ARTICLE

Dynamic Pricing Model of E-Commerce Platforms Based on Deep Reinforcement Learning

Chunli Yin[1,*] and Jinglong Han[2]

[1]College of Economics and Administration, Tonghua Normal University, Jilin, 130000, China

[2]Department of Administration Section, Tonghua Normal University, Jilin, 130000, China

[*]Corresponding Author: Chunli Yin. Email: TH_Yin@thnu.edu.cn

ABSTRACT

With the continuous development of artificial intelligence technology, its application field has gradually expanded. To further apply the deep reinforcement learning technology to the field of dynamic pricing, we build an intelligent dynamic pricing system, introduce the reinforcement learning technology related to dynamic pricing, and introduce existing research on the number of suppliers (single supplier and multiple suppliers), environmental models, and selection algorithms. A two-period dynamic pricing game model is designed to assess the optimal pricing strategy for e-commerce platforms under two market conditions and two consumer participation conditions. The first step is to analyze the pricing strategies of e-commerce platforms in mature markets, analyze the optimal pricing and profits of various enterprises under different strategy combinations, compare different market equilibriums and solve the Nash equilibrium. Then, assuming that all consumers are naive in the market, the pricing strategy of the duopoly e-commerce platform in emerging markets is analyzed. By comparing and analyzing the optimal pricing and total profit of each enterprise under different strategy combinations, the subgame refined Nash equilibrium is solved. Finally, assuming that the market includes all experienced consumers, the pricing strategy of the duopoly e-commerce platform in emerging markets is analyzed.

KEYWORDS

Deep reinforcement learning; e-commerce platform; dynamic evaluation; game model; pricing strategy

1 Introduction

With the development of the Internet and the popularization of e-commerce, it has become easier for people to obtain more comprehensive information on goods and services. Changes in the price of goods or services will also have an impact on consumers' shopping behavior in the shortest time, which directly affects corporate profits. To maximize efficiency, companies often adjust the prices of goods or services regularly or irregularly based on certain factors, which is also consistent with the goal of deep reinforcement learning in the field of artificial intelligence. The goal of deep reinforcement learning is to maximize long-term benefits. Therefore, the technical means of deep reinforcement learning can achieve the intelligent pricing of goods

or services. The e-commerce customer's purchase behavior prediction makes a real-time prediction of an online customer's purchase tendency behavior based on the behavioral laws contained in the consumer's historical access click operations, server logs, browsing records and product feedback information. Therefore, customers can recommend products, formulate marketing strategies, and determine the purchase and shipment of platform products.

Dynamic pricing is a strategy for enterprises to dynamically adjust commodity prices based on customer demand, their own supply capacity and other information to maximize revenues [1], and some scholars also call it personalized pricing [2]. With the continuous development of artificial intelligence technology, increasingly more scholars have sought to use intelligent methods to solve dynamic pricing problems. Deep reinforcement learning is one of the most widely used technologies. It is inspired by the ability of people and animals in nature to adapt to the environment effectively. Learning from the environment through continuous trial and error is an important branch of machine learning. It has a very wide range of applications in the fields of artificial intelligence problem solving, multiagent control, robot control and motion planning, and decision-making control [3,4], Learning from the environment is one of the core technologies of intelligent system design and decision-making, and it is also a key issue in dynamic pricing in strategy research. The development is of the Internet, increasingly fierce market competition, and the need for customer management have transformed the pricing model of commercial enterprises from fixed prices to dynamic pricing. This transformation relies heavily on the development of the Internet, market competition, and customer management needs. Dynamic pricing in an e-commerce environment is based on the customer's value of a subproduct or service [5,6] and a dynamic price adjustment strategy for different customers or commodities. Sellers can achieve the goal of dynamic pricing by integrating customer databases that meet specific standards of target customers [7,8]. When the quantity demanded is random and price sensitive, dynamic pricing becomes an effective method to maximize profits [9,10]. Varying, dynamic prices are an important feature of e-commerce pricing. Effectively formulating dynamic pricing strategies is an important factor for enterprises to succeed in the field of e-commerce [11,12]. E-commerce companies need to adopt four methods of dynamic pricing decision-making strategies, namely, a time-based pricing strategy, a market segmentation and limited rationing strategy, a dynamic marketing strategy, and comprehensive application based on dynamic pricing [13,14]. The time-based pricing strategy is implemented according to the price difference that consumers can bear at different times. The key is to grasp the psychological difference of customers' price tolerance at different times [15,16]. The basic principles of the market segmentation and limited rationing strategy are as follows: using different channels, different times, and different energy expenditures, customers have different price tolerance psychologies; companies have developed special product and service portfolios; and companies differentiate pricing based on different product configurations, channels, customer types, and times [17,18]. The dynamic marketing strategy takes advantage of the powerful advantages of the Internet to quickly and frequently implement price adjustments based on changes in supply and inventory levels to provide customers with different products, various promotional offers, multiple delivery methods, and differentiated products. In addition, in the actual application process, the enterprise may consider implementing a certain strategy individually or combining strategies. When formulating pricing strategies, the best approach is to experiment with specific customer groups, select the best pricing model [19,20], and then adjust the model accordingly. In dynamic pricing, companies can use some modeling methods, such as inventory models, data-driven models, game models, machine learning models, and simulation models, to assist analysis and decision-making [21,22]. Data-driven models use statistical or simulation techniques to effectively use customer data to calculate appropriate dynamic prices. Currently, dynamic pricing is

also one of the important research areas of customer relationship management and data mining technology [23,24]. Negotiation is a dynamic interactive process for the parties involved in the transaction to reach a transaction agreement. During the negotiation process, all parties to the negotiation exchanged proposals for negotiation reflecting their beliefs and intentions. In each round of negotiation, the agent proposes negotiation proposals based on its own negotiation strategy and evaluates the received proposals to determine whether to accept the other party's proposal [25,26]. The negotiation process is usually a dynamic process of learning and updating your beliefs.

