Title: AI Agent to help cab drivers maximize profit within a month

LITERATURE REVIEW

PAPER 1:

An Integrated Reinforcement Learning and Centralized Programming Approach for Online Taxi Dispatching.

Balancing supply and demand for ride-sharing companies is a challenge, especially with real-time requests and stochastic traffic conditions on large, congested road networks. To meet this challenge, this post proposes a robust and scalable approach that integrates reinforcement learning (RL) and centralized programming constructs (CP) to facilitate real-time taxi operations. Both real-time order matching decisions and vehicle relocation decisions at microscopic network scales are integrated into the framework of Markov's decision-making process. The RL component learns a decomposed state-value function representing taxi driver experience, historical offline demand patterns, and congestion in the transportation network. CP components collectively schedule non-myopic decision-making of drivers under prescribed system constraints to explicitly enable collaboration. Furthermore, a time-lagged learning algorithm using prioritized gradient descent and adaptive search techniques to avoid the problem of reward sparseness and sample imbalance across microscopic road networks. The simulator was built and trained using data from Manhattan's road network and New York City's yellow cabs to simulate a real-time vehicle handling environment. Both centralized and decentralized taxi dispatch policies are validated in the simulator. The approach can further improve profits for taxi drivers while reducing customer waiting times compared to some existing ride-hailing algorithms.

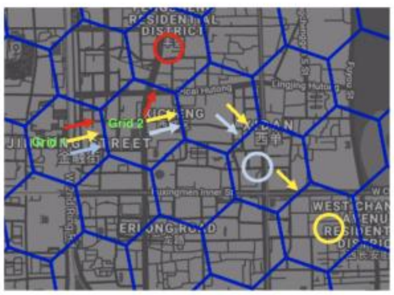
RELEVANT LITERATURE FOR THE VEHICLE DISPATCHING PROBLEM

Literature	Agent	Scale	State	Action	Reward	Training Algorithm
[1]	vacant vehicles	Gr	T+L	assign orders	discounted fare	PE
[2]	vacant vehicles	Gr	T+L+local CF	assign orders	discounted fare	PE
[3]	vacant vehicles	Gr	T+L+local CF	assign orders	fare	DRL
[4]	unmatched orders	Gr(1 km)	T+L+local CF	delay orders	weighted reward	DRL
[5]	vacant vehicles	Gr(1.2 km)&Co	T+L+local CF	assign orders	weighted reward	DRL
[6]	vacant vehicles	Gr(1.2 km)	L+local CF	assign orders	fare	DRL
[7]	bipartite graph	Co	batch length	delay orders	edge weight	RL
[8]	vacant vehicles	Gr(1.2 km)	T+L+local CF	reposition	averaged revenue	DRL
[9]	dispatch center	Gr(150 m)	global CF	reposition	profit	DRL
[10]	vacant vehicles	Gr(irregular)	global counts	reposition	profit	DRL
[11]	vacant vehicles	Ne	T+L+arriving direction	reposition	profit	DP
[12]	single vehicle	Gr(5 km)	L	reposition	profit	DP
[13]	vacant vehicles	Gr(700 m)	T+L+indicator	reposition	profit	DP
[14]	single vehicle	Co	T+L+time budget	assign orders+routes	profit	ADP
[15]	vacant vehicles	Gr(0.5 mile)	T+L+battery level	assign orders+reposition+recharge	profit	ADP
[16]	all vehicles	Gr(5 mile)	T+L+local CF	price+assign orders+routes	profit	ADP
[17]	whole system	Gr	T+global CF	assign orders + reposition	fare	DRL
[18]	hierarchical grids	Gr(1.2 km)	hierarchical CF	assign orders + reposition	hierarchical reward	DRL
This work	vacant (almost) vehicles	Ne	vacant T+L	assign orders + reposition	profit	DRL+CP

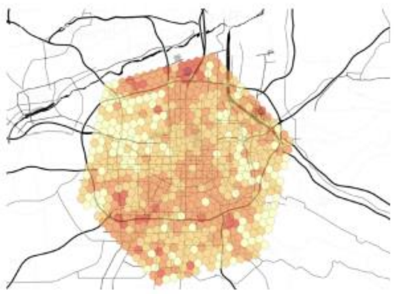
Scale: **Gr**: Grid level; **Co**: Coordinate level; **Ne**: Network level

State: **T**: Timestamps; **L**: Location; **CF**: Contextual features

Algorithm: **PE**: Policy Evaluation; **DRL**: Deep Reinforcement Learning; **DP**: Dynamic Programming; **ADP**: Approximate Dynamic Programming; **CP**: Centralized Programming



(a)



(b)

(c)

On investigating the online vehicle dispatching problem in practice with a platform to manage a fleet of thousands of taxis to serve tens of thousands of customers at a real-time operational level. To solve it, we then proposed an efficient RL-based algorithm that combines the data-driven RL components with a CP-based planning module. To leverage the real-time level operation and heterogeneous traffic network topology, we reformulate the online vehicle dispatching problem using a link-node-based micro-network representation and integrate both the order dispatching stage and the vehicle routing stage. Instead of decentralized/sequential decisions that lead to implicit and limited cooperation among vehicles, we adopt a centralized mathematical programming model, where the system objective function is decomposed and approximated by the sum of the parameterized agent’s action-value function. Two options of actions referring to order assignment and route selection are optimized collectively under system constraints to explicitly realize cooperation among agents and reduce the computational burden. To address the problem of sparse reward signals and transition sample imbalance when agents interact with large-scale micro-network environments, prioritized using an adaptive search strategy that speeds up the agent learning process. Proposed TD learning by gradient descent. Simulation experiments performed on a Manhattan street network demonstrated the computational feasibility and scalability of the proposed algorithm in large-scale applications. It used historical data to determine travel demand and simulated dynamic traffic conditions across the network to support online vehicle dispatch routes. The information learned was also used to direct free taxis to areas of high demand. Moreover, compared with some existing decentralized RL algorithms and pure optimization algorithms, trial results show that the proposed algorithm not only improves overall platform revenue and driver average profit, but also eliminates platform discrepancies. Rates and match delays, and overall customer wait times.

PAPER 2:

Autonomous Driving System based on Deep Q Learning

This abstract presents simulation results of an autonomous vehicle learning to drive in a simplified environment containing only lane markings and static obstacles. Training is done via a Deep Q Network. Giving input image of the road captured by the car's front-facing camera, the deep Q network computes Q-values (rewards) that correspond to actions available to self-driving cars. These actions are discrete angles the car can steer at a constant speed. In-car self-driving systems enforce behaviors that offer the highest rewards. Our simulation results show high accuracy in learning to drive by observing lanes and avoiding obstacles.

The application of artificial intelligence (AI) and machine learning (ML) techniques to the development of autonomous driving systems is currently a hotbed of research. Self-driving technology is seen as a safe and efficient means of transportation. Fully functional self-driving cars are still a long way off, but most of the world's major automakers have reached advanced stages of self-driving car development. Current autonomous driving typically relies on highly specialized infrastructure such as Light Detection and Ranging (LIDAR) for navigation, global positioning GPS for localization, and Laser Range Finder (LRF) for obstacle detection. It is very expensive because it depends on the structure. The research focus is on learning to drive by recognizing road features that depend only on the view of a single camera.

Q-learning is a type of reinforcement learning. In Q-learning, we have agents with states and corresponding actions. Agents are always in a manageable state. At the next time step, the state is transformed into another state by performing an action. This behavior is accompanied by a reward or punishment. The agent's goal is to maximize reward acquisition. The Q-learning algorithm is represented by the following update formula:

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha \left(r_t - Q(s_t, a_t) + \gamma \underbrace{max}_{a'} Q(s_{t+1}, a') \right)$$

where Q(st , at) represents the Q-value of an agent in state st and its action at time t rewarded with a reward rt. is the learning rate and γ is the discount factor. The γ parameter is in the range [0,1]. When γ is close to 0, the agent tends to consider only immediate rewards. On the other hand, the agent will consider future rewards with more weight and will be willing to delay rewards.

I’ve extracted a simulation study of an autonomous agent learning to drive in a simplified environment consisting only of lane boundaries and static obstacles. They’ve used Deep Q Network to train agents in a simulated environment. Future prospects for our research include validating the simulation results and driving a Robot car (a fully functional driving vehicle, 1/10 the size of a standard commercial vehicle) on the floor of a laboratory. Validation results should show that learning autonomous driving in a simulated environment is a step towards driving on real roads.

PAPER 3:

The Rich and the Poor: A Markov Decision Process Approach to Optimizing Taxi Driver Revenue Efficiency

Taxi services play an important role in the public transportation system of large cities. Improving the efficiency of taxi companies is an important social issue as it has the potential to improve drivers' income and reduce gas emissions and fuel consumption. Recent research on search strategies ignores the important impact of passenger destinations on future passenger searches, which may not be optimal for generating total revenue over time. To address these issues, this paper examines ways to improve revenue efficiency (revenue per unit of time) for taxi drivers, modeling the passenger search process as a Markov Decision Process (MDP). For each hourly timeframe, from the data he learns different sets of parameters for MDP and finds the optimal free taxi trip to maximize total revenue in that timeframe. Case studies and several experimental evaluations on real datasets from big cities in China show that the proposed approach improves the turnover efficiency of inexperienced drivers by up to 15% and is better than the baseline in all timeframes. It shows that it is better than the method.

A Markov decision process approach

- System States
- Actions
- State Transition and Objective Function
- Learning Parameters for MDP
- Solving MDP

They've explored how to learn from historical data the optimal taxi search strategy to improve the operational efficiency of taxi drivers over time. This paper proposed modeling the passenger search process as a Markov Decision Process (MDP) to find the best available train for taxis in each state. Case studies and experimental results showed that our approach effectively improved revenue efficiency for inexperienced drivers by up to 15%, up to 8.4% above baseline. They plan to improve the MDP model to recommend optimal strategies for improving total revenue efficiency, including total revenue and taxi shift time driving time. They plan to include more attributes such as weather and traffic conditions to improve the method's results.

