Research on supermarket pricing strategy based on deep learning and optimization model

Junchao Dai, Dinan Liu, Xiangjun Yue, Yin Liu*, Weijie Gao

School of computer and software, Chengdu Jincheng College, Chengdu, China, 611730

*Corresponding author: 13668348161@163.com

Abstract. To carry out reliable market demand analysis and improve the profit of supermarkets, this paper studied the automatic pricing and replenishment decision-making problem of vegetable products in fresh supermarkets. Firstly, the time series-LSTM prediction model is established. Secondly, the parameters of the Adam optimization algorithm and LSTM model are adjusted by grid search, Through the trained model, the daily sales volume, wholesale price, and sales price of each category in the next week are predicted. Finally, the pricing strategy of the total daily replenishment of each vegetable category in the next week is calculated.

Keywords: Time series-LSTM prediction model, Grid search, Adam optimization algorithm.

1. Introduction

In the fresh supermarket, there are many kinds of vegetables sold, and the producing areas are different. In the actual sales process of vegetable products, many factors such as market demand, supply, cost, competitor pricing, and damage need to be considered. At the same time, most vegetable products will have sales problems. At present, the research on the pricing and replenishment strategy of vegetable products has become a research hotspot. Lu Manjiao and others have chosen to use the dynamic matrix model to optimize the replenishment strategy. Such a replenishment strategy can improve the turnover rate of goods, increase the sales volume of goods, and also save the cost of replenishment operations[1]. Fang Jing et al.used the fusion model to optimize the replenishment strategy, which reduced the total inventory management cost of the replenishment plan by about 5 % and increased the inventory turnover rate by about 30 %[2]. Ding Guangwei et al.believed that the effect of combined model prediction is better than that of single model prediction[3]. Zhang Chen et al.selected the prediction method of statistical distribution determination and parameter fitting to achieve inventory optimization. The results showed that it is much higher than the single model[4]. Huang Guoxing et al.used random forests to predict the demand accordingly[5]. Luo Wei et al.established a two-level inventory decision model for spare parts by taking the service response time as the constraint condition and minimizing the inventory cost and out-of-stock cost as the objective[6].

In order to solve many problems in the sales process of vegetable commodities, this paper will establish a set of scientific and reasonable pricing models and replenishment decision models for fresh supermarkets based on historical data, so that fresh supermarkets can face such problems. According to the historical sales data and market demand for commodities, the daily commodities can be replenished to minimize losses and maximize profits.

2. Establish a Time series-LSTM model

2.1. The basic fundamental of LSTM network

LSTM network is an improved model based on recurrent neural networks. It uses the gate mechanism and memory cells to control the transmission of sequence information. When dealing with long sequence information, it can effectively solve the problem of memory loss and gradient disappearance [7]. The basic structure of LSTM and RNN is roughly the same, which is composed of three parts: input layer, hidden layer, and output layer, and both adopt cyclic feedback structure.

The difference is that there are ' gates ' and memory cells in the LSTM hidden layer that can selectively allow information to pass through, and the internal structure diagram is shown in Figure 1.

