Vol. 19 (2026): Proceedings of 2026 International Conference on Financial Technology, Mathematical Statistics, and Business Analytics (ICFMB 2026)
E-ISSN: 3105-0522 P-ISSN: 3105-0514
DOI: https://doi.org/10.70267/icfmb.202619170176

Dynamic Pricing Strategy for E-commerce Based on Reinforcement Learning

Hanglin Zhu*

SWUFE-UD Institute of Data Science, Southwestern University of Finance and Economics, Sichuan 611130, China

**Corresponding author: Hanglin Zhu.*

Abstract

Since 2025, with the rapid development and fierce competition of the e-commerce industry, the pricing strategy for e-commerce has already stepped into the stage of fine-tuning. Since the traditional fixed pricing strategy couldn't adapt to market requirements, dynamic pricing has become the key to increasing profit. The research selects sales data from the Amazon platform in 2025 as an object of the study and uses current data processing and modelling approaches. By data preprocessing, data analysis, and eigenvalue mining, the research selects the Q-learning reinforcement learning algorithm to build a dynamic pricing model and conducts a control experiment on a traditional fixed pricing strategy. The results show that factors like price and sponsorship status have a significant influence on e-commerce sales volume, and there is a striking negative correlation between profit and sales volume. The total profit of the established dynamic pricing model, which is based on the Q-learning reinforcement learning algorithm, is 12.15% higher than the traditional fixed pricing strategy, which effectively increases profits. The research not only provides practical operation ideas and model support, but also proves the feasibility of the reinforcement learning algorithm in the e-commerce field.

Keywords

e-commerce, dynamic pricing, Q-learning

1. Introduction

Since 2025, the competition in the e-commerce market has become increasingly fierce [1], and the same kind of goods are often sold simultaneously in multiple stores. The number of e-commerce stories and customers is increasing sharply [2], especially small-scale e-commerce businesses [3]. Due to the influence of multiple factors such as advertisement, discount, sponsorship [4], the traditional fixed pricing strategy [5] cannot adapt to the current market environment, which will reduce both profits and sales. Amazon, as the main e-commerce platform in the world, has great research value for the sales statistics in it. How to mine data through sales statistics and how to use an intelligence algorithm to achieve dynamic pricing [6] become the keys to increasing e-commerce businesses' profits. The research could clarify the impact of factors like prices, advertisements and discounts on sales which could provide a relevant strategy for businesses. It could also provide dynamic pricing suggestions for small-scale e-commerce businesses through the application of reinforcement learning dynamic pricing with low costs and convenience to achieve continuous optimization of profits.

The research is based on the data of yearly e-commerce data in 2025, extracts features like price, sales, discount, advertisement in the dataset, completes data preprocessing, data analysis, construction of reinforcement learning model and data visualization sequence and finally accomplishes the whole report.

The research is aimed at mastering the methods of e-commerce data processing, in order to achieve the application of the reinforcement learning dynamic pricing model in the e-commerce pricing environment. The research also compares and verifies the advantages of this model over traditional pricing strategies to provide relative pricing references for small-scale e-commerce businesses.

2. Method

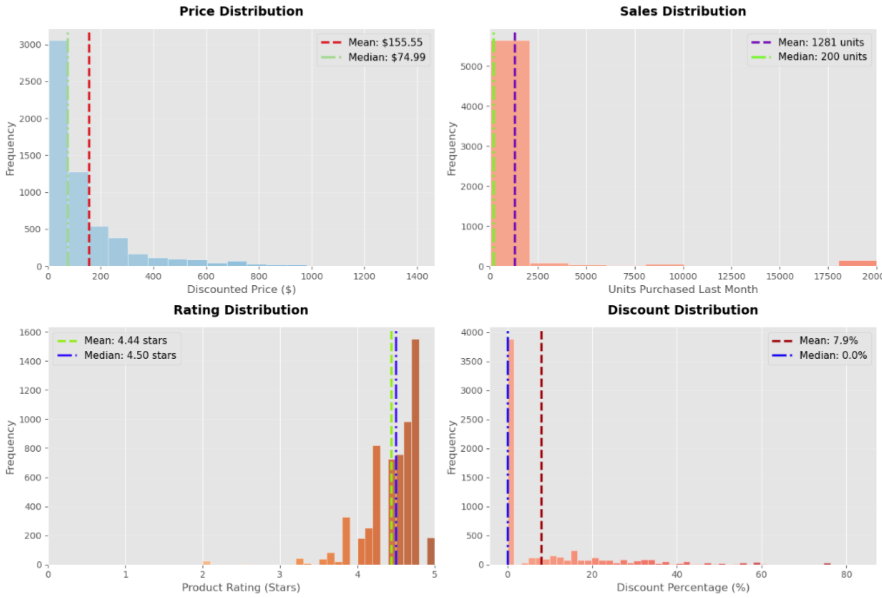
The research selects public dataset of Amazon e-commerce product sales as the object of the research. The dataset has 42675 data records and each data record includes core research fields such as discount, sponsorship, sales volume. This dataset matches the actual e-commerce operation situation and the data recorded in it has high quality and wide dimensions, which has strong research value. Since the original data doesn’t provide the ID of stores, the paper simulates the situation where different pricing strategies of the same category are regarded as a competitive environment between multiple businesses.

