International Journal of Research Publication and Reviews, Vol 4, no 12, pp 4522-4528 December 2023

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Dynamic Pricing Using Reinforcement Learning

B.V.S.S. Chandana, K. Kiranmai

GMR Institute of Technology, Rajam, India.

ABSTRACT—

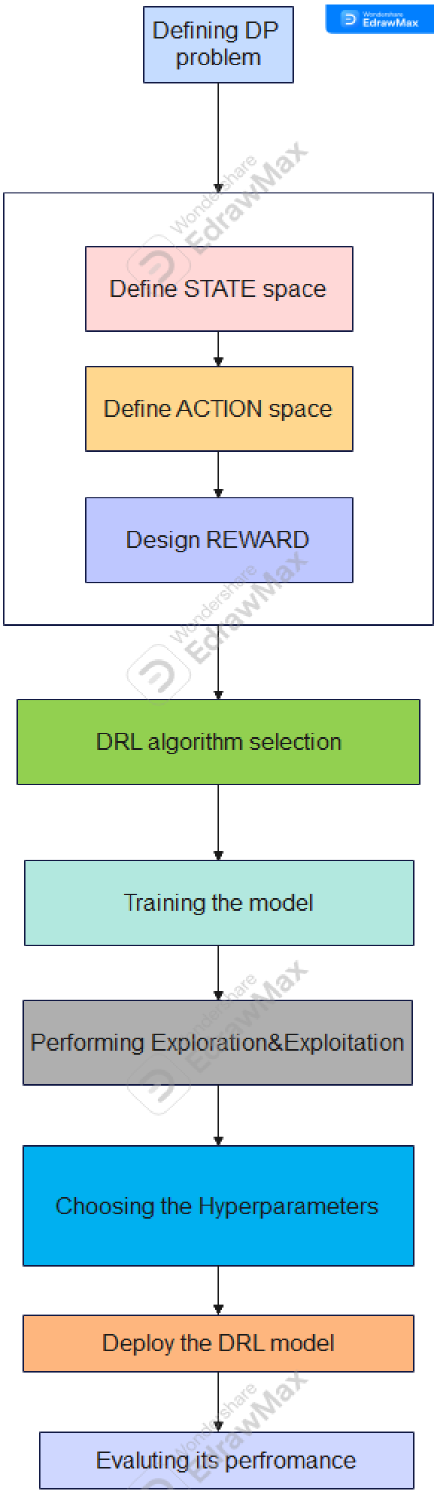
A dynamic pricing problem refers to the challenge of determining optimal prices for products or services in a constantly changing and unpredictable market environment. In this scenario, the demand for a product or service is often uncertain. Demand can vary due to factors like changing consumer behavior, external events, or market trends. Currently, most of the dynamic pricing schemes require full knowledge of the customer-side information. In last few years, researchers collected data for various problems such as, minimization of parking prices, maximization of revenue, managing and minimizing traffic congestion. This assumption helps to simplify complex real-world scenarios into manageable models. If the problem properties changes machine learning models can be trained on historical data to learn patterns and relationships. So, one of the methods is Deep reinforcement learning (DRL), it has a significant impact on dynamic pricing strategies by offering a flexible and adaptable approach to decision-making in complex and changing market environment. Despite the difficulties, effective dynamic pricing strategies can lead to increased revenue and company's ability to operate and thrive within markets.

Keywords— Machine learning algorithms, statistical modelling, and real time data analysis, Deep reinforcement learning (DRL) Framework, smart parking.

I. INTRODUCTION

Dynamic pricing (DP) problem involves deriving a pricing policy to maximize expected total revenue, but it is challenging due to the need to balance prices to attract buyers without negatively affecting revenue. Reinforcement learning (RL) is a suitable approach for solving DP problems due to the uncertainty and sequential decision-making nature of the problem. This proposes an RL framework based on deep reinforcement learning (DRL) to develop an efficient pricing strategy for DP problems. The proposed DRL pipeline automates the process of defining DRL components such as decision moments, Markov decision process (MDP) modeling, and setting hyperparameters. The pipeline starts with automatically defining states and reward function, followed by MDP formulation, algorithm selection, and hyperparameter optimization. The authors introduced a novel approach called Bayesian-genetic hyperparameter optimization, which integrates Bayesian optimization and the selection operator of the genetic algorithm. The integration of reinforcement learning (RL) algorithms in dynamic pricing represents a cutting-edge approach that leverages artificial intelligence to enhance pricing strategies. Unlike traditional static pricing models, dynamic pricing with RL allows businesses to adapt and optimize their pricing decisions in real-time based on changing market conditions, customer behavior, and other relevant factors. This introduction explores the transformative potential of RL algorithms, such as Deep Q Learning, Wolf PHC, and SAC, in reshaping the landscape of dynamic pricing. By employing these sophisticated techniques, businesses can strive for more agile and adaptive pricing strategies, leading to increased profitability and improved competitiveness in today's dynamic markets.

