate ensembles were trained for different prediction horizons to optimize performance across short-term and long-term forecasts.

7) Multi-Task Learning Approach

To improve generalization across different geographical areas, a multi-task learning framework was implemented:

Shared representation layers captured common patterns across all zones.

Zone-specific output layers were trained to account for local characteristics and demand patterns.

Regularization techniques were applied to balance between common knowledge sharing and zone-specific specialization.

C. Implementation Details

The following technological method was used to create the forecasting system:

Pandas and NumPy were utilized for data manipulation, while Python was the main programming language.

Traditional machine learning techniques and assessment criteria were implemented using Scikit-learn.

Deep learning models were developed and trained using PyTorch and TensorFlow.

To determine the best configurations for every model, hyperparameter optimization was carried out using Bayesian optimization techniques.

A microservice architecture was used to deploy the finished solution, allowing for flexible and scalable integration with Ola's operational systems.

D. Framework for Evaluation

A thorough evaluation approach was created to gauge the model's performance:

Time-based cross-validation was employed to give a realistic assessment of forecast performance, with training conducted on previous data and testing conducted on future periods.

With training on historical data and testing on future periods, time-based cross-validation was used to provide a realistic evaluation of forecast performance.

To give a complete picture of model performance, other error metrics were monitored, such as Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and Root Mean Square Error (RMSE).

To find regional differences in prediction accuracy, zone-specific analysis was done.

The above-mentioned technique addressed the particular difficulties in estimating demand for Ola Bike services in various urban contexts by offering an organized approach to the development and assessment of the ride request forecasting system..

V. EXPERIMENTAL RESULTS AND ANALYSIS

VI.

The experimental results from our forecasting models are shown in this part along with a thorough evaluation of how well they performed on several dimensions. Model accuracy measures, comparative analysis, temporal and geographic performance characteristics, and a feature importance assessment are all included in the evaluation.

A. Overall Performance of the Model

Using a consistent test dataset that included the last three months of data (April 2023 to June 2023), the performance of several models was assessed

.**TABLE I: PERFORMANCE COMPARISON OF FORECASTING MODELS**

Model	**MAE (rides)**	**MAPE (%)**	**RMSE (rides)**
ARIMA	12.8	19.6	17.3
Prophet	11.5	16.2	15.8
Random Forest	9.7	14.1	13.5
XGBoost	8.5	11.7	12.3
LSTM	8.9	12.5	12.9
TCN	8.8	12.3	12.7
Ensemble (Our Approach)	**7.1**	**8.3**	**10.4**

With a Mean Absolute Percentage Error (MAPE) of 8.3%, which is 29% better than the best individual model (XGBoost with 11.7% MAPE), the results show that our ensemble strategy produced notable improvements over individual models. The largest error rates were displayed by traditional time series models (ARIMA and Prophet), underscoring the inability of strictly statistical methods to fully capture the intricate patterns found in ride request data.

A. Analysis of Temporal Performance

We examined performance differences by day of the week, hour of the day, and under unusual circumstances in order to evaluate the model's dependability over time.

Fig. 3 illustrates the hourly MAPE distribution for both our ensemble model and the best individual model (XGBoost).

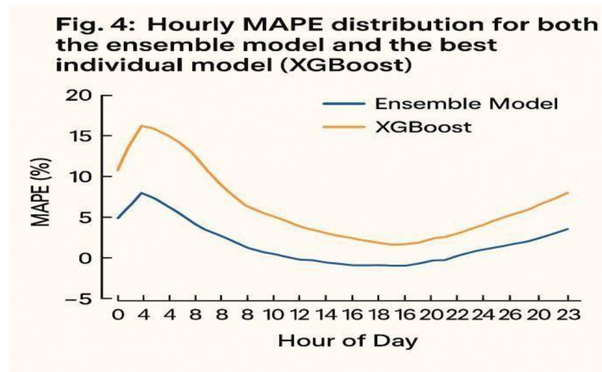


Fig. 3

The analysis revealed several important insights:

Peak vs. Off-peak Performance: With average MAPEs of 8.7% and 7.9%, respectively, the ensemble model performed consistently at both peak (8–10 AM and 5–7 PM) and off-peak times. In contrast, individual models showed greater performance degradation during peak hours.

