Dynamic pricing in Ride-Hailing intelligent transportation systems by using Deep Reinforcement Learning

Sajad Heydari, Elham Akhondzadeh

Abstract

Today, due to the advancement of technology and the development of cities and road networks in different countries, the issue of transportation has become an issue in the world and has received the attention of researchers. This problem has terrible impacts on the social and economic situation of countries. In this way, many countries are facing numerous crises because of the increase in the volume of passengers, the volume of traffic, environmental issues, and air pollution. For this reason, the importance of the intelligence of the transportation industry is felt. On the other hand, with the modernization of today's societies and the increasing use of the Internet, people have been led to use these tools, and the potential for intelligent transportation control has been created. These organizations have designed systems in two ways to help achieve the goals of intelligent transportation. These two include the use of Ride-Sharing and Ride-Hailing systems.

In recent years, a vital matter in Ride-Hailing systems is dynamic travel pricing, which is crucial for modern transportation services because it might cause in balancing supply and demand compared to the past that due to unstable environments, there was no suitable pricing policy. To handle this problem, researchers have proposed a dynamic pricing system aimed to minimize waiting time for passengers or maximized revenue. For this purpose, the new aspects and innovations that exist in this research are a new perspective on the problem of dynamic pricing to minimize waiting time and maximize earnings simultaneously by using reinforcement learning and deep reinforcement learning methods. in this research, the data of Uber Company's intelligent transportation service trips have been used, which includes the trips made with all types of vehicles available in that company in 2018. After collecting the data, they have been carefully scrutinized and analyzed according to the methods and techniques of reinforcement learning and deep reinforcement learning. For this purpose, with the two goals of maximizing income and minimizing waiting time, various methods have been used, among which the best result belongs to the Deep SARSA Network algorithm. To validate and rank the techniques used in the research, MARCOS multi-criteria decision-making methods have been utilized to obtain the best and most reliable results.

Keywords: *Dynamic pricing, Revenue management, Waiting Time, Reinforcement learning, Deep Reinforcement learning.*

1. Introduction

In recent years, with the advancement of technology and the increasing access to and use of cutting-edge smartphones by individuals, organizations in various industries have emerged and grown. Since transportation has become a vital issue in the world, it has received the attention of researchers. Whereas cities and road networks are improved, the problem has had a tremendous impact on the economic situations of different countries (Olayode et al., 2020). Ride-Hailing intelligent transportation systems serve a great many purposes. One of the most essential of them is dynamic pricing. Ride-Hailing owners believe pricing according to stochastic conditions helps them manage their revenues and handle unforeseen circumstances. Besides, revenue management can boost drivers' level of satisfaction, and they will be more motivated to accept trips. Thus, a vital matter in Ride-Hailing systems is dynamic travel pricing, which is crucial for modern transportation services because it might cause in balancing supply and demand compared to the past that due to unstable environments, there was no suitable pricing policy. To handle this problem, researchers have proposed a dynamic pricing system aimed to minimize waiting time for passengers or maximized revenue. In this research, firstly, a definition of the research subject is going to discuss. Secondly, we will state its importance and necessity. Recently, researchers have studied the coordination of requests in Ride-Hailing transportation systems and believe this is one of the most vital operational concerns in ride-hailing services. They think achieving a stable interaction requires special attention to the structure of supply, and demand (Qin et al., 2021). Another issue investigated in providing services in Ride-Hailing intelligent transportation systems is dynamic travel pricing. In the past, due to unstable environments, there was no suitable pricing policy (Rana and Oliveira, 2014). For this reason, researchers have presented different methods and ideas to solve this problem. For instance, the unbalanced distribution of supply and demand is controlled, and the prices are optimized using different approaches (Lei et al., 2019). Various Reinforcement Learning methods are widely used to make a balance between supply and demand. To handle it, researchers have proposed a dynamic pricing system aimed to minimize waiting time for passengers. Correspondingly, our effort will be examining research objectives and innovations. In addition, the other goals such as minimizing the idle time for drivers (O'Keeffe et al., 2021), minimizing waiting time (Haliem et al., 2021), minimizing travel time and distance (Al-Abbasi et al., 2019) and profit maximization (Pandey et al., 2019) employed by many researchers. It should be noted that because of the increasing volume of data and its complexity, neural network and deep learning methods (Feng et al., 2021) or a combination with reinforcement learning (Song et al., 2020) might be beneficial. Based on the main goal, the sub-goals of this research can be expressed in the form of the following:

- Analyzing and investigating factors affecting the price of travel to check the balance between supply and demand
- Estimating the best price dynamically according to a different time, places, and other characteristics
- Developing models to minimize waiting time and Maximizing Revenue.

According to the extensive and diverse research conducted by researchers in recent years, an attempt will be made to respect the innovation and new aspects that research should have by referring to them. Newer ideas and methods can verify this research. Any research, according to its goals, will need this aspect of being new. For this reason, it can be said that the new aspects and innovations that exist in this research are a new perspective on the problem of dynamic pricing to minimize waiting time and maximize earnings, as well as the balance between supply and demand using reinforcement learning and deep reinforcement learning methods. A set of methods should be examined and compared. Because the research in this field is often done with simple reinforcement learning methods and mathematical modeling, no significant research has been done. These methods cannot be acceptable in the long term because of the increase of the obtained data, as well as the greater complexity of the issues. There is a need for a method or methods to compensate for the gaps in these methods and recognize the prices of each trip dynamically with better accuracy and performance. Another approach that will cause innovation in this research is the use of other effective features in this category of problems. These features include consideration of weather conditions and specific conditions of different days, which have not been considered in previous research. In excess, the general structure of this research is as follows: in the second part related research will be discussed. Then the proposed method will be examined. After that, the results will be evaluated, and finally, conclusions and recommendations for future research will be presented.

2. Definitions and Related work

- **Intelligent Transportation Systems**

  In the not-so-distant past, there were established institutions that provided taxi services in a specific way. So that customers could use them by making a call and stating their destination. This service continued for a long time and it continues more or less now. But the important point is that this type of service has problems such as security, detection of the best route and traffic flow, variable costs and passive support (Devi and Neetha, 2017). Therefore, with the passage of time and the advancement of technology, and on the other hand, with today's comprehensive modernization and the use of the Internet (Rachbini et al., 2020), organizations were established to provide the aforementioned services in an intelligent manner. These services are offered under the title of Ride-Hailing intelligent transportation systems. Theoretically, Ride-Hailing systems are systems that can integrate the flexibility and speed of cars and reduce customer payment costs (Akbari et al., 2020). In these systems, the main goal is to obtain customer satisfaction (Habibah Arshad et al., 2008).

- **Pricing**

  Pricing is the process by which a business determines the price at which to sell its products and services, and may be part of the business's marketing plan. In general, in determining prices, the business considers the price at which it can purchase the product, manufacturing cost, market, competition, market conditions, brand and product quality (Smith, 2011). According to the literature, pricing has different types, which are generally divided into two categories: static and dynamic (Nagle et al., 2016).

- **Revenue Management**

  Revenue management is the application of systematic analysis that predicts consumer behavior at micro levels, optimizes product availability, and leverages price elasticity to maximize revenue growth and thus profits. The main goal of revenue management is to sell the right product to the right customer at the right time at the right price and with the right quality (Cross, 1997).

- **Machine Learning**

  Machine learning is a field of study that gives the machine the ability to learn and experience without explicit programming (Géron, 2019). In machine learning, how to learn and improve the performance of systems based on data is investigated. Also, the main research field in this field is automatic learning and identifying complex patterns in order to build a system that can make intelligent decisions based on data (Han et al., 2012). There are different types of machine learning systems. These systems include supervised, unsupervised and semi-supervised learning systems.

- **Reinforcement Learning**

  Reinforcement learning is different from the above. In this case, the learning system is called an agent that can observe the environment in a certain state, perform an action, and receive a reward or punishment based on the action performed. The learning system learns itself what is the best strategy to get the best reward. Strategy specifies the action that the agent should perform in certain situations (Géron, 2019). During this process, the agent's goal is to maximize the sum of the rewards he receives from performing actions in given situations. This means that the agent wants to maximize not only the immediate rewards, but also the cumulative rewards he receives over time. In other words, reinforcement learning evaluates performed actions instead of training with correct actions (Sutton and Barto, 2018).

- **Deep Learning**

  Deep learning is a class of machine learning algorithms that use multiple layers to progressively extract higher-level features from the raw input (Deng and Yu, 2014). In deep learning, each layer learns to transform its input data into a more abstract and composite representation. Deep learning can be used for the purpose of classification and prediction as well as clustering. Also, this group of algorithms can be combined with reinforcement learning methods and solve more complicated problems.

- **Deep Reinforcement Learning**

  Deep reinforcement learning methods perform parallel and independent computational processes. In fact, each agent acts on a different part of the environment with a different set of parameters. The updates of each agent are received by a global network and combined asynchronously and asynchronously to achieve a global policy (Farazi et al., 2021).

- **Research Background**

  i. **Intelligent Transportation Systems**

  To achieve the profitability in Ride-Hailing systems and intelligent transportation in general, as well as to achieve consistency in the competitive market and reduce operational costs, predictive methods based on machine learning have been used (Akyouz et al., 2020). Also, to optimize the performance of drivers in these systems and to monitor whether to choose or reject passenger requests, which may include additional costs for the service organization of these systems, decision-based machine learning methods such as decision trees are used. In this type of objective, the distance is the most important parameter for drivers to decide whether to approve or reject the request (Do et al., 2019). A vital point is that the performance of the mentioned methods will decrease over time as the amount of data increases, and it will make the success of developing ride-hailing intelligent transportation systems unclear. Because the goal of these systems is to use communication technologies to combine transportation systems based on people, cars, and roads, and this goal must be strictly followed. For this reason, more advanced methods are needed.

  ii. **Dynamic Pricing**

  The problem of pricing and capacity management has been modeled using the nonlinear stochastic integer programming technique and solved using a simulation-based method called Simulated Annealing. Due to the dynamic nature of this issue, decisions about prices and appropriate capacity allocation have been made dynamically (Kamandanipour et al., 2020). The time sequence is very important in intelligent transportation systems. For this purpose, researchers have tried using deep learning models to improve pricing dynamically by considering time sequences. Some have solved the problem by using data mining and machine learning techniques and methods such as Random Forest and XGBoost. In another study, researchers examined the issue of pricing trips in Ride-Hailing systems dynamically from the drivers' point of view. They believe that for a better distribution of drivers in smart transportation networks, various things such as the queuing system, their level of well-being and income, the average travel time, and the possibility of them leaving the system should be taken into consideration. Of course, there is another important factor that affects both customers and drivers. This is the invoice price. The right price should be considered once from the drivers' point of view, and once from the customers' point of view. Because it is considered essential for data analysis and reaching quality results (Mahshar et al., 2020). It will lead to accepting a price appropriate to different conditions (Wu et al., 2021).