Therefore, in-depth study of the application of deep reinforcement learning methods in the field of dynamic pricing is of great significance to the development of artificial intelligence, deep reinforcement learning methods and their applications in dynamic pricing and other fields. We will review two aspects of deep reinforcement learning technology and its specific application in the field of dynamic pricing. First, based on the existing dynamic pricing, the relevant key technologies of deep reinforcement learning are introduced. Then, the application of deep reinforcement learning in dynamic pricing is reviewed from different perspectives, and the advantages and disadvantages are analyzed. Next, we systematically review platform pricing theory and differential pricing theory, use game theory as the main research method to establish a competitive platform enterprise pricing game model, and analyze network externalities and consumer switching costs in mature and emerging markets as well as the impact of enterprise pricing strategies on market equilibrium to systematically analyze the dynamic pricing behavior of platform companies. The first section of this paper is the introduction, the second part introduces the construction of the e-commerce dynamic pricing model based on data mining, the third section studies the deep reinforcement learning transaction recognition model, and the fourth section studies the research on the e-commerce dynamic pricing model. The results and discussion are given in the fifth section, and the sixth section is a summary.

2 Construction of the E-Commerce Dynamic Pricing Model Based on Data Mining

At present, data mining should focus on customer relationship management in the application research of e-commerce tools. Although some scholars have also proposed the theory of applying data mining technology to e-commerce dynamic pricing tools, many of theories are scattered and general. Theoretical analysis, without comprehensive and systematic application analysis, lacks the overall grasp of the application of data mining in the dynamic pricing of e-commerce, and the effectiveness of data mining cannot be fully utilized. To this end, this article establishes a dynamic pricing model for e-commerce based on data mining and proposes applying data mining technology to dynamic pricing decisions, which will be of great help to e-commerce companies in pricing decisions. The model is composed of three layers, namely, the data layer, the analysis layer and the decision layer, from top to bottom [27]. These three levels are closely connected, and each level contains the application of related theories and technologies of data mining and dynamic pricing, which together achieve the goal of e-commerce dynamic pricing decisions. The model is shown in Fig. 1.

2.1 Data Layer

The task of the data layer is to collect data related to pricing decisions and preprocess these data to form a data warehouse to prepare for the next stage of data mining.

After the data source is selected, the data must be collected in a timely and high-quality manner and imported into a series of data files, usually in the form of database storage. This step

can be used to generate and obtain data in the form of network-free action, but it also requires enterprises to build a basic database in vain and update it in time according to inventory, market and sales reports. The data collected through various channels may have considerable redundancy, or there may be inaccurate, incomplete, and inconsistent data. This requires preprocessing the data if the data are extracted, verified, and cleaned. Conversion, integration and other processes to improve data quality, form a data collection suitable for data mining, and load it into the data warehouse.
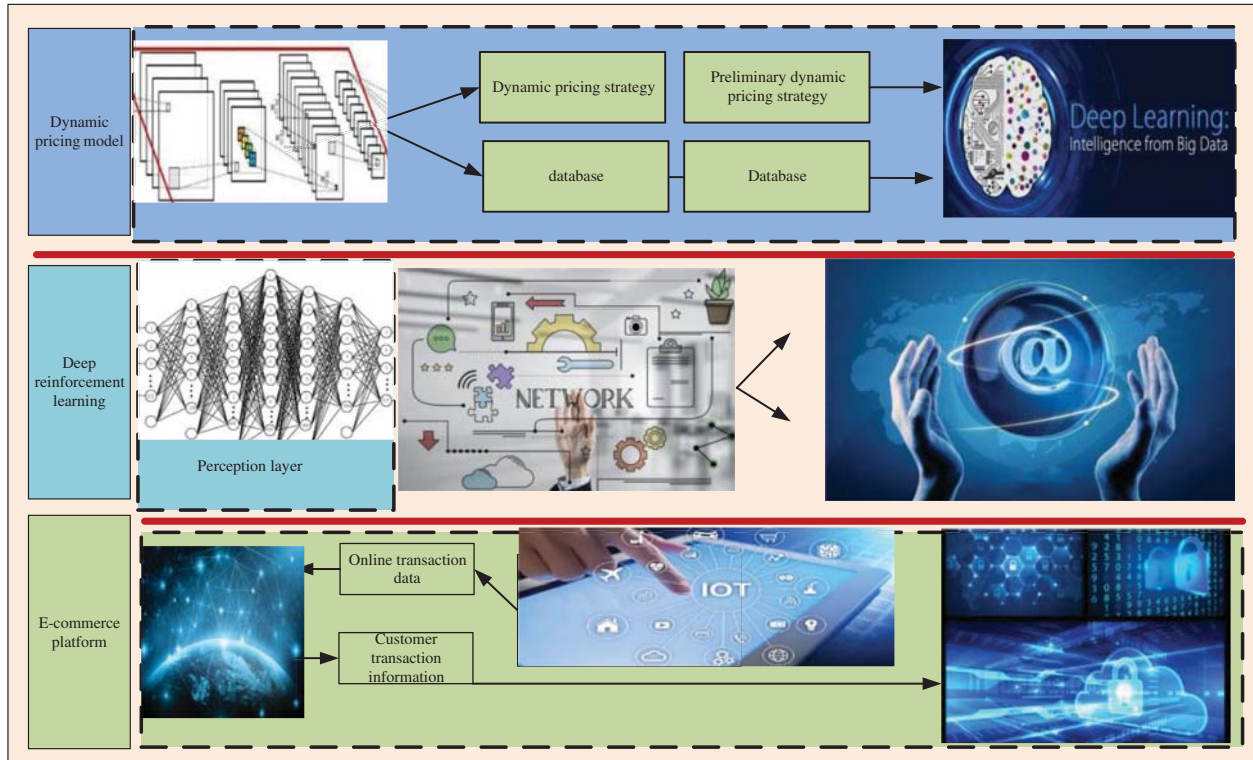
Figure 1: E-commerce dynamic pricing model based on deep reinforcement learning

2.2 Analysis Layer

The main tasks of the analysis layer are to use data mining models and related algorithms to analyze and process the data obtained, to mine knowledge useful for dynamic pricing decisions, and to form the initial knowledge base. The realization of this stage is the core of the whole model construction. In dynamic pricing-assisted decision-making tools, methods such as association rules, classification, clustering, and sequence pattern Analysis can be used.

Correlation analysis aims to mine the data relationships or rules hidden in the data (warehouse) database, that is, to discover the laws or knowledge of dependence or association between an event and other events. In e-commerce dynamic pricing tools, association analysis can be used to find customer's views on various product visits and purchases on a website, to determine various associations of customer buying behavior and to acquire information on customer buying behaviors and product prices and other product information The relationship between these types

of information can be used to further discover the relationship between demand and price, which is an important point for dynamic pricing decisions. The collected basic customer data and transaction data can use the Apriori algorithm to discover the details of the customers' purchase associations [28,29].