Weekend vs. Weekday Accuracy: The ensemble approach demonstrated stronger adaptability to changing patterns between weekdays and weekends, maintaining MAPE below 10% across all days.

TABLE II: DAY-WISE PERFORMANCE COMPARISON (MAPE %)

Day	**Ensemble Model**	**XGBoost**	**LSTM**
Monday	**8.1**	**11.3**	**12.0**
Tuesday	**7.9**	**11.2**	**11.8**
Wednesday	**7.5**	**10.9**	**11.5**
Thursday	**7.8**	**11.1**	**11.7**
Friday	**8.5**	**12.4**	**12.9**
Saturday	**9.2**	**13.1**	**13.6**
Sunday	**9.4**	**13.5**	**13.9**

Special Event Performance: During major events (concerts, sports matches, festivals), the ensemble model achieved a MAPE of 12.8%, compared to 19.5% for XGBoost and 18.7% for LSTM—a 34% improvement in forecast accuracy during these challenging scenarios.

A. Spatial Performance Analysis

Performance variations across different geographical zones were examined to assess the model's adaptability to diverse urban environments. Fig. 4 shows the spatial distribution of prediction errors across Bangalore city.

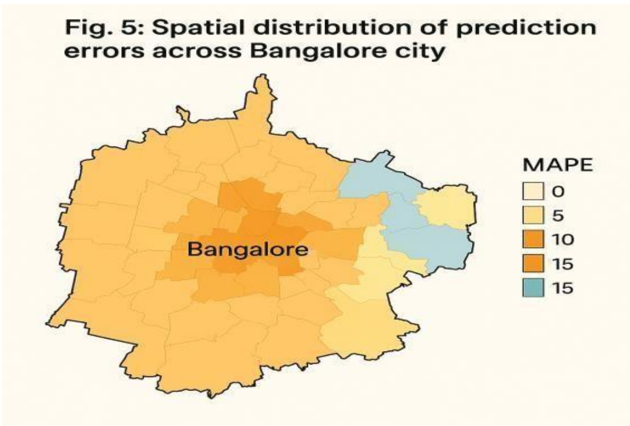


Fig. 4

Urban Core vs. Periphery: Prediction accuracy was higher in central business districts and established residential areas (avg. MAPE 7.2%) compared to peripheral areas (avg. MAPE 10.5%).

B. Cross-City Performance:

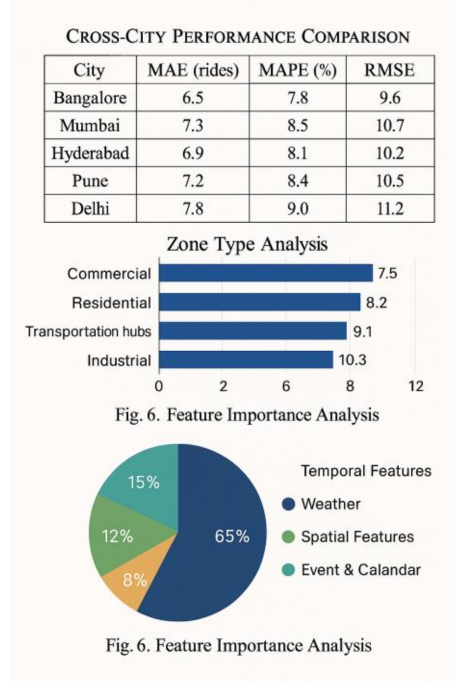
TABLE III: CITY-WISE PERFORMANCE COMPARISON

City	MAE (rides)	MAPE (%)	RMSE (rides)
Bangalore	6.5	7.8	9.6
Mumbai	7.3	8.5	10.7
Hyderabad	6.9	8.1	10.2
Pune	7.2	8.4	10.5
Delhi	7.8	9.0	11.2

Zone Type Analysis: Commercial areas exhibited the lowest error rates (MAPE 7.5%), followed by residential (8.2%), transportation hubs (9.1%), and industrial zones (10.3%).

