Received: 6 March 2024 | Revised: 1 April 2024 | Accepted: 28 April 2024 | Published online: 7 May 2024

RESEARCH ARTICLE

Journal of Comprehensive Business Administration Research
2024, Vol. 1(3) 171–178
DOI: 10.47852/bonviewJCBAR42022770

Dynamic Pricing Model for E-commerce Products Based on DDQN

Junyan Sun[1], Zihao Wang[1,*], Zheng Qiao[1] and Xiaopeng Li[1]

[1]*College of Mechanical and Electrical Engineering, Shaanxi University of Science and Technology, China*

Abstract: This paper addresses the challenges of intense price competition, price elasticity, and significant demand fluctuations in e-commerce product markets by adopting a dynamic pricing approach. Focusing on a product from the JD.com e-commerce platform, historical data spanning the past three years are analyzed, considering factors such as shipping costs, product inventory, product costs, and the impact of holidays. The study employs the Double Deep Q-Network (DDQN) for dynamic pricing optimization of the product and compares its performance with the Deep Q-Network (DQN) model. The results indicate that both the DQN algorithm and DDQN algorithm lead to varying degrees of profit improvement for the dynamic pricing of products. Specifically, the pricing profit with the DQN algorithm increased by an average of 1.925% compared with the original pricing profit, while the pricing profit with the DDQN algorithm increased by an average of 11.975% compared with the original pricing profit. These findings demonstrate practical significance.

Keywords: e-commerce products, dynamic pricing, deep reinforcement learning, DQN algorithm, DDQN algorithm

1. Introduction

With the rapid development of the Internet, the e-commerce industry has seen explosive growth. From 2012 to 2022, the market size of e-commerce platforms in China increased 4.36 times, with a total increase of 34.28 trillion yuan. The year-on-year growth of the e-commerce industry has strengthened economic development and driven the transformation of traditional industries. Currently, e-commerce platforms have become an indispensable part of people's lives, bringing significant convenience to daily life. However, fluctuations in the prices of goods or services can quickly influence consumer shopping behaviors, thereby directly impacting business profitability. It is evident that prices play a dual role in product sales, affecting both the production profits of enterprises and the daily lives of consumers. Product prices are closely related to various factors such as product value, inventory, shipping costs, holidays, etc. Decision-makers often need to consider and integrate multiple factors when pricing products. Therefore, constructing a dynamic pricing model for e-commerce products is of great practical significance for small and medium-sized e-commerce enterprises.

2. Literature Review

At present, scholars have conducted research on traditional pricing methods, including cost-plus pricing, demand-based pricing, and competitive pricing. Vives and Jacob [1] studied seven four-star hotels belonging to the same multinational hotel chain. The goal is to maximize revenue for hotels during peak seasons. They studied seven aspects of demand functions and implemented a deterministic dynamic pricing model to estimate the prices to be discovered, maximizing hotel revenue for each check-in date. Mohammed et al. [2] studied and analyzed dynamic pricing data of hotels in Hong Kong during the last week. They considered patterns and directions of room rate changes and their associations with hotel reputation, background factors, and other attributes. Surveys indicated that there was a higher likelihood of room rates increasing compared to decreasing or remaining unchanged. These results confirm the importance of differentiation in hotel room pricing and provide insights into how to implement differential pricing strategies. Krasheninnikova et al. [3] identified the contradictory problem of adjusting renewal prices in banks. In response to this issue, they proposed a new model for updating price adjustments as a sequential decision problem. In their results, they evaluated the real data using the insurance department of BBVA, one of the largest companies in the Spanish banking industry, to validate the feasibility of their approach. Dolgui and Proth [4] discussed various pricing methods such as price testing, cost-plus pricing, expert involvement, market analysis, and customer surveys. They also simulated the two-factor method using clustering algorithms, providing practical insights into pricing mechanisms. Hu et al. [5] discussed pricing strategies for electric vehicle rentals and established a game theory model to depict a simplified electric vehicle replenishment market. The results indicate that the choice of pricing strategy depends on the provider's operating costs, battery depreciation, and inventory scale, as well as the consumer's time sensitivity and opportunity costs. Ulmer [6] proposed the anticipatory pricing and routing policy method to incentivize customer purchases. This method addresses the issue of same-day delivery for electronic retailers and offers effective options for fleets to choose delivery deadlines. By modeling pricing and routing strategies, it enhances fleet flexibility and ultimately increases the revenue of electronic retailers. Traditional pricing methods are often static and do not

*Corresponding author: Zihao Wang, College of Mechanical and Electrical Engineering, Shaanxi University of Science and Technology, China. Email: 220511002@sust.edu.cn

adjust with market changes. They also tend to overlook consumer psychological factors. Given the rapid changes in the market, the adaptability of traditional pricing methods still needs improvement.

