Sales Forecasting and Consumption Recommendation System of E-commerce Agricultural Products Based on LSTM Model

Jianwei Wu
Zhengzhou College of Finance and Economics
China
wujianwei202407@163.com

Abstract: *With the rapid development of global e-commerce, the sales volume of agricultural products, as an important consumer product category, in e-commerce platforms is increasing. However, affected by seasonal changes and market demand fluctuations, the sales forecast of agricultural products has always been a key challenge faced by e-commerce platforms. With the increase of personalized consumption demand, how to use recommendation system to improve users' shopping satisfaction has become an important issue in the competition of e-commerce platforms. This study proposes a Long Short-Term Memory Network (LSTM) model based on e-commerce agricultural product sales forecasting method, combined with a recommendation system to provide users with personalized consumption suggestions. By analyzing the agricultural product sales data of a large e-commerce platform, the LSTM model can effectively capture complex patterns in time series data and provide high-precision sales forecast. In addition, the recommendation system designed in this paper realizes personalized commodity recommendation by combining users' historical behaviors with sales forecast results, and improves users' click-through rate and purchase rate. The experimental results show that the sales forecasting system based on LSTM model shows high forecasting accuracy in many agricultural products categories, and the user click-through rate after combining with the recommendation system increases by 1.3 percentage points, and the recommendation accuracy rate reaches 68.2%.*

Keywords: *LSTM model, sales forecast of agricultural products, e-commerce, personalized recommendation system.*

Received September 13, 2024; accepted April 26, 2025
https://doi.org/10.34028/iajit/22/4/9

1. Introduction

Since the 21st century, e-commerce in China has flourished due to rapid advancements in network technology and a growing user base. Online agricultural product sales have created an efficient platform connecting producers, suppliers, and consumers, expanding sales channels, accelerating agricultural transformation, promoting information flow, reducing costs, and driving the evolution of the rural economy [1]. Against this background, the support of national policies has made the online sales of agricultural products enter a new stage of diversified development, attracting many entrepreneurs to participate and forming a social atmosphere of "mass entrepreneurship and innovation" [2]. With its efficient information circulation ability, online sales mode reduces circulation links, successfully solves the problem of "difficult to sell and expensive to buy" agricultural products, and significantly improves farmers' income levels. By optimizing the production process and establishing the product quality traceability system, the quality of agricultural products is improved, and consumers' trust in them is enhanced. At the same time, the application of Internet technology has built a global information network for enterprises, expanded the scope of product dissemination, and achieved a fairer competitive environment. For consumers, online sales provide a rich and diverse choice of goods, bringing an unprecedented convenient shopping experience and becoming one of the critical changes in e-commerce. At present, B2C e-commerce platforms such as Tmall Fresh and JD.COM Fresh occupy a dominant position in the online sales market of agricultural products, leading the development trend of the industry. However, fresh agricultural products' perishable characteristics and limited shelf life pose challenges to precision marketing strategies [6]. Faced with this problem, it is necessary to deal with it through technological innovation and strategic optimization to ensure the stability and efficiency of online sales of agricultural products, further enhance consumer satisfaction and loyalty, and promote the healthy development of the whole industrial chain.

This paper optimizes the sales forecast and consumer experience of agricultural products by combining Long Short-Term Memory (LSTM) model and personalized recommendation technology. LSTM model can accurately predict the future sales trend by analyzing the historical sales data, and help e-commerce platforms carry out inventory management and supply chain optimization. At the same time, the personalized recommendation system provides customized recommendation of agricultural products according to consumers' purchase behavior and preferences, which improves the purchase conversion rate and user stickiness. Combined with sales forecast and personalized recommendation, the system effectively improves the accuracy of sales strategy and platform competitiveness.

The core of solving the precise marketing strategy in the online sales of agricultural products lies in the accurate forecast of the sales volume of agricultural commodities. This link needs to mine critical information from massive data to predict market trends and accurately locate target consumer groups and sales areas to help e-commerce platforms adjust strategies and optimize promotion methods promptly [11]. Building a sustainable prediction model to meet the challenge of precision marketing has become the central issue of online sales research of agricultural products. Especially under the background of "Internet Plus," the sales forecast of agricultural products is regarded as one of the most important research directions in "Internet Plus's agriculture."