PAPER 4:

Artificial Intelligence for Vehicle-to-Everything: A Survey

Advances in communications, intelligent traffic systems, and computer systems have opened up new possibilities for intelligent road safety, comfort, and efficiency solutions. Artificial intelligence (AI) is widely used in various fields of scientific research to optimize traditional data-driven approaches. Vehicle-to-everything (V2X) systems, along with AI, gather information from a variety of sources, augment driver awareness, and predict potential accident avoidance to improve driving comfort, safety, and efficiency. This paper provides a comprehensive overview of the research papers that used AI to address various research questions in his V2X system. This research contribution is summarized and categorized by application domain. Finally, we present open questions and research challenges that need to be addressed to unlock the full potential of AI to evolve V2X systems.

The latest developments in AI technology have opened up new possibilities for Intelligent Transport Systems (ITS). Vehicle sensors will also become more intelligent over time, allowing vehicles to better assess their surroundings. This progress has led to the possibility of realizing autonomous driving, based on the idea of mimicking human driving behavior while mitigating human error. A wealth of applications are being developed, from active and passive road safety to traffic optimization, autonomous vehicles to the Internet of Vehicles.

- Swarm Intelligence

  Swarm intelligence can be described as the collective behavior of self-organized and decentralized systems. In the context of the V2X paradigm, swarm intelligence is represented by a collection of vehicle agents interacting locally with each other and with their environment. Vehicles follow simple rules without a central control system. A result, we found that ant colonies collectively had a high probability of choosing the shortest route. The most common use cases for swarm intelligence are particle swarm optimization (PSO), ant colony optimization (ACO), swarm casting.

- Machine Learning

  ML covers most of AI. ML methods can be classified into three types: unsupervised learning, supervised learning, and reinforcement learning. There are several other types of ML schemes such as transfer learning and online learning that can be classified in terms of these three basic ML schemes. ML basically consists of two important phases: training and testing. The model is trained during the training phase based on realistic data. During the testing phase, predictions are made based on the trained model.

- Deep Learning

Deep learning is closely related to the three categories of ML mentioned above. This is a deeper network of neurons in several layers. It aims to extract knowledge from data representations that can be generated from the three categories of ML. A network consists of an input layer, a hidden layer, and an output layer. Each neuron has a nonlinear transform function such as ReLU, Tanh, Sigmoid, or Leaky-ReLU. Scaling the input data is very important as it can severely affect the prediction or classification of the network. As the number of hidden layers increases, so does the ability of the network to learn. However, beyond a certain point, adding more hidden layers does not improve performance. Training deeper networks is also difficult as it requires huge computational resources and the gradients of the network can explode or vanish. The deployment of these resource-intensive, deeper networks has increased the importance of edge computing technologies. Vehicles on the move can benefit from mobile edge computing servers. Also called multilayer perceptron (MLP).

- Expert Systems

  Expert systems mimic human decision-making abilities. Expert systems solve complex problems through reasoning extracted from human knowledge. This reasoning is expressed by his IF-THEN rule instead of procedural coding. The expert system is divided into his two parts, the knowledge base and the inference engine. Knowledge base: The knowledge base consists of rules derived from human knowledge. Inference Engine: The Inference Engine applies rules extracted from the knowledge base to known facts to derive new facts. It may also include explanation and debugging skills. The inference engine also has his two modes, forward chain and backward chain.

- Planning Scheduling Optimization

  This deals with implementing strategies or courses of action for execution by intelligent agents. Planning is also related to decision theory, and unlike traditional control and classification problems, complex solutions must be found from an n-dimensional space in an optimized way. Planning can be done offline in a known environment using available models and solutions evaluated prior to execution. In the highly dynamic and partially known environment of the V2X paradigm, strategies must be modified online. Models and associated policies should be adjusted accordingly.

This provides a comparative overview of relevant AI algorithms applied to the Vehicle-to-Everything paradigm. We introduced various AI technologies. AI-driven algorithms for V2X applications have improved performance over traditional algorithms. In general, all optimization problems face the uncertainty of surrounding participants' intentions. Different departments of AI can help each other find the best solutions that don't cause or generate problems in unintended areas. In most cases, AI algorithms require higher computing resources. These resources may not need to be in the vehicle. V2X, MEC and VEC technologies, AI computational load can be offloaded to edge compute servers located in nearby roadside infrastructure.

PAPER 5:

Reinforcement Learning for Optimizing Driving Policies on Cruising Taxis Services

As an important element of urban transportation, taxi services bring great convenience. However, reality is that it is not very efficient. Researchers have mainly aimed at optimizing policies by order dispatch for ride-hailing services, which are not applicable to ride-hailing services. In this paper, we developed a reinforcement learning (RL) framework to optimize driving policies for a cruise taxi service. First, we formulated a driver's behavior during the Markov decision process (MDP), considering the effects after the driver takes long-term actions. An RL framework with dynamic programming and data augmentation was used to compute the state action value function. According to the value function, drivers can make the best choices and quantify expected future rewards in a given state.

This white paper focuses on optimizing cruise taxi driving policies to increase driver benefits and opportunities for driver-passenger matching. Various goals aimed at increasing the efficiency of taxi services: B. Minimizing waiting times, minimizing distances between drivers and passengers, maximizing expected profits and probability of finding potential passengers. While these tasks seemed simple and effective, they ignored the spatio-temporal connection between cars and orders and yielded sub-optimal results in the long run. To build a taxi recommendation system, four factors were studied, including average waiting time, distance, fare, and transition probability for recommending potentially congested locations. The results showed that distance and latency to the next cruise point were the most important characteristics. Taxi monitoring methods have been presented to detect anomalous driving behavior and improve taxi services. A spatio-temporal data mining framework has been proposed to study mobility dynamics using taxi trajectories. This framework aimed to efficiently and timely forecast travel demand in different regions of the city. These works have helped us to better understand the factors that improve taxi service. Recent studies incorporating information and communication technology (ICT) tools have shown that reinforcement learning (RL) can achieve surprising results.

A new model-free deep reinforcement learning framework was defined to solve the taxi dispatch problem by predicting the number of assigned taxis. Numerical studies have shown that models without models perform better. Investigations using many Deep His learning models (Deep Q Network (DQN), Policy Gradients, Actor-Critic, etc.) are proposed to optimize the dispatch between drivers and orders. They provided reliable optimization guidelines and expected excellent results in real-world applications across the studied historical taxi trajectory and order datasets. However, this work is primarily focused on optimizing ride-hailing platform services.

In this paper, we propose a new method, classified as reinforcement learning, to optimize the driving policy of a tour taxi service to maximize the driver's profit and increase the driver-passenger match rate. By formulating the behavior of the driver of the MDP process, his TD update rule was applied using dynamic programming and data augmentation to compute the value of the state action function. The method is model-free and globally optimal, avoiding optimal policies becoming suboptimal. Results showed that the proposed method has reasonable improvements over other benchmarking methods that provide useful driver and system strategies. We also performed a sensitivity analysis to see how effectively the proposed method performed in different taxi number scenarios. We can draw the following main conclusions: However, ring roads have a higher value for action services than urban areas. • Optimal results are useful for studying transportation patterns in urban traffic. • The proposed RL model proved to be realistic for improving the overall profitability of the platform and increasing the order response rate. • Benefits over other approaches can be quantified through return differential and sensitivity analyses. However, this method can be further improved by examining more detailed datasets. Performance can be improved with more temporal and spatial features. Additionally, RLs for driving services and taxi dispatch can be developed in the future.

PAPER 6:

Optimize taxi driving strategies based on reinforcement learning

The efficiency of taxi services in big cities not only affects people's comfort, but also the benefits of city traffic and taxi drivers. To balance the demand and supply of taxis, spatio-temporal knowledge from past trajectories is recommended for both passengers finding available taxis and taxi drivers estimating the location of the next passenger. However, taxi trajectories are long sequences and single-step optimization cannot guarantee global optimization. Aiming for long-term profitability, we propose a new method based on reinforcement learning to optimize taxi driving strategies for global profit maximization. This optimization problem is formulated as a Markov decision process for the entire taxi driving sequence. The states set in this model are defined as the location and status of taxis. The set of actions includes the operational options idle, passenger transport or hold, and subsequent driving behavior. Reward as an objective function for evaluating driving policy is defined as the effective driving rate, which measures the taxi driver's total profit on the working day. The best choice for taxi his driver at each location is learned by Q-learning algorithm with maximum cumulative reward. Experiments were conducted to test the accuracy and efficiency of the method using historical trajectory data in Beijing. The results show that this method improves the profit and efficiency of taxi drivers and also increases the chances of passengers finding taxis. By replacing the reward function with other criteria, this method can also be used to discover and study new spatial patterns. This new model has no prior knowledge, is globally optimal, and outperforms previous methods.

