JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Advancements in Predictive Modeling for RideShare Systems with Optimal Features

[1]Sajith Sharwan BV, [2]Sandeep J, [3]Norton Stanley S A, [4]Smera C, [5]Sreeja C S

[1]Student, [2]Professor, [3]Professor
[1]Department of Computer Science,
[1]CHRIST (Deemed to be University), Bangalore, India

Abstract : This review explores how predictive modeling revolutionizes urban transportation by enhancing efficiency and resource allocation. By leveraging machine learning algorithms like RNNs and LSTM networks, transportation patterns are forecasted accurately. Spatial-temporal analysis aids in optimizing operations, while optimization techniques improve decision-making processes. Evaluation metrics ensure model accuracy. The study emphasizes the importance of integrating contextual factors and passenger behaviors for comprehensive urban mobility planning. Future research directions include refining prediction accuracy and exploring innovative algorithms for enhanced efficiency in transportation systems.

IndexTerms **- Predictive modeling, Transportation systems, Machine learning algorithms, Ride-sharing, Demand prediction, Urban mobility.**

I. INTRODUCTION

In today's urban landscape, predictive modeling is reshaping transportation by foreseeing patterns and optimizing efficiency. This review explores its impact on urban mobility, covering applications, methods, and outcomes. Advanced tech and diverse data now power the algorithms predicting demand and refining routes, from taxis to autonomous services. This study of recent trends aims to spotlight opportunities and challenges. We cover various topics, such as taxi and ride-sharing optimization, by the use of predictive modeling. The goal is to evaluate existing models, revealing insights into transformative potential and challenges.
Predictive modeling has the potential to improve transportation, creating more efficient, accessible, and sustainable urban spaces. It will also improve the user experience of customers using ride-share platforms and as well the drivers participating in the system. The next sections delve into the literature, methods, findings, and outline paths for future research. The literature review on advancements in predictive modeling for transportation systems delves into a dynamic and multidisciplinary field characterized by ongoing innovation, challenges, and opportunities. Across the papers analyzed, several key concepts and fundamental topics emerge consistently, shaping the landscape of predictive modeling in transportation.

1.1 Common Key Concepts

The common key concepts and ideas explored in this paper are as follows:

1.1.1 Predictive Modeling:

The central theme of the review, predictive modeling is extensively explored in each paper to forecast various aspects of transportation systems. For example, studies like "A Framework for Passengers Demand Prediction and Recommendation" [1] and "Real-Time Prediction of Taxi Demand Using Recurrent Neural Networks" [7] employ predictive modeling techniques to enhance efficiency and optimize resource allocation in urban transportation services.

1.1.2 Machine Learning Algorithm:

Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks are commonly utilized across the reviewed papers to capture sequential dependencies and improve prediction accuracy. The study "Sequence Learning Model with Recurrent Neural Networks for Taxi Demand Prediction" [2] highlights the effectiveness of LSTM networks in forecasting taxi demand patterns.

1.1.3 Spatial-Temporal Analysis:

The central theme of the review, predictive modeling is extensively explored in each paper to forecast various aspects of transportation systems. For example, studies like "A Framework for Passengers Demand Prediction and Recommendation" [1] and "Real-Time Prediction of Taxi Demand Using Recurrent Neural Networks" [7] employ predictive modeling techniques to enhance efficiency and optimize resource allocation in urban transportation services.

1.1.4 Optimization:

Optimization techniques are employed to enhance resource allocation, routing efficiency, and decision-making processes in transportation systems. The paper "Predictive Routing for Autonomous Mobility-on-Demand Systems with Ride-Sharing" [6] introduces a constrained optimization method for dynamic vehicle routing, improving fleet management and reducing travel times.

1.1.5 Evaluation Metrics:

Various evaluation metrics such as F-Measure and Exponentially Weighted Moving Average (EWMA) are used to assess the performance and accuracy of predictive models in transportation [1].

By bringing together these key concepts and methodologies, the review paper provides valuable insights into the evolving landscape of predictive modeling in transportation systems.

II. LITERATURE REVIEW

A review of the literature incorporating all the methodologies, key findings, and outcomes in the role of predictive modeling in urban transportation. Figure 1 gives an illustration of algorithms used in the literature in the past few years. In the following section, a thorough discussion of the literature review in this area is carried out. Table below gives a quick glimpse of key factors of the review conducted.