2.3 Decision-Making Layer

The decision layer is a key part of the realization of the entire model. The main task of this layer is to make dynamic pricing decisions based on the knowledge base that established by the analysis layer and combined with the business strategy of the enterprise.

Through the application of analysis layer data mining technology, one can obtain the characteristics of the access patterns, purchase patterns, habits and preferences of different customer groups; the correlation characteristics between price and demand and the sales of goods, as well as the number of people related to the goods and the amount of sales; the predicted value of time series data of inventory data; etc. Using this basic knowledge, the seller can make preliminary dynamic pricing decisions. In the time-based strategy, first determine the appropriate initial is determined, and factors such as historical sales data, cost information are comprehensively considered; then, given the initial maximum or minimum price, a double price change basis can be used to adjust the price by setting a time threshold on the quantity of goods or demand, and then controlling the time and range of the price changes [30]. When using market segmentation strategies to differentiate pricing based on customer information, the strategies must be understood by customers, and strategic consumers must adopt appropriate and targeted dynamic pricing strategies based on their purchase records and price sensitivity [31], thereby achieving customer satisfaction.

The ultimate goal of dynamic pricing for e-commerce companies is to maximize customer satisfaction or maximize corporate profits; moreover, companies have different goals in different periods of their operations and different requirements for pricing strategies. Therefore, the enterprise pricing decision is a multiobjective decision-making process. To this end, we must first establish a multiobjective function. Using various mined related information and forecast data, an appropriate demand function can also be established, and the price can be adjusted according to customer demand or corporate sales/inventory. When applying this traditional enterprise dynamic pricing strategy, there are many mature pricing models that can be referenced. For example, the pricing model based on inventory control uses dynamic programming to achieve dynamic pricing and the application of other mathematical models.

3 Deep Reinforcement Learning Transaction Recognition Model

The intelligent behavior between a group of autonomous and intelligent agents, and how they coordinate with each other to take action to achieve a certain goal forms Multi-Agent System (MAS) behavior. In MAS, the mutual coordination among agents includes the coordination of knowledge, goals, skills and planning directions. The goal they achieve may be a solution goal or a set of several solution goals. According to the definition, the multiagent collaborative solution model is shown in Fig. 2.

The input layer of the network has no calculation nodes, and is only used to obtain external input signals. The neurons of the hidden layer and the output layer are the calculation nodes. The basis function is a linear function and the activation function is a hard limit function. Suppose the MLP has only one hidden layer, and its input is $t_1, t_2, \ldots, t_n$. In addition, the hidden layer has m_1 neurons, and their outputs are $h_1, h_2, \ldots, h_n$. Finally, the network output is represented by δ_p.

Then, the output of the j-th neuron in the hidden layer is:

$$h_j = f\left(\sum_{i=1}^{n} \omega_{ij} t_i - \delta_i\right), \quad j = 1, 2 \ldots n \tag{1}$$
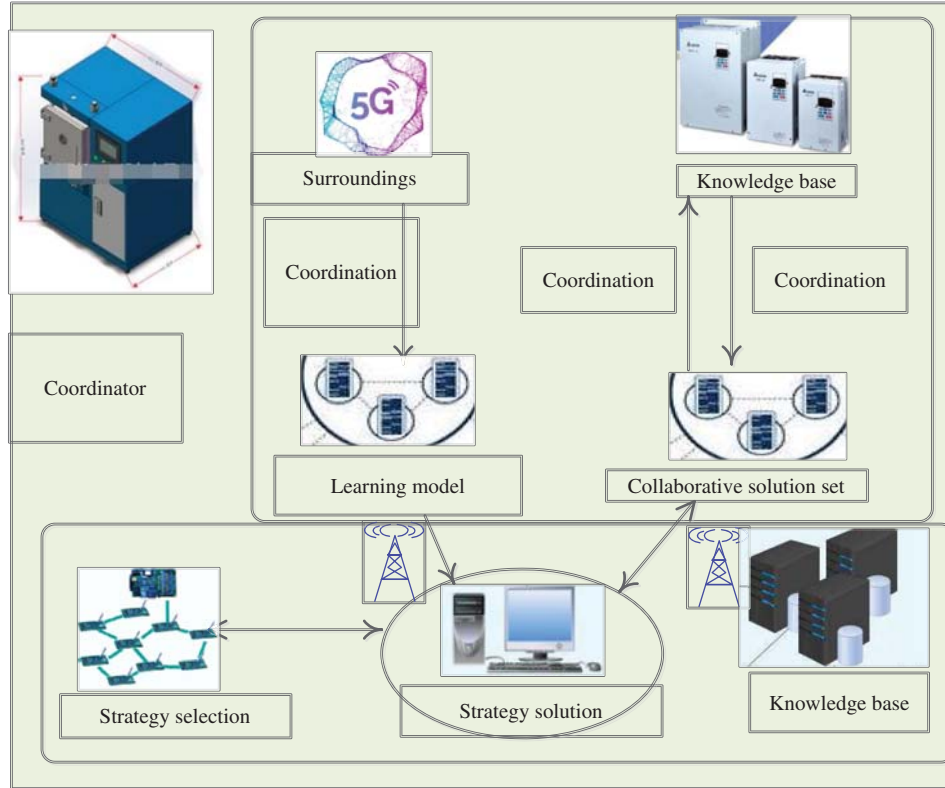

Figure 2: Multiagent collaborative solution model

When the multilayer perceptron is used to solve practical problems, it must first solve the problem of training the connection weight between the input and the hidden layer; however, because it is difficult to determine the expected output value of the hidden layer output, the network weight training cannot be achieved. Therefore, people seek other neural network solutions to solve the linear inseparable problem, and the BP network is such a network.