Reinforcement learning is a major approach to learning control strategies by considering the actions an autonomous agent should take to maximize its numerical reward signal. Reinforcement learning focuses on trial-and-error interactions between a target agent and a dynamic environment to learn the optimal sequence of actions with maximum cumulative reward. Reinforcement learning problems are usually formalized as Markov decision processes (MDPs) where the state transition process satisfies Markov properties. The environment is in a certain state with a set state. After the agent receives the current state input, executing an action from the action set changes the environment to the new state and determines the associated reward. Agents choose the appropriate action based on the new status. By maximizing the expected cumulative reward in each independent state of such iterations, the agent finds a set of actions that use the globally maximized cumulative reward as the optimal guideline. At each time step, the agent chooses an action based on its policy. An agent's policy π is a mapping from states s to probabilities of choosing an action a, denoted by

$$\pi(a|s) = P[A_t = a|S_t = s]$$

$$q_\pi(s,a) = E_\pi[G_t|S_t = s, A_t = a] = E_\pi\left[\sum_{k=0}^{\infty} \gamma^k R_{t+k+1}|S_t = s, A_t = a\right]$$

$$G_t = R_{t+1} + \gamma R_{t+2} + \gamma^2 R_{t+3} + \ldots = \sum_{k=0}^{\infty} \gamma^k R_{t+k+1}$$

$$q_\pi(s,a) = E_\pi[R_{t+1} + \gamma q_\pi(s',a') | S_t = s, A_t = a]$$

$$q_*(s,a) = \max_\pi q_\pi(s,a).$$

$$q_*(s,a) = \sum_{s',r} p(s',r|s,a)\left(r + \gamma \max_{a'} q_*(s',a')\right),$$

$$R(s,a) = E[R_{t+1}|S_t = s, A_t = a]$$

$$\hat{q}(s,a) \leftarrow \hat{q}(s,a) + \eta\left(R(s,a) + \gamma \max_{a'} \hat{q}(s',a') - \hat{q}(s,a)\right)$$

In order to balance taxi demand and supply and achieve maximum profit for taxi drivers, we propose a new method based on reinforcement learning to optimize taxi operation strategies. This optimization process is modeled as an MDP and solved by a Q-learning algorithm, using the cumulative effective drive rate as a reward. Past trajectories are used to determine the best choice of taxi driver to drive dry, carry passengers, or wait anywhere. By replacing the reward function with other criteria, the reinforcement learning method can make new operational decisions and also discover some new spatial patterns. This new model is superior to previous methods because it has no prior knowledge and is globally optimal. This method explores the potential of reinforcement learning for spatial decision support and intelligent transportation. However, this method can be further improved by examining some detailed features. A multi-agent approach can be introduced to model the interaction between taxis by distinguishing between different trajectories. Additionally, time parameters can be integrated into the model to dynamically adapt the optimal strategy. High-performance computing techniques are also needed to improve the efficiency of reinforcement learning algorithms for big data.

PAPER 7:

Towards Optimal Control of Air Handling Units using Deep Reinforcement Learning and Recurrent Neural Network

Optimal control of heating, ventilation, and air conditioning (HVAC) systems is aimed at minimizing equipment energy consumption while maintaining occupant thermal comfort. Traditional rule-based control methods are not optimized for his HVAC system with continuous sensor readings and actuator control. Recent developments in deep reinforcement learning (DRL) have enabled him to control the HVAC with continuous sensor inputs and actions while eliminating the need to create complex thermodynamic models. DRL controls include an environment close to his actual HVAC operation. An agent aimed at realizing optimal control of HVAC. Existing DRL control frameworks use simulation tools to create a DRL training environment with information about the HVAC system, but oversimplify the building geometry. In this study, we aim to achieve optimal control of the air handling unit (AHU) by implementing a Long-Short-Term-Memory (LSTM) network to approximate real-world HVAC operation and DRL construction training environment. This framework also implements state-of-the-art his DRL for optimal control of AHU. Three of his AHUs, each with two years of her building automation data (BAS), were used as the test bed for the evaluation.

This paper proposes a novel framework that combines LSTM and DRL to achieve optimal control of AHUs. The proposed DRL approach eliminates the need of engineering experience from the facility managers to decide static set-points for HVAC systems. Furthermore, they proposed a new method to create DRL training environments by using LSTM networks to approximate HVAC operational parameter values. LSTM-based training environments can provide a more accurate approximation of the real-world AHU operations using only historical BAS data when AHU schematics and other building information are unavailable. Moreover, even with the information readily available, simulations tools demand extensive parameter tuning during the calibration process, which requires extensive knowledge of thermal dynamics. Finally, using real BAS data provides the possibility of comparing the optimal policy achieved by the DRL agent to the actual current rule-based policy implemented in the buildings. They evaluated the framework using 3 AHUs of from a real environment. Our LSTM-based training environment achieved an average MSE of 0.0015 for all AHU parameters for 3 AHUs. In test deployments, trained DRL agents achieved an average 27% to 30% reduction in energy consumption over 3 months with 3 AHUs and improved thermal comfort during weekday occupied hours with 3 AHUs. Maintained level (mean PPD 10%). Thermal Meets the 5-20% PPD recommended by ASHRAE for building comfort. Finally, the proposed framework can be scaled to any HVAC system using historical sensor readings. This has become more common these days due to the widespread deployment of Internet of Things and accumulated BAS data.

Pre-trained DRL agents are perfectly tuned to our own HVAC system and provide optimal control given enough training data and training episodes.

PAPER 8:

Optimal charging of an electric vehicle using a Markov decision process

The combination of electric vehicles and renewable energy has emerged as a potential driver for a fossil fuel-free future. However, managing an electric vehicle fleet efficiently is not without its challenges. It seeks the participation of all stakeholders directly or indirectly associated with the energy and transport sectors, from governments, vehicle manufacturers, transmission system operators, to end consumers who are the ultimate beneficiaries of change. Since electric vehicles should be used primarily to meet driving needs, charging policies should be designed primarily for this purpose. However, the billing model presented in the literature overlooks the stochastic nature of driving patterns. Here, they present an efficient stochastic dynamic programming model for optimally charging an electric vehicle, taking into account the uncertainties associated with using it. Driving patterns are described by a non-uniform Markov model fitted to data from electric vehicle usage. We show that the randomness inherent in driving needs has a significant impact on the charging strategies implemented.

Electric vehicles (EVs) are emerging as a sustainable and environmentally friendly alternative to traditional vehicles when the energy used to charge them comes from renewable sources. However, the energy obtained from renewable sources such as sunlight, wind and waves is weather dependent. Therefore, power generation from these sources is inherently uncertain in terms of timing and quantity. Moreover, large-scale storage of generated energy is still very limited today, so electricity must be produced and consumed at the same time. As a result, energy generated from renewables can be wasted during times when electricity demand is not high enough to meet it, negatively impacting the profitability of renewables. Since electric vehicle batteries are essentially energy stores, the large-scale integration of electric vehicles into the transport sector can significantly increase the socio-economic value of energy systems with high renewable components. Dependence on liquid fossil fuels.

In this paper, we propose an optimal charging algorithm for electric vehicles considering the uncertainty of the user's driving behavior. The algorithm is based on a heterogeneous (hidden) Markov chain model, which provides vehicle usage probabilities at any time of the day and captures different travel times. It then uses stochastic dynamic programming to determine the optimal charging policy based on vehicle usage, end-user risk aversion, and electricity prices. The proposed billing model fits a training dataset that spans about two and a half months.

This survey will firstly be conducted with the following objectives:

(i) examine the impact of termination conditions on the resulting pricing policy,
(ii) highlight the benefits of a rolling horizon implementation of the algorithm,
(iii) model outputs that assess the impact of unavailability costs on services, and
(iv) Demonstrate the operation of electric vehicles in vehicle-to-grid mode.

Model performance is then evaluated off-sample using test datasets spanning the next two months. More precisely, the daily cost savings range is about 19-47% with different heuristic billing strategies. Running costs can be further reduced if vehicles can feed the power grid. In fact, the numerical results show that running costs can be converted into net profit under an optimized V2G scheme, and maximum Save 135%. The proposed probabilistic dynamic programming model for EV charging is versatile and can be easily adapted to each specific vehicle, offering customized charging policies to help EV users save operating costs or even make a profit per vehicle. A possible extension is to apply the proposed model to more Markov state data. This can be used to study the benefits of having more public charging stations versus charging at home, or to capture different driving states such as 'city' , "Countryside" or "Highway". As for practical applications, the proposed Markov decision model can be extended to handle time adaptively estimated transition probabilities. Adaptability is key to capturing structural changes in the driving behavior of EV users B. What may result from an EV user purchasing another vehicle or moving to a new location. Further research may also be directed to modeling fleets using mixed-effects models. The optimization scheme can be applied to each vehicle individually and the total occupancy load can be evaluated. This will show if and how EVs can be deployed to mitigate the increase in peak power demand when switching from ICE vehicles to EVs.

PAPER 9:

A Multi-Agent System for Acquired Brain Injury rehabilitation in Ambient Intelligence environments

Acquired brain injury (ABI) is becoming a prevalent problem in our society, especially among the elderly, known as the 'silent epidemic'. People with ABI seek solutions that provide a retraining process so that they can regain not only their physical abilities but also their cognitive abilities. The inherent transparency and intelligence of Ambient Intelligence (AmI) makes ABI one of the best approaches to combat the disruption it can cause. As AmI proposes the development of situational awareness systems that integrate various devices to perceive situations and act accordingly, these systems will meet the needs of people with ABI as they carry out the rehabilitation process can respond immediately. Moreover, leveraging a multi-agent architecture has proven to be a natural solution for developing AMI systems. This is because agents are reactive and proactive, exhibiting intelligent and autonomous behavior. Therefore, a multi-agent architecture (MAS) for AMI systems in healthcare is presented. Using specific devices to monitor patient movement and physiological responses contributes to the treatment of patients with ABI B. Controlling heart rate variability during the rehabilitation process. In this way, the natural relationship between AMI and MAS is exploited. Finally, they demonstrate how this system can be used to design and deliver treatments for people with ABI.