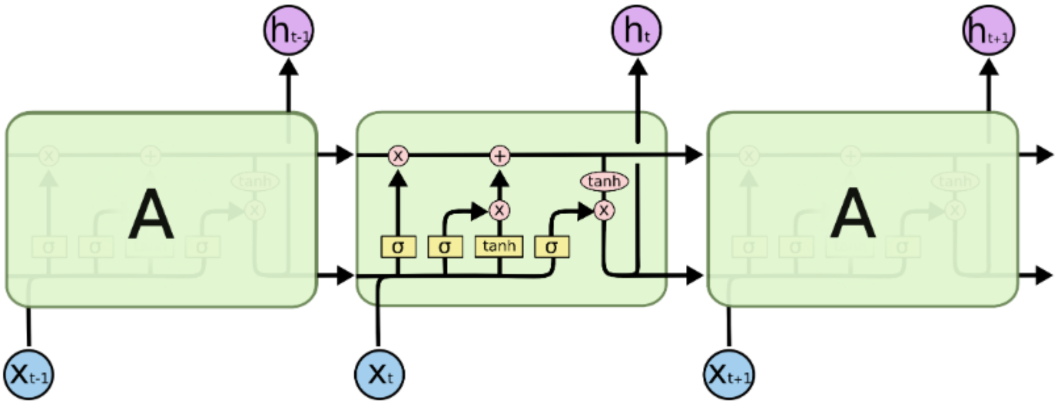

Figure 1. LSTM internal structure diagram

Where: xt: the current time input information. ht-1: The hidden state of the previous moment.ht: The hidden state passes to the next moment. σ: Sigmoid function, through which the data can be changed to a value in the range of 0-1. tanh: Tanh function, by which the data can be transformed into a value in the range [-1,1].

The state of the memory cell at the previous moment is determined by the control of each gate and the linear operation to determine whether the information is retained, and then the state of the memory cell at the current moment is generated, and then output to the next moment. The state of the memory cell is updated through the above process. To achieve long-term memory, in the whole long sequence, the state of memory cells in each moment is selected by three gates to select the information of the previous moment, and then the state of memory cells is updated. The relevant formula is as follows :

$$\begin{cases} i_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + b_i) \\ f_t = \sigma\left(W_{xf}x_t + W_{hf}h_{t-1} + b_f\right) \\ o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + b_o) \\ \tilde{c}_t = \sigma(W_{xc}x_t + W_{hc}h_{t-1} + b_c) \\ c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t \\ h_t = o_t \cdot \tan h(c_t) \end{cases} \tag{1}$$

Where: c represents the memory cell state. i denotes the input gate. o represents the output gate; W_{xi}, W_{xf}, W_{xo} and W_{xc} represent the weight matrix of the input layer to input gate, forgetting gate, output gate and cell state respectively. W_{hi}, W_{hf}, W_{ho} and W_{hc} represent the weight matrix of hidden layer to input gate, forgetting gate, output gate and cell state respectively. b_i, b_f, b_o and b_c represent the bias vectors of input gate, forgetting gate, output gate and cell state, respectively. $\sigma(\cdot)$ denotes the Sigmoid activation function . tanh($\cdot$) denotes the hyperbolic tangent activation function ; denotes the multiplication of the corresponding elements in two matrices of the same order, which is called Hadamard product[8].

By using h_{t-1} and x_t, sigmoid is used as the activation function to output a value between 0-1 for each element of c_{t-1}, which represents the degree of information retention or discarding in the unit state c_{t-1}. Next, information is added to the unitary state. h_{t-1} and x_t determine the value to be updated through a Sigmoid function, and h_{t-1} and x_t create a new value $\tilde{c}_t$ to be added to the unit state through a tanh function. Then, the unit state is updated to obtain a new unit state c_t. After the unit state is updated, the output is determined according to h_{t-1} and x_t[9].

To predict the sales data of each vegetable category in the next week, the training set is a multivariate time series, therefore, in this paper, the multivariate time series-LSTM model is used for prediction. The LSTM model is implemented by using the deep network designer of the Matlab toolbox. The three-layer neural network is used: LSTM layer, Dropout layer, and fully connected layer, and the loss function type and Adam optimizer are defined. The time series-LSTM prediction network is shown in Figure 2.