2.1 Demand Prediction and Optimization Studies Overview:

Paper	Key Finding	Factors used	Algorithm
[1]	Framework improves taxi efficiency with better prediction.	Location, Hotness, Timestamp	DBSCAN, EWMA.
[2]	LSTM outperforms in predicting taxi demand.	Taxi demand variables in different city areas. Historical taxi requests, Date, Time, City division.	LSTM, RNNs, MDNs.
[3]	Dispatch system reduces wait times effectively.	Passenger waiting time, Taxi idle distances.	LSTM, MDN, Mixed Integer Programming.
[4]	FFCS needed for urban mobility challenges.	Trip densities, Actual/predicted carsharing demand.	Not mentioned.
[5]	Method reduces travel and waiting times.	Service rate, Travel delays, Shared rides.	Prediction-based method from historical taxi data.
[6]	Sequence model accurately predicts future requests.	Taxi requests, Current demand, Weather.	LSTM, MDN.
[7]	Intelligent systems minimize wait and idle times.	Taxi demand and destination patterns.	RNNs, feed-forward neural network, Mixed Integer Programming.
[8]	Network effectively predicts taxi demand.	Taxi Origin-Destination Demand Data, Taxi trip geocoordinates, Trips between regions.	CSTN.
[9]	Mining data enhances demand prediction for taxis.	Demand distributions based on time, weather, and location.	K-means, DBSCAN, OPTICS.
[10]	Neural network forecasts taxi demands efficiently.	ID, Time, Region, Demands, Precipitation.	Neural network (MLP).
[11]	DESTPRE accurately predicts destinations.	Destination prediction accuracy, Algorithm efficiency.	FREQ, CLUSTER.

[12]	Ensemble model enhances taxi demand prediction.	Pick-up factors: Weather, Holiday, Time.	Prediction-based method from historical taxi data.
[13]	Co-prediction methods model demands effectively.	Taxi pick-up and drop-off demand.	Multi-task learning models, deep learning-based.
[14]	Ensemble model excels in destination prediction.	GPS coordinates, Timestamps, Day types.	Ensemble Learning Model (ELM), SVR, Deep Learning, Segment Estimation Classifier, kNN.

[1] introduces a new way to help taxis pick up passengers more efficiently. By analyzing GPS data from thousands of taxis in Beijing, researchers found better ways to predict where passengers are likely to be and suggest good pickup spots. They discovered that their method increased the accuracy of predictions by 15.21% and improved recommendation success by 79.6%. While the study didn't talk much about its limitations, it emphasized the importance of understanding passenger needs for making taxis work better. Future work could include adding traffic data to make scheduling even smoother.

In [2] the author introduces a powerful LSTM model for predicting taxi demand. By analyzing over 600 million taxi trips in New York City, researchers found that LSTM outperforms other methods in predicting future taxi demands with smaller errors. They identified challenges in predicting demand in small areas and capturing long-term patterns. Despite uncertainties, LSTM showed promise in predicting demand probabilities. The study suggests incorporating location-specific data and real-time demand insights for better taxi organization. While research gaps weren't specified, the study emphasizes the need for further exploration in optimizing taxi services based on demand predictions.

[3] presents an effective taxi dispatch system, reducing passenger wait times and taxi idle driving distances. Utilizing real-world data from New York City, the system proves practical in enhancing taxi services. The study measures success by evaluating the average waiting time for passengers and the idle driving distances of taxis. While the paper lacks explicit mention of limitations or future research, it highlights the need for more accurate prediction of taxi destination distributions and integrated studies on demand and destination patterns. The proposed system stands as a promising solution for improving taxi efficiency in urban areas.

The study [4] addresses urban mobility challenges by emphasizing the potential of flexible carsharing (FFCS). FFCS providers can benefit from decision support systems, reducing costs and aiding strategic decisions. Using data from a major carsharing provider in Amsterdam and Berlin, the study predicts and measures actual carsharing demand based on points of interest. The methodology involves analyzing rental data, points of interest, and employing predictive modeling. While limitations are not explicitly stated, the research suggests gaps in developing widely applicable carsharing demand models and enhancing decision support systems. The study encourages future research to validate proposed approaches in real-world settings and explore the balance between profitability and equal mobility access.