An e-commerce platform, the platform often needs to analyze and predict the customers' online shopping behavior. Based on the customer information database, the e-commerce platform completes real-time and targeted predictions of customers' online shopping behaviors, thus embodying intelligent predictions of customer behaviors. Therefore, as a complete predictive model system, we first need to use methods such as data mining, machine learning, and statistics to discover knowledge and extract features from the data. Based on this, we build a knowledge base of customer online shopping behavior as knowledge guidance, storage and representation and then establish a system from data input to prediction behavior. The main research contents are as follows:

(1) Consumer behavior data processing and feature construction

First, the interactive logs are extracted from the E-commerce interactive system to pre-pare data related to consumer behavior analysis and prediction. Then, data preprocessing, including data cleaning, filling missing values and removing outliers, is performed to ensure the uniqueness of the data to achieve consumer behavior prediction and provide a good basic guarantee.

(2) Construction of consumer behavior characteristics

Based on the original data, the user purchase behavior features are extracted. According to different classification methods, the features can be divided into original and extended or static and dynamic, or two or more categories of features can be combined into a new feature. To obtain a good prediction effect, the data and characteristics largely determine the upper limit of the model prediction. Therefore, how to construct suitable characteristics is the key factor to provide a good guarantee for the analysis of user behavior.

(3) Consumer behavior prediction model

The accuracy of the prediction model is the key to ensuring the prediction and analysis of consumer behavior. Although there are many prediction models at present, they are far from meeting the accuracy requirements under real conditions. How to use consumer static or dynamic data analysis to accurately predict consumer behavior is an extremely critical technology.

(4) Consumer shopping behavior analysis

In the representational learning of data, the goal is to seek better representation methods and create better models to learn these representation methods from large-scale unlabeled data. The workflow of consumer shopping behavior analysis based on deep learning is mainly divided into the following four steps.

Step 1: Prepare and process the data set. This step includes collecting user interaction information, data cleaning, etc.

Step 2: Feature construction is divided into three stages: feature selection, forming the sample training set and test set, and feature processing. Feature selection is the key to building a prediction model. It selects feature sets that are extremely important for classification from a large number of data sets, thereby improving the model's prediction accuracy and shortening the running time. The inconsistency of feature dimensions and units which selected for different dimensions will affect the weight of the assessment features, which in turn affects the model's estimated effect. Therefore, feature management is required to perform normalization.

Step 3: Design and train the prediction model. Select the basic model framework such as the convolutional neural network (CNN)+ recurrent neural network (RNN). Then, using the framework, randomly sample negative samples of the data, adjust the number of network layers, determine the loss function, and design the learning rate and other hyperparameters. The BP algorithm back-propagates using stochastic gradient descent (SGD) or the Adam algorithm to optimize model parameters.

Step 4: Model verification. Untrained data are used to verify the generalization ability of the model. If the prediction result is not ideal, you need to redesign the model and conduct a new round of training. There are several mature deep learning models to date, including deep neural networks (DNNs), convolutional neural networks (CNNs), deep confidence networks (DBNs), and recurrent neural networks (RNNs). These methods have been used in machine vision,

natural language processing, bioinformatics, speech recognition and other fields and have achieved remarkable results.

4 Research on E-Commerce Dynamic Pricing Model

4.1 Deep Reinforcement Learning

The working principle of deep reinforcement learning is similar to that of human learning. If an action of the agent obtains a positive reward from the environment, then the agent's future actions will be enhanced; conversely, if a negative reward is received, then the future actions will be weakened. The goal of deep reinforcement learning is to learn an action strategy, so that the system can obtain the largest cumulative reward. In deep reinforcement learning, the agent selects and executes an action a in the environment, the environment changes to s after accepting the action, and feeds back a reward signal r to the agent, and the agent selects the subsequent action according to the reward signal. In research related to dynamic pricing, the goal of deep reinforcement learning systems is to enable manufacturers to maximize their overall returns while ignoring the short-term benefits of a single transaction. A deep reinforcement learning architecture generally includes four elements: Strategy, reward and punishment feedback, the value function, and the environmental model. The environment-related factors of dynamic pricing are numerous and complex. Previous studies of dynamic pricing in deep reinforcement learning were mainly based on the following environmental frameworks.

Deep reinforcement learning can be divided into value-based deep reinforcement learning and policy-based deep reinforcement learning. In deep reinforcement learning based on value functions, commonly used learning algorithms include the Q-learning algorithm, SARSA algorithm and Monte Carlo algorithm. In dynamic pricing research based on deep reinforcement learning, these three algorithms are also frequently used algorithms. (1) Q-learning algorithm. The Q-learning algorithm is a model-free algorithm, and its iteration equation is expressed as:

$$\delta_i = m_{i+1} + \lambda \max Q(s_{i+1}, \alpha) - Q_t(s_t - \alpha_t) \tag{2}$$

where $Q(s_{t+1}, \alpha)$ is the state action value at time t, m is the reward value, λ is the discount factor, a is the learning rate, s_t is the time difference error, and a′ is the action that state s_{t+1} can perform.

4.2 SARSA Algorithm

SARSA is a strategy algorithm that can find the optimal strategy through iteration of the state action value function when the reward function and state transition probability are unknown. When the state action pair is accessed infinitely, the algorithm will converge to the optimal strategy and state action value function with a probability of 1. The SARSA algorithm adopts relatively safe actions in learning, so the convergence speed of the algorithm is slow. The iteration equation is expressed as:

$$Q(s, \alpha) = Q(s, \alpha) + \alpha \left\{ \lambda + \lambda Q(s_t, \alpha_y) \right\} \tag{3}$$

4.3 Monte Carlo Algorithm

The Monte Carlo algorithm does not require complete knowledge of the environment, and only requires experience to solve the optimal strategy. These experiences can be obtained online or according to some simulation mechanism. The Monte Carlo method keeps a count of the frequency of state actions and future rewards and establishes their values based on estimates. The Monte Carlo technique estimates the return of the average sample based on the sample. For each state, keep all the states obtained from state, and the value of one state is their average value.

Especially for periodic tasks, Monte Carlo technology is very useful, especially for periodic tasks. Since sampling depends on the current strategy, the strategy only evaluates the reward of the proposed action. The value function update rule is expressed as:

$$V(s_t) = V(s_{t+1}) + \alpha (\lambda_i - V(s_t)) \tag{4}$$

where λ_i is the reward value at time t, and a is the step parameter.

