

Real-time Taxi Demand and Supply prediction based on Specific Geo locations using Machine Learning - A Systematic Literature Review

AA Abeyssekara¹, SG Sumanasekara¹, SMDD Samaraweera¹, WAC Imasha¹, ASYH Mihiran¹,
ADAI Gunasekara¹, DU Vidanagama¹, SMM Lakmali¹, KD Madhubashini¹ and
HRWP Gunathilaka¹

¹Faculty Of Computing, General Sir John Kotelawala Defence University, Sri Lanka
38-dba-0020@kdu.ac.lk

Abstract—With the increasing popularity of ride-hailing services, the accurate prediction of taxi demand has become a crucial task for service providers. In recent years, the availability of large-scale geospatial data and the development of machine learning algorithms have led to significant advancements in taxi demand prediction. The aim of systematic literature review is to analyze the techniques and approaches for taxi demand and supply prediction using geospatial data and machine learning algorithms. A total of 21 research papers published between 2017 and 2023 were selected based on inclusion and exclusion criteria. The papers were analyzed based on their research objectives, methodology, datasets and evaluation metrics. The result of the literature review indicate that the accuracy of taxi demand prediction models depends on the quality and quantity of the data, the selection of learning algorithms, and the feature engineering techniques used. The systematic literature review highlights the potential of using geospatial data and machine learning algorithms for accurate taxi demand prediction and need for more standardized evaluation metrics and further research to address the challenges. Machine learning algorithms, such as linear regression, decision trees, and artificial neural networks, clustering have been widely used for prediction tasks, focusing on factors like real-time population data. Through a comprehensive analysis, it is determined that clustering emerges as the most suitable technique for the research.

Index Terms—Geo-location, Machine Learning, CNN

I. INTRODUCTION

The transportation industry is rapidly evolving with the emergence of ride-sharing services such as PickMe and Uber. These services offer a convenient and efficient way for passengers to arrange rides, order food and groceries through their smartphones. Demand for taxi drivers increases as most people are likely to use taxis for leisure, for business purposes and for quick rides. Although there is high demand for taxi drivers, there is a common issue they struggling with. The issue is the supply and demand imbalance. The following section presents the research findings pertaining to the research topic of "Live Demand and Supply Prediction based on Specific Geo-Locations."

Models Such as Joint Guidance Residual Network (JGNET) Author and Wu [2020], are proposed to forecast the demand value of an entire city at a future time based on historical data and Dmodel Author [2017a] is proposed to observe hidden contexts to infer passenger demand. Also, technologies such as Deep Learning Author [2022a], Author [2023], Author and F.Wang [2021], artificial neural networks Author [2017b], Recurrent Neural Networks Author and Turgut [2017], Machine learning Author and Grimshaw [2021], Author and P.Viveka [2017], cellular networks Author and Kawasaki [2018], and network modeling Author and Deflorio [2018] are used to predict taxi passenger demand and supply and improve the efficiency of taxi services and reduce the waiting time of the customers. Constructing an adaptive recommendation system Author [2018] based on supply and demand level where the hotspot Author [2021], Author [2020] with the highest value is recommended to the driver reduces driver searching time and customers' waiting time. Further, busy zones are analyzed Author and Prakash [2019] in interior cities and outskirts where more transactions take place. Another finding is identifying different types of travelers and tailoring mobility interventions to their needs Author and Gebhard [2019]. This is a cluster analysis approach using survey data.

The aim of conducting this review paper is to provide an overview of the current state of knowledge on the research topic, "Live demand and Supply prediction based on specific geo locations" to identify the strengths and weaknesses of the existing literature and identify areas for further research by providing recommendations for future research. Finally, identify the gaps in the existing literature related to the research topic. This paper describes the methodology in Section 2, and in Section 3, it provides answers to six major research questions during the discussion. The results of this study are presented in Section 4, while the conclusions drawn from the findings are discussed in Section 5.

II. METHODOLOGY

In systematic literature review method is a crucial factor. It includes the procedure of finding research

Real-time Taxi Demand and Supply prediction based on Specific Geo locations using Machine Learning - A Systematic Literature Review

papers, selecting suitable research papers, inclusion and exclusion criterias, research questions, quality assessment. We have to carry out these steps to conduct an unbiased review of the research papers. The steps of systematic literature review methods are discussed below.

A. Sources

In a systematic literature review we have to gather relevant studies conducted on the relevant topic which is "Predicting the live demand and supply of shared mobility service". To gather the relevant studies a diversified search through the resources such as Research Gate, Google scholar, ScienceDirect was conducted using specified and relevant key words such as live demand, prediction, demand and supply, shared mobility services etc. Furthermore, reference articles of the relevant articles were manually checked.

B. Research questions

Here we have defined six research questions to be studied furthermore.