A method for vehicle routing and request assignment, integrating future demand prediction to reduce travel and waiting times for passengers is proposed in [5]. Using New York City taxi data, the approach enhances fleet positioning for meeting future requests efficiently. Future research could focus on improving vehicle rebalancing and prediction accuracy. The methodology involves constrained optimization and probabilistic demand modeling. While limitations include operational cost-performance trade-offs, the method offers insights for optimizing autonomous mobility systems, considering service rates, travel delays, and computational efficiency.

In the study [6], a sequence learning model to forecast future taxi requests based on recent demand and relevant data. Utilizing New York City taxi data, the model surpasses traditional prediction methods. Key outcomes include improved prediction accuracy measured by sMAPE and RMSE. The dataset encompasses over 600 million taxi trips from January 2013 to June 2016. Methodologically, the city is divided into smaller areas, and a Long Short-Term Memory (LSTM) recurrent neural network is trained. While limited to New York City, the study underscores the potential for enhanced accuracy and generalizability in future research.

The paper [7] aims to enhance transportation efficiency by minimizing passenger waiting time and idle driving distances. Utilizing New York City taxi trip data, collected from over 15,000 taxis, the study focuses on one week of data in 2016. The methodology involves neural network modeling trained on 2015 data and validated using 2016 records. Future research suggests analyzing passenger behaviors and refining ride-sharing strategies. Although explicit limitations are absent, the study emphasizes the importance of accurate prediction models for integrated transportation systems, offering valuable insights into optimizing urban taxi dispatch.

In [8], a novel approach to predict taxi demand between various origin-destination pairs is proposed. Utilizing the NYC-TOD dataset, constructed from New York City yellow taxi trip records in 2014, the method employs a Contextualized Spatial-Temporal Network (CSTN) divided into spatial, temporal, and global correlation contexts. Future research could explore additional contextual factors' impact on prediction accuracy and real-time model implementation. Limitations include focusing solely on departure places and challenges with taxi preallocation systems. The study offers insights into improving taxi demand prediction models for urban transportation planning.

The study focuses on predicting taxi demand hotspots by analyzing historical data with consideration for time, weather, and taxi location. The dataset, likely comprising historical taxi demand info and contextual data, facilitates this analysis. The methodology involves four steps: data filtering, clustering, semantic annotation, and hotness calculation, with three clustering algorithms compared. The main outcome is improved fleet management through accurate predictions. However, limitations include single-company, area-specific data, potential accuracy issues due to data quality, and an exclusive focus on demand prediction. Addressing these issues in future research could enhance the broader applicability of the findings [9]. Explores forecasting taxi demands using taxi probe data and a neural network, emphasizing the role of the day of the week in predictions. It utilized taxi probe data from Tokyo over two months, configuring the neural network with 4-hour intervals and 50 neurons in the hidden layer. Findings highlight the neural network's efficacy in demand forecasting and the importance of considering weekdays due to periodic demands. Future research should integrate weather data and events like festivals. Dependent variables measured include taxi demands, with variables such as region, day, time, and precipitation considered [10].

DESTPRE, a practical method for predicting destinations in taxi rides based on historical trajectory data, offering promising results across various applications [11]. It directly analyzes trajectories without complex models, leveraging historical data for precise predictions. Results from experiments demonstrate the effectiveness of DESTPRE in accurately forecasting destinations compared to existing methods. The dataset comprises real GPS points from 12,000 Beijing taxis over a three-month period, meticulously cleaned to ensure reliability. Future research could enhance prediction accuracy by integrating additional factors like time and weather. Dependent variables measured include destination prediction accuracy and algorithm efficiency, with metrics such as AvgMinErr and DAV assessed.

Study done in [12] shows that combining different computer programs makes it better at guessing when people will need taxis in Bangkok. It says it's important to use a mix of these programs to guess well in different parts of the city. They used information from over 5,000 taxis in Bangkok for four months, paying attention to where they were and what the weather was like. However, they only looked at Bangkok and didn't check all the different ways the programs could be used. They looked at things like the number of people getting taxis and the time of day, but didn't look at other important things like special events or road closures.