4.4 E-Commerce Dynamic Pricing Model

Dynamic pricing in e-commerce is one of the fastest growing areas in Internet applications. By applying an online auction-style dynamic pricing model, companies can products based on the true market value of commodities. In most real markets, only the buyer himself knows exactly how many items he will be willing to buy at a specific price level. The seller does not have perfect knowledge of the market demand and cannot accurately understand the buyer's valuation. The seller only has statistical information about the market demand. This chapter mainly starts from the "individual valuation" model and discusses the "online auction" where a single seller provides auction items, multiple buyers bid on the auction items, and an auction-type dynamic pricing model exists.

4.4.1 Auction Dynamic Pricing Model and Analysis Based on Uncertain Demand

Suppose that the system is a market environment where a certain auctioneer on the Internet auctions many items, there are many demanders, and the quantity of demand is uncertain. Let the set of n demand-side agents sets be N, and let F be the set of all possible allocation combinations among them. Each distribution combination $\alpha \in F, Agent, j \in N$ is assigned a monetary amount $v_j(\alpha)$, and $v_j(\alpha)$ is private information, that is, an "independent individual valuation". Independence means that each buyer's personal information is independent of other bidders' personal information. Personal valuation means that once a buyer uses his own information to evaluate the value of the auction target, this valuation will not be subject to his follow-up knowledge of the impact of any other purchaser's personal information.

$$V(N) = Max \sum_{\alpha \in F} v_j(\alpha)$$

$$V(N \backslash j) = Max \sum_{\alpha \in F \backslash j} v_j(\alpha) \tag{5}$$

If the auction process is closed, the auction process is as follows: Agents submit their monetary amount function, and we temporarily assume that they are faithfully submitting their monetary function. Later, it will be explained that false reporting cannot improve the income of any agents. The auctioneer chooses the best distribution plan for all calculations of V(N) and V(N/j). In this way, the agent's payment is:

$$V(N \backslash j) = \sum_{i \neq j} v_i(\alpha^*) \tag{6}$$

The net income is:

$$v_j(\alpha^*) - \left\lfloor V(N \backslash j) - \sum_{i \neq j} v_i(\alpha^*) \right\rfloor \tag{7}$$

Suppose that the seller agent S has 5 indivisible commodities and that 5 bidders $A_1, A_2 \ldots A_5$ participate in the bidding. The possible demand and bid of each bidder $A_1, A_2 \ldots A_5$ are shown in Fig. 3, and the revenue of the seller's agent S is shown in Fig. 4.
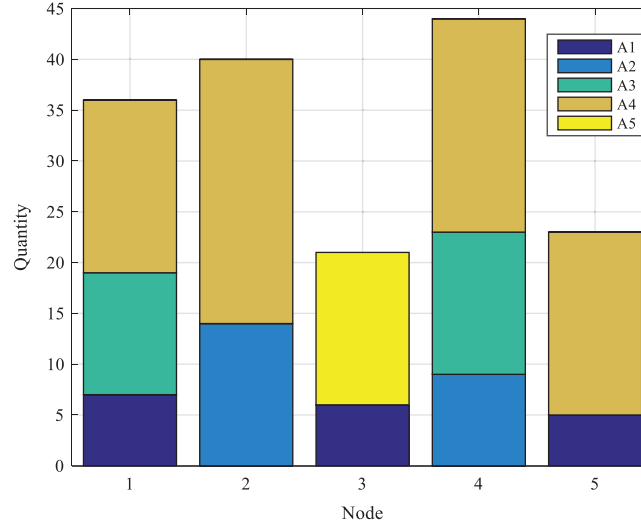

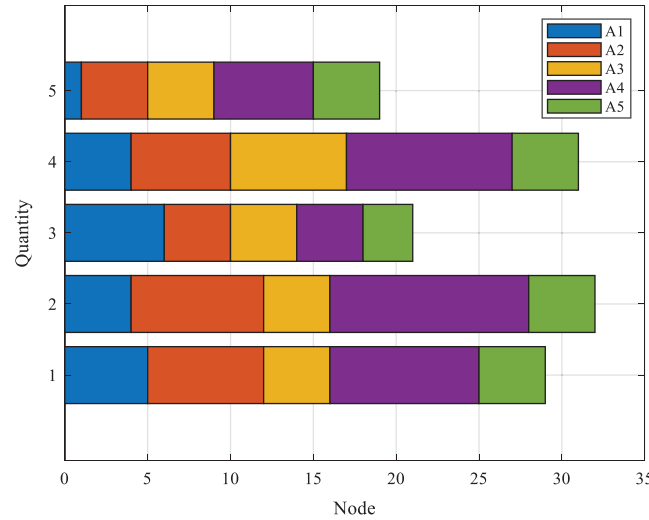**Figure 3:** Possible demand and bids of bidders

Figure 4: Income of the seller's agent S

4.4.2 Exchange-Based Dynamic Pricing Model Based on Uncertain Demand

In the online auction MDA market environment, it is assumed that there are m buyers and n sellers. The number of buyers and the number of sellers are arbitrary, and it is not assumed that there are more buyers than sellers or more sellers than buyers. Each buyer $i = 1, 2 \ldots m$ wants to buy X_i units of homogenous goods. Each seller $j = 1, 2 \ldots n$ has Y_j units of homogenous goods for sale. To simplify the analysis, it is assumed that X_i and Y_j are public information for all

participants, that is, the m buyers and n sellers know each other's quantity demanded for the commodity or quantity supplied of the commodity. However, the reserved price b_i of buyer i and the reserved price of seller j are s_j is private information, that is, each is an "independent individual valuation." The agent in the model assumes that their reserve price is static and remains unchanged during the auction.

When the auction is over (that is, market liquidation), assume that buyer i purchases t_{ij} units of goods from seller j, and m_{ij} is the transaction price of the transaction. In this way, the utility obtained by buyer i after the auction can be defined as:

$$ub_i = \sum_{j=1}^{N} (b_i - t_{ij}) m_{ij} \tag{8}$$

The utility obtained by seller j can be defined as:

$$us_j = \sum_{i=1}^{N} (t_{ij} - s_j) m_{ij} \tag{9}$$

If all information is public, the maximized total market value, that is, the aggregate utility of all agents participating in the auction, can be obtained through the following linear programming problem:

$$Max \sum_{i=1}^{m} \sum_{j=1}^{n} m_{ij} (b_i - s_j) \tag{10}$$

5 Results and Discussion

Since the third-party brushing platform uses exchange information for brushing customers and merchants as a profit method, to obtain false transaction information, the author entered the third-party brushing platform by pretending to be a brushing identity and released the billing information through the third-party platform. Then, the author collected comments and transaction records of fake trading products. In addition, to collect data on normal trading commodities, the author chose official flagship stores (such as Hailan House, ONLY, VERO, MODA, Uniqlo and other official Tmall flagship stores with a high reputation in reality) and combined these product reviews and transaction records are used as training sets for regular trading products. Based on this, the author collected the data of nearly 130,000 reviews data and the transaction record data of the most recent month of the product as the input data set of the recognition model. After normalizing the data, an independent sample t-test was performed, and the results are shown in Tab. 1.