  iii. **Dynamic Pricing and Revenue Management**

  Revenue management is the most important action of a business to stay in the competitive market. So it should be noted that revenue management depends on accurate forecasts. In order to reduce individual costs and minimize the distance a passenger is picked up by the driver of ride-hailing systems, researchers have taken advantage of discount policies and strategies. In such a way that every driver who travels a shorter distance to reach the passenger and pick him up will be included in the travel expenses and also from the passenger's point of view, a discount will be applied to the travel price. In this way, in the long-term horizon, it is possible to help maximize the revenue of the service providers of this type of service (Duan. et al., 2019). In recent years, in general, automated systems have played an important role in industries such as manufacturing and especially transportation. One of the most important effects that automatic systems can have in the intelligent transportation industry is tracking the route and movement completely automatically without any driver intervention (Hillebrand et al., 2020).

  iv. **Machine Learning in Dynamic Pricing**

  Several methods are used to optimize the service quality level of Ride-Hailing services. For this purpose, researchers have used a method based on multi-agent simulation (Bischoff et al., 2018). But these methods may not be efficient with the increasing complexity of data as well as the rapid development of communication and technology. For this reason, in order to increase the level of service quality in Ride-Hailing systems, the sensitive behaviors of passengers and drivers for trips are first investigated and then estimated using different classification and behavior prediction techniques (Tian et al., 2020). All these methods will be fruitful for service providers to explain the features of affordability, reliability and responsibility with their available technologies (Brown and LaValle, 2021). The concept of pricing can be applied in any industry. Rail transport service industry is one of the examples where proper pricing is used. In other words, for this group of organizations, ticket pricing and capacity management are the most important profit-making tools that should be paid attention to. For this reason, researchers have provided an advanced decision-making tool for the aforementioned problems. By considering the demand under conditions of uncertainty, they have modeled the pricing and capacity management problem using the nonlinear random integer programming technique and solved it using a simulation-based method called refrigeration simulation. Due to the dynamic nature of this issue, decisions about prices and appropriate capacity allocation have been made dynamically (Kamandanipour et al., 2020). In order to control and manage traffic, appropriate pricing can be a good measure. Pricing can increase service levels in transportation networks and reduce travel delays. Considering this issue, researchers have used a dynamic optimization system and a dynamic router for linear modeling. First, they used the Dijkstra algorithm to achieve the goal, and then they trained multilayer neural networks for optimal dynamic pricing. They finally arrived at a solution that recognizes the most complex behavior of customers and predicts the appropriate price according to that and manages traffic as a result (Genser and Kouvelas, 2022). For forecasting, regression methods are also used according to characteristics such as weather, traffic conditions, and the length and time of travel (Schwieterman and Smith, 2018). Time sequence is very important in intelligent transportation systems. For this purpose, researchers have tried using deep learning models to improve pricing dynamically by considering time sequence. To achieve the goal, they have predicted the appropriate prices by analyzing and investigating customer behavior and using algorithms such as long-short-term memory and convolutional neural network and evaluated the results with the method of mutual validation (Liang and Cai, 2022).

v. Reinforcement and Deep Reinforcement Learning in Dynamic Pricing

There are advanced methods for spatio-temporal dynamic pricing. Among these methods, we can point out the use of deep reinforcement learning methods such as approximate policy optimization, which either considers forward neural networks to examine the current situation and implement the optimal price strategy (Chen et al., 2021).

In recent years, researchers in the field of intelligent transportation systems and in particular matching the passenger with the driver to make a trip have completed numerous researches. For example, some researchers using convolutional neural networks have performed a proper classification of different samples that have not been seen before by the algorithm, and by using the concept of one-way learning and combining it with passenger clustering, they have performed their dynamic matching (Liu et al., 2019).

Moving people and goods in a safe, efficient and sustainable manner involves a wide range of decision-making tasks. Deep reinforcement learning, by integrating the power of deep learning and reinforcement learning, provides a general and flexible framework for sequential decision making that is applicable to many transportation operations and planning problems (Farazi et al., 2021). One of the methods that can be very helpful is to use the combination of deep learning and reinforcement learning, which is used under the title of deep reinforcement learning to solve existing challenges. In other words, deep learning and reinforcement learning methods are part of the necessary factors to achieve artificial intelligence systems at the human or superhuman level. So, their combination can be very fruitful (Matsuo et al., 2022).

The success of deep reinforcement learning could be an imminent entry into the industrial world. However, like many machine learning algorithms, reinforcement learning algorithms suffer from a lack of explainability. This shortcoming can be very problematic because many applications related to these methods require a model that can explain their decisions and actions to users as a condition of their full acceptance by the community. Furthermore, deep reinforcement learning models are complex for developers to debug because they rely on many factors: the environment (especially the design of the reward function), the encoding of observations, large deep learning models, and the algorithm used to train the policy. Therefore, an explainable model can help solve problems faster and greatly accelerate new development in reinforcement learning methods (Heuillet et al., 2021).

vi. Other methods in Dynamic Pricing

According to these conditions, prices can be estimated with different regression methods and optimized according to factors such as time of day and day of the week (Zhang et al., 2020). In a different research, game theory concepts have been used to achieve the goal of dynamic pricing. Thus, based on rational strategic consumers, a dynamic game is constructed to construct a two-period dynamic pricing model for two alternative brands. The concept of dynamic game solution is Nash equilibrium. In this model, according to the consumer and his choice, according to the expected value of the products, which are clearly divided into several consumption classes. Two competing firms enter into a pricing game and eventually reach a Nash equilibrium (Bi et al., 2014). On the other hand, the same problem has been done by using a bipolar game theory model in the form of dynamic pricing (Bi et al., 2014).

In another research, using a multi-period game theory model, the two objectives of "dynamic pricing" and "car allocation" in Ride-Hailing smart transportation systems have been investigated. In this issue, dynamic planning is used to advance the decision-making process in order to reach the two mentioned goals (Lei et al., 2019). The proliferation of ride-sharing systems is a major driver in the advancement of self-driving and electric vehicle technologies. To this end, the researchers consider the pricing problem posed by a profit-maximizing transportation service provider operating a fleet of autonomous electric vehicles. First, they have considered a static policy for this purpose so that they can get the longest waiting time for passengers and the lowest amount of profit for drivers. After that, they have modeled the problem by using different dynamic programming methods and solved it with multiple deep reinforcement learning methods to reach a near-optimal policy. The obtained results will cause a stability as a result of using these systems in the service delivery structure (Turan et al., 2020). In a similar research, optimization methods have been used to solve the traffic timing problems of self-driving cars and optimize it. Its results have been used for the purpose of dynamic pricing and investigating passenger behavior to make revenue management decisions (Yang et al., 2021).

The research conducted in the field of intelligent transportation systems will be summarized in Table 1.

Table1. Classification of research done in the field of intelligent transportation systems.

Index	Source	Techniques used	The field of study in intelligent transportation systems
1	Do et al., (2019)	Classification and prediction	Prediction of drivers' behaviour in systems Ride-Hailing
2	Akbari et al., (2020)	Statistical Methods	Analysis of the relationship between different factors for customer satisfaction and retention
3	Zhankaziev et al., (2020)	Statistical Methods	Examining the relationship between security and customer satisfaction
4	Akyouz et al., (2020)	Classification and prediction	Analyzing the relationship between customer satisfaction and performance costs
5	Olayode et al, (2020)	Classification and prediction	Traffic analysis and control in road transport
6	Lan, (2020)	Classification and prediction	Examining customer demand in the systems Ride-Hailing
7	Kaffash et al., (2020)	Classification and prediction	Travel planning and traffic flow forecasting
8	Kuswanto et al., (2019)	Optimization Methods	Analysis of important factors on customer loyalty in the systems Ride-Hailing
9	Gumilar et al., (2019)	Game Theory Methods	Examining the relationship between customer satisfaction and loyalty
10	Mao et al., (2021)	Statistical Methods	Analysis of the impact of different factors on customer satisfaction in systems Ride-Hailing
11	Mahshar et al., (2020)	Statistical Methods	Analysis of the impact of different factors on customer satisfaction in systems Ride-Hailing

The new aspects and innovation that exist in this research is a novel perspective on the problem of dynamic pricing to minimize waiting time and maximize profit, as well as the balance between supply and demand using various methods, which should be examined and compared. For instance, there are several methods to evaluate the efficiency of ride-hailing intelligent transportation systems, especially when introducing a new feature in service organizations. One of these methods is formulating related issues in different scenarios by considering aspects such as comfort, speed and cost of travel. After that, according to the available data, regression and classification decision trees are used to advance the goals and analysis (de Sa, 2021).

In real-world problems, it is sometimes possible that the data set examined by methods related to machine learning in a project may encounter a concept called the curse of dimensions. In this situation, to reduce the impact of this concept on the data set and obtain better results, techniques and methods related to deep learning are combined with the machine learning algorithm or algorithms used in it. So, it can be said that to overcome this problem, we can use the combination of reinforcement learning and deep learning, which is called deep reinforcement learning (Farazi et al., 2021). Researchers also suggested to use the combination of convolutional neural network or recurrent neural network methods with reinforcement learning methods to solve pricing issues or challenges related to rail transportation (Shalev-Schwartz et al., 2016).

Xu and his colleagues in 2022 have used statistical error evaluation methods such as MAE, MSE, MAPE and RMSE to check the algorithms used for pricing and ranking them so that they can find the best method for pricing in their problem. In a specific study, Chan et al. presented methods in 2020. They believe that unlike other machine learning and deep learning algorithms that have multiple performance evaluation methods, there are limited methods for reinforcement learning and deep reinforcement learning that can be used depending on the problem conditions. These methods can be divided into the following categories (Chan et al., 2020):

1. Risk-based methods
2. Scattering based methods

In different research in the energy industry, May and Huang in 2023 have investigated the effect of reinforcement learning algorithms on energy consumption based on a fixed price and a dynamic price. They showed that dynamic pricing will be effective in the first degree and the use of reinforcement learning to minimize energy consumption in the second degree.

The MARCOS method is one of the new multi-criteria decision-making methods presented by Stivic et al. (2019). This method is used to rank research options.

According to the extensive and diverse researches that have been conducted by researchers in recent years in the fields studied in this research and according to the review of the literature, the summary of the categories of the studied researches is as follows:

1. Pricing with reinforcement learning and deep reinforcement learning
2. Pricing with reinforcement learning and deep reinforcement learning in intelligent transportation systems
3. Pricing with Reinforcement Learning and Deep Reinforcement Learning in Intelligent Ride-Hailing Transportation Systems

According to the above classification, the research that was done in the field of pricing with the mentioned techniques was mainly in the fields of energy and electricity supply, in supply chains, in the railway industry and in the airline industry. After studying the literature in these fields and familiarizing with other applications, this issue was investigated in intelligent transportation systems. In intelligent transportation systems, mostly articles focused on self-driving cars or traffic control, and very few researches focused on internet taxi systems. For this reason, pricing with reinforcement learning and deep reinforcement learning in Ride-Hailing intelligent transportation systems seemed to be a suitable case as a very important research and a research topic. On the other hand, by reviewing the articles and few researches in this field in Ride-Hailing systems, there are two points of interest:

1. In the research, it was proposed to consider different weather conditions as a proposal for the mentioned problem.
2. In researches, pricing was stated either with the aim of maximizing income or with the aim of minimizing waiting time.