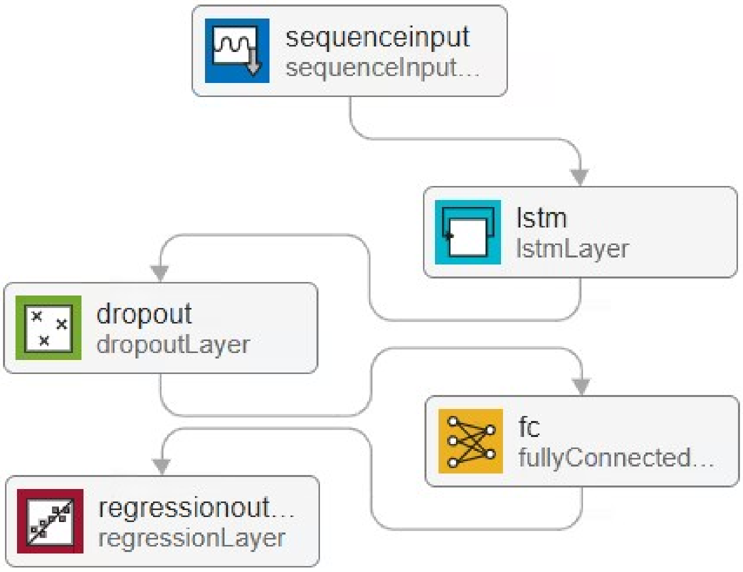

Figure 2. time series-LSTM prediction network

2.2. Algorithm steps

2.2.1 Data preprocessing

In deep learning, it is usually necessary to standardize or normalize the data to remove the dimension get better fitting data, and prevent training divergence. Therefore, this paper performs Z-score Normalization of the data, and the processing method is shown in the following formula.

$$\boldsymbol{X}^{*} = \frac{\boldsymbol{x}-\boldsymbol{u}}{\sigma} \tag{2}$$

Here u and σ are the mean and standard deviation of this set of data, respectively.

2.2.2 Train the network

(1)Important theoretical parameters

In deep learning, the learning rate and the number of hidden units often have a great impact on the accuracy of the model. The learning rate is an important hyperparameter, which is used to control the speed at which the model updates the parameters in each iteration. In general, a larger learning rate can speed up the convergence of the model, but it may cause the model to fail to converge or converge to a poor local optimal solution. A smaller learning rate can improve the stability of the model, but it may lead to slow convergence.

The hidden unit is mainly used to nonlinearly disrupt the input data, so that the categories corresponding to the input data become linearly separable at the last layer, to improve the representation ability of the neural network and the processing ability of complex data. For a certain number of hidden units, the computational cost of the network will increase accordingly, and there may be poor learning. Therefore, in deep learning, it is necessary to determine the number of hidden units and the location of hidden layers according to specific tasks.

(2)Grid search

To improve the accuracy of the model and avoid over-fitting and under-fitting, this paper splits the cleaned data obtained by Step 1 into a training set and a small number of test sets for grid search to adjust the parameters of the Adam algorithm. Mainly by adjusting the learning rate of the model and the number of hidden units, the specific operation is to define the learning rate set { 0.0001,0.0005,0.001,0.002 } and the number of hidden units set { 50,100,200,500,800 }.

In the above two sets, the optimal combination is found by grid search, and the grid search results are shown in Figure 3.