It is not difficult to see the convergence of the algorithm in Fig. 5. In the case that the specified number of iterations is 80, the scale of the problems involved in our discussion can converge well to the optimal value. As the scale of the problem continues to increase, the maximum number of iterations can be adjusted according to the specific situation.

Taking the dynamic bidding market as an example, in K transaction cycles, there are N transaction agents bidding on M brand cars, and the matching agent calculates and matches the bids based on the matching transaction model and algorithm. Trading agents are risk-neutral, and all participate in bidding in a random optimal way. According to the microstructure and

dynamic trading mechanism, the market equilibrium easily forms for the same type of commodity bidding; however, when multiple types of goods are matched at the same time, the market status will become very complicated. Therefore, we designed market price dynamic fluctuations and equilibrium experiments for single commodities and multiple types of commodities.

Table 1: Summary statistics of the characteristics

Is it false	N	Mean	Standard deviation	Standard error of the mean
Store registration time 0	125	0.479	0.587	0.0534
Store registration time 1	85	−0.865	0.927	0.1023
Refund dispute rate 0	116	−0.4789	0.0167	0.00145
Refund dispute rate 1	78	0.8675	1.234	0.1356
Product review rate 0	118	−0.2689	0.0428	0.00256
Product review rate 1	78	0.4456	1.267	0.1543
Single product review Ratio 0	119	−0.578	0.0278	0.00267
Single product review Ratio 1	77	1.036	1.0387	0.1156
Collection rate 0	116	−0.234	0.3768	0.03345
Collection rate 1	75	0.367	1.467	0.1678
Repeat review rate 0	115	−0.3567	0.265	0.0243
Repeat comment rate 1	82	0.6754	1.675	0.1864
Average comment Length 0	122	−0.4876	0.421	0.0375
Average comment Length 1	82	0.7894	1.234	0.1365
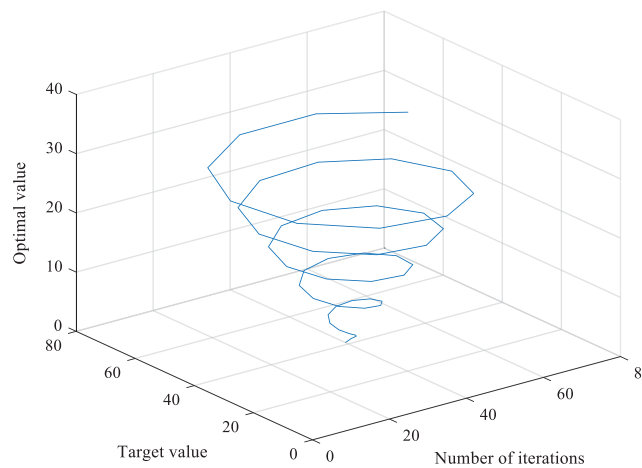
Figure 5: Variation curve of individual target value for 80 iterations under 80 periods

In experiment 1, set $N = 26$, $K = 2$, and $M = 1$; and a total of 120 bids were made. In experiment 2, set $N = 26$, $K = 36$, and $M = 5$; and a total of 140 bids were made; Matching

and matching will be performed according to the bid price, and the matching transaction price will be calculated. In addition, based on the actual transaction data of 400 groups of an automobile trading market, the standard deviation of the matching transaction price is calculated.

Let EquTe represent the degree of equilibrium of market prices. Then, according to the trading entropy and Walrasian equilibrium, EquTe can be defined as the probability of the occurrence of an equilibrium trading price.

$$EquTe = \frac{1}{K} \sum_{k=1}^{K} \frac{p_k^* - E_k(V_k)}{E_k(V_k)} \tag{11}$$

Here, $E_k(V_k)$ is the expectation of the final value of the commodity in trading cycle k, p_k^* is the equilibrium price of market liquidation, K is the number of rounds of the trading cycle, Time represents the trading cycle, and Price Diff represents the current market transaction price correction. The experimental results are shown in Figs. 6 and 7.
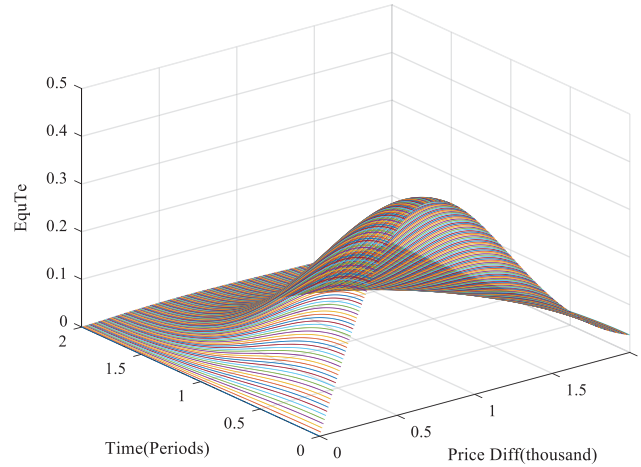
Figure 6: Price fluctuation and equilibrium of a single commodity in 2 rounds of bidding transactions

Fig. 6 shows that in the trading cycle, when the price correction value of the commodity market bidding is at a medium level, the probability that the transaction price reaches an equilibrium is greater. In fact, the market price dispersion measure is close to the ideal value. After multiple trading cycles (L = 2), the peak sequence of price fluctuations forms a Walrasian equilibrium curve that matches the trading market. Fig. 7 shows that when there are multiple types of commodities (M = 80) participating in multiple rounds of bidding transactions (L = 80), the equilibrium point sequence of various types of commodity transactions forms multiple peaks, which better reflects the market competition and is balanced. The above experimental results show that in the multiagent matching trading model, the price correction 8 is insensitive to individual trading agents, and the entire market has good sensitivity to 8. Through multiagent bidding that continuously adjusts the transaction price, the market can achieve an equilibrium with better efficiency. The experimental results show that the matching transaction model is cost effective and has good market efficiency. According to the market prediction model, the transaction price fluctuation trend of a certain brand car in the auto trading market is predicted. There are