For this reason, as a research and innovation conversation in this research, while considering different weather conditions for dynamic pricing, the goals of maximizing income and minimizing waiting time have been taken into consideration. Also, the complexity of the issues is the need for a method or methods that can compensate for the gaps in these methods and dynamically recognize the prices of each trip with better accuracy and performance. For this reason, reinforcement learning and deep reinforcement learning techniques and methods will be used to achieve the goals of the research. On the other hand, in the research conducted in this field, none of the multi-criteria decision-making methods were used. For this reason, these methods have been used to check the methods and validate the results in this research.

3. Research Method

The primary purpose of the research is to employ Reinforcement Learning Deep Learning techniques for dynamic pricing in an intelligent Ride-Hailing transportation system to maximize revenues. We investigate and find the best price according to the changing conditions in a way that the expected utility can calculate. In other words, our aims are as follows:

- Analysing and investigating factors affecting the price of a trip to check the equilibrium between supply and demand.
- Estimating the most reasonable price dynamically according to a different time, location, and other characteristics using Deep Learning and Reinforcement Learning to prove that other methods are inefficient.

Generally, to achieve research purposes, some questions must be answered appropriately during the research that includes the following:

- What aspects should consider for dynamic pricing in Ride-Hailing systems?
- What RL and DRL-based models perform better for dynamic pricing in ride-hailing systems by simultaneously optimizing waiting time and revenue?
- How can we boost the probability of trip acceptance from drivers, enhance their satisfaction level, and create loyalty?

To achieve our goals, we utilize the CRISP-DM method for the project. The Cross Industry Standard Process for Data Mining (CRISP-DM) is a process model that serves as the base for a data science process. It has six sequential phases:
Business understanding – What does the business need?
Data understanding – What data do we need? Is it clean?
Data preparation – How do we organize the data for modeling?
Modeling – What modeling techniques should we apply?
Evaluation – Which model best meets the business objectives?
Deployment – How do stakeholders access the results?

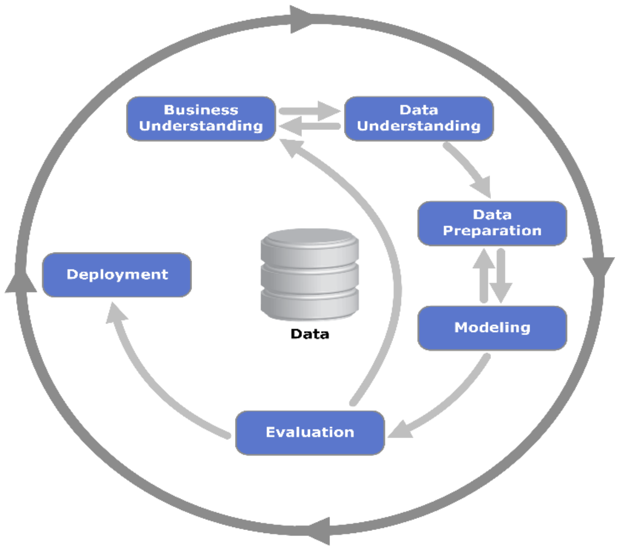

Figure1. CRISP-DM Methodology

In the First step, it is better to first get a brief understanding of the data and how it is collected and its organizational affiliation (Business and Data Understanding). The data used in this research was obtained from the website of the Uber intelligent transportation service company (www.uber.com). This dataset contains 50 features and 360,000 trips sent by different types of Uber vehicles from an origin to a destination in the city of Boston, USA. Also, this data set belongs to travel data in 2018 and in the last two months of the year (November and December), which has been provided.
In general, the features of the aforementioned data set can be divided into two parts: features related to travel information and features related to weather, which are summarized in the following two sectors:

1. Related to Trip: Number of each trip, Travel time, initial point of travel, travel destination, The type of service (car) chosen by the passenger, and length of trip
2. Related to weather: Air temperature, Apparent temperature, Precipitation intensity, Chance of rain, humidity, wind speed, wind time, intensity of vision, air pressure, The amount of cloud cover, sunrise time, sunset time, Moonrise rate

In the Data preparation step, we are going to utilize various techniques. Mainly, in every data project, important things should be paid attention to such as Data Transformation, Data Cleansing, and Feature Selection. Examining these cases will make the project get favorable results in the end.

One of the most important topics in data preparation and pre-processing is the selection of appropriate features for problem solving. For this purpose, there are different methods. These methods vary according to the type of algorithm that is chosen. Because reinforcement learning methods and techniques have been used in this research, the feature selection method *Principal Feature Analysis (PFA)* was chosen, and finally the selected features are according to the table below. There are different steps to implement this method. These steps are divided into 5 categories. In the first step, the covariance matrix of the sample should be calculated. Of course, it should be noted that in some cases the use of a correlation matrix is preferred instead of the covariance matrix. In the second step, the calculation of the main components and eigenvalues of the covariance/correlation matrix should be considered. In the next step, the subspace q must be selected and the matrix A_q can be made from A. This can be chosen by deciding how much data variability to preserve. The conserved variable is the ratio between the sum of the first eigenvalues q and all eigenvalues. In the fourth step, the obtained vectors are clustered using the K-Means algorithm so that the scores obtained by the PCA algorithm can be concluded by clustering. In the last step, for each cluster, the corresponding V_i vector should be found, which is closest to the cluster mean. The corresponding feature, X_i, is selected as the main feature. This step represents the selection of p features. The reason for choosing the closest vector to the average of the two cases. This feature can be considered the central feature of that cluster—a large "spread" in lower dimensional space, and a good representation of the original data.
Furthermore, For the analysis of attributes in our Dataset (as a part of EDA), we decided to use Correlation for this purpose. The results illustrated that the correlation between main numerical variables is good and they are appropriate for the problem. Additionally, for categorical-based values we employed spearman method and proved the importance of them.

In accordance with the various research conducted by researchers in recent years in the fields related to our research, most of the methods used to solve the problem of dynamic pricing in Ride-Hailing intelligent transportation systems have been statistical techniques or mathematical optimization methods. Furthermore, techniques related to game theory have been used in different research. As a research and innovation chat, in this research, while considering different weather conditions for dynamic pricing, the goals of maximizing income and minimizing waiting time have been taken into consideration. Additionally, the complexity of the issues is the need for a method or some methods that can compensate for the gaps in these methods and dynamically recognize the prices of each trip with better accuracy and performance. For this reason, reinforcement learning and deep reinforcement learning techniques will be used to achieve the goals of the research (Modelling Step in CRISP).
In contemporary years, due to the great importance of dynamic pricing as well as the growing importance of proper service of intelligent transportation systems, more modern and advanced approaches have been utilized to obtain the satisfaction of customers and drivers. Those are called Reinforcement Learning.
In some articles in the literature, those are also combined with Neural Networks and Deep Learning concepts and have created Deep Reinforcement Learning. They have been operated to boost the quality of service.

Table2. Reinforcement Learning and Deep Reinforcement Learning methods used in the literature.

Index	Relevant Topic	Year	Method	Algorithm	Usage Count
1	No	2021	DRL	SQDDPG	1
2	No	2021	DRL	IDDPG	1
3	No	2021	DRL	MADDPG	1
4	Yes	2021	DRL	DDPG	2
5	No	2022	RL	MDP	1
6	Yes	2021	DRL	DQN	6
7	Yes	2021	DRL	DDQN	4
8	No	2021	DRL	HDQN	1
9	Yes	2021	RL	A2C	4
10	Yes	2021	RL	A3C	3
11	Yes	2020	RL	PPO	7
12	No	2018	DRL	TPRO	2
13	Yes	2019	DRL	DDPG	8
14	Yes	2018	DRL	DQN	18
15	Yes	2021	RL	QL	2
16	Yes	2012	RL	SARSA	2
17	Yes	2019	RL	POMDP	1
18	Yes	2022	DRL	SAC	1
19	No	2016	RL	MDP	29

In the problem of dynamic pricing, RL and DRL methods employed are negligible which is shown in **Table2**. For this reason, we decided to utilize these techniques to reach the goals of our research.
Finally, from all of algorithms mentioned in the table of literature, we choose SARSA, and PPO as a RL. In addition, the DRL methods are DSN, and DDPG.

3.1. SARSA

SARSA algorithm is an on-policy reinforcement learning algorithm. In this algorithm, the learning agent first observes the state of the system (S) and selects action (A) based on a specific policy. Next, after selecting the action, the environment determines the next state of the system and the reward. By observing the next state of the system and the received reward, the agent calculates and updates the value of the action-value function. This process will continue until the value of the action value function converges to its optimal value.
The implementation steps of this algorithm are as follows:
1. Determining the initial value of α and ε
2. Determining the initial values of $Q(S, a)$
3. Repeat the following steps for each learning episode:
Select an S status
Choosing action A in state S with a specific policy (e.g. ε-greedy concept)
Execute action A and view R and change state to S'

$$Q_\pi(s,a) \leftarrow Q_\pi(s,a) + \alpha\,[R - \gamma Q_\pi(s',a) - Q_\pi(s,a)]$$

As shown in the relation above, in the first step, the initial values of alpha and epsilon as well as the values of the matrix Q(S, a) should be determined. Then a state is chosen randomly and with the considered policy the action in this state is adopted and the reward (R) and the next value of the system state are received. At the end, using the observed values, the value of Q(s, a) will be updated.

3.2. PPO

Proximal Policy Optimization will simplify this by using an alternative cut target while maintaining the same performance. First, it is better to determine the probability ratio between the old and new policies as follows:

$$r(\theta) = \frac{\pi_\theta(a|s)}{\pi_{\theta_{old}}(a|s)}$$

Then the objective function will be as follows:

$$J(\theta) = E[r(\theta)A_{\theta_{old}}\widehat{(s,a)}]$$

When applying PPO to a network architecture with common parameters for both policy (performance) and value (critic) functions, in addition to the truncated reward, the objective function is augmented with an error term in the value estimation and a sufficient exploration number called entropy.

$$J(\theta)\_new = E[J(\theta) - c_1\left(V_\theta(s) - V_{target}\right) + c_2 H(s, \pi_\theta(0))]$$

3.3. DSN

As mentioned, SARSA is an on-policy reinforcement learning algorithm. In this algorithm, the learning agent first observes the state of the system (S) and selects the action (A) based on the specific policy and determines the environment of the next state of the system and the reward. Also, the way this process was carried out in mathematical language was also investigated. Now, this algorithm may not be used in some problems due to its inability to solve complex and high-volume problems. For this reason, to solve this problem, the combination of deep learning and this algorithm is used to update the Q value function.

$$Q_\pi(s,a) \leftarrow Q_\pi(s,a) + \alpha\,[R - \gamma Q_\pi(s',a) - Q_\pi(s,a)]$$

This method has several advantages over SARSA itself, including the fact that it can support a very large amount of data in each iteration of the experiment and avoid overfitting during data training.