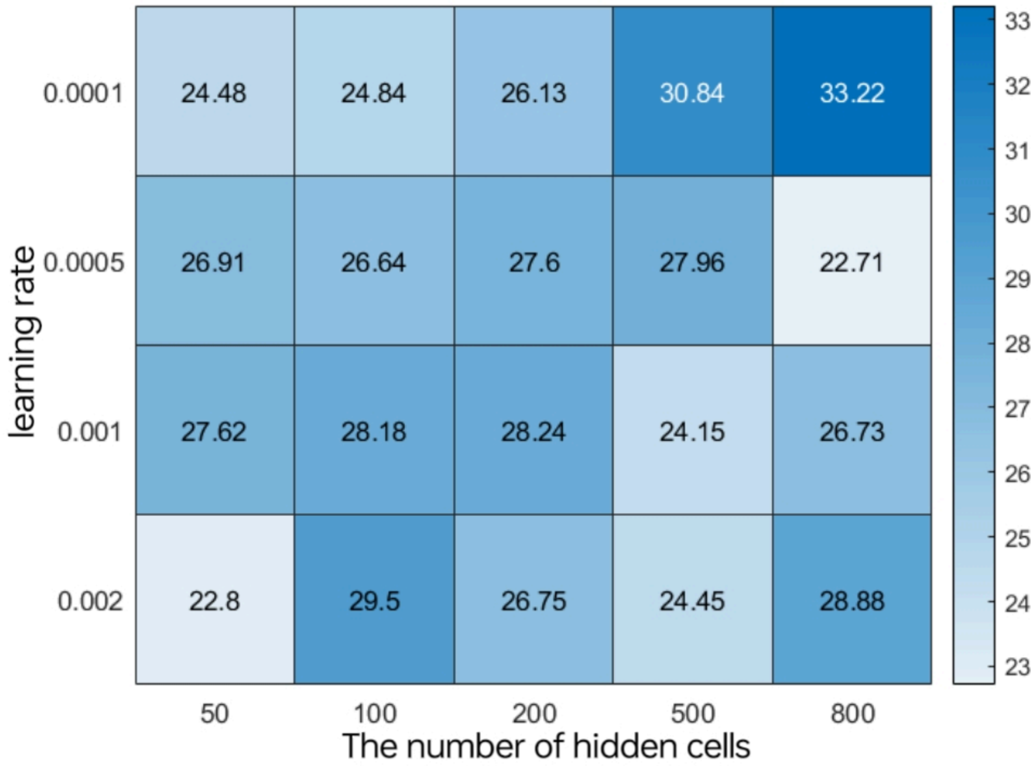

Figure 3. Grid search RMSE thermal map

2.2.3 Adam optimization algorithm

Because there are too many parameters in the deep learning network, ordinary optimization algorithms such as genetic algorithms and simulated annealing algorithms have slow calculation speeds, so the gradient descent algorithm is needed. However, the traditional gradient descent algorithm has two drawbacks.

(1)Learning rates remain unchanged

The excessive learning rate will cause the algorithm to oscillate near the optimal solution and fail to converge. Too small a learning rate will lead to slow convergence speed and fall into a local optimal solution. So it is difficult to find the optimal learning rate.

(2)It oscillates at the saddle point and is easy to fall into the local optimum

To solve the above two problems, adaptive gradient descent algorithm can be used, Adam is one of the algorithms[10]. The Adam algorithm can adaptively adjust the learning rate of each parameter according to the estimation of the first and second moments of the gradient, avoiding the use of fixed or manually adjusted learning rates, and improving optimization efficiency and stability. At the same time, the Adam algorithm can use the first-order moment estimation of the gradient to add an inertia term to the update direction of each parameter, making the update direction smoother and more stable, to avoid the problem of oscillation or deviation from the optimal solution caused by the gradient descent algorithm in the case of noise or inconsistent curvature.

In this paper, the initial learning rate obtained by the grid search algorithm is 0.0005, the number of iterations is 1000, the exponential decay rate of the first-order moment estimation (Beta1), and the exponential decay rate of the second-order moment estimation (Beta2) are the default values 0.9 and 0.999, respectively. In addition, this paper also uses the method of segmented learning rate to update the learning rate. The specific operation is to attenuate the learning rate to 0.2 per 500 rounds.

The segmented learning rate can dynamically adjust the learning rate according to the performance of the model in different training stages, to maintain the convergence speed of the model while ensuring the stability of the model.

3. Results

3.1. Data prediction

The goodness of fit (R^2) and root mean square error (RMSE) in the time series-LSTM prediction model are indicators which used to evaluate the performance of the model. The closer the value of R^2 is to 1, the better the model can explain the change of observation data and the fitting effect. The smaller the RMSE value, the higher the prediction accuracy of the model, and the smaller the difference between the predicted value and the actual observation value. The R2 and RMSE for each category are shown in Table 1.