RQ1. What are the key factors that contribute to the supply and demand imbalance in taxi services?

RQ2. How accurate are the current methods used to predict high demand areas in taxi services?

RQ3. What is the impact of reducing supply in high demand areas and over supply in low demand areas on overall profitability for taxi service providers?

RQ4. What is the current methodology employed to predict high demand areas in taxi services and its limitations?

RQ5. How will the proposed new methodology improve the efficiency of taxi services?

RQ6. What are the methodologies, tools and techniques used throughout the research?

With respect to RQ1, the objective is to identify the factors which affects the variability of the demand and supply time to time. According to those factors we can identify the patterns of the demand and supply. With respect to RQ2, We have to consider the accuracy of the current methods used to predict the demand and supply to select the most suitable methods to apply. With respect to RQ3, main purpose of the research is to increase the profitability of the taxi service. So, we have to consider how the research affects the profitability of the taxi service providers. With respect to RQ4, We have to identify the limitations of the current methods and how to improve those methods. So, we have a clear insight into the current methods. With respect to RQ5, we have to contribute to the improvement of the efficiency of the taxi service using our results. So, in our study we have to focus on efficiency. With respect to RQ6, We need to study the methods, tools and techniques used to predict the demand and supply of the taxi service. We have to identify the pros and cons of those methods and select the most appropriate methods which can be used to predict demand and supply.

C. Data extraction

Here we have to select the appropriate research papers to be studied further. Therefore, we have to build a protocol to avoid adding unnecessary data into the literature review. So, we have defined some inclusion and exclusion criterias. Here we have to check whether the objective clearly connected with the the literature review objectives and the data collection and analysing part are clearly defined. If not, we have to exclude the paper.

1) Inclusion criteria:

- Papers published in last 10 years
- Papers focused on more on the prediction of demand and supply
- Studies which used machine learning and deep learning techniques more than statistical methods
- Papers that are based on real world scenarios
- Papers that are implemented in the world

2) Exclusion criteria :

- Papers that are using stram processing
- Papers that are not using geo-locations
- Papers that are only focusing on either demand or supply
- Papers which are not predicting live demand and supply
- Papers which are not implemented well

III. RESULTS

Through this process, we identified 21 articles by this search process, and they are 21 unique studies. Based on the research articles were identified, prospective significant results for a systematic review on forecasting taxi supply and demand prediction and mobility behavior. These studies suggest a variety of approaches and techniques have been proposed for predicting taxi demand, including machine learning algorithms, neural networks, spatiotemporal analysis and real time population data from cellular networks. Advanced data analytics techniques are used to predict taxi passenger demand, which can help transportation providers optimize their services and reduce wait times for passengers. An example of the techniques are machine learning algorithms such as decision trees, neural networks, and support vector machines. As well as the use of roving sensors to collect data on passenger demand in real-time. Geospatial data is also frequently used to predict demand, which involves analyzing factors such as the time of day, weather conditions, and the locations of popular destinations. Author and Wu [2020], Author [2018], Author [2017a], Author and Grimshaw [2021], Author [2022b], Author [2022a], Author [2022a], Author and P.Viveka [2017]

Cluster analysis is employed to identify different types of users based in their mobility behavior Author and Gebhard [2019] based on their transportation preferences, frequency of travel, and preferred modes

Real-time Taxi Demand and Supply prediction based on Specific Geo locations using Machine Learning - A Systematic Literature Review

of transportation. Tailoring transportation services to different types of users, and better understand the factors that influence travel behavior. Mobility interventions can be tailored to different types of travelers based on their unique and intermodal mobility behavior, as identified by user typologies developed through cluster analysis. To optimize transportation services, wireless sensors and networks Author and Turgut [2017], Author and Hajela [2021], Author and Kawasaki [2018] are used to monitor traffic flow. For example, one paper uses wireless sensor networks to monitor traffic flow and optimize signal timing, while another paper uses network analysis to optimize routes for air taxis. Additionally, several proposed systems are based on real-time prediction models for taxi demand using recurrent neural networks (RNN) and hybrid model that combines linear regression with k-nearest neighbor (KNN) algorithms.

Big data Author and Prakash [2019] and quantitative histograms are utilized to predict taxi passenger demand. The research paper uses a novel approach to predict taxi passenger demand, by analyzing large volumes of data and creating quantitative histograms to visualize the patterns of demand over time. This can provide transportation providers with a better understanding of how demand for their services fluctuates throughout the day, and help them optimize their services accordingly. Data mining techniques and machine learning algorithms Author [2018], Author and P.Viveka [2017], Author [2020] used to predict taxi demand, which helps transportation providers optimize their services and reduce wait times for passengers.