Some scholars have conducted research on dynamic pricing. Lobel [7] considered the scenario where consumers maintain a certain valuation over a period of time and will immediately make a purchase once the price equals or falls below their valuation. They also allowed for arbitrary joint distributions of patience levels and valuations and proposed a dynamic programming approach to solve for the optimal pricing strategy. Branda et al. [8] discovered a significant increase in demand for collective mobility services. They proposed a method called DA4PT (Data Analytics for Public Transport) to discover factors influencing passengers' booking and purchasing of tickets. Based on data from a bus ticketing platform with a cumulative user base of 3.23 million users, they revealed the correlation between booking factors and ticket purchases. Their method for pricing license plate prices predicted ticket purchases with an accuracy of 95%. den Boer and Keskin [9] studied a class of dynamic pricing problems with unknown discontinuous variables. They constructed a dynamic estimation and pricing strategy considering the discontinuity in demand and demonstrated that it can achieve close-to-optimal profitability performance. Najafi et al. [10] investigated the dynamic pricing problem for multiple products with limited inventory under a cascade click model. They considered that customer click-and-search behaviors could lead to different structures of optimal pricing strategies. Pricing based on the classic selection model, which ignores customer click-and-search behaviors, may significantly affect profitability. Hu et al. [11] integrated consumer strategic behavior into omnichannel retailing, establishing a two-stage omnichannel advertising and dynamic pricing model for retailers under three advertising decision modes: no advertising in both stages, normal advertising in the first stage, and discounted advertising in the second stage. They investigated the optimal response strategy for omnichannel retailers and provided numerical examples to validate the model. Dynamic pricing allows prices to be adjusted in real time based on market demand, supply conditions, competitor pricing, and other factors, providing higher flexibility and adaptability. By dynamically adjusting prices, businesses can more accurately gauge market demand and consumer willingness to pay, thereby maximizing profits.

In terms of algorithms, many scholars have utilized traditional algorithms for pricing research. Sun and Han [12] constructed a two-tier decision model for equipment ordering with multi-factor incentive pricing. Combining the rapid search capability of the particle swarm optimization (PSO) algorithm with the global search capability of the taboo search (TS) algorithm, a taboo search particle swarm optimization algorithm with test factors (TS-PSO-TF) was designed. The effectiveness of the model and algorithm was validated through numerical examples. Zhao et al. [13] studied the dynamic pricing method for airline cabin upgrade services, constructing a balanced utility model for both passengers and airlines. They used the Weibull distribution function to characterize the probability of passengers' acceptance of specific prices. Through numerical examples, they analyzed the problems of upgrading single-class cabins and multi-class cabins, respectively, and determined the optimal prices for offering cabin upgrade services for different cabin classes under both scenarios. Indeed, there are scholars who utilize machine learning methods to study pricing problems. Guo et al. [14] considered consumers' dynamic preferences for products and prices and proposed a dynamic recommendation model based on deep learning. The research findings can help retailers understand consumers' price preferences and make informed decisions regarding pricing, discounts, and bundled sales strategies. Pandey et al. [15] proposed a dynamic pricing framework based on deep reinforcement learning for managed lanes with multiple access points and travelers' time value. The superiority of a heuristic algorithm based on minimizing the total system travel time was demonstrated compared to using feedforward neural networks with feedback control methods.

Kamandanipour et al. [16] proposed a data-driven dynamic pricing method for passenger railway service providers. They utilized a multilayer perceptron artificial neural network as the model and used a regression model as the price elasticity function to quantify the impact of prices, seasonal conditions, and competition on company sales. Lee et al. [17] studied the problem of time-varying electricity pricing and proposed a method based on deep reinforcement learning. They utilized nonparametric density estimation methods and used the estimated density function for sampling variables related to charger usage patterns. Their analysis focused on charging costs and load-shifting effects. The final simulation results demonstrate that this method exhibits good adaptability. Lu et al. [18] proposed a dynamic pricing algorithm for airline tickets based on policy learning. Its core lies in no longer predicting the demand for each fare class but modeling the dynamic pricing problem for airline tickets as an offline reinforcement learning problem. Aljafari et al. [19] addressed the problem of managing charging stations for electric vehicles to minimize waiting times and reduce electricity prices for electric vehicle customers. They proposed a constrained model for dynamic pricing and optimization scheduling and utilized deep neural networks for optimization. Poh et al. [20] addressed the issue of traffic congestion caused by insufficient parking spaces during peak hours. They proposed a method called dynamic pricing based on deep reinforcement learning, where parking utilization and profit are considered as incentives for price control. Through experiments, they demonstrated the rationality of the algorithm. Fang and Le [21] addressed the pricing problem of airline tickets by modeling the dynamic pricing problem of flights as a Markov game process and establishing a logit choice model for mixed-type passengers. They used multi-agent reinforcement learning algorithms to solve the instances, and the results indicated that while the WoLF-PHC (Win or Learn Fast-Policy Hill Climbing Algorithm) algorithm required more iterations to converge compared with the Nash-Q algorithm, the WoLF-PHC algorithm showed significant advantages in terms of computational speed and demonstrated strong adaptability.