Here a system is presented for monitoring treatment delivery in an ambient intelligence environment using a multi-agent system. This system was designed to provide rehabilitation physical therapy to people with ABI. These treatments are designed from the ground up by therapists, defining not only the basic activities, but also the rules that govern the operation of the system. Their design is based on a meta-model that defines the rehabilitation environment and the ability to define a fuzzy reasoning system that allows the system to coordinate treatments and determine the execution order of activities, tasks and steps at runtime. MAS was developed to support these functions. Use AmI to enrich the environment, improve decision-making, and provide treatments tailored to environmental conditions and people using the system. Therefore, MAS monitors environmental characteristics: oxygen levels, stress, heart rate, and the person's current emotional state. All of these characteristics are captured using dedicated hardware that provides the necessary data for MAS. Additionally, the system provides therapists with feedback from performance statistics so that they can monitor treatment delivery and patient progress. The main advantage of our proposal comes directly from the design of the MAS architecture itself. As mentioned, it is designed to provide an expandable system, so new sensors can be easily integrated to make new capabilities available on top of existing ones. Additionally, Vi-SMARt also seeks a better separation of concerns about the inference process to facilitate integration with other inference engines. One that supports neural networks. Moreover, our system decouples treatment regulation from treatment delivery, allowing the regulation mechanism to be changed without changing treatment delivery at all, and vice versa.

PAPER 10:

Incentive-based demand response for smart grid with reinforcement learning and deep neural network

Balancing power generation and consumption is essential to smoothing the power grid. Any mismatch between energy demand and supply increases costs for both service providers and customers, and can even bring down entire networks. This paper proposes a novel real-time incentive-based demand response algorithm for smart grid systems using reinforcement learning and deep neural networks. It is intended to smooth energy fluctuations and reduce energy fluctuations by purchasing energy resources from service providers subscribed customers. Improve network reliability. In particular, they use deep neural networks to predict unknown prices and energy demand to overcome future uncertainties. Reinforcement learning is then applied to obtain the best incentive rates for different customers, taking into account the interests of both service providers and customers. Simulation results demonstrate that this proposed incentive-based demand response algorithm can induce demand-side participation, drive profitability for service providers and customers, and improve system reliability by balancing energy resources indicates that this can be seen as a win-win strategy for service providers and operator customers.

This paper proposes a novel real-time incentive-based DR algorithm for smart grid systems with RL and DNN, aiming to assist the SP in purchasing energy resources from its various CUs to balance energy fluctuations and enhance grid reliability. Due to the in-herent nature of real-time electricity market, the SP can only access the price from wholesale electricity market and energy demand from its CUs for the current hour to overcome the future uncertainties, DNN is used to predict the unknown prices and energy demands. After that, RL is adopted to derive the optimal incentive rates for different CUs considering the profitabilities of both SP and CUs. By employing RL, the SP can adaptively determine the incentives without knowing a predefined complete model about how to select the incentive rates, but instead, it discovers the optimal incentive rates by learning them from direct interaction with the CUs. Simulation results show that this proposed incentive-based DR algorithm, can induce demand side participation, promote SP and CUs profitabilities, and improve system reliability by balancing energy resources, which can be regarded as a win-win strategy for both SP and CUs. In the future, further analysis will be performed towards the weighting factor to determine the optimal between the CU incentive income and incurred dissatisfaction cost from CU side. This work can be extended to a wholesale capacity resource trading framework involving the GO and multiple SPs. And, another interesting direction for extending the current work is to apply RL on both price and incentive-based DR to provide integrated DR solution for the smart grid systems.

Cite

For Papers,

1) Liang, E., Wen, K., Lam, W. H., Sumalee, A., & Zhong, R. (2021). An integrated reinforcement learning and centralized programming approach for online taxi dispatching. *IEEE Transactions on Neural Networks and Learning Systems*.
2) Okuyama, T., Gonsalves, T., & Upadhay, J. (2018, March). Autonomous driving system based on deep q learnig. In *2018 International conference on intelligent autonomous systems (ICoIAS)* (pp. 201-205). IEEE.
3) Rong, H., Zhou, X., Yang, C., Shafiq, Z., & Liu, A. (2016, October). The rich and the poor: A Markov decision process approach to optimizing taxi driver revenue efficiency. In *Proceedings of the 25th ACM International on Conference on Information and Knowledge Management* (pp. 2329-2334).
4) Tong, W., Hussain, A., Bo, W. X., & Maharjan, S. (2019). Artificial intelligence for vehicle-to-everything: A survey. *IEEE Access*, *7*, 10823-10843.
5) Jin, K., Wang, W., Hua, X., & Zhou, W. (2020). Reinforcement Learning for Optimizing Driving Policies on Cruising Taxis Services. *Sustainability*, *12*(21), 8883.
6) Gao, Y., Jiang, D., & Xu, Y. (2018). Optimize taxi driving strategies based on reinforcement learning. *International Journal of Geographical Information Science*, *32*(8), 1677-1696.
7) Zou, Z., Yu, X., & Ergan, S. (2020). Towards optimal control of air handling units using deep reinforcement learning and recurrent neural network. *Building and Environment*, *168*, 106535.
8) Iversen, E. B., Morales, J. M., & Madsen, H. (2014). Optimal charging of an electric vehicle using a Markov decision process. *Applied Energy*, *123*, 1-12.
9) Roda, C., Rodríguez, A. C., López-Jaquero, V., Navarro, E., & González, P. (2017). A multi-agent system for acquired brain injury rehabilitation in ambient intelligence environments. *Neurocomputing*, *231*, 11-18.
10) Lu, R., & Hong, S. H. (2019). Incentive-based demand response for smart grid with reinforcement learning and deep neural network. *Applied energy*, *236*, 937-949.

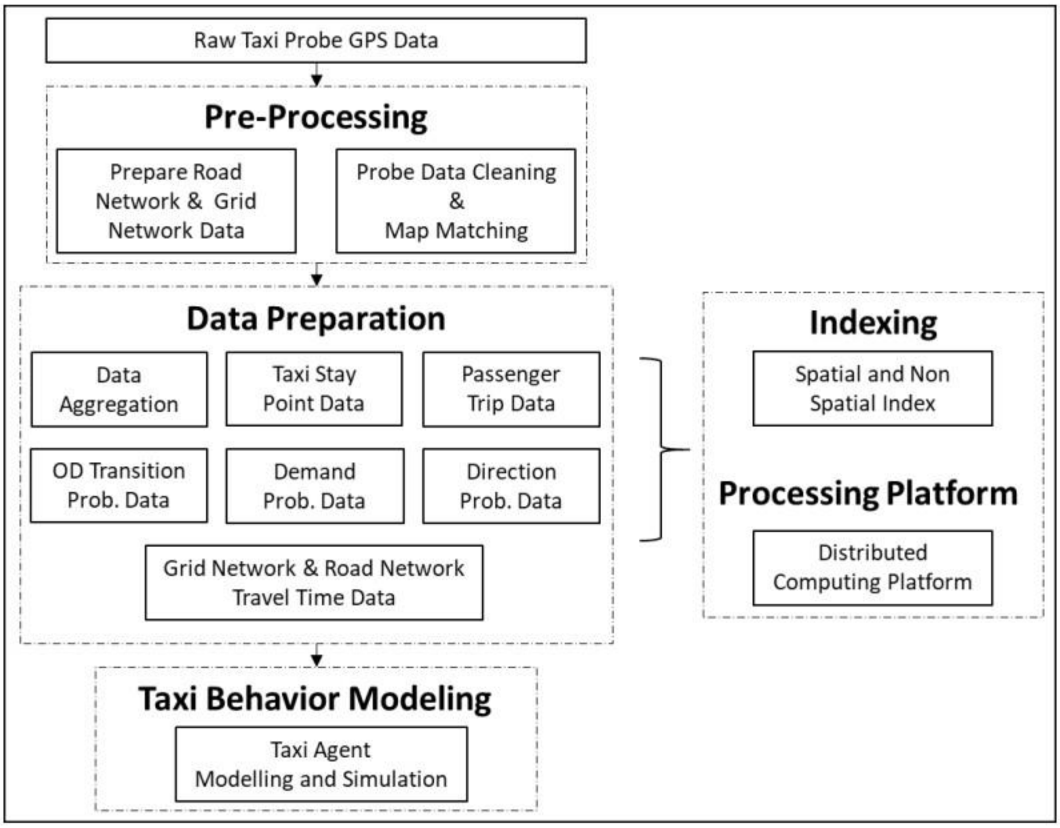
Artificial Intelligence
Literature Review
<u>AI Agent to help cab drivers maximize profit within a month</u>

Name : Lokeswar M
Reg.No. : 20BCE1825

Paper 1:
Agent-Based Modelling of Taxi Behaviour Simulation with Probe Vehicle Data