In [13], the author proposes a new way to predict both when people will need taxis and when they'll be dropped off, saying it's crucial for making taxi services better and faster. It found that predicting pick-ups and drop-offs together works well. Future research could look into how weather and events affect taxi demand and use similar methods for other traffic predictions. They measured how well the model could predict both pick-ups and drop-offs at the same time using data from New York City taxis in 2016 and 2018. They used a special kind of computer model and compared it to others to see how well it worked. They found that pick-ups and drop-offs are related to each other. Limitations include that the findings might only apply to New York City taxis and that the computer models used could be improved. They looked at how well the model could predict when people would get picked up and dropped off by taxis.

Paper [14] presents a new method using machine learning to predict where taxis will go, outperforming other methods. It found that the closer a taxi is to its destination, the more accurate the prediction. Future studies might explore different ways to guess where a taxi is on its route and predict when it will arrive. They tested their method using real data from taxis in Porto, Portugal, and Chengdu, China. They used a special technique to process the data and combined different computer models to make predictions. They measured how well their method worked by looking at how close their predictions were to the actual taxi destinations.

III. KNOWLEDGE OUTCOME FROM THE LITERATURE

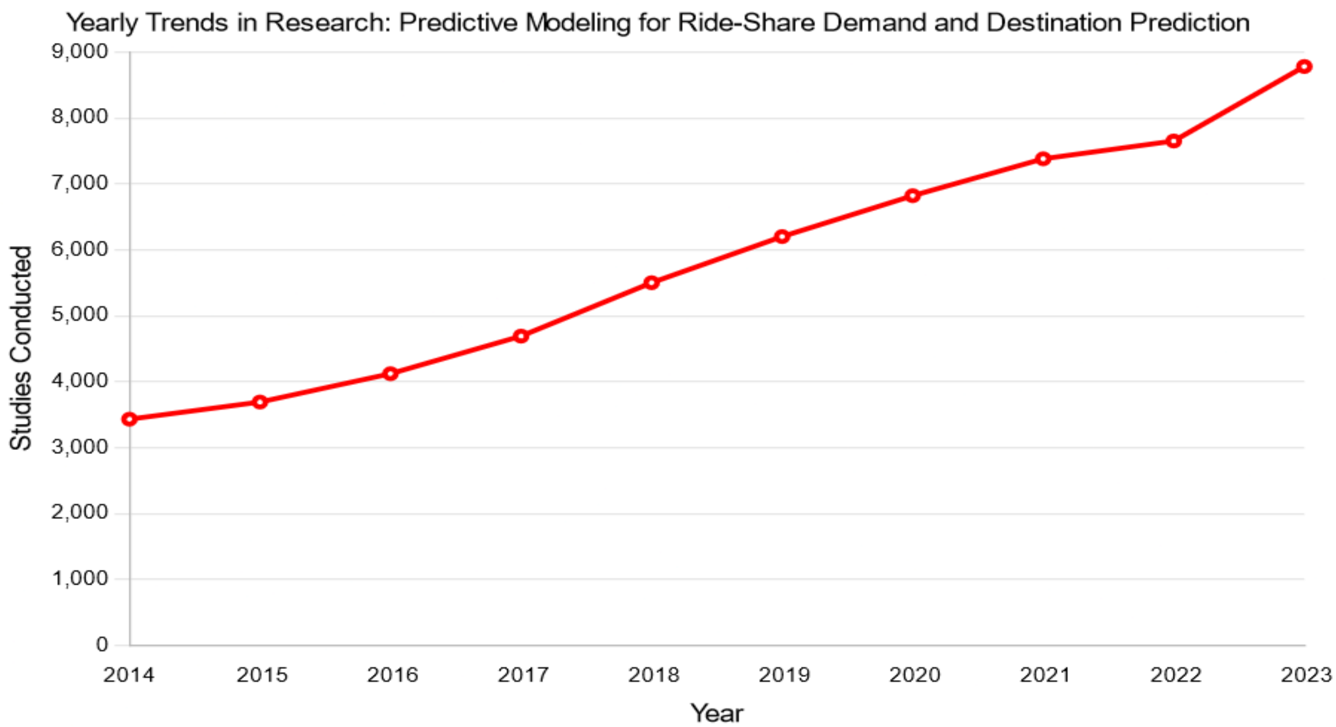
Many of these studies are found to use common algorithms to arrive at their result as depicted in the table below:

Table 3.1: Trends in Algorithms Used for Ride-Share Predictive Modeling

Algorithm	Used in	Year (Range)
LSTM (Long Short-Term Memory)	[2], [3], [6], [15], [16]	2017 - 2020
MDN (Mixture Density Network)	[2], [3], [6], [14], [17]	2017-2023
DBSCAN (Density-Based Spatial Clustering of Applications with Noise)	[1], [9]	2010-2016
RNNs (Recurrent Neural Networks)	[2], [7], [18], [19]	2017-2023
Mixed Integer Programming	[3], [7], [20], [21]	2017-2019

LSTM and MDN are consistently popular for ride-share predictive modeling. RNNs are also widely used, showing ongoing relevance. Mixed Integer Programming was popular earlier, hinting at a shift in focus lately. DBSCAN, less common recently, may have influenced new clustering methods.

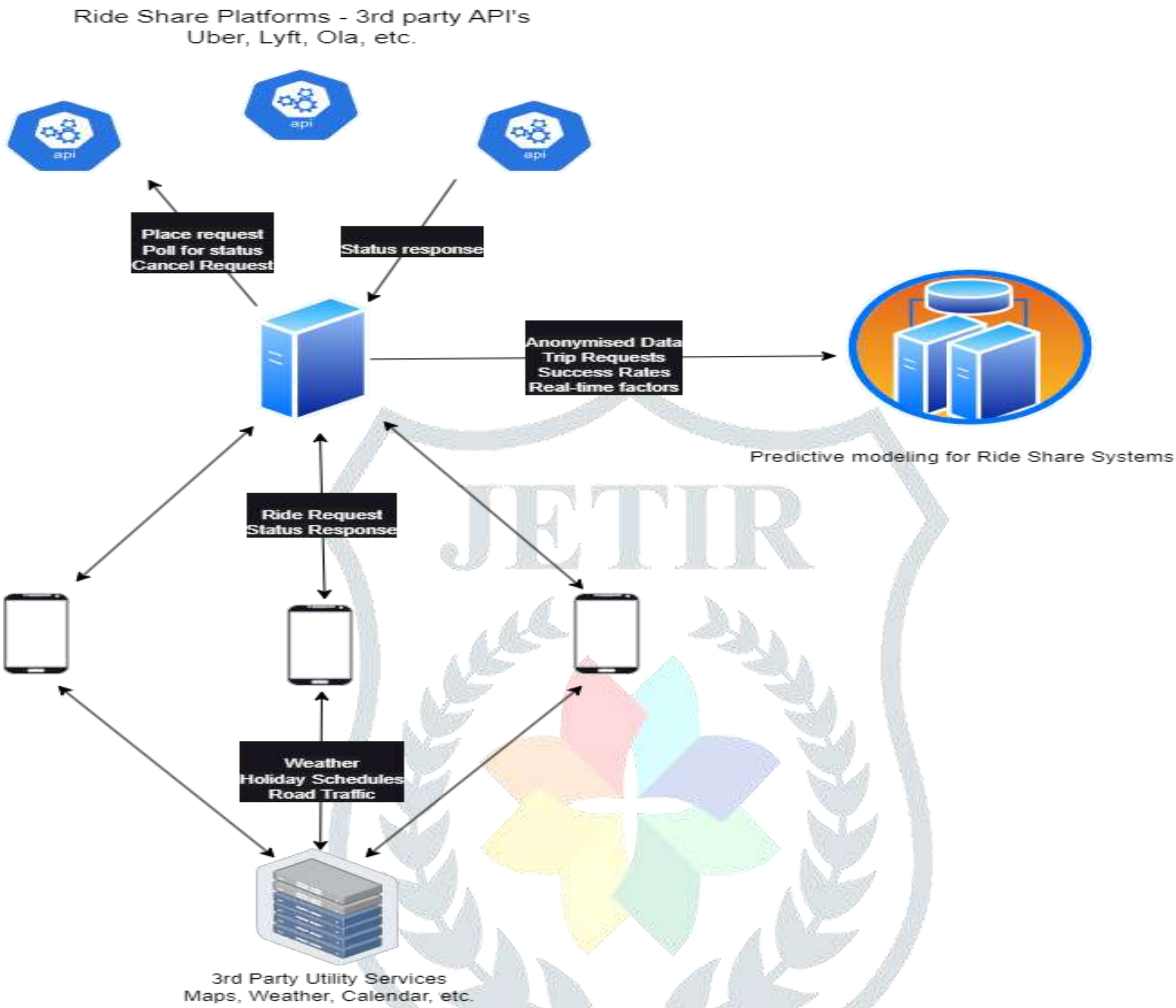
Figure 3.2: Recent trends in RideShare predictive modeling.