In summary, a large number of scholars have traditionally addressed pricing problems using conventional methods, while the utilization of deep learning methods by scholars is relatively limited at present. While machine learning algorithms have been applied in the pricing domain, they are mostly used for static pricing or multi-supplier pricing where consumer types are determined. There are also studies focusing on dynamic pricing for special commodities like airline tickets and hotel rooms, but their applicability is limited to specific types of goods. Currently, there is little research on dynamic pricing for e-commerce products. This paper addresses the dynamic pricing problem for e-commerce products by crawling relevant pricing data of a certain electronic product from the JD.com platform. The problem is formulated as a Markov decision process (MDP), considering factors such as product value, competition costs, shipping costs, and holiday effects. The Double Deep Q-Network (DDQN) method from deep reinforcement learning is employed to build the dynamic pricing model for e-commerce products. This model possesses strong practical significance and applicability.

3. Research Methodology

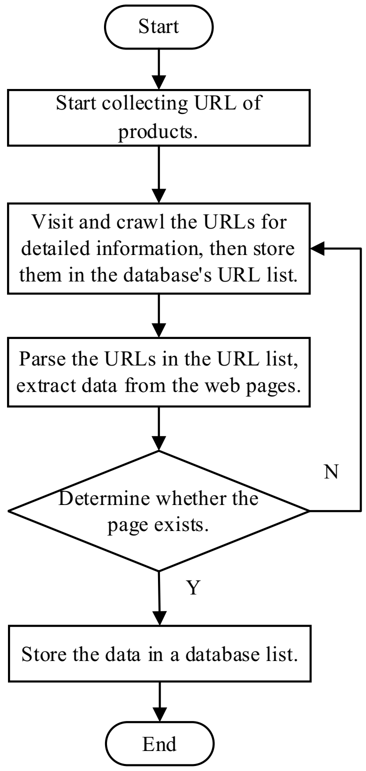
3.1. Dataset

The data to be collected primarily include the following: sales date of the product, shipping fee of the product, price of the product, inventory of the product, cost of the product, and other relevant data. The data collection process is illustrated in Figure 1. The specific process is outlined as follows: (1) The primary objective of data crawling is to collect the URLs of products, thereby obtaining the webpage's source code which encapsulates a substantial amount of valuable information. (2) Subsequently, each link within the URL list is iterated through, and an HTTP request is sent to acquire the webpage content. Simultaneously, the HTTP response status code is inspected to ensure the successful loading of the page. (3) The URLs within the list are subjected to webpage parsing to extract pertinent data. This involves a meticulous examination of each link, parsing the webpage, and extracting vital product information such as name, price, sales volume, and reviews using HTML tags or CSS selectors. (4) Should the page not exist, the process reverts to step two; if it does, it proceeds to the subsequent step. (5) The gleaned data is then stored in a database.

After importing the collected data into an Excel sheet, it was observed that many entries exhibited redundancy. Therefore, we performed data cleansing to address invalid, missing, and duplicate values. We also handled missing entries appropriately. The cleaned dataset, depicted in Table 1, encompasses information such as order timestamps, shipping fees, product prices, inventory levels, and product costs.

Moreover, for enhanced accuracy in learning outcomes, we factored in the influence of special holidays. During such occasions, sales of certain products are significantly affected. Hence, it was deemed necessary to consider such factors. We employed one-hot encoding to encode holidays such as Christmas, Singles' Day, and Mother's Day. When a product coincides with these holidays, its sales are notably impacted.

Figure 1
Data crawling process diagram



3.2. Construction of reinforcement learning algorithm

For businesses, the primary objective is invariably to maximize their total revenue. Revenue maximization encompasses both short-term and long-term revenue optimization.

Short-term revenue maximization focuses on optimizing revenue within specific time frames or contexts, such as maximizing revenue for a particular season, promotion period, or sales event. This involves adjusting pricing strategies dynamically to exploit immediate opportunities, respond to short-term fluctuations in demand, and outperform competitors in the short run. On the other hand, long-term revenue maximization involves developing strategies to sustainably increase overall revenue over extended periods. This includes implementing pricing strategies that consider broader market trends, customer behaviors, and competitive dynamics to drive consistent revenue growth over time. Additionally, it involves fostering customer loyalty, enhancing brand value, and investing in innovation and long-term business development initiatives.