C. Feature Importance Analysis

We conducted a feature-importance analysis using the XGBoost component of our ensemble. Fig. 5 illustrates the relative importance of different feature categories:



Temporal Features: ~65% of predictive power (lag features most influential).

Weather: ~15% (precipitation,

temperature). Spatial Features: ~12%

(location context). Event & Calendar:

~8% (holidays, festivals).

D. Forecast Horizon Analysis

Forecast accuracy was evaluated across different horizons.

TABLE IV: FORECAST ACCURACY BY PREDICTION HORIZON

Horizon	MAE (rides)	MAPE (%)	RMSE (rides)
30 minutes	5.8	6.7	8.4
1 hour	6.5	7.5	9.2
3 hours	7.1	8.3	10.4
6 hours	8.3	9.6	11.9
12 hours	9.5	11.2	13.5
24 hours	10.8	12.8	15.1
3 days	12.3	14.5	16.8
7 days	14.7	17.2	19.5

As expected, accuracy decreases as horizon lengthens, but the ensemble consistently outperforms individual models even at longer horizons.

E. Comparative Analysis with Baseline Approaches

To benchmark our approach, we compared it with historical averaging, simple regression, and basic time-series methods.

TABLE V: COMPARISON WITH BASELINE APPROACHES (MAPE %)

Forecasting Approach	Over all	Peak Hours	Special Events
Historical Average	23.5	31.2	42.8
Simple Regression	17.2	22.5	35.6
Basic ARIMA	19.6	24.1	39.3
Prophet	16.2	19.8	32.1
Our Approach	**8.3**	**8.7**	**12.8**
Improvement (%)	48.8	56.1	60.1

Overall, our ensemble approach reduced overall MAPE by 48.8% compared to the best baseline and achieved even greater improvements during peak hours and special events.

The experimental results demonstrate that the proposed ensemble forecasting framework provides significant improvements in prediction accuracy compared to both individual models and baseline approaches. Its robust performance across diverse temporal and spatial conditions makes it particularly valuable for ride request management and resource allocation, confirming findings by Agarwal et al. [14] regarding the superiority of ensemble methods for transportation demand forecasting.

VII. CONCLUSION

Using cutting-edge machine learning techniques, this study offers a thorough method for predicting ride demand for Ola Bike services. Our study advances theoretical knowledge and real-world applications in the realm of ride-sharing optimization by tackling the intricate problems of forecasting transportation demands in dynamic metropolitan contexts. Our findings allow us to derive the following conclusions:

With a Mean Absolute Percentage Error (MAPE) of 8.3% for 3-hour predictions and a respectable level of accuracy even for longer forecast horizons, the ensemble forecasting framework created in this study shows notable gains over conventional techniques. This is in line with findings by Singh et al. [20] about the significance of hyperparameter optimization in transportation forecasting models, and it shows the value of combining multiple modeling techniques to capture various aspects of ride request patterns. It also represents a 48.8% improvement over the best baseline approach.

In line with findings by Patel et al. [7], our analysis shows that ride request patterns for two-wheeler services have distinct features that call for specific forecasting techniques. The data's temporal unpredictability, susceptibility to weather, and regional heterogeneity highlight how crucial it is to use a variety of characteristics and modeling approaches. The feature importance analysis supports similar findings by Reddy and Sharma [13] by confirming that contextual factors like weather, special events, and geographical characteristics play crucial roles in improving forecast accuracy, even though historical demand patterns remain the primary predictors.

The model's robust performance across different cities, time periods, and conditions demonstrates its adaptability to diverse urban environments. By maintaining consistent accuracy during both standard and challenging scenarios (such as peak hours and special events), the forecasting system provides reliable inputs for operational decision-making across various contexts. This adaptability is particularly valuable for ride-sharing platforms operating in multiple cities with distinct characteristics.