3.4. DDPG

DDPG (Lillicrap et al., 2015) or Deep Deterministic Policy Gradient, is a model-free out-of-policy algorithm. As stated, DQN stabilizes Q-function learning by replaying the experience and frozen target network. DQN works in discrete space, and DDPG extends it to continuous space with an actor-critic framework while learning a deterministic policy.

$$\acute{\mu}(s) = \mu_\theta(s) + N$$

3.5. Q-Learning

This Algorithm is a model-free, value-based, off-policy algorithm that will find the best series of actions based on the agent's current state. The core of the algorithm is a Bellman equation as a simple value iteration update, using the weighted average of the current value and the new information.

3.6. DQN

Deep Q Networks (DQN) are neural networks that utilize deep Q learning in order to provide models such as the simulation of intelligent video game play. Rather than being a specific name for a specific neural network build, Deep Q Networks may be composed of convolutional neural networks and other structures that use specific methods to learn about various processes. In a general sense, deep Q networks train on inputs that represent active players in areas or other experienced samples and learn to match those data with desired outputs.

3.7. SAC

Soft Actor Critic (SAC) is an algorithm that optimizes a stochastic policy in an off-policy way, forming a bridge between stochastic policy optimization and DDPG-style approaches. It isn't a direct successor to TD3 (having been published roughly concurrently), but it incorporates the clipped double-Q trick, and due to the inherent stochasticity of the policy in SAC, it also winds up benefiting from something like target policy smoothing. SAC is an off-policy algorithm. An alternate version of SAC, which slightly changes the policy update rule, can be implemented to handle discrete action spaces. The Spinning Up implementation of SAC does not support parallelization.

Unlike other machine learning and deep learning algorithms that have multiple performance evaluation methods, there are limited methods for reinforcement learning and deep reinforcement learning that can be used depending on the problem conditions. These methods can be divided into the following categories (Chan et al., 2020):
1. Risk-based methods
2. Scattering-based methods

In this research, by reviewing the literature and considering the nature of the research problem, risk-based assessment methods are used. The reason for choosing these methods is that dynamic pricing in Ride-Hailing intelligent transportation systems is a changeable thing according to different conditions. In other words, this action takes place under conditions of uncertainty. For this reason, there is a risk of choosing a suitable price to obtain customer satisfaction and maximize revenue.

In this research, two risk-based assessment methods have been used. The first is the conditional value at risk. Conditional value at risk (CVaR) is an expected shortfall. CVaR is derived by taking the weighted average of the "lots" of losses at the tail of the distribution of possible returns beyond the value-at-risk (VaR) cutoff point. Conditional value at risk is used for effective risk management. In general, if the price is stable over time, the value at risk may be sufficient to manage risk. However, the lower the price stability, the greater the chance that VaR does not provide a complete picture of risks, as it is indifferent to anything beyond its threshold.

Conditional value at risk (CVaR) attempts to address the shortcomings of the VaR model. This method is a statistical technique used to measure the level of price risk in a company in a specific time frame.

Finally, the results of the research problem must respond to the objective function of the two objectives of the following mathematical model:

$$\max E_{predicted}(x) - E_{actual}(x)$$
$$\min E_{predicted}(x) - E_{actual}(x)$$

The limitations of this model are the changes in different weather conditions or travel-related characteristics, which can be achieved by considering them and solving the model.

Of course, it should be noted that the use of a creative method validated in the literature can be a seal of approval on the results obtained from the research. For this reason, in this research, one of the latest Marcus multi-criteria decision-making methods is used to validate the results of the algorithms and rank the options. Benefiting from this method will be done by forming a decision matrix of 7 criteria and 24 alternatives in order to rank the methods for the goals of maximizing income and minimizing waiting time.

4. Results

In conducting any research, special attention should be paid to the way of doing it, especially because this way is unique for each research.

In the preliminary analysis of this data set, it was clear that the highest amount of travel occurred during the times when the weather was cloudy. Also, the least amount of travel has been done during fog. In the comparison made between partly cloudy weather during the day and at night, it is clear that the willingness of drivers to travel and accept it is more at night. Also, the tendency to travel on rainy days is higher than on sunny days. It should be noted that sunny days have fewer trips than even nights without rain.

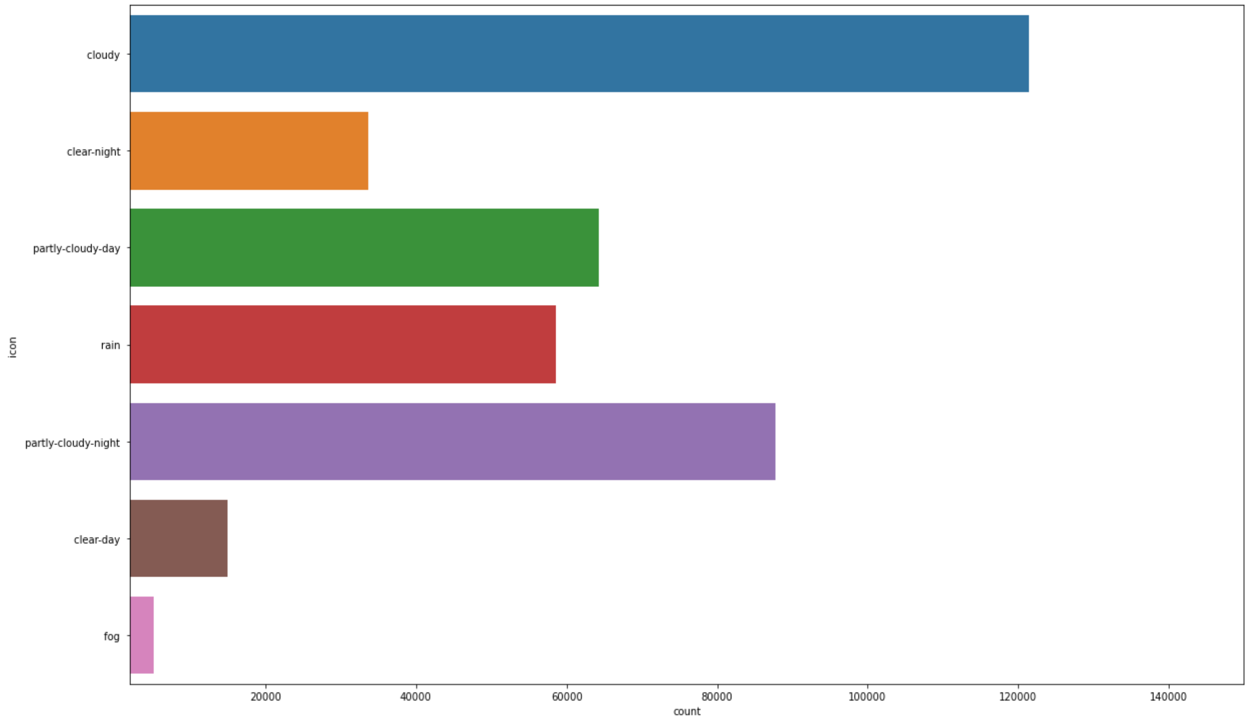

Figure2. Frequency of Uber trips in Boston, USA in different weather conditions

Another point that is important is that most of the trips made in the city of Boston under different conditions were short-term. In other words, and more precisely, it can be said that almost 62% of the trips made have travelled a short distance. After that, approximately 37% of the trips were regular distance trips and finally only approximately 2% of the trips were long distance.

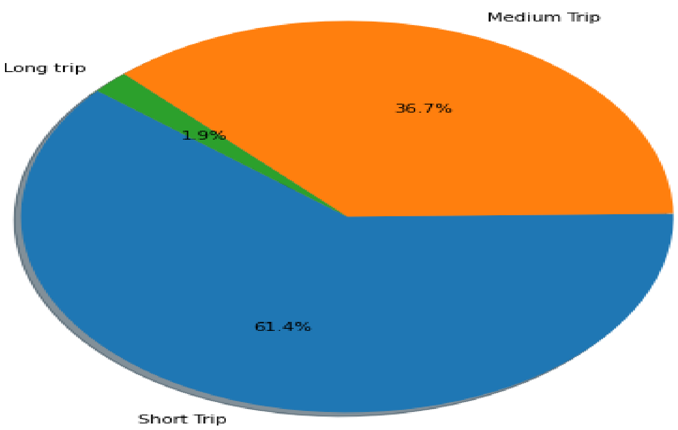

Figure3. Distribution of trips made in the city of Boston under different conditions

As we get closer to the last hours of a day, the waiting time for passengers to arrive from an origin to a steep destination decreases. This shows that as it gets dark, the supply is higher and the time each passenger has to wait for the driver to accept their request decreases. Also, this issue can affect the pricing method. It should be noted that peak hours on different days are also important. For example, the amount of waiting for passengers reaches its peak in the period from 16 to 18 when all administrative centres, universities and schools are closed. The reason is that the demand has increased sharply and drivers are sensitive to the price of trips. Of course, it should be noted that these cases are different from considering different weather conditions.