2.1 Data Preprocessing

Before the research, the paper underwent data preprocessing. Firstly, the paper cleans the data, including eliminating duplicate values and abnormal values, filtering invalid data records, to ensure the reliability of the data. Secondly, for different types of missing fields, the research uses the mean value filling method for numerical values and the mode filling method for categorical values, like sponsorship to ensure the integrity of data. Thirdly, the research encodes textual data, translating it into numerical value that can be identified by the model, to complete the standardization of data formats [7]. Lastly, the research constructs derivative features based on research requirements, including profits per unit, actual sales volume after discount, and screens out core features highly related to pricing and sales volume, in order to prepare for building reinforcement learning pricing model.

Through data preprocessing, the research finally obtained 41845 experimental data. The miss rate is 9.17% and the abnormal proportion is about 1.94%. The research also counted the core features, including distribution of price, sales, rating and discount. The distribution of the eigenvalues is shown in Figure 1.

Figure 1: Distribution of eigenvalues of Amazon products (Picture credit: Original)



2.2 Pricing Strategy Design

The research builds two pricing strategy respectively: fixed pricing and dynamic pricing. These two pricing strategies are based on the same dataset and pricing standard to ensure the justice of the comparison.

2.2.1 Fixed Pricing Method

In this research, based on the regular pricing methods for small-scale e-commerce businesses on Amazon in 2025, the fixed pricing strategy is priced at a fixed rate for the entire year based on the average selling price of the same type of commodities at each e-commerce store within the training set. The price doesn't change in the test set and the pricing logic is easy to operate which fits small-sized stories' real operational capabilities.

2.2.2 Dynamic Pricing Method

Based on reinforcement learning, data-driven logic, the dynamic pricing strategy combines the core features like competitive price and sales volume fluctuation of similar goods in different stores and designs self-adaptive learning-driven adjustment rules [9]. The competitive price-driven rules are as follows: when the average price of similar competing products is more than 5% lower than that of this store currently, the store will reduce the price to a small extent (3%-5%), to grab market share; when it is higher, the store will increase the price within a small margin (2%-4%).

2.3 Model Selection

In the reinforcement learning modeling stage, the research compared common reinforcement algorithms used in the e-commerce pricing field by other researchers: DQN algorithm, PPO algorithm, and Q-learning algorithm [10]. The DQN algorithm is the representative of deep learning. It has high precision but it is needed to construct neural network which requires high computing power and needs high cost of time and material resources [11]. It is suitable for application in high-precision pricing scenario. The PPO algorithm has strong stability, which is suitable for application to continuous action space. It is usually applied to pricing research about multi-brand joint pricing and commodities with large price fluctuation [12]. Neither of them is suitable for small-scale e-commerce businesses due to the flexibility and high costs. The Q-learning algorithm is the representative of traditional reinforcement learning algorithms. It has simple principle, fast training speed, low time cost and not complicated neural network architectures [13], which is widely used in lightweight decision-making scenarios. Obviously, the Q-learning algorithm is more adaptive to the research purpose and the demand of small-scale e-commerce businesses. Therefore, the research chooses the Q-learning algorithm as the core algorithm.

2.4 Establishment of Q-learning Model

Core factors that are used in the Q-learning algorithm model are shown as follows:

State Space: The research selects core features of commodities such as categories, prices, sales and discounts and quantifies the range of these feature values to catch decisive factors which will influence e-commerce pricing strategy.