Achieving revenue maximization requires businesses to strike a balance between short-term revenue gains and long-term profitability. It demands a comprehensive understanding of market dynamics, robust pricing strategies, and agile decision-making processes that can adapt to evolving market conditions while

Table 1
Partial raw data

Date	Offer	Price	Freight_Value	C-Price	Stock	Others
2017/7/3	931.1843	931.1843	25.73	931.1843	0	0
2017/7/4	931.1843	931.1843	25.73	931.1843	0	0
2017/7/5	907.2738	907.2738	25.73	907.2738	0	0
2017/7/6	907.2738	907.2738	25.73	907.2738	0	0
2017/7/7	907.2738	907.2738	25.73	907.2738	0	0
……	……	……	……	……	……	……
2020/2/13	926.0964	926.0964	26.4728	958.7011	1465	0
2020/2/14	930.0148	930.0148	27.16841	964.8637	446	0
2020/2/15	933.6833	933.6833	26.82645	969.4615	7591	0
2020/2/16	934.2676	933.7343	27.27459	968.5057	7593	0
2020/2/17	933.1509	933.1509	26.32039	968.2658	7675	0

aligning with long-term business objectives. Moreover, leveraging advanced analytics, data-driven insights, and optimization techniques can enable businesses to optimize revenue across both short and long horizons effectively. In real-world pricing scenarios, considering the MDP, the pricing of a product on a given day not only influences the sales volume for that day but also affects future sales volumes. However, the impact on today's sales is evidently the most significant. Therefore, we formulate the revenue function for the product as the accumulation of all revenues, incorporating a decay factor $\lambda(\lambda < 1)$. We express this by raising λ to increasing powers, indicating that the pricing of the product on a given day still has the greatest impact on the profit for that day, with the influence on future profits gradually decaying. Therefore, we provide the following reward formula:

$$G_t = R_{t+1} + \lambda R_{t+2} + \ldots \lambda R_n = \sum_{k=0}^{n} \lambda^k R_{t+k+1} \tag{1}$$

To address the MDP problem of dynamic pricing using reinforcement learning as the fundamental model, we first employ Q-Learning to find the optimal pricing strategy. Q-Learning is a value iteration method for computing the optimal policy. It starts with randomly initialized Q-values and iteratively improves them using transitions to obtain the optimal Q-table and policy:

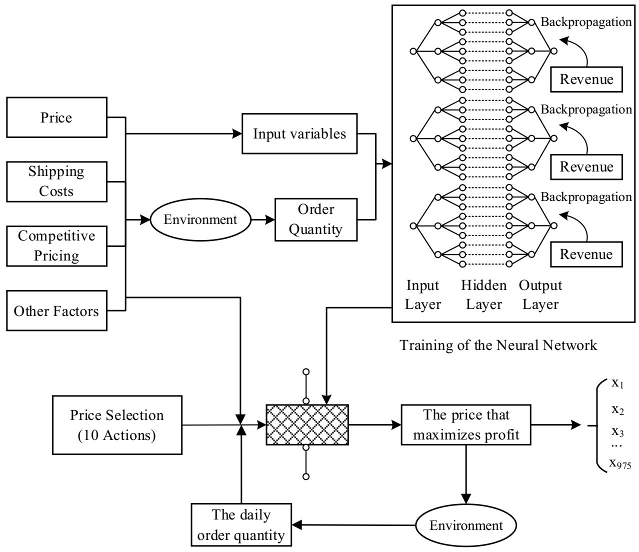
$$Q(s,a) \leftarrow Q(s,a) + \alpha[r + \gamma \max_{a'} Q(s',a') - Q(s,a)] \tag{2}$$

In the Q-Learning algorithm of reinforcement learning, before any iterations, the Q-table values are initialized to zero, effectively creating a zero-order matrix. The iterative process begins by employing an ε-greedy policy at the initial state to select actions. After choosing an action, the agent transitions to a new state and receives a reward. The agent then records this reward and updates the Q-table values using the iteration formula mentioned earlier. At this point, the agent needs to determine if it has reached a terminal state. If so, the iteration process ends; otherwise, it continues.

In the DQN (Deep Q-Network) algorithm, to address the storage problem associated with a two-dimensional Q-table, we replace the Q-table with a neural network to store the relevant historical data we have collected and organized. The iteration algorithm and process still follow the flow of the Q-Learning algorithm. In comparison to DQN, DDQN involves searching for action Q-values in the target network. It uses the Q-estimate neural network to estimate the maximum action value of Max(s', a') in Q-display and then utilizes this estimated action from the Q-estimate neural network to select Q(s') in Q-real. The logical process is illustrated in Figure 2.

The price, shipping cost, bidding, and other factors are input variables that we collect or pre-set. Price refers to the historical price of the product on the given day, shipping cost refers to the cost of shipping the product, bidding represents the price of similar products that compete with the product, and other factors include the holiday impact mentioned earlier. All variables are input into another known three-layer neural network model to predict the daily order quantity of the product. Then, these order quantities and all other input variables are input into the neural network's experience pool and stored. The neural network collects the features of this data based on statistical principles and provides the profit for each dataset. The profit is updated through backpropagation and gradient descent. After several rounds of training, the neural network is trained. The specific operation process of the model is as follows: (1) First, select the price. Ten price actions are provided for pricing the product, ranging from 1.1 times the cost to 1.55

Figure 2
Basic logic flowchart of the calculation example



times the cost. These, along with other historical data input variables, are fed into the trained model. (2) The model selects the optimal price based on the Q-Learning algorithm's iterative formula. After selecting the optimal price, it interacts with the environment to obtain the order quantity, which represents the current profit of selling a single product. (3) On the next day, the order quantity from the previous day, the price action selection for the next day, and other historical data for the next day are input into the model again. The model then selects the optimal price based on the Q-Learning algorithm's iterative formula. (4) Repeat the above steps to determine dynamic pricing for 975 days.