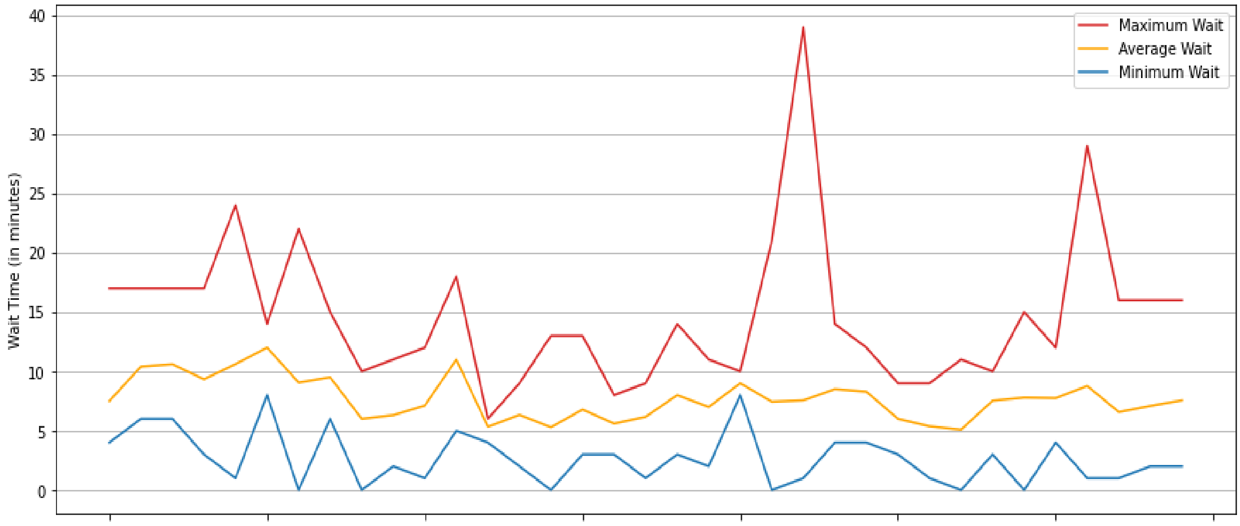

Figure4. Busy and demanding hours in the city of Boston to benefit from the services of the Uber transportation system

As mentioned, the data set of this research includes 50 features and 360,000 trips sent by different types of Uber vehicles from one origin to one destination in the city of Boston, USA. This dataset needs deep preparation to train the algorithm. According to researchers and experts in the field of data science, data preparation and pre-processing need to spend 60-80% of the entire project time (Acuna, 2011). For this reason, special attention has been paid to data pre-processing in this research. According to the nature of the data set used, the need for initial changes in its form is felt. In such a way that the features related to the date were in the form of a time stamp. For this purpose, changes have been made to these features to obtain hours, minutes, and seconds. After that, the spatial coordinates of latitude and longitude are converted into x, y, and z geometric coordinates so that the origin and destination coordinates of the trip can be easily used in the next steps. In the next step, considering that the trips were made at different times and may have been during rush hours due to working days or holidays, changes were made to the data set to determine daily shifts. With the investigation that was done from the city of Boston, it was found that the high traffic hours include the hours in the morning from 6 to 10, noon from 12 to 14, and evening from 15 to 19. In other words, the dataset is divided into three shifts: morning, noon, and evening.

In the next stages of data pre-processing, non-numerical data should be examined. In this dataset, the values of these features are encrypted to become features with binary values. Then the missing data were analysed. In the investigations, it was observed that according to the data collected from the Uber website and the possibility of the data being real, some missing data were recorded, and this type of data was replaced by the nearest neighbour imputation method according to the data of each feature. With the results obtained from the box plot and the way the data is scattered according to the relationship presented in the previous chapter, it was found that there is no data outlier. But because the obtained results are unreliable and may have a negative effect on the final results after the implementation of the algorithms. For this reason, another method called isolated forest was used to get the final results from outlier data. As a result of the implementation of this method, there were 178 outlier data, and according to the type of data set and machine learning methods, it was decided to remove these data from the investigated data set. After determining the assignment of outlier data, the assignment of duplicate data should be determined. As mentioned, duplicate data is data that occupies two rows of the data set exactly. In this case, one of them should be removed to integrate the data set as much as possible. In this data set, it was determined that there are 518 duplicate data, all of which were removed.

One of the most important topics in data preparation is the selection of appropriate features for problem-solving. For this purpose, there are different methods. These methods vary according to the type of algorithm that is chosen. Because reinforcement learning methods and techniques have been used in this research, the feature selection method of 'main feature analysis' (PFA) was chosen, and finally the selected features are according to the table below.

Table3. Selected Features from PFA Method

Index	Feature Names
1	temperature
2	distance
3	windSpeed
4	cloudCover
5	srt_month
6	srt_day
7	precipIntensity
8	srt_minute
9	windGust
10	X
11	Y
12	Z
13	apparentTemperature
14	visibility

In the tables below, the performance of each of the algorithms has been checked by considering each of the characteristics of the data set. According to the method mentioned in the third part for evaluation, the algorithm that can maximize and minimize. By optimizing, we can make sure that the revenue is also maximized and waiting time is minimized by the considered algorithm. To achieve this goal, the MinMax method was first used to scale the predicted and actual data related to the price. It should be noted that according to the central limit theorem in the science of statistics and according to the amount of data under investigation, the distribution of data tends to be almost normal, which can be said that the mathematical expectation is equal to the average value of each of the prices (in dollars).

Table4. Decision matrix based on mathematical hope.

Method	Price						Waiting Time					
	$E_{Safar}(x)$	$E_{Aftabi}(x)$	$E_{Abri}(x)$	$E_{NimeAbri}(x)$	$E_{Meh}(x)$	$E_{Barani}(x)$	$E_{Safar}(x)$	$E_{Aftabi}(x)$	$E_{Abri}(x)$	$E_{NimeAbri}(x)$	$E_{Meh}(x)$	$E_{Barani}(x)$
Q-Learning	27.116	15.531	17.489	8.30	15.073	14.35	9.0039	10.72	8.23	8.30	10.45	14.35
DDPG	25.008	15.431	15.508	8.12	15.143	13.82	8.38	10.47	8.15	8.12	10.77	13.82
SARSA	23.971	16.028	15.160	8.13	15.292	12.91	8.55	10.2086	8.35	8.13	12.02	12.91
DSN	27.680	17.727	15.054	8.34	15.096	12.28	8.52	10.0012	8.13	8.34	10.74	12.28
DQN	25.816	17.472	15.836	8.20	15.499	11.17	8.54	10.35	8.05	8.20	11.40	11.17
DPO	26.341	15.190	14.851	8.35	15.859	12.78	8.26	10.83	8.28	8.35	11.26	12.78
SAC	26.577	17.270	16.773	8.07	15.052	12.10	8.59	10.93	8.26	8.07	12.27	12.10

Table5. Reward-based decision-making matrix.

Method	Price						Waiting Time					
	R_{Safar}	R_{Aftabi}	R_{Abri}	$R_{NimeAbri}$	R_{Meh}	R_{Barani}	R_{Safar}	R_{Aftabi}	R_{Abri}	$R_{NimeAbri}$	R_{Meh}	R_{Barani}
Q-Learning	0.595	0.49	0.511	0.687	0.5001	0.397	0.301	0.262	0.329	0.384	0.452	0.71
DDPG	0.267	0.268	0.289	0.472	0.365	0.408	0.393	0.32	0.429	0.50	0.536	0.437
SARSA	0.307	0.229	0.118	0.122	0.397	0.238	0.331	0.42	0.361	0.381	0.408	0.602
DSN	0.961	0.659	0.582	0.947	0.815	0.849	0.299	0.36	0.326	0.29	0.31	0.4236
DQN	0.872	0.621	0.671	0.773	0.595	0.564	0.233	0.326	0.254	0.45	0.328	0.43
DPO	0.528	0.156	0.563	0.635	0.634	0.493	0.355	0.254	0.387	0.30	0.317	12.78
SAC	0.727	0.662	0.233	0.229	0.312	0.385	0.2406	0.262	0.262	0.612	0.48	0.64

On the other hand, for validation, you can check the rewards received from repeating the experience and implementing each of them. In the basic rules of reinforcement learning and deep reinforcement learning, the algorithm reaches the answer when it can receive the maximum reward. For this purpose, it can be said that the price prediction and also the maximization of income by that algorithm are done optimally and its results can be relied on and used in the real world to receive the maximum reward. Similarly, to minimize the waiting time, this reward must be minimized because our objective function is defined as such. The following table describes the learning rates according to the conditions of each algorithm obtained by the network search method:

Table6. Learning algorithms used in research under different scenarios.

Index	Method	Learning rate under the scenario with weather	Learning rate under the no-weather scenario
1	DSN	0.175	0.257
2	DQN	0.056	0.012
3	DPO	0.11	0.1507
4	Q-Learning	0.118	0.194
5	SAC	0.752	0.683
6	DDPG	0.175	0.190
7	SARSA	0.056	0.0163

At this stage, according to the mentioned learning rates, the reward for each algorithm will be obtained by training the algorithms. The data in this problem are divided in the ratio of 70 to 30 between training and testing data. It should also be mentioned that the optimizer for the implementation of reinforcement learning algorithms of the ADAM method and its cost function is also the mean square error method according to the network search parameter setting method.

In fact, in this research, to accurately evaluate the results, there is a decision matrix with 7 criteria and 24 alternatives that should be examined in the next steps of the algorithm to finally get the best answer.

The entropy weighting method is used in the second step of the MARCOS method, where criteria should be weighted. This method was presented in 1974 by Shannon and Weaver. Entropy expresses the amount of uncertainty in a continuous probability distribution. The main idea of this method is that the greater the dispersion in the values of an index, the more important that index is. Weight is usually assigned to each of the indicators. So that the total weight of the indicators is equal to one. There are various methods to determine the weight of indicators, the entropy method, Linmap method, eigenvector method, and least squares method are the most important methods for determining the weight of indicators. In this research, the entropy method is used because entropy is one of the most practical multi-criteria decision-making methods for calculating the weight of the criteria, which is weighted by the criteria-option matrix.

In the next step, the options should be scaled.

The most common and widely used technique is rescaling, also known as min-max normalization. In this way, in addition to equating the scale of the data, the limits of their change will also be placed in the [0, 1] range. The calculation relationship of this method is given in the data pre-processing section in the third section.

Finally, with the MARCOS the final result will be the ranking of the algorithms. Thesis method, by going through these steps, ranks methods are as follow:

Table7. The final ranking of algorithms according to the MARCOS method.

Method	Rank
DSN	1
DQN	2
DPO	3
Q-Learning	4
SAC	5
DDPG	6
SARSA	7

According to the above table, the DSN algorithm will be known as the best algorithm in optimizing the two goals of maximizing revenue and minimizing waiting time according to different scenarios and the conditions of the research problem. Because it provides more accurate and reliable results.

5. Conclusion and Future Research

According to the results, the proposed method can reduce waiting time robustly and quickly. This would be due to the learning frequency (i.e., an agent learning at every time step) and the small dimensionality of state and action spaces (i.e., distributed learning). However, in comparison to the single bottleneck case, the multiple bottleneck case was slightly inefficient. This might be due to insufficient coordination among the time zones in the current definitions of behavior and reward. In addition, more examination in the scalability of the proposed method is necessary. Future expansion will enable in this research. Firstly, a deep review of the literature with the keywords of dynamic pricing, revenue management, machine learning, reinforcement learning, and deep reinforcement learning in ride-hailing intelligent transportation systems was carried out, as a result of which it was determined what the general purpose of the research is. Follows In the end, after introducing and describing the methods of evaluating and validating the results obtained from the problem-solving techniques, they were implemented in the mentioned software to finally provide the best prices to improve the pricing system dynamically as well as better revenue management. The results of this research are as follows:

1. Discovering and identifying factors affecting dynamic pricing such as factors related to travel and weather conditions.
2. Developing the pricing system.
3. Optimizing revenue management.
4. Optimizing Passenger waiting time

As a suggestion for future research, it can be said that there is a third goal in order to reach a suitable price in the conditions of uncertainty, and that is dynamic matching. Another suggestion that can be made is to combine mathematical optimization methods with machine learning to reach a suitable price under conditions of uncertainty.