II. Related Work



1. State representation: The first step is to define a state representation that captures all of the relevant information about the current state of the system. This information could include things like the current inventory levels, the prices of competitors, and the historical demand data.

2.Reward function: The next step is to define a reward function that specifies the reward that the agent will receive for taking a particular action in a given state. The reward function should be designed in a way that encourages the agent to learn a pricing policy that maximizes the overall revenue.

3.DRL algorithm: The third step is to choose a DRL algorithm to train the agent. There are many different DRL algorithms available, each with its own strengths and weaknesses. Some popular DRL algorithms for dynamic pricing include Q-learning, policy gradients, and actor-critic methods.

4.Hyperparameter optimization: Once a DRL algorithm has been chosen, it is important to optimize the hyperparameters of the algorithm. Hyperparameters are parameters that control the behavior of the DRL algorithm, such as the learning rate and the exploration-exploitation trade-off.

5.Deployment: Once the agent has been trained, it can be deployed to the production environment. The agent will then use its learned pricing policy to set prices in real time.

III. LITERATURE SURVEYS

Maestre, R., Duque, J., Rubio, A., & Arévalo, J. (2019). Reinforcement learning for fair dynamic pricing. In *Intelligent Systems and Applications: Proceedings of the 2018 Intelligent Systems Conference (IntelliSys) Volume 1* (pp. 120-135). Springer International Publishing.

In this it mentioned about the importance of fairness in dynamic pricing to build trust with customers and avoid financial losses. It also highlighted how it also emphasizes the use of Reinforcement Learning (RL) in pricing optimization, but it also makes note of the need to balance revenue objectives with fairness and ethical considerations, with a focus on allocating prices equally among customer groups for improved fairness and trust.The emphasis on fairness in dynamic pricing reflects a broader understanding that customers value transparency and equitable treatment. Unfair pricing practices can erode customer trust, leading to reputational damage and potential long-term financial consequences. Acknowledging this, businesses are increasingly recognizing the need to strike a balance between revenue objectives and the ethical implications of their pricing strategies. The study underscores the use of RL not only as a powerful tool for pricing optimization but also as a means to incorporate fairness considerations into the decision-making process. RL algorithms, when trained on diverse datasets that represent various customer segments, have the potential to learn and adapt pricing policies that align with fairness objectives. This includes allocating prices equally among different customer groups, mitigating the risk of discriminatory practices that could alienate certain segments of the customer base.

Könönen, V. (2006). Dynamic pricing based on asymmetric multiagent reinforcement learning. *International journal of intelligent systems*, *21*(1), 73-98.

In this article, the use of asymmetric multiagent reinforcement learning techniques to resolve a dynamic pricing issue involving two rival brokers selling the same products is covered in the article. It mentioned the difficulties of using reinforcement learning in multiagent environments and mentions a number of strategies, such as Q-learning and policy gradient techniques. The article also provides background information on Markov decision processes and game theory, as well as an analysis of the effectiveness of the suggested learning techniques in resolving the price issue. The difficulties associated with applying reinforcement learning in multiagent environments are multifaceted. Traditional RL models may struggle to capture the strategic interactions and interdependencies that emerge when multiple agents, in this case, the rival brokers, are simultaneously adapting their pricing strategies. The competitive nature of the market introduces complexities that demand advanced learning techniques to enable agents to respond dynamically to the actions of their rivals. The article provides a comprehensive overview of several strategies employed to tackle these challenges. Q-learning, a foundational reinforcement learning algorithm, is discussed in the context of its application in a multiagent setting. The model's ability to learn optimal pricing decisions by iteratively adjusting strategies based on past experiences is explored as a means to navigate the complexities of the competitive pricing environment. Simultaneously, the study investigates the application of actor-critic algorithms, a class of reinforcement learning models that combine aspects of both value-based and policy-based methods. Actor-critic models provide a more nuanced understanding of the pricing landscape by concurrently learning the optimal policy (pricing strategy) and estimating the value of different actions in various market scenarios.