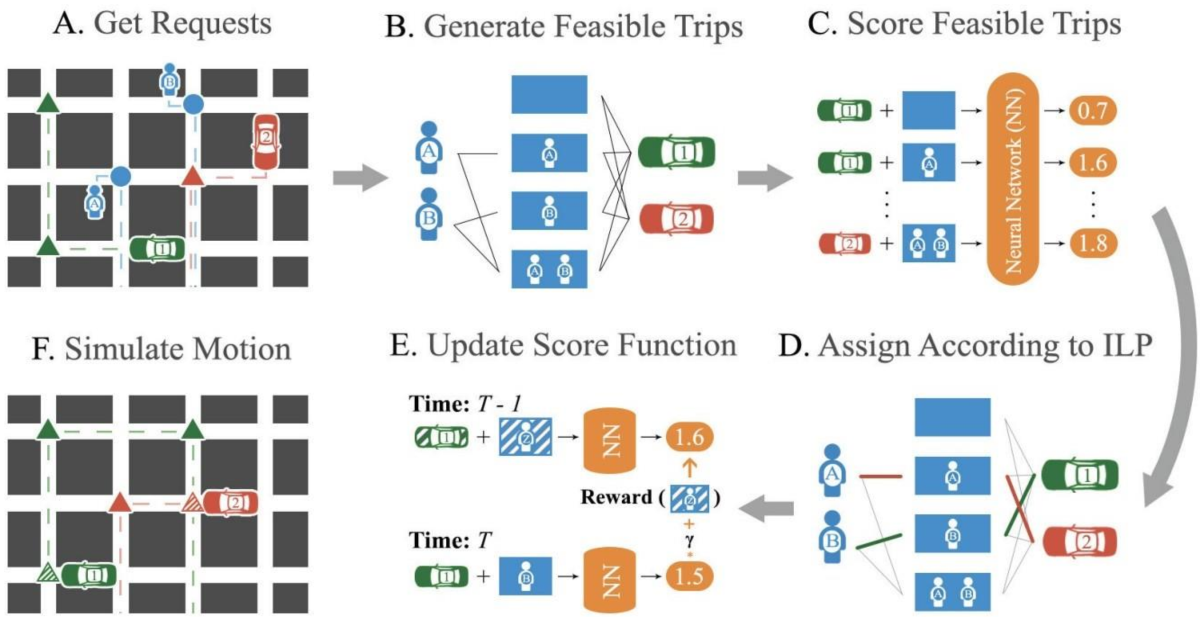
Taxi behaviour is a discrete time-dependent spatiotemporal dynamic process
Events such as customer attraction, drop off, cruising, and parking. simulation model,
This is a simplification of the actual system and is useful for understanding the impact of changing the system
Dynamic behaviour. In this paper, agent-based modelling and simulation are proposed.
Determined by the agent's dynamic behaviour, i.e. behavioural rules and characteristics that mimic a taxi
taxi behaviour. A taxi behavioural simulation is always running to optimize the service
Both taxi driver and passenger levels. In addition, simulation techniques can be such
It can also be transferred to another application area where getting the actual raw data is somewhat difficult
B. people's movement data or call detail record data. This paper describes
the development of an agent-based simulation model which is based on multiple input parameters (taxi stay point cluster; trip information (origin and destination); taxi demand information; free taxi movement; and network travel time) that were derived from taxi probe GPS data. As such, agent's parameters were mapped into grid network, and the road network, for which the grid network was used as a base for query/search/retrieval of taxi agent's parameters, while the actual movement of taxi agents was on the road network with routing and interpolation. The results obtained from the simulated taxi agent data and real taxi data showed a significant level of similarity of different taxi behaviour, such as trip generation; trip time; trip distance as well as trip occupancy, based on its Distribution. As for efficient data handling, a distributed computing platform for large-scale data was used for extracting taxi agent parameter from the probe data by utilizing both spatial and non spatial indexing technique.
The overall system overview is shown in Figure 1, which has preprocessing, data preparation, and taxi behaviour modelling as its stages, and indexing and processing platform as its tool to handle the big data. The preprocessing stage consists of preparing for the grid network and road network, with conducting cleaning and map matching of the raw probe GPS data.

Paper 2:
Neural



Approximate Dynamic Programming for On-Demand Ride-Pooling

On-demand ride-pooling (e.g., UberPool, Lyft Line, GrabShare) has recently become popular because of its ability to lower costs for passengers while simultaneously increasing revenue for drivers and aggregation companies (e.g., Uber). Unlike in Taxi on Demand (ToD) services – where a vehicle is assigned one passenger at a time – in on-demand ride-pooling, each vehicle must simultaneously serve multiple passengers with heterogeneous origin and destination pairs without violating any quality constraints. To ensure near real-time response, existing solutions to the real-time ride pooling problem are myopic in that they optimise the objective (e.g., maximise the number of passengers served) for the current time step without considering the effect such an assignment could have on assignments in future time steps. However, considering the future effects of an assignment that also has to consider what combinations of passenger requests can be assigned to vehicles adds a layer of combinatorial complexity to the already challenging problem of considering future effects in the ToD case. A popular approach that addresses the limitations of myopic assignments in ToD problems is Approximate Dynamic Programming (ADP). The existing ADP method for ToD is However, we treat linear programming (LP)-based assignments as follows: Updating the value depends on the double value from LP. The ride-pooling mapping problem requires integer linear programming (ILP) with poor LP relaxation. therefore, Our main technical contribution is to provide a generic ADP. How we can learn from the ILP-based mappings we find with ride pooling. In addition, it handles additional combinatorial complexities due to combinations of passenger requests using a neural network-based approximation function Hence, we show a connection to deep reinforcement learning You can learn this value function with increased stability and sample efficiency. Our approach proves to be far superior to the leading approaches to on-demand ride pooling. Up to 16% of real dataset, significant improvement When a traffic accident occurs in an urban area.



Paper 3:

Artificial Intelligence for Vehicle-to-Everything: A Survey

Recent advances in communications, intelligent traffic systems, and computer systems are opening up new possibilities for intelligent road safety, comfort, and efficiency. solution. Artificial intelligence (AI) is widely used to optimize traditional data-driven approaches in various fields of scientific research. By combining a Vehicle-to-Everything (V2X) system and AI, Receives information from various sources to enhance driver awareness and anticipate avoidance Prevent potential accidents and improve driving comfort, safety and efficiency. This paper Comprehensive overview of research papers addressing various research questions using AI V2X system. The contributions of this study have been summarized and classified accordingly to the application domain. Finally, we present open problems and research questions to be resolved. It aims to enhance his V2X system with the full potential of AI. This article provides a comparative overview of related products. AI algorithms applied to the Vehicle-to-Everything paradigm. We introduced various AI technologies. AI-controlled Algorithms for V2X applications have improved performance over traditional algorithms. In general, all optimization problems face uncertainty about intent. surrounding participants. Different areas of AI can help with this to bring together the best solutions does not cause or create problems within the domain is determined for AI algorithms almost always need more computing resources. You may not need these resources Get in the car Thanks to V2X, MEC and VEC Technology, AI computational load can be outsourced Edge computing servers in nearby roadside infrastructure. Future V2X applications will benefit greatly From the emerging field of edge computing.

Paper 3:

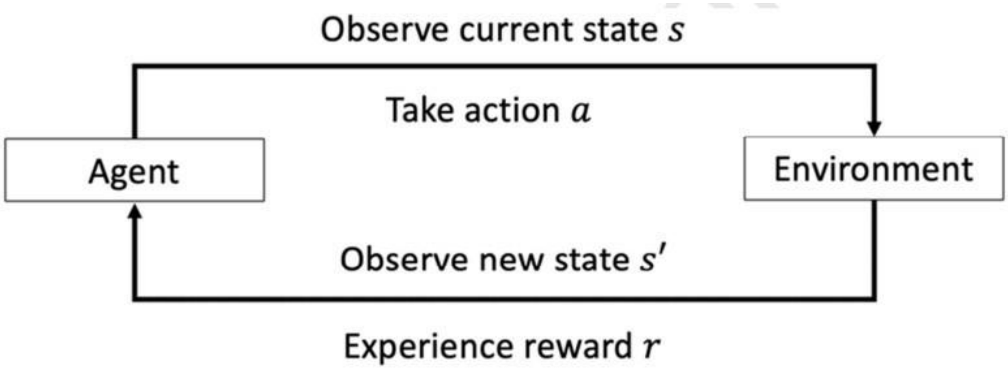
Reinforcement Learning for Optimizing Driving Policies on Cruising Taxis Services

As an important element of urban transport, taxis make a significant contribution Convenience and comfort for resident travel. However, the reality is not very efficient. Until now, researchers have mainly aimed at optimizing policies by order dispatch in ride-hailing services. Not applicable for taxis. In this work, reinforcement learning was developed (RL) Framework for optimizing driving guidelines for cruising taxi services. first formalized Driver behaviour during the Markov decision process (MDP) considering subsequent effects act long-term. RL framework with dynamic

programming and data augmentation Used to compute the state action value function. According to the value function, the driver You can determine the optimal choice and quantify the expected future reward in a given state. Historical order data from Chengdu was used to analyse the spatial distribution of function values We have shown how the model can optimize driving guidelines. Finally, a realistic simulation An on-demand platform was built. Results are validated against other benchmarking methods New model outperforms in overall sales, higher response rates, and lower wait times time. Relative percentages are up to 4.8%, 6.2%, and -27.27%.
This paper proposed a new method, classified as reinforcement learning, for optimizing driving Ride-on Taxi Service Guidelines to Maximize Driver Profits and Increase Compliance Allocation of drivers and passengers. TD is updated by formulating the driver behaviour of the MDP process Rules using dynamic programming and data expansion were applied to compute state actions. function value. The proposed method is model-free and globally optimal, avoiding this optimization The policy is suboptimal. This paper uses historical order data to recommend optimal choices Serving drivers by idling at specific locations and times. A simulation environment was also included. Developed to perform detailed modelling of the physical world for on-demand platforms. Result Noting that the proposed method shows reasonable improvements over other benchmarking methods, This provided a useful strategy for drivers and systems. We also performed a sensitivity analysis We examine how effectively the proposed method works in different taxi number scenarios. From the case studies and simulation results, the main conclusions are:

• Drivers generally benefit more in quiet urban areas.
Suburbs. However, ring roads have a higher value for action services than urban areas.
• Optimal results are useful for studying transportation patterns in urban traffic. • The proposed RL model proved realistic in improving the total revenue of the platform as a whole.
Increase your order response rate.
•Benefits over other approaches can be quantified by the difference in returns.
sensitivity analysis.