The graph shows how many research studies were done each year about predicting things in ride-share services. It started with fewer studies in 2014, then increased every year, reaching the most in 2023. This shows that more and more people are interested in using data to make ride-sharing better.

The study shows that there has been a lot of interest from researchers in the domain of predictive modeling, specific to ride renting and sharing applications.

IV. IDENTIFICATION OF STANDARD ARCHITECTURE FOR THE PREDICTIVE MODEL BASED TRANSPORTATION APPLICATION.



The standard architecture for the proposed application involves 3 major entities. Starting with an app that integrates the common functionality of all the major ride-share platforms, the 3rd party services that provide the influencing data factors such as weather, traffic, etc., and the anonymised data collection and storage for the ride success measures, availability, along with the influencing factors.

V. IMPLEMENTATION

The implementation proposal is a mobile application that simplifies the process of ride-sharing by consolidating ride requests and broadcasting them to multiple platforms such as Uber, Ola, and Lyft simultaneously. The app takes user's ride preferences once and sends them as requests in multiple platforms, and polls for the request status. Once a request is accepted by one platform, the app automatically cancels redundant requests. While the ride is being requested by the app, it also registers the influencing factor data such as time and location, as well as temperature, weather, road condition, traffic analysis, etc. On the whole, the app gathers such anonymous data on requested trips and also integrates information from third-party utility services, such as weather conditions, holiday schedules, and real-time traffic conditions.

By analyzing spatial and temporal patterns along with the success rates of different platforms, our goal is to forecast future ride demand accurately, thus enhancing the overall user experience. This approach not only provides convenience for passengers but also offers valuable insights for predictive modeling, benefiting both users and service providers in the ride-share industry.

VI. CONCLUSION

The document delves into the realm of predictive modeling in urban transportation systems, highlighting the utilization of machine learning algorithms such as LSTM and MDN, spatial-temporal analysis, and optimization techniques. It emphasizes the importance of forecasting transportation patterns and behaviors to optimize infrastructure, reduce congestion, and enhance overall transportation services.

Future research directions include integrating additional contextual factors into predictive frameworks, analyzing passenger behaviors to inform ride-sharing strategies, and improving prediction accuracy by incorporating location-specific information and weather data.

The document highlights the potential of predictive modeling to create more efficient, accessible, and sustainable urban spaces, benefiting both customers and drivers in the transportation system. By exploring various methodologies and key concepts in predictive modeling for transportation systems, the document aims to deepen understanding and make way for innovative approaches in urban mobility and transportation services.

In addition to this, an implementation proposal for an app that centralizes ride requests across different ride-share platforms and collects anonymous trip details, along with other data factors used in predictive modeling, such as weather and traffic at the time of the trip. This implementation will serve as a valuable data source for predictive modeling in the ride-share transportation sector.

REFERENCES

[1] K. Zhang, Z. Feng, S. Chen, K. Huang and G. Wang, "A Framework for Passengers Demand Prediction and Recommendation," 2016 IEEE International Conference on Services Computing (SCC), San Francisco, CA, USA, 2016, pp. 340-347

[2] J. Xu, R. Rahmatizadeh, L. Bölöni and D. Turgut, "A Sequence Learning Model with Recurrent Neural Networks for Taxi Demand Prediction," 2017 IEEE 42nd Conference on Local Computer Networks (LCN), Singapore, 2017, pp. 261-268

[3] Xu, Jun, et al. "A taxi dispatch system based on prediction of demand and destination." Journal of Parallel and Distributed Computing 157 (2021): 269-279.