Kong, D., Kong, X., Xiao, J., Zhang, J., Li, S., & Yue, L. (2019, August). Dynamic pricing of demand response based on elasticity transfer and reinforcement learning. In *2019 22nd International Conference on Electrical Machines and Systems (ICEMS)* (pp. 1-5). IEEE.

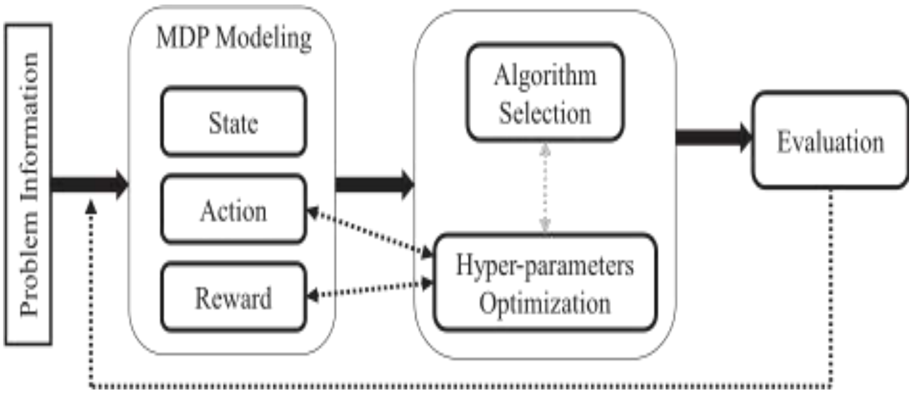
The paper discusses the existing literature on price-based demand response and its focus on user elasticity modeling and optimal load scheduling. It mentioned the drawbacks of requiring highly precise models in unpredictable and dynamic contexts**.** It also presents the idea of knowledge transfer and suggests a dynamic pricing strategy that combines elasticity transfer and reinforcement learning to speed up learning in unknown regions. The drawbacks associated with requiring precise models in unpredictable environments are elucidated in the paper. The dynamic nature of energy systems, influenced by factors such as weather conditions, unexpected events, and evolving consumer behavior, introduces challenges in maintaining the accuracy of predictive models. The limitations of traditional approaches highlight the need for innovative solutions that can adapt and learn in real-time, accommodating the inherent uncertainties of complex systems. To address these challenges, the paper introduces the concept of knowledge transfer. Knowledge transfer involves leveraging insights gained from one context or domain to inform decision-making in another, potentially less-explored, context. In the context of dynamic pricing for demand response, the paper suggests that transferring knowledge about user elasticity from known regions to unknown or less-explored regions can significantly enhance the adaptability and performance of pricing strategies.

IV. METHODOLOGIES

Base Paper: An Automated Deep Reinforcement Learning Pipeline for Dynamic Pricing

Methods used in this paper, are DRL pipeline

Architecture:



1.clearly defining the dynamic pricing problem (like knowing parameters)

2.defining the **state space** means knowing information about current market conditions

3.defining the **action space** means finding the possible pricing decisions

4.designing the **reward design** which means grading the performance of the pricing strategy

5.choosing the suitable DRL algorithm like deep Q-learning.

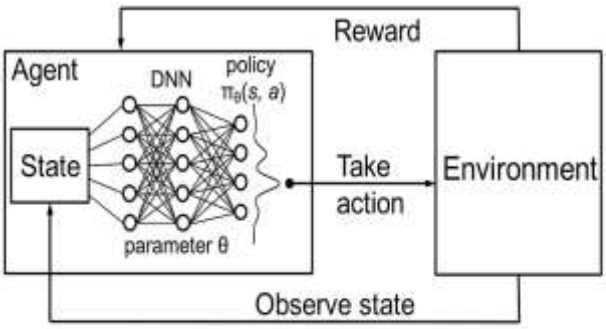
6.Training the DRL model on the historical data using the chosen algorithm

7.Exploration (trying new pricing strategies for balancing exploration) and exploitation (choosing known effective pricing strategies)

8.deploying the trained DRL model.

9.monitoring the model for tracking the performance and testing the model with new data according to market conditions.

The next step is choosing the hyperparameters of the DRL pipeline using the method BAYESIAN OPTIMIZATION and the selection operator of the GENETIC ALGORITHM



Bayesian optimization is a probabilistic model-based approach to global optimization that is particularly useful in scenarios where the objective function is expensive to evaluate. In the context of dynamic pricing, Bayesian optimization can be applied to find the optimal pricing strategy for maximizing revenue or achieving other business objectives.