Table 1. R^2 and RMSE indicators of the model

	Cauliflower	Mosaic class	Capsicum	Eggplant	Edible fungi	Aquatic rhizomes
RMSE	14.02	48.62	26.46	7.539	25.38	15.98
R^2	0.6182	0.6818	0.7547	0.687	0.7259	0.7401

From the above table, the average RMSE of different categories of vegetables on the model is about 23.0, and the average R2 is about 0.701. The model performs well. Through the time series-LSTM prediction network that performs well on the test set, the total daily sales volume from July 1,2023 to July 7,2023 is predicted. The daily sales volume of each category is shown in Table 2.

Table 2. Sales volume forecast of each category (kg)

Date	Cauliflower	Mosaic class	Capsicum	Eggplant	Edible fungi	Aquatic rhizomes
2023-7-1	21.532	152.257	91.171	23.480	56.549	23.559
2023-7-2	21.794	151.030	90.516	23.819	54.711	22.247
2023-7-3	21.411	146.419	88.119	23.681	51.878	20.257
2023-7-4	20.894	141.591	85.547	23.449	49.358	18.385
2023-7-5	20.502	137.962	83.552	23.318	47.596	17.020
2023-7-6	20.329	135.873	82.325	23.355	46.589	16.203
2023-7-7	20.365	135.086	81.743	23.550	46.127	15.801

The daily replenishment quantity can be calculated by the algebraic relationship between the daily sales volume and the average loss rate. The calculation formula is as follows:

$$Q = (1 + n\%) * N \tag{3}$$

Among them, Q is the total daily replenishment quantity, n % is the average loss rate, and N is the daily sales volume. The total daily replenishment for each category is shown in Table 3.

Table 3. Forecast of total replenishment quantity of each category (kg)

Date	Cauliflower	Mosaic class	Capsicum	Eggplant	Edible fungi	Aquatic rhizomes
2023-7-1	24.577	167.909	98.935	25.153	61.147	26.380
2023-7-2	24.876	166.556	98.224	25.515	59.160	24.911
2023-7-3	24.439	161.471	95.623	25.367	56.096	22.683
2023-7-4	23.848	156.147	92.832	25.119	53.371	20.586
2023-7-5	23.401	152.145	90.667	24.979	51.466	19.059
2023-7-6	23.204	149.841	89.335	25.018	50.377	18.143
2023-7-7	23.245	148.973	88.704	25.227	49.878	17.694

The established time series-LSTM prediction model can predict the wholesale prices and pricing of various categories from july 1st to july 7th,2023. The wholesale prices for each category are shown in Table 4.

Table 4. Wholesale price forecast of each category (Yuan / kg)

Date	Cauliflower	Mosaic class	Capsicum	Eggplant	Edible fungi	Aquatic rhizomes
2023-7-1	7.965	3.458	5.881	4.625	4.818	11.968
2023-7-2	7.863	3.380	5.834	4.651	4.797	12.046
2023-7-3	7.754	3.322	5.815	4.687	4.773	12.117
2023-7-4	7.639	3.277	5.814	4.725	4.744	12.188
2023-7-5	7.517	3.244	5.829	4.759	4.711	12.264
2023-7-6	7.387	3.219	5.854	4.786	4.674	12.342
2023-7-7	7.248	3.201	5.887	4.806	4.636	12.425

Pricing for each category is shown in Table 5.

Table 5. Pricing prediction of each category (Yuan / kg)

Date	Cauliflower	Mosaic class	Capsicum	Eggplant	Edible fungi	Aquatic rhizomes
2023-07-01	11.251	4.953	7.566	8.150	11.613	15.352
2023-07-02	10.870	4.940	7.793	8.247	11.719	15.480
2023-07-03	10.481	4.919	7.932	8.289	11.768	15.597
2023-07-04	10.100	4.900	8.077	8.333	11.802	15.724
2023-07-05	9.741	4.881	8.225	8.377	11.817	15.859
2023-07-06	9.412	4.864	8.377	8.420	11.813	15.995
2023-07-07	9.121	4.852	8.533	8.463	11.789	16.122

3.2. Calculate profit

The final profit can be calculated by income minus cost. The specific calculation formula is as follows:

$$W = N * S - Q * P \tag{4}$$

Among them, S is the pricing, and P is the wholesale price. The returns for each category are shown in Table 6.