Paper 4:
Towards Optimal Control of Air Handling Units using Deep 2 Reinforcement Learning and Recurrent Neural Network
Optimal control of heating, ventilation, and air conditioning (HVAC) systems is aimed at minimizing equipment energy consumption while maintaining occupant thermal comfort. Traditional rule-based control methods are not optimized for his HVAC system with continuous sensor readings and actuator control. Recent developments in deep reinforcement learning (DRL) have enabled him to control HVAC with continuous sensor inputs and actions while eliminating the need to create complex thermodynamic models. DRL controls include an environment close to his real HVAC operation. An agent aimed at realizing optimal control of HVAC. Existing DRL control frameworks use simulation tools (e.g. EnergyPlus) to create a DRL training environment with information about the HVAC system, but oversimplify the building geometry. In this study, we aim to achieve optimal control of the air handling unit (AHU) by implementing a Long-Short-Term-Memory (LSTM) network to approximate real-world HVAC operation and DRL construction training environment. Suggest a target framework. This framework also implements state-of-the-art his DRL algorithms (such as deep deterministic policy gradients) for optimal control of AHU. Three AHUs, each with two years of building automation system 'BAS' data, were used as a testbed for the evaluation. The LSTM-based his DRL training environment built using first-year BAS data achieved a mean squared error of 0.0015 on 16 normalized AHU parameters. When deployed in a test environment built using his BAS data from his second year of the same AHU, the DRL agent achieved energy savings of 27% to 30% compared to actual energy consumption. , the predicted percentage of discomfort remained at 26 (PPD). % at 10 o'clock.



. Interactions between an agent and an environment in RL

Paper 5:
Optimize taxi driving strategies based on reinforcement learning

The efficiency of taxi services in big cities is Urban transport and benefits as well as the convenience of people's mobility Taxi driver. To balance the demand and supply of taxis, Spatio-temporal knowledge obtained from historical trajectories Both passengers are advised to find an available taxi A taxi driver estimates the location of the next passenger. However, Taxi Lane is a long sequence where he is one step. Optimization cannot guarantee global optimization. With the aim of generating long-term returns, new methods have been proposed based on. Reinforcement learning to optimize taxi driving strategies for global profit maximization. This optimization problem is formulated As a Markov decision process for the entire taxi driving sequence. The states specified in this model are defined as taxi locations, working state. Action sets contain the following operation options: Running in the sky, carrying people, waiting, etc. Reward as an objective function for evaluating driving behavioural driving guidelines is defined as the effective driving rate. Measure a taxi driver's total profit for the day. Of Best Choice for Taxi Drivers Everywhere Q-learning algorithm with maximum cumulative reward. Use Experiments were conducted to test the accuracy and efficiency of this method on historical trajectory data in Beijing. Of Results show that this method improves profit and efficiency Increases chances of being found by taxi drivers and passengers taxi too. By replacing the reward function with other criteria, This method can also be used to discover and explore new spatial pattern. This new model requires no prior knowledge, This has advantages over the previous method.

Paper 6:

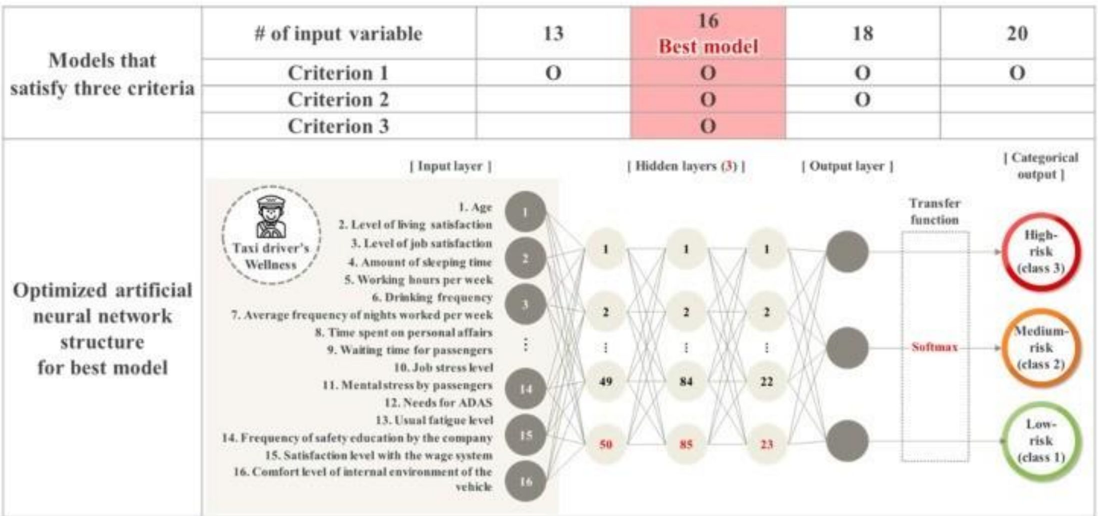
Deep-Learning-Based Prediction of High-Risk Taxi Drivers Using Wellness Data

To identify at-risk drivers based on human factors, factors related to the health status of taxi drivers are important. The purpose of this study is to predict high-risk taxi drivers based on deep learning techniques by identifying driver well-being that reflects the driver's personal characteristics. Method: In-depth interviews are conducted with taxi drivers to collect health data. A prioritization of factors influencing accident severity is derived by a random forest model. In addition, based on the derived variable priorities, various combinations of inputs are set as scenarios, and optimal artificial neural network models are derived for each scenario. Finally, the best model for predicting risky taxi drivers is selected based on three criteria. Result: The model with up to priority 16 variables as input is selected as the best model. The classification accuracy is 86% and the F1 value is 0.77. Conclusion: The health-based model for predicting high-risk taxi drivers presented in this study can be used to develop taxi driver management systems. It is also expected to help establish measures to improve traffic safety tailored to commercial vehicle drivers.

- Criterion (1) The F1-score is 0.77 or higher.

- Criterion (2) The classification accuracies of each class are 53% or higher.

- Criterion (3) Scenarios that satisfy criteria 1 and 2, with fewer input variables.



Paper 7:

Multi-Agent Coordination: A Reinforcement Learning Approach

Multi-agent coordination: Reinforcement learning approaches provide a comprehensive, insightful and unique process for developing multi-robot coordination algorithms with minimal computational effort and reduced memory footprint compared to traditional algorithms. Skilled scholars, engineers, and authors provide the reader with both a general introduction and overview of multi-robot coordination, as well as a detailed analysis of learning-based planning algorithms. Learn alternative approaches to accelerate TMAQL convergence by identifying ways to accelerate team goal exploration and desirable joint behaviours for teams. The authors also propose a new approach to consensus Q-learning that addresses the problem of equilibrium selection and a new method for evaluating empire consolidation thresholds without a significant computational effort. Finally, the paper concludes with a review of possible directions for future research in this rapidly developing field. Readers will discover state-of-the-art techniques for multi-agent coordination such as: An introduction to multi-agent coordination with reinforcement learning and evolutionary algorithms, including topics such as Nash equilibrium and correlation equilibrium Improve Convergence Speed of Q-Learning with Multiple Agents for Collaborative Task Scheduling Consensus Q-Learning for Multi-Agent Collaborative Planning Efficient Correlation Equilibrium Computation for Multiagent Scheduling Based on Collaborative Q-learning A Modified Imperial Competition Algorithm for Multi-Agent Applications Carrying Sticks Perfect for academics, engineers, and professionals who regularly use multi-agent learning algorithms, multi-agent tuning. Reinforcement learning approaches also belong on the bookshelf of anyone with a high degree of interest in machine learning and artificial intelligence in the field of collaborative or competitive robotics.

Paper 8:

Multiagent Reinforcement Learning-Based Taxi Pre dispatching Model to Balance Taxi Supply and Demand

With the improvement of people's living standards, people's demand of traveling by taxi is increasing, but the taxi service system is not perfect yet; taxi drivers usually rely on their operational experience or cruise randomly to find passengers. Without macro guidance, the role of the taxi system cannot be fully utilized. Many scholars have studied taxi behaviours to find better operational strategies for drivers, but their researches rely on local optimization methods to improve the profit of drivers, which will lead to imbalance between supply and demand in the city. To solve this problem, we propose a Multiagent Reinforcement Learning- (MARL-) based taxi pre dispatching model through analysing the running data of 13,000 taxis. Different from other methods of scheduling taxis based on the real-time location of orders, our model first predicts the demand for taxis in different regions in the next period and then dispatches taxis in advance to meet the future requirement; thus, the number of taxis needed and available in different regions can be balanced. Furthermore, to reduce the computational complexity, we propose several ways to reduce the state space and action space of reinforcement learning. Finally, the comparison of our method with other taxi dispatch methods shows that the proposed method significantly improves vehicle utilization and passenger demand satisfaction. In the future, we plan to make more detailed plans. Specifically, consider which taxis to dispatch to each grid, how to choose the route for each taxi, and where to find passengers once they reach a given grid. We will continue to solve these problems and improve the efficiency of taxi services.