4. Examples and Results

4.1. The result of DQN algorithm

First, using the DQN algorithm, we loaded the known demand prediction model and price action selection, constructed the neural network, and designed the relevant parameters. In the DQN model, the unchanged parameters were as follows: hidden layer node quantity (hidden size) = 50, input state size (input size) = 2 + random parameters, output action quantity (output size) = 10, learning rate (LR) = 0.001, current neural network update frequency (train frequency) = 10, discount factor (gamma) = 0.9, and initial exploration probability (epsilon) = 1.

In the DQN algorithm, the results were outputted to the final Excel table, including the year, month, day, pricing, predicted order quantity, total revenue, and other relevant data. Partial training results are provided in Table 2 below.

As shown in Figures 3 and 4, the profit results obtained from the DQN algorithm (with gamma = 0.9 and epsilon = 1.0) are plotted as a line graph. Comparing the two pricing methods, it can be observed that the profit learned by the DQN algorithm is not high initially. However, as time progresses and dynamic pricing is applied to the product on a daily basis, the profit shows an upward trend. At certain time periods, there may be significant fluctuations in profit due to the introduction of random factors such as Double 11, Mother's Day, and shopping festivals. These factors significantly affect the number of orders for the product, thus impacting its profit.

Table 2
DQN algorithm (gamma = 0.9, epsilon = 1.0) partial output results

Date	Base-prices	Base-orders	Base-rewards	RL-prices	RL-orders	RL-rewards
2017/6/30	931.1843	0	0	782.1948	0.803023	50.57231
2017/7/1	931.1843	0.438547	92.95734	782.1948	2.625046	165.3185
2017/7/2	931.1843	0.167312	35.4647	782.1948	2.353812	148.2369
2017/7/3	931.1843	1.369577	290.305	782.1948	3.556076	223.9523
……	……	……	……	……	……	……
2020/2/23	922.4432	4.550854	957.1679	940.8921	4.280107	979.1856
2020/2/24	934.21	7.327245	1559.583	953.0982	7.050051	1633.746
2020/2/25	939.4614	7.098918	1522.11	958.2507	6.823176	1591.189
2020/2/26	930.8749	10.45958	2175.074	953.1029	10.13337	2332.483
2020/2/27	929.7077	11.07159	2317.965	947.631	10.80856	2456.621

Figure 3
Comparison graph between DQN model and original pricing

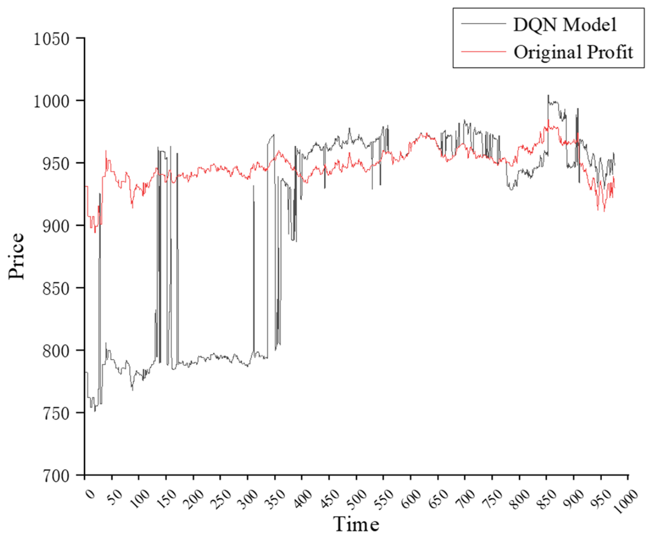
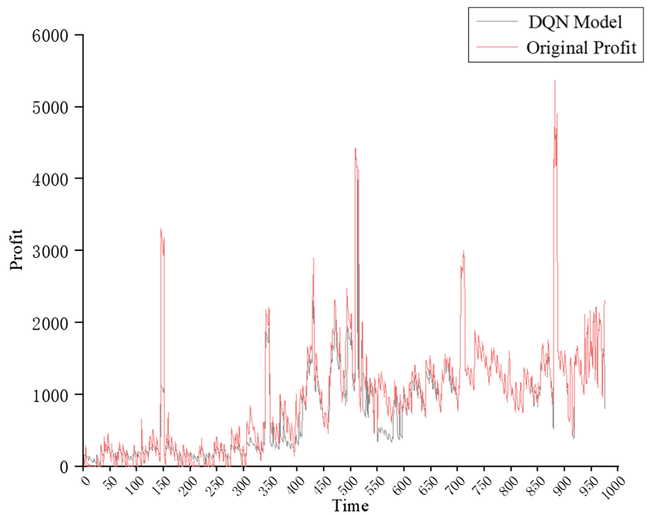


Figure 4
Comparison graph between DQN model and original profit



According to the profit comparison graph, it can be seen that overall, the profit obtained from the DQN algorithm is higher than the profit from the original pricing. In the statistical example, the total profit of the product is approximately 931145.2 yuan, compared with the total profit of 912692.8 yuan from the original pricing, indicating an increase in profit of 92790.9 yuan, or a 2.0% increase.