Conditions of uncertainty, and that is dynamic matching. Another suggestion that can be made is to combine mathematical optimization methods with machine learning to reach a suitable price under conditions of uncertainty.

6. References

[1] Sutton, R. S., Barto, A. G. (2018), Reinforcement Learning: An Introduction, Cambridge, MA: The MIT Press.

[2] Deng, L.; Yu, D. (2014), Deep Learning: Methods and Applications. Foundations and Trends in Signal Processing. 7 (3–4): 1–199.

[3] Bengio, Y.; Courville, A.; Vincent, P. (2013), Representation Learning: A Review and New Perspectives. IEEE Transactions on Pattern Analysis and Machine Intelligence. 35 (8): 1798–1828.

[4] Zhan, Z. H., Li, J. Y., Zhang, J. (2022). Evolutionary deep learning: A survey, Neurocomputing, 483, pp. 42-58.

[5] Yan, C., Zhu, H., Korolko, N., Woodard, D. (2019), Dynamic Pricing and matching in rid-hailing platforms, Operations Innovations in the Age of Data Analytics, 67(8), pp. 705-724.

[6] Smith, T., Pricing Strategy: Setting Price Levels, Managing Price Discounts and Establishing Price Structures, Cengage Learning, 2011, pp 270-272.

[7] Nagle, T., Hogan, J. and Zale, J., The Strategy and Tactics of Pricing: A Guide to Growing More Profitably, Oxon, Routledge, 2016, p. 1 and 6.

[8] Ziari, M., Ghomi-Avili, M., Pishvaee, M. S., Jahani, H. (2022), A review on competitive pricing in supply chain management problems: models, classification, and applications, 29, pp. 2082-2115.

[9] Taleizadeh, A.A., Akhavizadegan, F., Ansarifar, J. (2019). Pricing and quality level decisions of substitutable products in online and traditional selling channels: game-theoretical approaches. International Transactions in Operational Research, 26, pp. 1718–1751.

[10] Larson, R.B. (2014). Psychological pricing principles for organizations with market power, Journal of Applied Business and Economics, 16, pp. 11–25.

[11] Gongbing B., Lechi, L., Feng, Y., Liang, L. (2014), Dynamic Pricing Based on Strategic Consumers and Substitutes in a Duopoly Setting, Discrete Dynamics in Nature and Society, pp. 1-9.

[12] Chen, C., Yao, F., Mo, D., Zhu, J., Chen, X. (2021), Spatial-temporal pricing for ride-sourcing platform with reinforcement learning, Transportation Research Part C, 130, pp. 1-17.

[13] Cho, Y., Song, J., Kang, M., Hwang, K. (2021), An Application of a Deep Q-Network Based Dynamic Fare Bidding System to Improve the Use of Taxi Services during Off-Peak Hours in Seoul, Sustainability, pp. 1-17.

[14] Ladosz, P., Weng, L., Kim, M., Oh, H. (2022). Exploration in deep reinforcement learning: A survey, Information Fusion, 85, pp. 1-22.

[15] Farazi, N. P., Zou, B., Ahamed, T., Barua, L. (2021). Deep reinforcement learning in transportation research: A review, Transportation Research Interdisciplinary Perspectives, 11, pp. 1-21.

[16] Cross, R. (1997) Revenue Management: Hard-Core Tactics for Market Domination. New York, NY: Broadway Books
[17] Talluri, K., and van Ryzin, G. (1999) Revenue Management: Research Overview and Prospects. Transportation Science 33:233-256.

[18] Legoherel, P. (2013). Revenue Management for Hospitality and Tourism. Oxford, UK: Goodfellow Publishers Ltd. p. 33.

[19] Wang, H., Yang, H. (2019), Ride sourcing systems: A framework and review, Transportation Research Part B, pp. 1-22.

[20] Matsuo, Y., LeCun, Y., Sahani, M., Precup, D., Silver, D., Sugiyama, M., Uchibe, E., Morimoto, J. (2022), Deep Learning, reinforcement learning, and world models, Neural Networks, 22, pp. 1-20.

[21] Heuillet, A., Couthouis, F., Diaz-Rodriquez, N. (2021). Explainability in Deep Reinforcement Learning, Knowledge-Based Systems, 214, pp. 1-14.

[22] Shalev-Shwartz, S., Shammah, S., Shashua, A. (2016). Safe, multi-agent, reinforcement learning for autonomous driving, arXiv preprint arXiv:1610.03295, pp. 1-13.

[23] Wu, Y., Tan, H., Peng, J., Ran, B. (2019). A Deep Reinforcement Learning Based Car Following Model for Electric Vehicle, Appl. Energy, 2(5), pp. 1-18.

[24] Wang, P., Chan, C.Y. (2017). In: Formulation of deep reinforcement learning architecture toward autonomous driving for on-ramp merge, IEEE, pp. 1–6.

[25] Nishi, T., Doshi, P., Prokhorov, D. (2019). Merging in congested freeway traffic using multipolicy decision making and passive actor-critic learning. IEEE Trans, 4 (2), pp. 287–297.

[26] Ye, F., Cheng, X., Wang, P., Chan, C. Y. (2020). Automated Lane Change Strategy using Proximal Policy Optimization-based Deep Reinforcement Learning. arXiv preprint arXiv:2002.02667, pp. 1-22.

[27] Islam, M. A., Mohammad, M. M., Sarathi Das, S. S., Ali, M. E. (2022). A survey on deep learning-based Point-of-Interest (POI) recommendations, Neurocomputing, 472, pp. 306-325.

[28] Chen, X. M., Wu, C. X., Wu, Y., Xiong, N. X., Han, R., Ju, B. B. (2019). Design and Analysis for Early Warning of Rotor UAV Based on Data-Driven DBN, Electronics, 8, pp. 1-22.

[29] Li, P., Chen, Z., Yang, L. T., Gao, J., Zhang, Q., Deen, M. J. (2018). An Improved Stacked Auto-Encoder for Network Traffic Flow Classification, IEEE, 18, pp.1-6.

[30] Chandrakala, S., Daniel, R. M., Deepak, K., Deepan, P., Francis, S. A. J., Isravel, D. P., Issac, T., Jones, T. A., Karunakaran, A., Krishnanugrah, P. U., Kumar, G. P., Kumar, R. J. S. J., Mahalingam, M., Mahesh, K. (2020). The Cognitive Approach in Cloud Computing and Internet of Things Technologies for Surveillance Tracking Systems, Elsevier Inc., chapter 11, pp. 162-171.

[31] Le, X. H., Lee, G., Jung, S. (2019). Application of LSTM neural network for Flood Forecasting, Water, 11, pp. 1-19.

[32] Ngo, N. S., Gotchi, T., Clark, B. Y. (2021). The effects of ride-hailing services on bus ridership in a medium-sized urban area using micro-level data: Evidence from the Lane Transit District, Transport Policy, 105, pp. 44-53.

[33] Qin, G., Yin, Y., Sun, J., Ye, J. (2021). Optimizing matching time intervals for ride-hailing services using reinforcement learning, Transportation Research Part C, 129, pp. 1-18.

[34] Jiao, Y., Tang, X., Qin, Z. T., Li, S., Zhang, F., Zhu, H., Ye, J. (2021). Real-world ride-hailing vehicle repositioning using deep reinforcement learning, Transportation Research Part C,130, pp. 1-25.

[35] Ke, J., Xiao, F., Yang, H., Ye, J. (2019). Optimizing Online Matching for Ride-Sourcing Services with Multi-Agent Deep Reinforcement Learning, IEEE, pp. 1-10.

[36] Shukla, A., Bhattacharya, P., Tanwar, S., Kumar, N., Guizani, M. (2020) DwaRa: A Deep Learning-Based Dynamic Toll Pricing Scheme for Intelligent Transportation Systems, IEEE transactions on vehicular technology, 69 (11), pp. 12510-2520.

[37] Lu, C. C., Mahmassani, H. S. (2009) Dynamic Pricing with Heterogeneous Users, Journal of the Transportation Research Board, 2090, pp. 75-85.

[38] Olayode, I. O., Tartibu, L. K., Okwu, M. O. and Uchechi, U.F. (2020) Intelligent transportation systems, un-signalized road intersections and traffic congestion in Johannesburg: a systematic review, Procedia CIRP 91 (CIRP design Conference), pp. 844-850.

[39] Cabrera, R. S., de la Cruz, A. P., Molina, and J. M. M. (2020) Sustainable transit vehicle tracking service, using intelligent transportation system services and emerging communication technologies: A review, Journal of traffic and transportation engineering, 7 (6), pp. 729-747.

[40] Do, M., Byun, W., Shin, D. K., Jin, H. (2019). Factors Influencing Matching of Ride-Hailing Service Using Machine Learning Method, Sustainability, 11, pp. 1-13.

[41] Lei, C., Jiang, Z., Ouyang, Y. (2019) Path-based Dynamic Pricing for Vehicle Allocation in Ridesharing Systems with Fully Compliant Drivers, Transportation Research Procedia, 38, pp. 77-97.

[42] Yan, C., Zhu, H., Korolko, N., Woodard, D. (2019). Dynamic pricing and matching in ride-hailing platform, Naval Res Logistics, 67,

pp. 705-724.

[43] Sun, Z., Xu, Q., Shi, B. (2020). Dynamic Pricing of Ride-Hailing Platforms considering Service Quality and Supply Capacity under Demand Fluctuation, Hindawi: Mathematical Problems in Engineering, pp. 1-26.

[44] O'Keeffe, K., Anklesaria, S., Santi, P., Ratti, C. (2021). Using reinforcement learning to minimize taxi idle times, journal of intelligent transportation systems, pp. 1-12.

[45] Guizzardi, A., Pons, F. M. E., Angelini, Giovanni, Ranieri, E. (2021). Big data from dynamic pricing: A smart approach to tourism demand forecasting.

[46] Haliem, M., Mani, G., Aggarwal, V., Bhargava, B. (2021). Dynamic pricing and matching in ride-hailing platform, IEEE transactions on intelligent transportation systems, pp. 1-12.

[47] Al-Abbasi, A. O., Ghosh, A., Aggarwal, V. (2019). DeepPool: Distributed Model-Free Algorithm for Ride-Sharing Using Deep Reinforcement Learning, IEEE transactions on intelligent transportation systems, 20 (12), pp. 705-724.

[48] Feng, J., Gluzman, M., Dai, J. G. (2021). Scalable Deep Reinforcement Learning for Ride-Hailing, IEEE control systems letters, 5 (6), pp. 2060-2065.

[49] Song, J., Cho, Y. J., Kang, M. H., Hwang, K. Y. (2020). An Application of Reinforced Learning-Based Dynamic Pricing for Improvement of Ridesharing Platform Service in Seoul, Electronics, 9, pp. 1-14.