Action Space: Based on multiple-level discrete pricing adjustment, the research designs the idea of multiple-level pricing adjustment (-6%, -3%, -1%, 0%, 1%, 3%, 6%), which covers reduction, maintenance and increasement.

Reward Function: The research refers to the principle that takes into account both profits and sales. The research uses "the weighted value of increasement of profits per unit + sales volume" as a core reward and quantifies the pros and cons of pricing actions to avoid resulting in the pricing deviations of singly pursuing profits or sales.

The dataset, which is after preprocessing, is divided into training set and test set in a 7:3 ratio. The research sets core parameters such as the learning rate $\alpha = 0.1$, the discount factor $\gamma = 0.92$, and the exploration rate $\varepsilon_{min} = 0.03$, based on other researches.

3. Results and Discussions

3.1 Data Analysis Results

Through the dataset after data preprocessing, the research could find that sales of commodities with a low price level (<200$) are much higher than commodities with a high price level (>500$). The research also finds that sponsorship has a great influence on average sales and profits. The average profits and sales of commodities that are sponsored far exceed those of commodities without sponsorship. The data shows that the average profits of commodities with sponsorship are 60.25% higher and the average sales are actually 277.89% higher.

The research finds that discounts will also impact the final profits. Although discounts will cause a decrease in profit per unit, the average sales will increase sharply by 245.54% and the final total profit will be higher than goods without discounts.

3.2 Q-learning Modelling Results

According to the heat map in Figure 2 of the Q-table, the research observes that the optimal actions(price adjustments) of the intelligent agent vary significantly in different states. This illustrates that the agent has learned how to price differently: The agent will decrease the price in a small area in a high-value situation; It will increase the price in a small range in a medium-value situation and avoid huge adjustments in a low-value situation.

Figure 2: Q-table heatmap (Picture credit: Original)

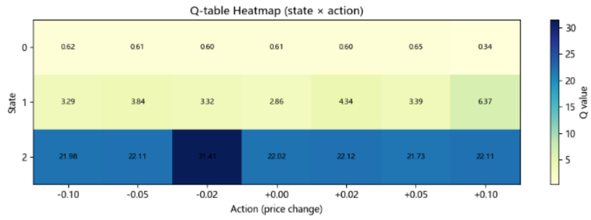


Figure 3: Key indices' curve during the model training process (Picture credit: Original)

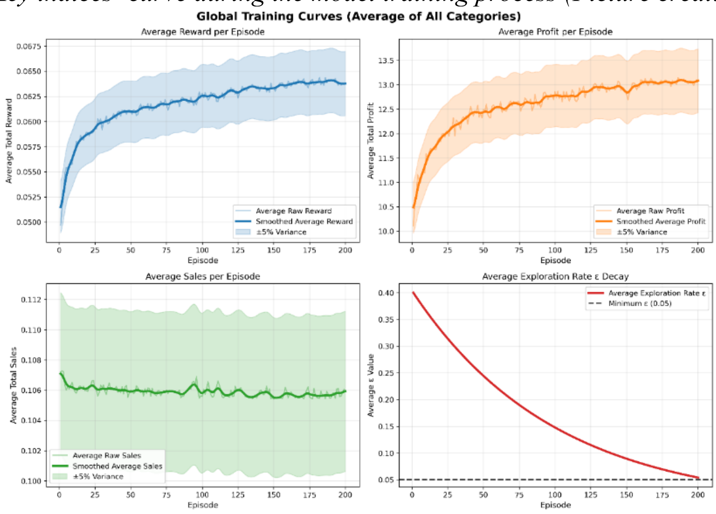
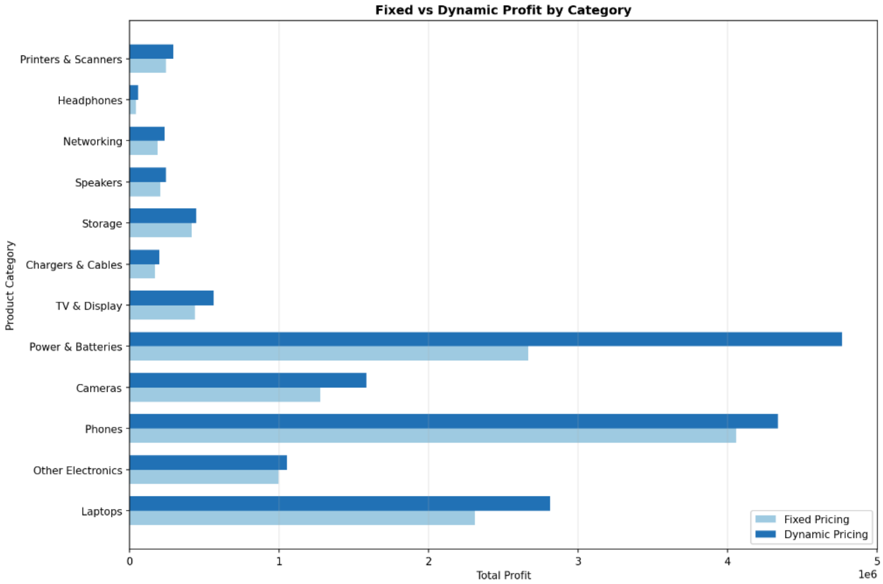