4.2. The result of DDQN algorithm

To solve the dynamic pricing problem of e-commerce products using the DDQN algorithm, we load the known demand prediction model and price action selection, construct the neural network, and design relevant parameters. In the DDQN model, we slightly adjusted the unchanged parameters, including the hidden layer node size (hidden size) of 30, the input size of 2 plus random parameters, the output action size of 10, the memory size of 640, the LR of 0.002, the minimum exploration probability (epsilon min) of 0.1, the current neural network update frequency (train freq) of 5, the discount factor (gamma) of 0.9, and the initial exploration probability (epsilon) of 1.

The results are output to the final Excel table, including time, pricing, predicted order quantity, total revenue, and other relevant data. Some training results are shown in Table 3.

Based on the DDQN algorithm (gamma = 0.9, epsilon = 1.0), the profit results are computed as shown in Figures 5 and 6. It can be observed that the DDQN model generates larger price fluctuations, and starting from the mid-term, its pricing is consistently higher than the original prices. This indicates that the agent has identified a clear strategy, which is to increase prices. The agent determines that selling at prices significantly above the current cost price can result in better profits.

Comparison between the DDQN algorithm and original pricing in terms of profit: Initially, the DDQN algorithm performs similarly to the DQN algorithm. Due to the randomness of actions and the need for continuous learning and exploration, its current profit remains slightly lower than the profit from the original prices. However, as the algorithm converges, the superiority of the DDQN algorithm begins to manifest, with prices consistently higher than the original prices in the mid to later stages. The total profit for the statistical example is approximately 1,029,185.1 yuan, representing an increase of 116,492.3 yuan compared with the total profit from the original pricing, indicating a profit increase of 12.8%.

Table 3
DDQN algorithm (gamma = 0.9, epsilon = 1.0) partial output results

Date	Base-prices	Base-orders	Base-rewards	RL-prices	RL-orders	RL-rewards
2017/7/3	931.1843	1.369577	290.305	819.4422	3.009452	158.5338
2017/7/4	931.1843	1.098343	232.8124	819.4422	2.738218	131.8683
2017/7/5	907.2738	0.286028	59.26058	798.4009	1.883794	105.2028
2017/7/6	907.2738	0.014794	3.065015	798.4009	1.61256	78.53722
……	……	……	……	……	……	……
2020/2/23	922.4432	4.550854	957.1679	1033.136	2.926374	939.4252
2020/2/24	934.21	7.327245	1559.583	1158.668	4.0332	1763.741
2020/2/25	939.4614	7.098918	1522.11	1164.932	3.790016	1667.172
2020/2/26	930.8749	10.45958	2175.074	1158.674	7.116503	3101.014

Figure 5
Comparison graph between DDQN model and original pricing

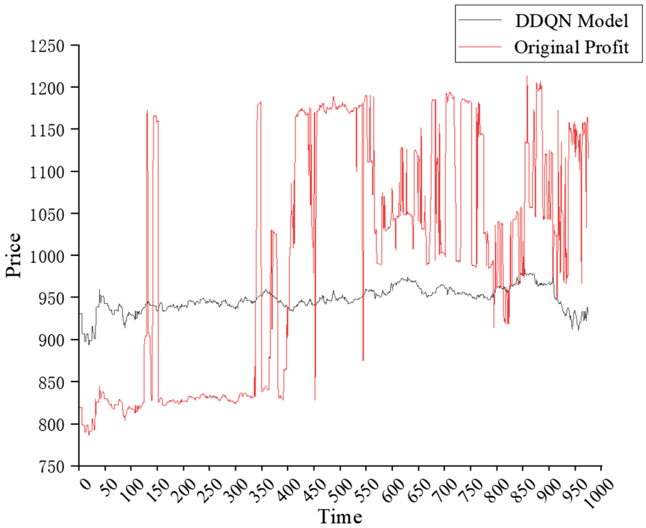
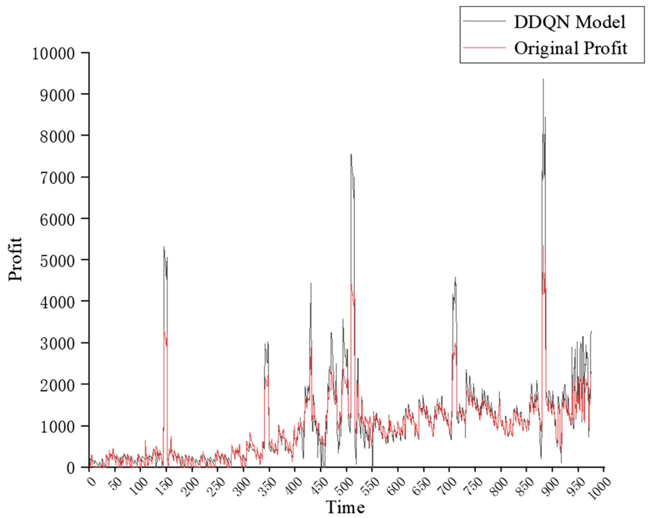