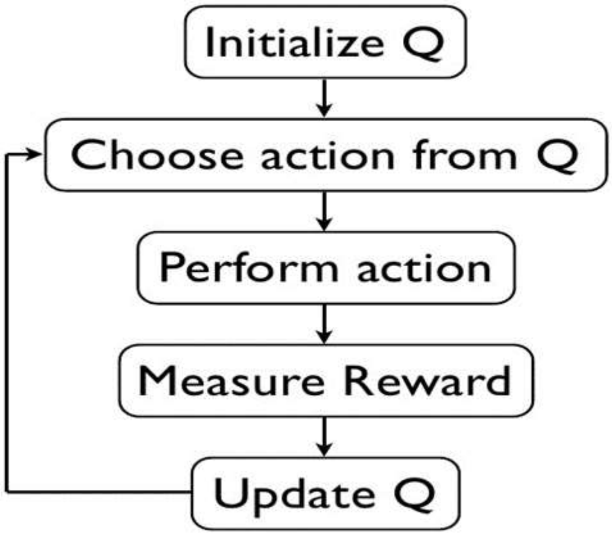
Genetic algorithms (GAs) are optimization algorithms inspired by the process of natural selection. Here in this paper Selection operator is used.

Selection Operator: Individuals (pricing strategies) in the population are selected for reproduction based on their fitness. Strategies that perform well are more likely to be selected. This mimics the idea of "survival of the fittest."

Reference Paper 1: A Dynamic Pricing Algorithm by Bayesian Q- Learning

Method: Q-learning method for exploration and exploitation tradeoff

Flowchart:



Q-learning can be applied to dynamic pricing as a reinforcement learning technique to find an optimal pricing strategy that maximizes revenue over time.

1)**State representation**

2)**Action representation**

3)**Q-table initialization**: Initialize a Q-table: Create a Q-table to store Q-values, which represent the expected cumulative reward for each state-action pair.

4) **Exploration-Exploitation Strategy:** The epsilon-greedy strategy is commonly used, where the agent chooses a random action with probability ε and the action with the highest Q-value with probability $1-\varepsilon$.

5) **Reward Definition**

6) **Q-Value Update:** After each pricing decision, update the Q-value for the chosen state-action pair using the Q-learning update rule.

The exploration-exploitation tradeoff is a fundamental challenge in reinforcement learning, and Q-learning addresses this by balancing the exploration of new actions and exploiting known actions that have yielded positive outcomes.

Exploration is the process of trying out new actions to discover their effects on the environment. It is necessary because the agent initially has limited knowledge about the environment and needs to explore to learn which actions lead to favorable outcomes and which do not.

Exploitation is the process of choosing actions that are known to yield high rewards based on the agent's current knowledge. Once the agent has learned which actions are most rewarding, exploitation allows it to maximize its immediate reward by consistently choosing those actions.

Method: Bayesian model-based approach for dynamic pricing

Bayesian model: Dynamic pricing involves adjusting the prices of products or services in response to changing market conditions, demand fluctuations, competitor pricing, and other relevant factors. Bayesian approaches can provide a framework for updating pricing strategies based on new information and observations

Step-by-Step Process:

1) Prior knowledge
2) Observations and likelihood
3) Posterior distribution
4) Adaptive pricing
5) Incorporation seasonality and trends
6) Uncertainity quantification
7) A/B testing

8) Sequential decision making

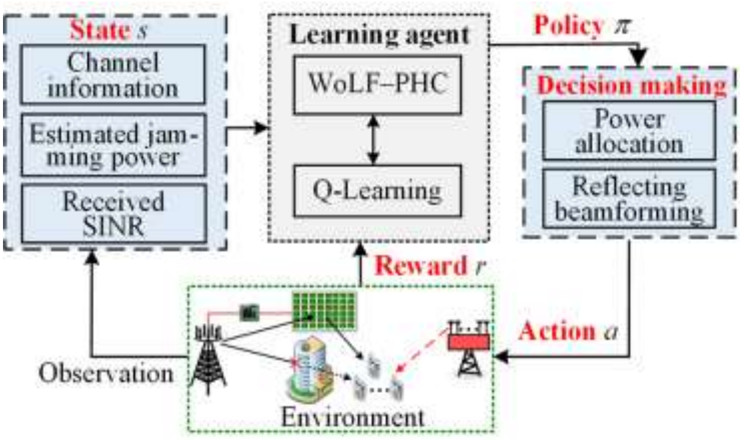
9) Hierachial Bayesian models

10) Model evalution and iteration

Reference Paper 2: Application of Reinforcement Learning in Dynamic Pricing Algorithms.