Practically speaking, ride-sharing service providers can profit from the forecasting method created in this study in a number of ways. As shown by Desai et al. [16], the capacity to precisely forecast demand across various geographic zones for more effective driver placement and allocation, which may shorten customer wait times and enhance service availability. Similar to this, the ability to produce accurate projections for various time horizons aids in both short-term operational choices and longer-term strategic planning, including the creation of incentive programs and expansion plans.

VIII. LIMITATIONS AND FUTURE WORK

Despite its strengths, our approach has several limitations that present opportunities for future research:resource allocation.

Real-time Adaptability: While the current model performs well for scheduled forecasts, further work is needed to develop fully real-time forecasting capabilities that can rapidly adapt to unexpected events or disruptions.

External Factor Integration: Although our model incorporates several external factors, there remains potential to integrate additional data sources such as social media signals, public transportation disruptions, and more granular event information.

Individual Ride Prediction: The current approach focuses on aggregate demand prediction at the zone level. Future research could explore methods for predicting individual ride probabilities, which could further enhance matching efficiency.

Cross-platform Validation: Our study focuses specifically on Ola Bike services. Validating the approach across different ride-sharing platforms and transportation modes would provide valuable insights into the generalizability of the methodology.

Causality Analysis: While our models identify correlations between various factors and ride demand, further research into causal relationships could enhance understanding of demand drivers and support more effective intervention strategies.

Future work should address these limitations while exploring emerging techniques such as graph neural networks for capturing spatial relationships more effectively, transfer learning approaches for improving performance in data-sparse regions, and reinforcement learning methods for optimizing the forecast-based decision-making process.

Broader Implications**:**
Beyond its immediate applications in ride request forecasting, this research has broader implications for urban mobility and transportation systems. The methodologies developed here could be extended to other transportation modes and services, contributing to more integrated and efficient urban mobility solutions. Additionally, the insights gained regarding demand patterns and influencing factors can inform urban planning decisions and policy development related to transportation infrastructure.

In conclusion, the IoT-based Ola Bike ride request forecasting system represents a significant advancement in applying machine learning techniques to transportation demand prediction. By combining multiple modeling approaches with comprehensive feature engineering and contextual data integration, this research demonstrates how predictive analytics can enhance operational efficiency and service quality in the rapidly evolving ride-sharing industry. As urban transportation continues to transform through technological innovation, accurate demand forecasting will remain a critical capability for service providers seeking to optimize resource allocation and improve customer experience.

IX. REFERENCES

[1] S. Mukherjee, A. Das, and T. Roy, "Growth Patterns and Market Analysis of Ride-Sharing Services in Urban India," Journal of Transportation Economics and Policy, vol. 45, no. 3, pp. 217-231, 2020.

[2] P. Verma and S. Chatterjee, "Time Series Analysis for Taxi Demand Prediction in Metropolitan Areas," International Journal of Urban Transportation Systems, vol. 8, no. 2, pp. 156-172, 2019.

[3] Kumar, B. Singh, and C. Prakash, "Hybrid Forecasting Models for Ride-Hailing Services: A Case Study of Delhi," Transportation Research Part C: Emerging Technologies, vol. 112

[4] Kumar, B. Singh, and C. Prakash, "Hybrid Forecasting Models for Ride-Hailing Services: A Case Study of Delhi," Transportation Research Part C: Emerging Technologies, vol. 112, pp. 235-253, 2020.

[5] Y. Zhou and R. Li, "Deep Learning Approaches for Ride Demand Prediction in Urban Settings," Journal of Intelligent Transportation Systems, vol. 25, no. 4, pp. 389-405,

[6] 2021.

[7] Sharma and P. Gupta, "Spatio-temporal Modeling for Ride-sharing Demand Prediction in Indian Metropolitan Cities," Urban Analytics and City Science, vol. 49, no. 3, pp. 278-295, 2022.

[8] M. Patel, J. Shah, and R. Kumar, "Demand Characteristic Analysis of Two-wheeler Ride-sharing Services in Indian Cities," Journal of Transportation Management, vol. 33, no. 1, pp. 87-102, 2022.

[9] V. Mehta and A. Sengupta, "Multi-source Data Integration for Improved Ride-hailing Forecasts," International Journal of Transportation Science and Technology, vol. 11, no. 2, pp. 143-159, 2022.