Figure 6
Comparison graph between DDQN model and original profit



4.3. Comparison of models

Based on the computation and results analysis under different gamma coefficients and epsilon values, a summary is provided in Table 4. A comparison reveals that as the discount factor increases, the profit rises, whereas increasing the initial exploration probability leads to a decrease in profit. This summary suggests that adjusting the gamma coefficient can positively impact profit, while higher initial exploration probabilities tend to decrease profit.

Simultaneously, the average profit improvement rates for both algorithm models under different parameter settings are provided.

Table 4
Comparison of results under different parameters and algorithms

Parameters	Model	Total sales volume	Average price per unit in yuan	Total profit in yuan
gamma 0.9 epsilon 1.0	Original pricing	4167	948.9	912692.8
	DQN model	5226	870.4	931145.2
	DDQN model	4637	911.0	1029185.1
gamma 0.9 epsilon 1.2	Original pricing	4167	948.3	912692.8
	DQN model	4866	896.4	946081.5
	DDQN model	4628	912.2	1053166.3
gamma 0.95 epsilon 1.0	Original pricing	4167	948.8	912692.8
	DQN model	4051	957.9	899360.2
	DDQN model	4667	909.7	1064301.3
gamma 0.95 epsilon 1.2	Original pricing	4167	948.7	912692.8
	DQN model	4270	941.2	943367.8
	DDQN model	4632	904.3	1047394.3

Additionally, based on the trained neural networks and pricing results from the previous 975 days, a price prediction for day 976 is presented in Table 5. Using the DQN model for dynamic pricing resulted in a profit increase of 1.925% compared with the original pricing, while using the DDQN model led to a profit increase of 11.975%. Thus, the DDQN model demonstrated superior effectiveness.

Table 5
Profit comparison under different algorithms

Model	Average profit increase rate
DQN model	1.925%
DDQN model	11.975%

5. Conclusion

The application of deep reinforcement learning algorithms in dynamic pricing for e-commerce products was explored in this study. Both the DQN and DDQN algorithms were employed for continuous pricing over a period of 975 days. Throughout the pricing process, parameters were continuously adjusted to fine-tune the models in order to achieve better pricing results. The final outcomes indicate that under different parameter settings, both the DQN and DDQN algorithms yielded significantly higher profits compared with the original pricing strategy. Specifically, under the same parameters, the DDQN algorithm outperformed the DQN algorithm by 10.05% in terms of final profit and demonstrated greater stability and convenience in parameter settings and adjustments. While this study achieved certain results, challenges such as long algorithm runtime and low efficiency persist. Furthermore, the results are currently limited to a theoretical level. Future work should focus on optimizing the algorithms further and conducting real-world applications in enterprises to validate the effectiveness of the approach.

Funding Support

This work is sponsored by the Key R&D Project of the Science and Technology Department of Shaanxi Province (2023-YBGY-408); Social Science Foundation of Shaanxi Province, No. 2020R043; and Science and Technology Project of Xi'an, Shaanxi Province (23RKYJ0026).

Ethical Statement

This study does not contain any studies with human or animal subjects performed by any of the authors.

Conflicts of Interest

The authors declare that they have no conflicts of interest to this work.

Data Availability Statement

The data that support this work are available upon reasonable request to the corresponding author.

Author Contribution Statement

Junyan Sun: Conceptualization, Supervision, Project administration, Funding acquisition. **Zihao Wang:** Methodology, Software, Formal analysis, Investigation, Writing – original draft, Writing – review & editing. **Zheng Qiao:** Data curation, Visualization. **Xiaopeng Li:** Validation, Resources, Writing – review & editing.

References

[1] Vives, A., & Jacob, M. (2021). Dynamic pricing in different Spanish resort hotels. *Tourism Economics*, *27*(2), 398–411. https://doi.org/10.1177/1354816619870652

[2] Mohammed, I., Guillet, B. D., Law, R., & Rahaman, W. A. (2021). Predicting the direction of dynamic price adjustment in the Hong Kong hotel industry. *Tourism Economics*, *27*(2), 346–364. https://doi.org/10.1177/1354816620903900

[3] Krasheninnikova, E., García, J., Maestre, R., & Fernández, F. (2019). Reinforcement learning for pricing strategy optimization in the insurance industry. *Engineering Applications of Artificial Intelligence*, *80*, 8–19. https://doi.org/10.1016/j.engappai.2019.01.010