[50] Rana, R., Oliveira, F. S. (2014). Real-time dynamic pricing in a non-stationary environment using model-free reinforcement learning, Omega, 47, pp. 116-126.

[51] Liu, J., Gibson, S. J., Mills, J., Osadchy, M. (2019). Dynamic spectrum matching with one-shot learning, Chemometrics and Intelligent Laboratory Systems, 184, pp. 175-181.

[52] Pahwa, D., Starly, B. (2021). Dynamic matching with deep reinforcement learning for a two-sided Manufacturing-as-a-Service (MaaS) marketplace, Manufacturing Letters, 29, pp. 11-14.

[53] Qiu, J., Haung, K., Hawkins, J. (2022). The taxi sharing practices: Matching, routing and pricing methods, Multimodal Transportation, 1, pp. 1-11.

[54] Guo, W., Atasoy, B., Van Blokland, W. B. (2021). Global synchromodal transport with dynamic and stochastic shipment matching, Transportation Research Part E, 152, pp. 1-25.

[55] Cadas, A., Doncel, J., Busic, A. (2022). Analysis of an optimal policy in dynamic bipartite matching models, Performance Evaluation, 154, pp. 1-30.

[56] Wang, J., Wang, X., Yang, S., Zhang, X., Gao, Z. (2021). Predicting the matching probability and the expected ride/shared distance for each dynamic ridepooling order: A mathematical modeling approach, Transportation Research Part B, 154, pp. 125-146.

[57] Palma, A., Stokkink, P., Geroliminis, N. (2022). Influence of dynamic congestion with scheduling preferences on carpooling matching with heterogeneous users, Transportation Research Part B, 155, pp. 479-498.

[58] Yan, P., Lee, C. Y., Chen, C., Chu, C., Luo, Z. (2021). Matching and pricing in ride-sharing: Optimality, stability, and financial sustainability, Omega, 102, pp. 1-18.

[59] Tian, R., Wang, C., Ma, Z., Liu, Y., Gao, S. (2022). Research on vehicle-cargo matching algorithm based on improved dynamic Bayesian network, Computers & Industrial Engineering, 168, pp. 1-14.

[60] Wei, S., Feng, S., Ke, J., Yang, H. (2022). Calibration and validation of matching functions for ride-sourcing markets, Communications in Transportation Research, 2, pp. 1-14.

[61] Yan, C., Zhu, H., Korolko, N., Woodard, D. (2019). Dynamic Pricing and Matching in Ride-Hailing Platforms, Uber Technologies, pp. 1-36.

[62] Viglia, G., Canio, F. D., Stoppani, A., Invernizzi, A. C., Cerutti, S. (2021). Adopting revenue management strategies and data sharing to cope with crises, Journal of Business Research, 137, pp. 336-344.

[63] Kyparisis, G. J., Koulamas, C. (2018). Optimal Pricing and seat allocation for a two-cabin airline revenue management problem, International Journal of Production Economics, 201, pp. 18-25.

[64] Snoeck, A., Merchan, D., Winkenbach, M. (2020). Revenue management in last-mile delivery: state-of-the-art and future research directions, Transportation Research Procedia, 46, pp. 109-116.

[65] Luo, Y., Yan, H., Zhang, S. (2021). An interactive booking-limit control for passenger railway revenue management, Journal of Rail Transport Planning & Management, 20, pp. 1-11.

[66] An, J., Mikhaylov, A., Jung, S. U. (2021). A linear Programming approach for robust network revenue management in the airline industry, Journal of Air Transport Management, 91, pp. 1-7.

[67] Shi, Y., Yu, Y., Dong, Y. (2021). Warehousing platform's revenue management: A dynamic model of coordinating space allocation for self-use and rent, European Journal of Operational Research, 293, pp. 167-176.

[68] Selcuk, A. M., Avsar, Z. M. (2019). Dynamic Pricing in airline revenue management, Journal of Mathematical Analysis and Applications, 478, pp. 1191-1217.

[69] Xiao, B., Yang, W. (2021). A Bayesian learning model for estimating unknown demand parameter in revenue management, European Journal of Operational Research, 293, pp. 248-262.

[70] Ballestero, P. T., Garcia, M. N., Serrano, L. G. (2022). The wheel of dynamic pricing: Towards open pricing and one to one pricing in hotel revenue management, International Journal of Hospitality Management, 102, pp. 103-184.

[71] Rennie, N., Cleophas, C., Sykulski, A. M., Dost, F. (2021). Identifying and responding to outlier demand in revenue management, European Journal of Operational Research, 293, pp. 1015-1030.

[72] Webb, T., Schwartz, Z., Xiang, Z., Singal, M. (2020). Revenue management forecasting: The resiliency of advanced booking methods given dynamic booking windows, International Journal of Hospitality Management, 89, pp. 1-9.

[73] Saito, T., Takahashi, A., Koide, N., Ichifuji, Y. (2019). Application of online booking data to hotel revenue management, International Journal of Information Management, 46, pp. 37-53.

[74] Luo, Y., Yan, H., Zhang, S. (2020). Simulation-based integrated optimization of nesting policy and booking limits for revenue management, Computers & Industrial Engineering, 150, pp. 1-15.

[75] Seo, D. W., Chang, K., Cheong, T., Baek, J. G. (2021). A Reinforcement Learning approach to distribution-free capacity allocation for sea cargo revenue management, Information Sciences, 571, pp. 623-648.

[76] Talluri, K. T., Karaesmen, I. Z., Vulcano, G. J.., Ryzin, G. J. V. (2008). Revenue management: models and methods, Proceedings of the 2008 Winter Simulation Conference, pp. 145-156.

[77] Agrawal, S., Yin, S., Zeevi, A. (2021). Dynamic Pricing and Learning under the Bass Model, Proceedings of the 22nd ACM Conference on Economics and Computation, pp. 1-44, https://doi.org/10.1145/3465456.3467546.

[78] Kamandanipour, K., Nasiri, M. M., Konur, D., Yakhchali, S. H. (2020). Stochastic data-driven optimization for multi-class dynamic pricing and capacity allocation in the passenger railroad transportation, Expert Systems with Applications, 158, pp. 1-16.

[79] Bi, G., Li, L., Yang, F., Liang, L. (2014). Dynamic Pricing Based on Strategic Consumers and Substitutes in a Duopoly Setting, Discrete Dynamics in Nature and Society, pp. 1-10.

[80] Chen, Q. G., Lei, Y. M., Jasin, S. (2020). Real-Time Spatial-Intertemporal Dynamic Pricing for Balancing Supply and Demand in a Ride-Hailing, SSRN, pp.1-53. Available at SSRN: https://ssrn.com/abstract=3610517

[81] Genser, A., Kouvelas, A. (2022). Dynamic optimal congestion pricing in multi-region urban networks by application of a multi-layer-neural network, Transportation Research Part C, 134, pp. 1-23.

[82] Liang, L., Cai, X. (2022). Time-sequencing European options and pricing with deep learning Analyzing based on interpretable ALE method, Expert Systems With Applications, 187, pp. 1-22.

[83] Buchholz, N. (2021). Spatial Equilibrium, Search Frictions, and Dynamic Efficiency in the Taxi Industry, Review of Economic Studies, pp. 1-36.

[84] Branda, F., Marozzo, F., Talia, D. (2020). Ticket Sales Prediction and Dynamic Pricing Strategies in Public Transport, Big data and cognitive computing, 36 (4), pp. 1-17.

[85] Feng, G. (2018). Analysis of On-Demand Ride-Hailing Systems, A DISSERTATION SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL OF THE UNIVERSITY OF MINNESOTA.

[86] Wang, J., Chua, G., Sainathan, A., Vijayendiran, A. (2018). Tipping Point in Ride-Hailing Service Systems with Sharing Option, pp. 1-39, Available at SSRN: https://ssrn.com/abstract=3138458 or http://dx.doi.org/10.2139/ssrn.3138458.

[87] Afeche, P., Liu, Z., Maglaras, C. (2018). Ride-Hailing Networks with Strategic Drivers: The Impact of Platform Control Capabilities on Performance, pp. 1-56, Available at SSRN: https://ssrn.com/abstract=3120544 or http://dx.doi.org/10.2139/ssrn.3120544.

[88] Fatemipour, E., Madanizadeh, S. A., Joshaghani, H. (2020). Modeling Online Taxi Drivers' Decision Making in Accepting or Rejecting Ride Offers, Sharif University of Technology, pp. 1-44.

[89] Schroder, M., Storch, D. M., Marszal, P., Timme, M. (2020). Anomalous supply shortages from dynamic pricing in on-demand mobility, Nature Communications, 11, pp. 1-93.

[90] Du, J., Rakha, A. H., Breuer, H. (2020). An In-Depth Analysis of Ride-Hailing Travel Using a Large-scale Trip-Based Dataset, Economics, pp. 1-15.

[91] Gillan, J. M. (2017). Dynamic Pricing, Attention, and Automation: Evidence from a Field Experiment in Electricity Consumption, 36 (4), pp. 1-17.

[92] Castillo, J. C. (2020). Essays on the economics of ride hailing, A DISSERTATION SUBMITTED TO THE DEPARTMENT OF ECONOMICS AND THE COMMITTEE ON GRADUATE STUDIES OF STANFORD UNIVERSITY.

[93] Nair, V. J., Marozzo, F., Talia, D. (2021). Estimation of Cumulative Prospect Theory-based passenger behavioral models for dynamic pricing & transactive control of shared mobility on demand, Department of Mechanical Engineering in partial fulfillment of the requirements for the degree of Master of Science in Computational Science and Engineering at the MASSACHUSETTS INSTITUTE OF TECHNOLOGY.

[94] Skochdopole, N. (2019). Modeling competitive markets: retailers and ride-hailing platforms, COMPUTATIONAL AND MATHEMATICAL ENGINEERING AND THE COMMITTEE ON GRADUATE STUDIES OF STANFORD UNIVERSITY.

[95] Tervo, E. (2020). Dynamic pricing or not? — Pricing models of Finnish taxi dispatch centers under the Act on Transport Services, University of Oulu Department of Information Processing Science.

[96] Zha, L. (2017). Modeling and analysis of on-demand ride-sourcing markets, A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL OF THE UNIVERSITY OF FLORIDA.

[97] Heeks, R., Morantes, J. E. G., Graham, M., Howson, K., Mungai, P., Nicholson, B., Van Belle, J. P. (2021). Digital Platforms and institutional voids in developing countries: The case of ride-hailing markets, World Development, 145, pp. 1-17.

[98] Turan, B., Pedarsani, R., Alizadeh, M. (2020). Dynamic pricing and fleet management for electric autonomous mobility on demand systems, Transportation Research Part C, 121, pp. 1-20.