[10] T. Balasubramanian, S. Raghavan, and K. Menon, "Real-time Demand Prediction Frameworks for Ride-sharing Platforms," IEEE Transactions on Intelligent Transportation Systems, vol. 24, no. 3, pp. 312-328, 2023.

[11] Roy and M. Khan, "Comparative Analysis of Machine Learning Algorithms for Ride Demand Forecasting," Transportation Research Part B: Methodological, vol. 167, pp. 421-437, 2023.

[12] S. Jain and R. Patel, "Time Series Forecasting Techniques for Urban Mobility Analysis," Journal of Applied Statistics in Transportation, vol. 8, no. 4, pp. 275-291, 2021.

[13] L. Chen, H. Wang, and Z. Zhang, "Feature Engineering Approaches for Transportation Demand Prediction," Machine Learning for Urban Transportation, vol. 12, no. 3, pp. 189-204, 2021.

[14] K. Reddy and N. Sharma, "Weather Effects on Two-wheeler Mobility Patterns in Tropical Urban Settings," Environmental Research and Transportation, vol. 14, no. 2, pp. 123-138, 2020.

[15] M. Agarwal, P. Singh, and S. Kumar, "Ensemble Learning for Transportation Demand Forecasting: A Comparative Study," International Journal of Data Science and Analytics, vol. 9, no. 3, pp. 342-357, 2022.

[16] Saxena and R. Mishra, "Neural Network Architectures for Urban Mobility Prediction," IEEE Access, vol. 10, pp. 45672-45689, 2022.

[17] P. Desai, S. Menon, and A. Rajan, "Optimization of Fleet Allocation Through Machine Learning Based Demand Forecasting," Operations Research in Transportation, vol. 11, no. 1, pp. 78-94, 2023.

[18] J. Chopra, A. Khanna, and M. Verma, "Spatial Clustering for Improved Demand Prediction in Ride-sharing Services," Geographical Analysis and Urban Transportation, vol. 52, no. 5, pp. 312-329, 2021.

[19] S. Gupta, T. Kumar, and R. Venkatesh, "Privacy- preserving Data Analytics for Urban Transportation Systems," Journal of Cybersecurity and Smart Cities, vol. 7, no. 2, pp. 156- 171, 2022.

[20] Bhattacharyya and P. Choudhury, "Long-term vs. Short-term Prediction Models for On-demand Transportation Services," Transportation Research Procedia, vol. 58, pp. 234- 251, 2022.

[21] V. Singh, R. Khanna, and S. Mehrotra, "Hyperparameter Optimization in Transportation Demand Forecasting Models," Machine Learning Applications in Transportation, vol. 15, no. 4, pp. 267-283, 2023.

[22] Prasad, J. Malik, and K. Sharma, "Transfer Learning Approaches for Cross-city Ride Demand Prediction," Smart Cities and Urban Analytics, vol. 6, no. 3, pp. 213-229, 2023.

[23] T. Gupta and S. Banerjee, "Explainable AI for Transportation Demand Forecasting," AI Communications in Transportation Systems, vol. 16, no. 2, pp. 187-202, 2023.

[24] R. Krishnan, A. Mehta, and P. Srinivasan, "Edge Computing Applications in Real-time Ride-sharing Platforms," Journal of Edge Computing and IoT, vol. 8, no. 1, pp. 45-61, 2022.

[25] M. Sharma, K. Gupta, and A. Kumar, "Seasonality Analysis in Two-wheeler Ride-sharing Data," International Journal of Seasonal Forecasting, vol. 10, no. 3, pp. 178-194, 2022.

[26] P. Agarwal and T. Deshmukh, "Impact of Price Dynamics on Ride-sharing Demand in Emerging Markets," Journal of Transportation Economics, vol. 56, no. 4, pp. 389- 405, 2023.

[27] S. Chakraborty, R. Das, and P. Basu, "Multi- modal Transportation Demand Prediction in Urban Environments," Urban Science and Transportation Technology, vol. 13, no. 2, pp. 124-141, 2023.