[4] Dolgui, A., & Proth, J. M. (2010). Pricing strategies and models. *Annual Reviews in Control*, *34*(1), 101–110. https://doi.org/10.1016/j.arcontrol.2010.02.005

[5] Hu, X., Yang, Z., Sun, J., & Zhang, Y. (2023). Optimal pricing strategy for electric vehicle battery swapping: Pay-per-swap or subscription? *Transportation Research Part E: Logistics and Transportation Review*, *171*, 103030. https://doi.org/10.1016/j.tre.2023.103030

[6] Ulmer, M. W. (2020). Dynamic pricing and routing for same-day delivery. *Transportation Science*, *54*(4), 1016–1033. https://doi.org/10.1287/trsc.2019.0958

[7] Lobel, I. (2020). Technical note—Dynamic pricing with heterogeneous patience levels. *Operations Research*, *68*(4), 1038–1046. https://doi.org/10.1287/opre.2019.1951

[8] Branda, F., Marozzo, F., & Talia, D. (2020). Ticket sales prediction and dynamic pricing strategies in public transport. *Big Data and Cognitive Computing*, *4*(4), 36. https://doi.org/10.3390/bdcc4040036

[9] den Boer, A. V., & Keskin, N. B. (2020). Discontinuous demand functions: Estimation and pricing. *Management Science*, *66*(10), 4516–4534. https://doi.org/10.1287/mnsc.2019.3446

[10] Najafi, S., Duenyas, I., Jasin, S., & Uichanco, J. (2024). Multiproduct dynamic pricing with limited inventories under a cascade click model. *Manufacturing & Service Operations Management*, *26*(2), 554–572. https://doi.org/10.1287/msom.2021.0504

[11] Hu, J., Li, L., Zhu, X., Zhang, H., & Yang, W. (2023). Two-period advertising and dynamic pricing strategy of omni-channel considering consumer strategy behavior. *Operations Research and Management Science*, *32*(8), 114–121.

[12] Sun, S., & Han, S. (2020). Multi-factor incentive pricing model and algorithm for military equipment ordering based on bi-level decision-making. *Systems Engineering and Electronics*, *42*(6), 1338–1347. https://doi.org/10.3969/j.issn.1001-506X.2020.06.18

[13] Zhao, G., Ma, C., Zhao, Q., & Li, J. (2023). Dynamic pricing of airline upgrade service considering revenue management. *Journal of Transportation Systems Engineering and Information Technology*, *23*(3), 76–84. https://doi.org/10.16097/j.cnki.1009-6744.2023.03.009
[14] Guo, W., Tian, J., & Li, M. (2023). Price-aware enhanced dynamic recommendation based on deep learning. *Journal of Retailing and Consumer Services*, *75*, 103500. https://doi.org/10.1016/j.jretconser.2023.103500
[15] Pandey, V., Wang, E., & Boyles, S. D. (2020). Deep reinforcement learning algorithm for dynamic pricing of express lanes with multiple access locations. *Transportation Research Part C: Emerging Technologies*, *119*, 102715. https://doi.org/10.1016/j.trc.2020.102715
[16] Kamandanipour, K., Yakhchali, S. H., & Tavakkoli-Moghaddam, R. (2023). Learning-based dynamic ticket pricing for passenger railway service providers. *Engineering Optimization*, *55*(4), 703–717. https://doi.org/10.1080/0305215X.2022.2030324
[17] Lee, J., Lee, E., & Kim, J. (2020). Electric vehicle charging and discharging algorithm based on reinforcement learning with data-driven approach in dynamic pricing scheme. *Energies*, *13*(8), 1950. https://doi.org/10.3390/en13081950
[18] Lu, M., Zhang, Y., & Lu, C. (2021). Approach for dynamic flight pricing based on strategy learning. *Journal of Electronics & Information Technology*, *43*(4), 1022–1028.
[19] Aljafari, B., Jeyaraj, P. R., Kathiresan, A. C., & Thanikanti, S. B. (2023). Electric vehicle optimum charging-discharging scheduling with dynamic pricing employing multi agent deep neural network. *Computers and Electrical Engineering*, *105*, 108555. https://doi.org/10.1016/j.compeleceng.2022.108555
[20] Poh, L. Z., Connie, T., Ong, T. S., & Goh, M. K. O. (2023). Deep reinforcement learning-based dynamic pricing for parking solutions. *Algorithms*, *16*(1), 32. https://doi.org/10.3390/a16010032
[21] Fang, Y., & Le, M. (2020). Dynamic pricing of parallel flights based on the reinforcement learning. *Journal of East China Jiaotong University*, *37*(1), 47–53. http://dx.doi.org/10.16749/j.cnki.jecjtu.2020.01.007

How to Cite: Sun, J., Wang, Z., Qiao, Z., & Li, X. (2024). Dynamic Pricing Model for E-commerce Products Based on DDQN. *Journal of Comprehensive Business Administration Research*, *1*(3), 171–178. https://doi.org/10.47852/bonviewJCBAR42022770