27 risk-neutral agents participating on the bidding of a certain type of car (using the buyer's market as an example). Then the actual transaction price, average bid price and transaction forecast price fluctuation trend of the 7 matching transaction cycles are shown in Fig. 8.
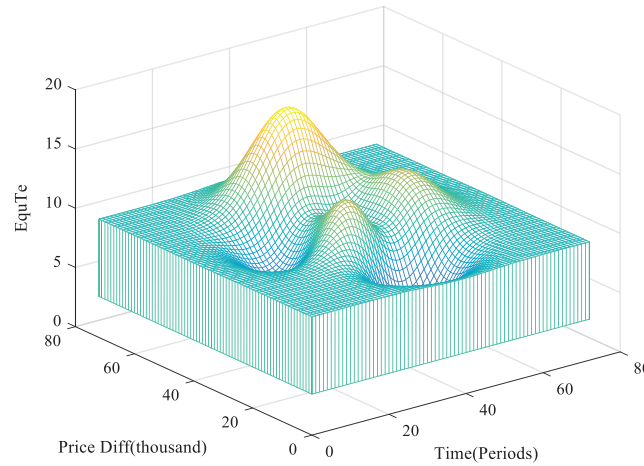

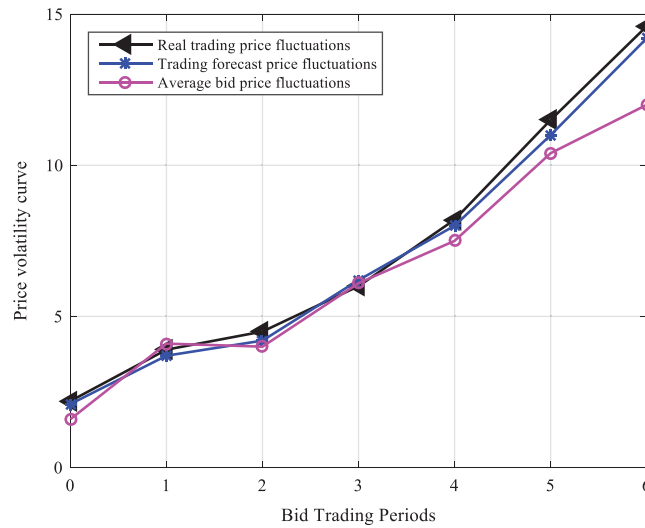**Figure 7:** Price fluctuation and equilibrium of commodities in 80 rounds of bidding transactions

Figure 8: Fluctuation curve of actual transaction price, average bid price and transaction predicted price

Fig. 8 shows that the predicted price fluctuations is basically between the actual price level and the average bidding level, which better reflects the price trend of this type of car brand in the market matching transaction. Through market price prediction, the trading agent further adjusts the bidding strategy to form their initial trading price and belief.

6 Conclusion

The development of Internet technology and the popularization of the networks have expanded the application range of data mining, and the application of data mining in e-commerce tools has become increasingly extensive. This article uses data mining theory and methods and dynamic pricing-related strategies to establish an e-commerce dynamic pricing model based on data mining. Based on the mechanism of the model, the auction mechanism is analyzed and discussed and suggestions for improving pricing strategies are proposed. The comprehensive data mining of the model system in the application of e-commerce dynamic pricing tools has a relatively general applicability to e-commerce enterprises, which can help enterprises improve customer satisfaction and economic efficiency. The E-commerce platform integrates the production and sales of the enterprise, and the production and sales are mutually restricted. In the study of the specific substitution effect of the multiproduct dynamic pricing research, we simply considered the production constraints, but did not closely integrate production planning and sales and combine them together. How to adjust commodity prices according to changes in production plans is a question that requires further study.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

1. Hsu, L. F. (2016). E-commerce model based on the internet of things. *Advanced Science Letters, 22(10),* 3089–3091. DOI 10.1166/asl.2016.7992.

2. Huang, P. (2016). Research on the construction mode of e-commerce business platform in higher vocational colleges based on the government purchase of public service theory. *Electronic Test, 16(8X),* 170–171. DOI 10.16520/j.cnki.1000-8519.2016.16.093.

3. Zhang, H., Tian, Y., Zhang, G. (2016). Dynamic option pricing model based on the realized-GARCH-NIG approach. *Open Journal of Social Sciences, 4(3),* 66–71. DOI 10.4236/jss.2016.43011.

4. Kraines, S., Koyama, M., Weber, C. (2017). A collaborative platform for sustainable building design based on model integration over the internet. *International Journal of Environmental Technology & Management, 5(2),* 135–161. DOI 10.1504/IJETM.2005.006847.

5. Kamalapurkar, R., Klotz, J. R., Walters, P., Dixon, W. E. (2018). Model-based reinforcement learning in differential graphical games. *IEEE Transactions on Control of Network Systems, 5(1),* 423–433. DOI 10.1109/TCNS.2016.2617622.

6. Hu, C. (2016). Application of e-learning assessment based on AHP-BP algorithm in the cloud computing teaching platform. *International Journal of Emerging Technologies in Learning, 11(8),* 27. DOI 10.3991/ijet.v11i08.6039.

7. Oliveira, S. M. D., Häkkinen, A., Lloyd-Price, J., Tran, H. Kandavalli, V. et al. (2016). Temperature-dependent model of multi-step transcription initiation in Escherichia coli based on live single-cell measurements. *PLoS Computational Biology, 12(10),* e1005174. DOI 10.1371/journal.pcbi.1005174.

8. Yun, Q. J., Fei, Z., Yue, Z. (2016). Change and prediction of the land use/cover in Ebinur Lake Wetland Nature Reserve based on CA-Markov model. *Journal of Applied Ecology, 27(11),* 3649–3658. DOI 10.13287/j.1001-9332.201611.027.