Figure 3 shows the key indices during the training process based on the Q-learning algorithm. Both the reward curve and profit curve present a trend of fluctuating upward, and reach the peak in the later stage of training. This illustrates that the agent's dynamic pricing strategy is improved continuously. Though the sales(normalization) curve exists fluctuation during the training process, the overall level remains at a relatively high level. The result illustrates that while striving for higher profits, the agent maintains a good sales performance. The exploration rate decreases from the initial 0.40 to 0.05. This guarantees the full exploration of the agent during the earlier training stage and concentration on the optimal strategies that have been learnt. All of the above data indicate that the training period of the agent is healthy and effective. The agent is growing up gradually from random exploration and becoming a model that could be applied stably.

3.3 Comparison between Dynamic Pricing and Fixed Pricing

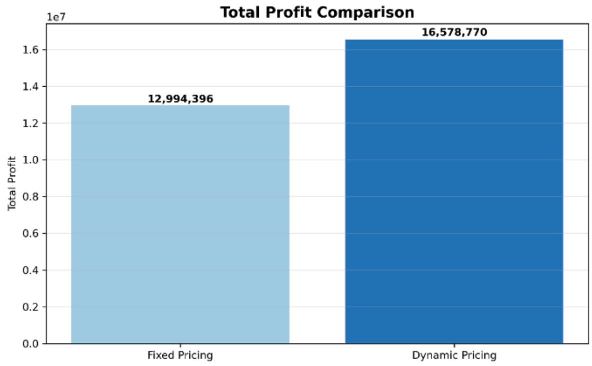
Figure 4 shows the comparison of different kinds of commodities' profits between the dynamic pricing strategy and the fixed pricing strategy. The research could clearly find that the dynamic pricing profits for different types of commodities are all higher than those of fixed pricing.

Figure 4: Profits comparison of different categories between fixed and dynamic pricing (Picture credit: Original)



In the premise of reducing sales volume, the total profit of dynamic pricing is 27.69% higher than that of fixed pricing. The results are shown in Figure 5. The result provided an intuitive proof that dynamic pricing strategy based on the Q-learning algorithm is better than traditional fixed pricing strategy.

Figure 5: Total profit comparison between fixed pricing and dynamic pricing (Picture credit: Original)



3.4 Discussions

Through the analysis of different categories in Amazon platform, the research finds that sales of e-commerce commodities are influenced by factors like prices, discounts and ratings: Commodities in low price will satisfy the requirement of the general public so that sales volume is higher; Discounts could reduce consumption threshold which could increase sales; Ratings and views are important decision bases, and a commodity with higher views and rating is easier to attract consumers to reach higher sales.

The reinforcement learning dynamic pricing model based on the Q-learning algorithm could capture market dynamics more comprehensively by setting the state space based on multidimensional sales data records in the Amazon platform. The application of this model is more flexible than traditional fixed pricing models and simpler and easier to operate than complex reinforcement algorithms like DQN and PPO, which will provide easier and more flexible pricing strategies for small-scale e-commerce businesses.