[99] De Sa, A. L. S., Pitombo, C. S. (2021). Methodological proposal for stated preference scenarios regarding an exploratory evaluation of ride-hailing implications on transit: A Brazilian context analysis, Case Studies on Transport Policy, pp. 1-10.

[100] Schwieterman, J., Smith S. (2018). Sharing the ride: A paired-trip analysis of UberPool and Chicago Transit Authority services in Chicago, Illinois, Research in Transportation Economics, 71, pp. 9-16.

[101] Bischoff, J., Kaddoura, I., Maciejewski, M., Negel, K. (2018). Simulation-based optimization of service areas for pooled ride-hailing operators, Procedia Computer Science,130, pp. 816-823.

[102] Yang, T., Gou, Q., Xu, L., Sun, H. (2021). Dynamic pricing for integrated energy-traffic systems from a cyber-physical-human perspective, Renewable and Sustainable Energy Reviews, 136, pp. 1-11.

[103] Tian, Q., Yang, Y., Wen, J. Ding, F., He, J. (2020). How to Eliminate Detour Behaviors in E-hailing? Real-time Detecting and Time-dependent Pricing, IEEE transactions on intelligent transportation systems, pp. 1-10.

[104] Yuan, E., Hentenryck, P. V. (2021). Real-Time Pricing Optimization for Ride-Hailing Quality of Service, Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence (IJCAI-21), pp. 3742-3748.

[105] Wu, Y., Li, W., Yu, Q., Chen, J. (2021). Modeling the Impacts of Driver Income Distributions on Online Ride-Hailing Services, Mathematical Problems in Engineering, pp. 1-9.

[106] Kooti, F., Grbovic, M., Aiello, L. M., Djuric, N., Radosavljevic, V., Lerman, K. (2017). Analyzing Uber's Ride-sharing Economy, International World Wide Web Conference Committee (IW3C2), pp. 574-582.

[107] Wang, Q., Lownes, N. E. (2019). All-Links-Based E-Hailing Pricing and Surcharge Mechanism for Transportation System Performance Improvement, Transportation Research Record, 2673 (12), pp. 103-114.

[108] Jin, Y. M., Ye, X. F., Liu, W. L., Wang, T., Wang, H. (2019). Dynamic pricing model for cruising taxicab based on system dynamics, Advances in Mechanical Engineering, 11 (2), pp. 1-8.

[109] Li, S., Fei, F., Ruihan, D., Yu, S., Dou, W. (2016). A Dynamic Pricing Method for Carpooling Service Based on Coalitional Game Analysis, IEEE 18th International Conference on High Performance Computing and Communications, pp. 78-85.

[110] Duan, Y., Gao, G., Xiao, M., Wu, J. (2019). A Privacy-Preserving Order Dispatch Scheme for Ride-Hailing Services, IEEE 16th International Conference on Mobile Ad Hoc and Sensor Systems (MASS), pp. 118-126.

[111] Yang, T., Guo, Q., Sun, H. (2020). A Three-stage Incentive Scheme for Integrated Energy-Traffic Systems Using Deep Q-Network, Applied Energy Symposium, pp. 1-6.

[112] Hidayat, F., Anki, P. (2021). Application of Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) Models to Determine Prices on Ride Hailing, 7th International Conference on Education and Technology (ICET), pp. 304-309.

[113] Xu, K., Sun, L., Liu, J., Wang, H. (2018). An empirical investigation of taxi driver response behavior to ride-hailing requests: A spatio-temporal perspective, PLOS ONE, pp. 1-17.

[114] Zhang, W., Xi, Y., Ukkusuri, S. V. (2020). Understanding Spatiotemporal Variations of Ridership by Multiple Taxi Services, International Journal of Geo-Information, 9, pp. 1-18.

[115] Chaudri, H. A., Byers, J. W., Terzi, E. (2020). Learn to Earn: Enabling Coordination Within a Ride-Hailing Fleet, IEEE International Conference on Big Data (Big Data), pp. 1127-1136.

[116] Chroder, M., Storch, D. M., Marszal, P., Timme, M. (2020). Anomalous supply shortages from dynamic pricing in on-demand mobility, NATURE COMMUNICATIONS, pp. 1-8

[117] Guan, Y., Annaswamy, A. M., Tseng, H. E. (2020). Towards Dynamic Pricing for Shared Mobility on Demand using Markov Decision Processes and Dynamic Programming, 23rd IEEE International Conference on Intelligent Transportation Systems, pp. 1-7.

[118] Lei, C., Jiang, Z., Ouyang, Y. (2019). Path-based Dynamic Pricing for Vehicle Allocation in Ridesharing Systems with Fully Compliant Drivers, 23rd International Symposium on Transportation and Traffic Theory, pp. 77-97.

[119] Chao, J. (2019). Modeling and Analysis of Uber's Rider Pricing, International Conference on Economic Management and Cultural Industry (ICEMCI 2019), pp. 693-711.

[120] Li, S., Tavafoghi, H., Poolla, K., Varaiya, P. (2019). Regulating TNCs: Should Uber and Lyft set their own rules?, Transportation

Research Part B, 129, pp. 193-225.

[121] Chen, X. M., Zheng, H., Ke, J., Yang, H. (2020). Dynamic optimization strategies for on-demand ride services platform: Surge pricing, commission rate, and incentives, Transportation Research Part B, 138, pp. 23-45.

[122] Zhong, Y., Yang, T., Cao, B,. Cheng, T. C. E. (2022). On-demand ride-hailing platforms in competition with the taxi industry: Pricing strategies and government supervision, International Journal of Production Economics, pp. 1-27.

[123] Rangel, T., Gonzalez, J. N., Gomez, J., Romero, F., Vassallo, J. M. (2021). Exploring ride-hailing fares: an empirical analysis of the case of Madrid, Transportation, 49, pp. 373–393.

[124] Wang, Y., Tang, T., Zhang, W., Sun, Z., Xiong, Q. (2021). The Achilles tendon of dynamic pricing — the effect of consumers' fairness preferences on platform's dynamic pricing strategies, Journal of Internet and Digital Economics, pp. 1-21.

[125] Brown, A., Lavalle, W. (2021). Hailing a change: comparing taxi and ridehail service quality in Los Angeles, Transportation, 48, pp. 1007-1031.

[126] Yang, C., Feng, Y., Whinston, A. (2021). Dynamic Pricing and Information Disclosure for Fresh Produce: An Artificial Intelligence Approach, Production and Operations Management Society, 0 (0), pp. 1-17.

[127] Feng, Y., Niazadeh, R., Saberi, A. (2020). Two-stage Stochastic Matching and Pricing with Applications to Ride Hailing, Available at SSRN: https://ssrn.com/abstract=3613755 or http://dx.doi.org/10.2139/ssrn.3613755, pp. 1-53.

[128] Zhang, L., Gao, Y., Zhu, H., Tao, L. (2022). Bi-level stochastic real-time pricing model in multi-energy generation system: A reinforcement learning approach, Energy, 239, pp. 1-19.

[129] Sato, K., Seo, T., Fuse, T. (2021). A reinforcement learning-based dynamic congestion pricing method for the morning commute problem, Transportation Research Procedia, 52, pp. 347-355.

[130] Nguyen, T., Gao, W., Zhong, X., Agarwal, S. (2021). Reinforcement Learning and Adaptive Optimal Control of Congestion Pricing, IFAC PapersOnLine, pp. 221-226.

[131] Bi, W., Sun, Y., Liu, H., Chen, X. (2014). Dynamic Nonlinear Pricing Model Based on Adaptive and Sophisticated Learning, Mathematical Problems in Engineering, pp. 1-11.

[132] Pandey, V., Wang, E., Boyles, S. D. (2019). Deep Reinforcement Learning Algorithm for Dynamic Pricing of Express Lanes with Multiple Access Locations, Systems and Control, pp. 1-39.

[133] Xu, H., Wu, Q., Wen, J., Yang, Z (2022). Joint bidding and pricing for electricity retailers based on multi-task deep reinforcement learning, International Journal of Electrical Power and Energy Systems, 138, pp. 1-17.

[134] Deng, Y., Chen, H., Shao, S., Tang, J., Pi, J., Gupta, A. (2021). Multi-Objective Vehicle Rebalancing for Ridehailing System using a Reinforcement Learning Approach, Journal of Management Science and Engineering, pp. 1-30.

[135] Zhu, Z., Ke, J., Wang, H., (2021). A mean-field Markov decision process model for spatial-temporal subsidies in ride-sourcing markets, Transportation Research Part B, pp. 540-565.

[136] Levina, T., Levin, Y., McGill, J., Nediak, M. (2009). Dynamic Pricing with Online Learning and Strategic Consumers: An Application of the Aggregating Algorithm, Operations Research, 57(2), pp. 327-341.

[137] Yin, C., Han, J. (2021). Dynamic Pricing Model of E-Commerce Platforms Based on Deep Reinforcement Learning, Computer Modeling in Engineering & Sciences, 127(1), pp. 291-307.

[138] Collins, A., Thomas, L., (2013). Learning competitive dynamic airline pricing under different customer models, Journal of Revenue and Pricing Management, 12(5), pp. 416-430.

[139] Raju, C. V. L., Narahari, Y., Ravikumar, K., (2006). Learning dynamic prices in electronic retail markets with customer segmentation, Ann Oper Res, 143, pp. 59-75.

[140] Cho, Y., Song, J., Kang, M., Hwang, K. (2021). An Application of a Deep Q-Network Based Dynamic Fare Bidding System to Improve the Use of Taxi Services during Off-Peak Hours in Seoul, Sustainability, 13, pp. 1-17.

[141] Acuna, E. (2011). Preprocessing in Data Mining, Springer, 15, pp. 1–5.

[142] Sato, K., Seo, T., Fuse, T. (2021). A reinforcement learning-based dynamic congestion pricing method for the morning commute problem, Transportation Research Procedia, 52, pp. 347-355.

[143] Maestre, R., Duque, J., Rubio, A., Arevalo, J. (2018). Reinforcement Learning for Fair Dynamic Pricing, Intelligent Systems Conference, pp. 1-7.

[144] Collins, K., Thomas, T., Fuse, T. (2012). Comparing reinforcement learning approaches for solving game theoretic models: a dynamic airline pricing game example, Journal of the Operational Research Society, 63, pp. 1165-1173.

[145] Qin, G., Yin, Y., Sun, J., Ye, J. (2021). Optimizing matching time intervals for ride-hailing services using reinforcement learning, Transportation Research Part C, 129, pp. 1-18.

[146] Xu, H., Wu, Q., Wen, J., Yang, Z (2022). Joint bidding and pricing for electricity retailers based on multi-task deep reinforcement learning, International Journal of Electrical Power and Energy Systems, 138, pp. 1-17.

[147] Lee, S., Choi, D. (2021). Dynamic pricing and energy management for profit maximization in multiple smart electric vehicle charging stations: A privacy-preserving deep reinforcement learning approach, Applied Energy, 304, pp. 1-16.