Moreover, the modeling of air taxi demand using machine learning algorithm Author and Grimshaw [2021] such as support vector regression, enables air taxi providers to optimize their services and reduce wait times for passengers. This is the main objective of this systematic literature review. That find a solution to reduce the waiting time of passengers and also drivers by addressing the supply and demand imbalance in the taxi service industry.

IV. CONCLUSION

This section describes the related research approaches to this research topic and the gaps in general. Taxi demand prediction is productive for both drivers and passengers by reducing passengers waiting time and drivers' idle time. This review suggests that more researchers do research based on Machine Learning. Most findings used Machine learning algorithms, such as linear regression, decision tree, random forest, Artificial Neural Networks (ANN) and Knearest neighbors (KNN) for the prediction task. Real-time population was the primary parameter for prediction in most studies. As discussed through research questions the main focus of these studies is to

improve the efficiency of taxi services and reduce wait times for customers.

In addition, going through the review process, a research gap has been identified when comparing the studies. The study Author and Kawasaki [2018] suggests leveraging realtime population generated from cellular networks for taxi demand forecast, but there is not much information on how to use this data effectively. Future research could concentrate on creating more complex models that can precisely forecast taxi demand based on current population statistics. The paper Author [2020] suggests a combination forecasting model for estimating taxi demand in hotspots. Lack of comparison between various forecasting models. There aren't many comparisons between these model and other ones that are already in use, though. Future research could evaluate the performance of various forecasting models and highlight the advantages and disadvantages of each strategy. Regional variability is only partially taken into account in the research Author and F.Wang [2021], which suggests a multi-level deep learning model for forecasting taxi demand. On how to successfully include regional variability into models for predicting taxi demand, there is, however, limited study. Future research could concentrate on creating more complex models that can precisely forecast taxi demand based on data related to geographical variation.

By the results of this study, future work will be implementing a model to predict live demand and supply based on specific geo-locations for taxis with a good accurate level than these discussed studies' findings. In addition, it will be composed using the best algorithms that can be used to accurate prediction. Overall, taxi demand prediction is a crucial step that is helpful to gain profit for the taxi industry and by doing prediction, passengers receive a better service.

REFERENCES

- L. Xiang C. Wu Author, I J. Yan and H. Wu. City-scale taxi demand prediction using multisource urban geospatial data. 2020.
- T. He S. Lin S. Munir J. Stankovic Author, . Zhang. Taxipassenger-demand modeling based on big data from a roving sensor network. 2017a.
- C. Morency Author, A. Mehdizadeh Dastjerdi. Bike-sharing demand prediction at community level under covid-19 using deep learning. 2022a.
- M. A. Perez-Juarez Author, J. M. Aguiar-Perez. An insight of deep learning based demand forecasting in smart grids. 2023.
- F.Zhu Y.Lv-P.Ye Author, C.Zhang and F.Wang. Taxi demand prediction based on multi level deep learning and regional heterogeneity analysis. 2021.

Real-time Taxi Demand and Supply prediction based on Specific Geo locations using Machine Learning - A Systematic Literature Review

- G. Zander Author. Predicting taxi passenger demand using artificial neural networks". 2017b.
- R. Rahmatizadeh L. Boloni Author, J. Xu and D. Turgut.
Real-time prediction of taxi demand using recurrent neural networks. 2017.
- S.Srinivas Author, S. Rajendran and T. Grimshaw. Predicting demand for air taxi urban aviation services using machine learning algorithms. 2021.
- A. Batool M. Ramavath Author, A. V. Ramana and P.Viveka. Taxi demand prediction using ml. 2017.
- Y.Fukazawa Author, S.Ishiguro and S. Kawasaki. Taxi demand forecast using real-time population generated from cellular networks. 2018.
- A. Carboni Author, M. Pirra and F. Deflorio. Monitoring urban accessibility for freight delivery services from vehicle traces and network modelling. 2018.
- H. Zhang L. Wang Z. Ning Author, X. Wang. A demandsupply oriented taxi recommendation system for vehicular social networks. 2018.
- J. Han Y. Mei Y. Li Author, M. Liu. Dynamic balance between demand and supply of urban taxis over trajectories. 2021.
- H.Chen Y.Li and Q.Zhang Author, Z.Liu. Taxi demand prediction based on a combination forecasting model in hotspots. 2020.
- A.Bharathi Author and S.Surya Prakash. An approach to predict taxi-passenger demand using quantitative histogram on uber data. 2019.
- S. Nieland Author, R. Oostendorp and L. Gebhard. Developing a user typology considering uni- and intermodal mobility behavior. 2019.
- J. Kuchu Author. Spatial temporal pattern analysis on ondemand cab services in urban areas using gis. 2022b.
- P. Pratyush Author, A.Gangrad and G. Hajela. Taxi-demand forecasting using dynamic spatiotemporal analysis. 2021.