9. Nuan, W., Zheng, H. L., Ling, P. Z. (2017). Deep reinforcement learning and its application on autonomous shape optimization for morphing aircrafts. *Journal of Astronautics, 38(11),* 1153–1159. DOI 10.3873/j.issn.1000-1328.2017.11.003.

10. Wan, C., Li, T., Guan, Z. H. (2017). Spreading dynamics of an e-commerce preferential information model on scale-free networks. *Physica A Statistical Mechanics & Its Applications, 467,* 192–200. DOI 10.1016/j.physa.2016.09.035.

11. Li, C., Cao, L., Chen, X. (2018). Cloud reasoning model-based exploration for deep reinforcement learning. *Dianzi Yu Xinxi Xuebao/Journal of Electronics & Information Technology, 40(1),* 244–248. DOI 10.11999/JEIT170347.

12. Zeigheimat, F., Ebadi, A., Rahmati-Najarkolaei, F., Ghadamgahi, F. (2016). An investigation into the effect of health belief model-based education on healthcare behaviors of nursing staff in controlling nosocomial infections. *Journal of Education & Health Promotion, 5(1),* 23–35. DOI 10.4103/2277-9531.184549.

13. Li, H. H., Cang, Y. C. (2016). GM (0,N) model-based analysis of the influence factors of network english learning platform. *Journal of Grey System, 19(1),* 31–40. DOI 10.30016/JGS.

14. Alladio, E., Giacomelli, L., Biosa, G., Corcia, D. D. Gerace, E. et al. (2018). Development and validation of a partial least squares-discriminant analysis (PLS-DA) model based on the determination of ethyl glucuronide (EtG) and fatty acid ethyl esters (FAEEs) in hair for the diagnosis of chronic alcohol abuse. *Forensic Science International, 282,* 221–234. DOI 10.1016/j.forsciint.2017.11.010.

15. Zare, M., Ghodsbin, F., Jahanbin, I. (2016). The effect of health belief model-based education on knowledge and prostate cancer screening behaviors: A randomized controlled trial. *International Journal of Community Based Nursing & Midwifery, 4(1),* 57–68. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4709816/.

16. Qin, R., Zeng, S., Li, J. J. (2017). Parallel enterprises resource planning based on deep reinforcement learning. *Zidonghua Xuebao/Acta Automatica Sinica, 43(9),* 1588–1596. DOI 10.16383/j.aas.2017.c160664.

17. Liang, M., Wang, B., Yan, T. (2017). Dynamic optimization of robot arm based on flexible multi-body model. *Journal of Mechanical Science and Technology, 31(8),* 3747–3754. DOI 10.1007/s12206-017-0717-9.

18. Li, L., Han, Y., Chen, W., Lv, C., Sun, D. et al. (2016). An improved wavelet packet-chaos model for life prediction of space relays based on Volterra series. *PLoS One, 11(6),* e0158435. DOI 10.1371/journal.pone.0158435.

19. Ivan, C. F., Jones, E. S., Thiago, V. C. (2016). Development of a predictive control based on Takagi-Sugeno model applied in a non-linear system of industrial refrigeration. *Chemical Engineering Communications, 204(1),* 39–54. DOI 10.1080/00986445.2016.1230850.

20. Wei, W. Z. (2018). Research on social responsibility of e-commerce platform. *IOP Conference Series: Materials Science and Engineering, 439(3),* 32063. DOI 10.1088/1757-899X/439/3/032063.

21. Ge, F., Ding, X. (2016). Uncertain type of multiple-attribute electronic commerce investment decision model based on the close degree of the scheme and its applications. *iBusiness, 8(2),* 31–35. DOI 10.4236/ib.2016.82004.

22. Hartadiyati, E., Rizqiyah, K., Wiyanto, Rusilowati, A., Prasetia, A. P. B. (2017). The integrated model of sustainability perspective in spermatophyta learning based on local wisdom. *Journal of Physics Conference, 895(1),* 12051. DOI 10.1088/1742-6596/895/1/012051.

23. Shen, S., Zhu, D. H. (2017). Chinese place name recognition based on deep learning. *Beijing Ligong Daxue Xuebao/Transaction of Beijing Institute of Technology, 37(11),* 1150–1155. DOI 10.15918/j.tbit1001-0645.2017.11.08.

24. Gang, J. T., Huang, L., Zhao, Z. W. (2015). Dynamic simulation of a SEIQR-V epidemic, model based on cellular automata. *Numerical Algebra Control & Optimization, 5(4),* 327–337. DOI 10.3934/naco.2015.5.327.

25. Nikolaos, A., Christodoulou, N. E., Tousert, E. C. (2016). A modular repository-based infrastructure for simulation model storage and execution support in the context of in silico oncology and in silico medicine. *Cancer Informatics, 2016(15),* 219–235. DOI 10.4137/CIN.S40189.

26. Larson, D. B., Chen, M. C., Lungren, M. P., Halabi, S. S. Stence, N. V. et al. (2018). Performance of a deep-learning neural network model in assessing skeletal maturity on pediatric hand radiographs. *Radiology, 287(1),* 313–322. DOI 10.1148/radiol.2017170236.

27. Ding, P., Li, Y. (2016). An electromechanical transient model of VSC and DC grid based on multi-rate simulation method and simplified discrete newton method. *Proceedings of the Chinese Society of Electrical Engineering, 36(24),* 6809–6819. DOI 10.13334/j.0258-8013.pcsee.160398.

28. Qu, S., Xi, Y., Ding, S. (2018). Image caption description of traffic scene based on deep learning. *Journal of Northwestern Polytechnical University, 36(3),* 522–527. DOI 10.1051/jnwpu/20183630522.

29. Luo, N., Wang, X., Van, F. (2015). Integrated simulation platform of chemical processes based on virtual reality and dynamic model. *Computer Aided Chemical Engineering, 37,* 581–586. DOI 10.1016/B978-0-444-63578-5.50092-X.

30. Salah, E. B., Jamila, E. A., Youssef, L. (2015). Learners' attitudes towards extended-blended learning experience based on the S2P learning model. *International Journal of Advanced Computer Science & Applications, 6(70),* 78. DOI 10.14569/IJACSA.2015.061010.

31. Hsu, P. S. (2015). The design concept of self-determination based on-linelearning platform. *Information Japan, 16(9),* 6531–6538. https://www.researchgate.net/publication/289210371_The_design_concept_of_self-determination_based_on-line_learning_platform.