[4] Willing, Christoph, et al. "Moving in time and space–Location intelligence for carsharing decision support." Decision Support Systems 99 (2017): 75-85.

[5] Alonso-Mora, Javier, Alex Wallar, and Daniela Rus. "Predictive routing for autonomous mobility-on-demand systems with ride-sharing." 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, 2017.

[6] Xu, Jun, et al. "Real-time prediction of taxi demand using recurrent neural networks." IEEE Transactions on Intelligent Transportation Systems 19.8 (2017): 2572-2581.

[7] J. Xu, R. Rahmatizadeh, L. Bölöni and D. Turgut, "Taxi Dispatch Planning via Demand and Destination Modeling," 2018 IEEE 43rd Conference on Local Computer Networks (LCN), Chicago, IL, USA, 2018, pp. 377-384

[8] Qiu, Zhilin, et al. "Taxi origin-destination demand prediction with contextualized spatial-temporal network." 2019 IEEE International Conference on Multimedia and Expo (ICME). IEEE, 2019.

[9] Chang, Han-wen, Yu-chin Tai, and Jane Yung-jen Hsu. "Context-aware taxi demand hotspots prediction." International Journal of Business Intelligence and Data Mining 5.1 (2010): 3-18.

[10] Mukai, Naoto, and Naoto Yoden. "Taxi demand forecasting based on taxi probe data by neural network." Intelligent Interactive Multimedia: Systems and Services: Proceedings of the 5th International Conference on Intelligent Interactive Multimedia Systems and Services (IIMSS 2012). Berlin, Heidelberg: Springer Berlin Heidelberg, 2012.

[11] Xu, Mengwen, Dong Wang, and Jian Li. "DESTPRE: a data-driven approach to destination prediction for taxi rides." Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing. 2016.

[12] Vanichrujee, Ukrish, et al. "Taxi demand prediction using ensemble model based on rnns and xgboost." 2018 International Conference on Embedded Systems and Intelligent Technology & International Conference on Information and Communication Technology for Embedded Systems (ICESIT-ICICTES). IEEE, 2018.

[13] Zhang, Chizhan, et al. "Taxi demand prediction using parallel multi-task learning model." IEEE Transactions on Intelligent Transportation Systems 23.2 (2020): 794-803.

[14] Zhang, Xiaocai, et al. "Prediction of taxi destinations using a novel data embedding method and ensemble learning." IEEE Transactions on Intelligent Transportation Systems 21.1 (2019): 68-78.

[15] Jin, Guangyin, et al. "Urban ride-hailing demand prediction with multiple spatio-temporal information fusion network." Transportation Research Part C: Emerging Technologies 117 (2020): 102665.

[16] Zhao, Jie, et al. "Unifying Uber and taxi data via deep models for taxi passenger demand prediction." Personal and Ubiquitous Computing (2023): 1-13.

[17] Li, Xiaoming, et al. "Demand Density Forecasting in Mobility-on-Demand Systems Through Recurrent Mixture Density Networks." 2023 IEEE 26th International Conference on Intelligent Transportation Systems (ITSC). IEEE, 2023.

[18] Wang, Ning, et al. "Ride-hailing origin-destination demand prediction with spatiotemporal information fusion." Transportation Safety and Environment (2023): tdad026.

[19] Nejadettehad, Alireza, Hamid Mahini, and Behnam Bahrak. "Short-term demand forecasting for online car-hailing services using recurrent neural networks." Applied Artificial Intelligence 34.9 (2020): 674-689.

[20] Jin, Guangyin, et al. "Urban ride-hailing demand prediction with multiple spatio-temporal information fusion network." Transportation Research Part C: Emerging Technologies 117 (2020): 102665.

[21] Tsao, Matthew, et al. "Model predictive control of ride-sharing autonomous mobility-on-demand systems." 2019 International conference on robotics and automation (ICRA). IEEE, 2019.

[22] Guo, Zhen, et al. "Data-driven robust optimization for contextual vehicle rebalancing in on-demand ride services under demand uncertainty." Transportation Research Part C: Emerging Technologies 154 (2023): 104244.