Paper 9:

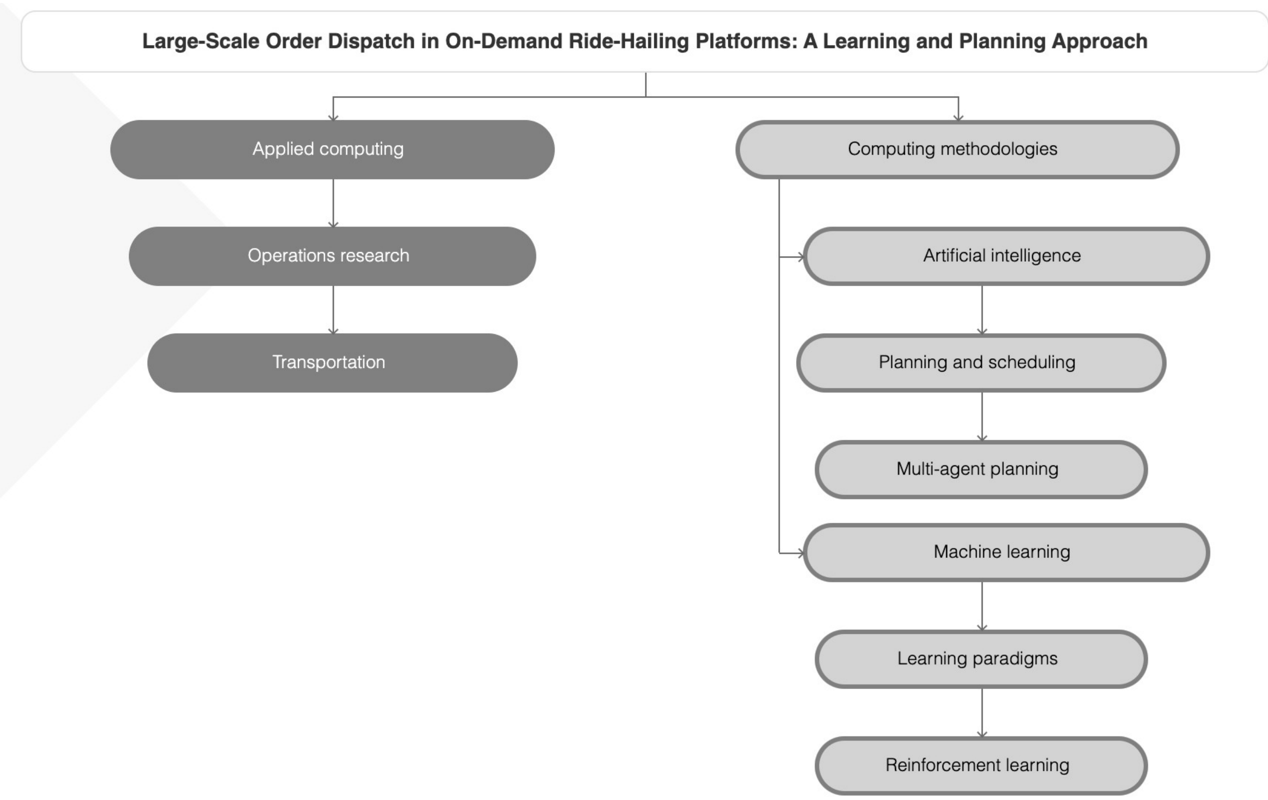
A predictive model for the passenger demand on a taxi network

In the last decade, the real-time vehicle location systems attracted everyone attention for the new kind of rich spatio-temporal information. The fast processing of this large amount of information is a growing and explosive challenge. Taxi companies are already exploring such information in efficient taxi dispatching and time-saving route finding. In this paper, we propose a novel methodology to produce online short term predictions on the passenger demand spatial distribution over 63 taxi stands in the city of Porto, Portugal. We did so using time series forecasting techniques to the processed events constantly communicated for 441 taxi vehicles. Our tests - using 4 months of real data - demonstrated that this model is a true major contribution to the driver mobility intelligence: 76% of the 86411 demanded taxi services were accurately forecasted in a 30 minutes time horizon.

Paper 10:

Large-Scale Order Dispatch in On-Demand Ride-Hailing Platforms: A Learning and Planning Approach

We present a novel order dispatch algorithm in large-scale on-demand ride-hailing platforms. While traditional order dispatch approaches usually focus on immediate customer satisfaction, the proposed algorithm is designed to provide a more efficient way to optimize resource utilization and user experience in a global and more farsighted view. In particular, we model order dispatch as a large-scale sequential decision-making problem, where the decision of assigning an order to a driver is determined by a centralized algorithm in a coordinated way. The problem is solved in a learning and planning manner: 1) based on historical data, we first summarize demand and supply patterns into a spatiotemporal quantization, each of which indicates the expected value of a driver being in a particular state; 2) a planning step is conducted in real-time, where each driver-order-pair is valued in consideration of both immediate rewards and future gains, and then dispatch is solved using a combinatorial optimizing algorithm. Through extensive offline experiments and online AB tests, the proposed approach delivers remarkable improvement on the platform's efficiency and has been successfully deployed in the production system of Didi Chuxing.



Citations:

Paper 1:

Ranjit, S.; Witayangkurn, A.; Nagai, M.; Shibasaki, R. Agent-Based Modeling of Taxi Behavior Simulation with Probe Vehicle Data. ISPRS Int. J. Geo-Inf. ***2018****, 7, 177.*
https://doi.org/10.3390/ijgi7050177

Paper 2:

Shah, Sanket, Meghna Lowalekar, and Pradeep Varakantham. "Neural approximate dynamic programming for on-demand ride-pooling." Proceedings of the AAAI Conference on Artificial Intelligence. Vol. 34. No. 01. 2020.

Paper 3:

Tong, Wang, et al. "Artificial intelligence for vehicle-to-everything: A survey." *IEEE Access* 7 (2019): 10823-10843.

Paper 4:

Zou, Zhengbo & Yu, Xinran & Ergan, Semiha. (2019). Towards Optimal Control of Air Handling Units using Deep Reinforcement Learning and Recurrent Neural Network. Building and Environment. 168. 106535. 10.1016/j.buildenv.2019.106535.

Paper 5:

Mingyue Xu, Peng Yue, Fan Yu, Can Yang, Mingda Zhang, Shangcheng Li, Hao Li. (2022) Multi-agent reinforcement learning to unify order-matching and vehicle-repositioning in ride-hailing services. International Journal of Geographical Information Science 0:0, pages 1-23.

Paper 6:

Lee S, Kim JH, Park J, Oh C, Lee G. Deep-Learning-Based Prediction of High-Risk Taxi Drivers Using Wellness Data. Int J Environ Res Public Health. 2020 Dec 18;17(24):9505. doi: 10.3390/ijerph17249505. PMID: 33353012; PMCID: PMC7766844.

Paper 7:

Sadhu, A.K. and Konar, A. (2020) Multi-agent coordination: A reinforcement learning approach, Wiley.com. Available at: https://www.wiley.com/en-in/Multi+Agent+Coordination:+A+Reinforcement+Learning+Approach-p-9781119699033 (Accessed: January 9, 2023).

Paper 8:

Yang, Y., Wang, X., Xu, Y., & Huang, Q. (2020). Multiagent Reinforcement Learning-Based Taxi Predispatching Model to Balance Taxi Supply and Demand. Journal of Advanced Transportation, 2020, 8674512. doi:10.1155/2020/8674512

Paper 9:

L. Moreira-Matias, J. Gama, M. Ferreira and L. Damas, "A predictive model for the passenger demand on a taxi network,"2012 15th International IEEE Conference on Intelligent Transportation Systems, 2012, pp. 1014-1019, doi: 10.1109/ITSC.2012.6338680

Paper 10:

Xu, Z., Li, Z., Guan, Q., Zhang, D., Li, Q., Nan, J., Liu, C., Bian, W., & Ye, J. (2018). Large-Scale Order Dispatch in On-Demand Ride-Hailing Platforms: A Learning and Planning Approach. In Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (pp. 905–913). Association for Computing Machinery.

CSE3013 -ARTIFICIAL INTELLIGENCE
Literature Review

NAME:AKSHAY KUMARAN
REG NO: 20BCE1433

1. AI, Skill, and Productivity: The Case of Taxi Drivers

Author: Kyogo Kanazawa,Hitoshi Shigeoka,Daiji Kawaguchi,Yasutora Watanabe
Publish date: October 2022
Issue:IZA Institute of Labor Economics
Link:https://docs.iza.org/dp15677.pdf

SUMMARY:

In the context of taxi drivers, we investigate the effect that artificial intelligence has on productivity. By suggesting routes along which it is anticipated that there will be a lot of demand, the AI we are studying helps drivers locate customers. The productivity gap between high- and low-skilled drivers is reduced by 14% thanks to AI's ability to shorten the cruising time, which only benefits low-skilled drivers. The finding suggests that AI's impact on human labor is more nuanced and intricate than the story of job displacement, which was the primary focus of previous research.And moreover there are few limitations to this study. First, even though we demonstrate that low-skilled drivers benefit from AI, the fact that even low-skilled drivers do not use it is a puzzle. However, the shorter time it takes to locate a taxi benefits customers. The market may expand, and social welfare may improve, to the extent that this increased convenience increases taxi demand.

2. Deep Reinforcement Learning: A Brief Survey

Author: Kai Arulkumaran,Marc Peter Deisenroth,Miles Brundage,Anil Anthony Bharath

Publish date: October 2022

Issue: IZA Institute of Labor Economics

Link: https://docs.iza.org/dp15677.pdf

SUMMARY:

Deep reinforcement learning (DRL) is ready to change the field of man-made reasoning (simulated intelligence) and addresses a stage toward building independent frameworks with a more significant level comprehension of the visual world. At present, deep learning is empowering reinforcement learning (RL) to scale to issues that were already recalcitrant, for example, learning to play computer games straightforwardly from pixels. DRL calculations are likewise applied to mechanical technology, permitting control strategies for robots to be advanced straightforwardly from camera inputs in reality. In this review, we start with a prologue to the general field of RL, then progress to the standards of significant worth based and strategy based techniques. Our review will cover focal calculations in deep RL, including the deep Q-network (DQN), trust area strategy advancement (TRPO), and offbeat benefit entertainer pundit. In equal, we feature the unique benefits of deep brain networks, zeroing in on visual comprehension by means of RL. To close, we depict a few momentum areas of examination inside the field.

3. Optimize taxi driving strategies based on reinforcement learning

Author: Yong Gao, Dan Jiang, Yan Xu
Publish date: April 2018
Issue: International Journal of Geographical Information Science
Link:https://www.researchgate.net/publication/324189753_Optimize_taxi_driving_strategies_based_on_reinforcement_learning

SUMMARY:

The effectiveness of taxi administrations in enormous urban communities impacts the comfort of people groups' movement as well as metropolitan traffic and benefits for cabbies. To adjust the requests and supplies of taxis, spatio-worldly information mined from verifiable directions is suggested for the two travelers finding an accessible taxi and cab drivers assessing the area of the following traveler. Be that as it may, taxi directions are long sequences where single-step enhancement can't ensure the worldwide ideal.