Table 6. Earnings forecast for each category (Yuan)

Date	Cauliflower	Mosaic class	Capsicum	Eggplant	Edible fungi	Aquatic rhizomes	Total profit
2023-7-1	46.478	173.541	107.930	75.028	362.104	45.963	811.045
2023-7-2	41.304	183.174	132.336	77.743	357.373	44.300	836.230
2023-7-3	34.903	183.897	142.904	77.376	342.751	41.106	822.936
2023-7-4	28.849	182.015	151.206	76.723	329.317	38.167	806.276
2023-7-5	23.793	179.793	158.762	76.467	320.008	36.205	795.028
2023-7-6	19.932	178.531	166.648	76.909	314.871	35.234	792.125
2023-7-7	17.257	178.589	175.344	78.079	312.566	34.918	796.752
Total profit	212.516	1259.540	1035.130	538.325	2338.989	275.892	5660.392

From the prediction data, the time series-LSTM prediction model is of high precision, the error is relatively small, it can fully meet the demand, and the prediction speed is fast and the operation is convenient.

4. Conclusions

With the intensification of competition in the commercial super industry, how to maintain the sustained growth of profits in the fierce market competition has become the focus of the attention of merchants. This paper studies the commercial super pricing strategy based on deep learning and optimization models. Establish a time series-LSTM prediction model, the Adam optimization algorithm and the LSTM model are adjusted by grid search. Find the optimal combination of learning rate and number of hidden units, to improve the accuracy and stability of the model. Forecast the daily sales volume, wholesale price, and sales price of each category of vegetables in the next week. Finally, according to the prediction results, the pricing strategy of daily replenishment of each vegetable category in the next week is calculated, to minimize losses and maximize benefits. The pricing strategy and replenishment decision model studied in this paper can provide a scientific and reasonable reference for supermarkets, It is helpful to solve many problems existing in the sales process of vegetable commodities and improve the profit and competitiveness of supermarkets. At the same time, this study also provides a useful reference for other similar supermarkets, which has certain practical significance and theoretical value.

References

[1] Lu Manjiao,zhang wei,Xu Tao.Optimized replenishment strategy based on dynamic matrix model[J].Computer Engineering and Applications,2021,57(07):263-268.

[2] Fang Jing,Xu Ming,Wan JiaLe,etc.Automatic replenishment strategy of unmanned warehouse based on fusion model [J].Journal of Wuhan University of Technology (Traffic Science and Engineering Edition),2022,46(05):812-815+820.

[3] Ding Guangwei, Liu Xuning, Wang Daozhong, etc.Summary of research status of spare parts demand forecasting methods [J].Value engineering,2020,39(04):296-298.

[4] Zhang Chen, Li Jia, Wang Haining, etc.Application of big data in equipment health prediction and spare parts replenishment [J]. China Mechanical Engineering,2019,30(02):183-187.

[5] Huang Guoxing, Cao Xianhuai, Qian Xiaofei.A spare parts prediction model based on random forest [J].Operation and management,2021,30(10):165-168.

[6] Luo Wei, Fu Zhuo, Dong Wei.Inventory optimization of equipment spare parts with dependent demand [J].Industrial engineering,2019,22(02):57-66.

[7] HOCHREITER S,SCHMIDHUBER J. Long short-term memory [J]. Neural Computation,1997,9(8):1735-1780.

[8] Xuan Yinglv, Wan Yuan, Chen Jiahui.LSTM time series classification based on multi-scale convolution and attention mechanism [J].Computer application,2022,42(08):2343-2352.

[9] Jin Chenyang. Research on time series classification algorithm based on deep learning [D].Yunnan University,2022.DOI:10.27456/d.cnki.gyndu.2022.002215.

[10] Kingma P D ,Ba J . Adam: A Method for Stochastic Optimization.[J]. CoRR,2014,abs/1412.6980.