4. Conclusion

The research finds that sales of e-commerce commodities are affected by multiple factors such as prices, sponsorship, discounts and categories. Among these factors, there is a remarkable negative correlation between price and sales. This result is consistent with other mainstream e-commerce research results, which verifies the effectiveness and reliability of this research.

The research also finds that reinforcement learning pricing model based on Q-learning algorithm could effectively adapt to e-commerce pricing scenario. By drawing on design thought and the factors system of existing research, the model could learn pricing strategy in different market states through the dataset to achieve dynamic pricing. On the one hand, compared with the traditional fixed pricing strategy, this model has seen a significant increase in profits. On the other hand, the Q-learning algorithm is more convenient and economical than complex algorithms like DQN and PPO, which are more suitable for satisfying the pricing demand of small-scale e-commerce businesses.

Based on existing research methods, the method of dynamic pricing strategy for e-commerce in this research has a good practical value. This method provides a low-cost, handy pricing reference model for small-scale e-commerce businesses. The research is aimed at helping small-scale e-commerce businesses get out of the situation where they can only adopt the experiential pricing method when the price fluctuation frequency is high and the amplitude is large and achieve data-driven intelligent dynamic pricing strategies based on reinforcement learning. The research could also provide microscopic practical experience for delicacy management in the e-commerce industry.

The research hopes to introduce multi-dimensional data like inventory and operation cost and combine with factors like holiday in the future work. In this way, the model will be closer to daily e-commerce operating scenario.

References

[1] Akin, M. S. (2024). Enhancing e-commerce competitiveness: A comprehensive analysis of customer experiences and strategies in the Turkish market. Journal of Open Innovation: Technology, Market, and Complexity, 10(1), 100222.

[2] Gao, J., Siddik, A. B., Khawar Abbas, S., Hamayun, M., Masukujjaman, M., & Alam, S. S. (2023). Impact of E-commerce and digital marketing adoption on the financial and sustainability performance of MSMEs during the COVID-19 pandemic: An empirical study. *Sustainability*, *15*(2), 1594.

[3] Costa, P., & Rodrigues, H. (2024). The ever-changing business of e-commerce-net benefits while designing a new platform for small companies. *Review of managerial science*, *18*(9), 2507-2545.

[4] Li, H., & Chen, Z. (2024). How do e-commerce platforms and retailers implement discount pricing policies under consumers are strategic?. *Plos one*, *19*(5), e0296654.

[5] Nagle, T. T., Müller, G., & Gruyaert, E. (2023). *The strategy and tactics of pricing: A guide to growing more profitably*. Routledge.

[6] Ferreira, K. J., Simchi-Levi, D., & Wang, H. (2018). Online network revenue management using thompson sampling. *Operations research*, *66*(6), 1586-1602.

[7] Qian, W., & Wang, Y. (2024). Analyzing e-commerce market data using deep learning techniques to predict industry trends. *Journal of Organizational and End User Computing (JOEUC)*, *36*(1), 1-22.

[8] Suo, L., Lin, X., Chung, Y., & Kim, E. (2026). Layered Feature Engineering for E-Commerce Purchase Prediction: A Hierarchical Evaluation on Taobao User Behavior Datasets. *Computers, Materials, & Continua*, *87*(1).

[9] Chen, W., Tang, X., Qi, J., & Liu, H. (2026). Dynamic pricing strategy design for manufacturing service providers in manufacturing platforms. *Journal of Industrial and Management Optimization*, *22*(2), 1140-1167.

[10] Guo, X., & Zhang, L. (2025). Dynamic pricing models in e-commerce: Exploring machine learning techniques to balance profitability and customer satisfaction. *IEEE Access*.

[11] Li, L. (2026). The Application of Adaptive Reinforcement Learning in Dynamic Pricing Strategies. *Informatica*, *50*(10). https://doi.org/10.31449/inf.v50i10.11974

[12] Yildiz, S. A., Şahinaslan, Ö., Borandag, E., & Yücalar, F. (2026). Comparing the Performance of PPO and DQN Algorithms in Different Game Environments Using a Reinforcement Learning Approach. *Journal of Innovative Science and Engineering (JISE)*, *10*(1).

[13] Liu, H. (2023, August). Reinforcement Learning for E-Commerce Dynamic Pricing. In *International Conference on Economic Management and Green Development* (pp. 2051-2060). Singapore: Springer Nature Singapore.

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgment

This paper is an output of the science project.