The state set in this model is characterized as the taxi area and activity status. The activity set incorporates the activity decisions of void driving, conveying travelers or pausing, and the subsequent driving ways of behaving. The prize, as the goal capability for assessing driving strategies, is characterized as the viable driving proportion that actions the complete benefit of a cab driver in a functioning day.

The ideal decision for cab drivers at any area is advanced by the Q-learning calculation with greatest total prizes. Using verifiable direction information in Beijing, the trials were led to test the precision and proficiency of the technique. The outcomes show that the strategy further develops benefits and proficiency for cab drivers and expands the open doors for travelers to track down taxis too. By supplanting the award capability with different standards, the technique can likewise be utilized to find and research novel spatial examples. This new model is earlier information free and around the world ideal, which enjoys upper hands over past strategies.

4.Decision Support for Agent Populations in Uncertain and Congested Environments

Author: Pradeep Varakantham , Shih-Fen Cheng , Geoff Gordon ,Asrar Ahmed
Publish date: April 2018
Issue:AAAI Conference on Artificial Intelligence
Twenty-Sixth AAAI Conference on Artificial Intelligence
Link:https://www.aaai.org/ocs/index.php/AAAI/AAAI12/paper/view/5030/5485
SUMMARY:

The large-scale issues of urban transportation and labor mobility, where resources are congested and movement is uncertain, serve as the impetus for this study. Even though the individual agents in these domains lack an identity and do not explicitly interact with other agents, they have an effect on other agents. Although there has been a lot of research into dealing with such implicit effects, the majority of it has assumed deterministic agent movements. Taking into account the agents' implicit interactions is the main obstacle in these problems. For instance, vehicles attempting to share a road are competing implicitly.

5. Optimizing Taxi Driver Profit Efficiency: A Spatial Network-Based Markov Decision Process Approach

Author: Xun Zhou , Huigui Rong , Chang Yang, Qun Zhang, Amin Vahedian Khezerlou , Hui Zheng, Zubair Shafiq, and Alex X. Liu
Publish date: January 2020

Issue:IEEE TRANSACTIONS ON BIG DATA
Twenty-Sixth AAAI Conference on Artificial Intelligence
Link:https://www.aaai.org/ocs/index.php/AAAI/AAAI12/paper/view/5030/5485

SUMMARY:

In large cities, the public transportation system relies heavily on taxi services. A significant issue facing society is how to improve taxi business efficiency. When recommending seeking routes, the majority of the most recent analytical approaches to this issue only considered how to maximize the likelihood of pickup, energy efficiency, or profit for the upcoming trip. As a result, they may not be optimal for the overall profit over an extended period of time due to ignoring potential passengers' choice of destination. For better driving directions, we propose a novel Spatial Network-based Markov Decision Process (SN-MDP) with a rolling horizon configuration.

We determine the best course of action for a vacant taxi in order to maximize profit in the near future using a set of historical taxi records and the current situation (such as the time and road segment). To prevent drivers from competing, we propose statistical models that use data to estimate the necessary time-variant parameters of SN-MDP. In addition, rather than focusing solely on income, we consider the cost of fuel when determining profit. Our proposed method outperforms baseline methods in every time slot, as demonstrated by a case study and a number of experimental evaluations on a real taxi dataset from a major Chinese city.

6. The Automation of the Taxi Industry – Taxi Drivers’ Expectations and Attitudes Towards the Future of their Work

Author: Christina Pakusch, Alexander Boden, Martin Stein & Gunnar Stevens
Publish date: September 2021
Issue: Computer Supported Cooperative Work (CSCW)
Link: https://link.springer.com/article/10.1007/s10606-021-09408-1

SUMMARY:

Shared autonomous vehicles (SAVs), according to advocates of autonomous driving, may eventually render taxi drivers obsolete. We conducted interviews with German taxi drivers to find out how they see the changes brought on by growing automation for their business's future. Our research sheds light on how taxi drivers' jobs might evolve in light of autonomous vehicles: Taxi drivers are certain that other aspects of their work, such as providing additional services and assistance to passengers, would constitute a limit to such forms of automation. However, this would probably entail a shifting role for the taxi drivers, one that is focused on the sociality of the work. Although SAVs could take over driving for standard trips, taxi drivers are also certain that these other aspects of their work would constitute a limit to such forms of automation. The significance of including taxi drivers in the co-design of future SAV services and taxis is illustrated by our findings, which also suggest design implications for tools that take into account various forms of assistance.

7. Optimizing Efficiency of Taxi Systems: Scaling-up and Handling Arbitrary Constraints

Author: Jiarui Gan,Bo An, Chunyan Miao
Publish date: May 2015
Issue:14th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2015)
Link:https://personal.ntu.edu.sg/boan/papers/AAMAS15pricing.pdf

SUMMARY:

In today's cities, taxi services are an essential component of public transportation. However, taxi services are inefficient in many cities due to their decentralized operation model. In addition, the decentralized nature of taxi services makes it difficult to analyze and regulate them. There are two significant drawbacks to current computational methods for maximizing taxi market efficiency: 1) they cannot be effectively scaled up; Secondly, they are unable to deal with complicated real-world market situations that necessitate taking into account additional scheduling constraints. To address the deficiencies, we propose two novel algorithms—FLORA and FLORA-A—in this paper. FLORA scales up more quickly than other algorithms because it uses techniques for convex polytope representation to create a compact representation of the strategy space of taxi drivers. By gradually expanding the strategy space, FLORA-A avoids enumerating the entire exponentially large pure strategy space. It is the first known method for maximizing the efficiency of the taxi system by dealing with arbitrary scheduling constraints. The results of the experiments indicate that the speed that FLORA provides has increased by orders of magnitude, and that taxi drivers' operational strategies have changed in response to various market conditions, indicating the necessity of using FLORA-A.

8. AGENT-BASED SIMULATION FRAMEWORK FOR THE TAXI SECTOR MODELING

Author: Josep Maria Salanova Graua , Miquel Estradab , Panagiotis Tzenosa , Georgia Aifandopouloua
Publish date: September 2018
Issue: The 9th International Conference on Ambient Systems, Networks and Technologies (ANT 2018)
Link: https://www.sciencedirect.com/science/article/pii/S1877050918303958/pdf?md5=6c035366cf977a1cd3cef2464fb6ee6d&pid=1-s2.0-S1877050918303958-main.pdf

SUMMARY:

In the majority of global cities, taxi services account for a significant portion of daily travel. These services are governed by a centralized authority, which typically sets policies for the taxi industry and keeps an eye on taxi service providers' performance. The development of models that comprehend the behavior of these markets is required in order to provide assistance to policymakers, fleet managers, and taxi drivers on an individual basis. The majority of models developed for analyzing the taxi market do not take into account the spatial distribution of both supply and demand for taxis because they are based on econometric measurements. The taxi market's operational characteristics can only be better understood by a small number of simulation models. A framework for the creation of agent-based taxi simulation models is presented in this paper. Its purpose is to assess policymakers, taxi fleet managers, and individual drivers regarding the definition of the optimal mode of operation and vehicle count.

9. Applications of Artificial Intelligence in Transport: An Overview

Author: Saeed Asadi Bagloee,Sohani Liyanage,Hussein Dia
Publish date:2 January 2019
Issue: Department of Civil and Construction Engineering; Swinburne University of Technology, Hawthorn, VIC 3122, Australia
Link:https://www.mdpi.com/2071-1050/11/1/189

SUMMARY:

The transportation industry is one of many that stand to benefit from the unprecedented opportunities presented by the rapid pace of advancement in artificial intelligence (AI). High-tech computational techniques that are modeled after the human brain are among the AI innovations. The goal of using AI in transportation is to solve problems like a growing demand for travel, CO2 emissions, safety issues, and the degradation of the environment. In the digital age, the possibility of addressing these issues in a more efficient and effective manner has increased due to the abundance of quantitative and qualitative data and AI. Artificial Neural Networks (ANN), Genetic Algorithms (GA), Simulated Annealing (SA), Artificial Immune System (AIS), Ant Colony Optimiser (ACO), Bee Colony Optimization (BCO), and Fuzzy Logic Model (FLM) are a few examples of AI techniques that are finding their way into the transportation industry.

For AI to be effective, it is necessary to have a solid understanding of the connections that exist between AI and data and transportation system characteristics and variables on In addition, it looks promising for transportation authorities to figure out how to use these technologies to quickly reduce congestion, increase the dependability of travel times for their customers, and boost the profitability and productivity of their most important assets. The topics of traffic management, traffic safety, public transportation, and urban mobility are the primary areas in which this paper provides an overview of the AI strategies that are utilized worldwide to address transportation issues. The difficulties and limitations of AI applications in transportation are discussed at the conclusion of the overview.

10. Improving Taxi Revenue using Reinforcement Learning

Author: Shahil Subham, Saurabh Singh, Anusha Sunil Kumar, Farheen Fatima, Geetha G
Publish date: August 2020
Issue:INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY (IJERT)
Link:https://www.ijert.org/improving-taxi-revenue-using-reinforcement-learning

SUMMARY:

The transportation system has undergone a number of improvements recently, and online taxi services like Uber, Ola, and others have significantly increased the use of taxis, which are an important part of urban transportation. Uber, Ola, and other aggregation systems rely on these advancements. was able to activate more taxis, resulting in increased availability and shorter wait times for customers and an improved customer experience. To provide a better customer experience, numerous studies were conducted from the perspective of the customer. However, the sole focus of this paper is on maximizing the driver's long-term revenue by enhancing performance from the driver's perspective by utilizing current and previous movement trajectories and trips.