1)Method: WoLF-PHC algorithm

Architecture:



WoLF-PHC algorithm based on performance potentials.

Multi-agent learning has been brought to Q-learning. It does not, however, take into account the activities taken by the other agents. WoLF-PHC is a Q-learning extension that uses variable **policy hill-climbing** (PHC). learning rates , which don't always imply that a machine can Observe the behaviors and incentives of the other agents. It changes to environmental changes more effectively than Q-learning since it employs varying learning rate in accordance with the policies of the opponents, and is able to resolve the situation of giving little information. PHC employs a variety of tactics and enhances the present multifaceted policy by raising the likelihood that the if it is greddy with the present Q-value, act at the current state, else going down.

2)Method used: SA-Q Learning

SA-Q in dynamic pricing aims to find an optimal pricing strategy by learning and adapting to the changing environment over time.

However, the tradeoff between exploration and exploitation in Q-learning is difficult. Excessive exploration will drastically decrease the learning performance, but excessive exploitation will lead to local optimal solutions. In order to acquire balance between exploration and exploitation, the Metropolis criterion from simulated annealing algorithm can be applied to the Q-learning

V. Conclusion

In conclusion, the integration of reinforcement learning algorithms such as Deep Q Learning, Wolf PHC, and SAC in dynamic pricing research offers promising avenues for optimizing pricing strategies. These advanced techniques enable businesses to adapt in real-time to market dynamics, customer behavior, and other influential factors. The use of these algorithms enhances the efficiency and effectiveness of pricing decisions, leading to improved revenue generation and customer satisfaction. The application of reinforcement learning algorithms like Q-learning, WOLF-PHC, and SAC in dynamic pricing is a promising avenue. Each algorithm has its strengths and challenges, and the choice depends on the specific characteristics of the pricing problem at hand. The dynamic nature of pricing environments calls for adaptive, quick-learning algorithms, and the ongoing research in reinforcement learning continues to refine these approaches for effective real-world applications in dynamic pricing scenarios.

VI. REFERENCES

- Afshar, R. R., Rhuggenaath, J., Zhang, Y., & Kaymak, U. (2022). An Automated Deep Reinforcement Learning Pipeline for Dynamic Pricing. IEEE Transactions on Artificial Intelligence.
- Kastius, A., & Schlosser, R. (2021). Dynamic pricing under competition using reinforcement learning. Journal of Revenue and Pricing Management, 1-14.
- Raju, C. V. L., Narahari, Y., & Ravikumar, K. (2003, June). Reinforcement learning applications in dynamic pricing of retail markets. In EEE International Conference on E-Commerce, 2003. CEC 2003. (pp. 339-346). IEEE.

- Maestre, R., Duque, J., Rubio, A., & Arévalo, J. (2019). Reinforcement learning for fair dynamic pricing. In Intelligent Systems and Applications: Proceedings of the 2018 Intelligent Systems Conference (IntelliSys) Volume 1 (pp. 120-135). Springer International Publishing.
- Könönen, V. (2006). Dynamic pricing based on asymmetric multiagent reinforcement learning. International journal of intelligent systems, 21(1), 73-98.
- Kim, B. G., Zhang, Y., Van Der Schaar, M., & Lee, J. W. (2014, April). Dynamic pricing for smart grid with reinforcement learning. In 2014 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS) (pp. 640-645). IEEE.
- Kong, D., Kong, X., Xiao, J., Zhang, J., Li, S., & Yue, L. (2019, August). Dynamic pricing of demand response based on elasticity transfer and reinforcement learning. In 2019 22nd International Conference on Electrical Machines and Systems (ICEMS) (pp. 1-5). IEEE.
- Radner, R., Radunskaya, A., & Sundararajan, A. (2014). Dynamic pricing of network goods with boundedly rational consumers. Proceedings of the National Academy of Sciences, 111(1), 99-104.
- Rana, R., & Oliveira, F. S. (2015). Dynamic pricing policies for interdependent perishable products or services using reinforcement learning. Expert systems with applications, 42(1), 426-436.
- Abdalrahman, A., & Zhuang, W. (2020). Dynamic pricing for differentiated PEV charging services using deep reinforcement learning. IEEE Transactions on Intelligent Transportation Systems, 23(2), 1415-1427.
- Yang, X., Zhou, Z., Wallace, S., Lan, Z., Tang, W., Coghlan, S., & Papka, M. E. (2013, November). Integrating dynamic pricing of electricity into energy aware scheduling for HPC systems. In Proceedings of the International Conference on High Performance Computing, Networking, Storage and Analysis (pp. 1-11).