Developed countries in Europe and America are significantly ahead of China in practice and theoretical research in agricultural e-commerce, and the application of electronics and networking in agriculture has been widely implemented [12]. This research focuses on four core directions:

1. Online sales mode of agricultural products and construction of e-commerce platform: This direction deeply analyzes the factors affecting agricultural products transactions under B2B mode, the performance evaluation system of B2B e-commerce, and the quality evaluation standard of consumer-oriented B2C website.
2. E-commerce adoption behavior of agriculture-related enterprises: Through empirical research, it is found that large agriculture-related enterprises actively adopt e-commerce strategies and regression analysis is used to identify the key factors driving them to adopt e-commerce [15].
3. The interaction between e-commerce and agricultural supply chain: Relevant studies reveal that e-commerce can effectively promote the vertical integration of the supply chain, but to improve the adoption rate, strengthening the trust mechanism has become an important issue.
4. E-commerce application of agricultural means of production: this field explores the economic benefits of local agricultural enterprises through e-commerce, with particular emphasis on improving the quality of express delivery services and enhancing consumer loyalty. In addition, the development of agricultural e-commerce must focus on addressing farmers' usage barriers, including developing functions such as information transmission, data recording, and production model formulation to promote its wider popularization and deepening application.

In e-commerce, unmeasured influencing factors, including buyer evaluation and product sales, can be displayed to consumers through quantitative methods. This operation significantly reshapes consumers' shopping behavior path, thus greatly improving the complexity of sales forecasting. Therefore, using data mining technology to build appropriate models to identify and weigh key influencing variables is vital for merchants to predict the market performance of similar products accurately.

LSTM network and recommendation system have made significant progress in the field of e-commerce. LSTM, as an effective time series forecasting model, it has been widely used in sales forecasting, demand forecasting and other tasks. It can capture the long-term dependence in the data, especially in dealing with the seasonal and trend changes of agricultural product sales data. At the same time, the application of personalized recommendation system in e-commerce platforms has also been deepened. By analyzing users' historical behavior, preferences and consumption patterns, the recommendation system can provide more accurate and customized product recommendations, improve user experience and purchase conversion rate. In recent years, the combination of LSTM and recommendation system has received more and more attention. By combining the sales forecasting ability of LSTM with the personalized recommendation function of recommendation system, it can more accurately predict the demand for agricultural products, and provide customized shopping recommendations for consumers, which further promotes the development of e-commerce for agricultural products.

2. Theoretical Basis of Sales Forecast of Agricultural Products in E-Commerce

2.1. Theories Related to Deep Learning

As a specific application of pattern recognition in information science, online sales forecast of agricultural products aims to quantitatively predict future market trends and demand by analyzing historical sales data. This field has various forecasting methods, including gray theory, artificial neural network, time series analysis, and combined forecasting strategies combining multiple technologies. However, the widely used regression-based forecasting models often need more forecasting accuracy when dealing with large-scale data sets and have limitations in the effectiveness and scalability of long-term forecasting [3]. Although this model may perform well in short-term forecasting, its forecasting ability to accurately predict demand over a more extended period is significantly weakened, which limits its application effectiveness in business decision support. Therefore, exploring and developing forecasting models that can effectively process big data with high precision and long-term stability has significant academic and practical value for improving the accuracy and practicality of online sales forecasting of agricultural products. The forgetting gate formula for the LSTM unit is shown in Equation (1).

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (1)$$

where f_t denotes the forgetting gate output vector, σ denotes the Sigmoid activation function, W_f denotes the weight matrix of the forgetting gate, $[h_{t-1}, x_t]$ denotes the connected vector, and bf denotes the bias term of the forgetting gate. As a strategy to efficiently manage high-dimensional data, data-driven technology was compared and affirmed in a study by Alzghoul *et al.* [3] in 2014. This study emphasizes that compared with traditional knowledge discovery methods, the application of data-driven technology in data analysis shows significant advantages. Meanwhile, Zhang [26] proposed an innovative data-driven method in the same year, focusing on using motion data to drive the face animation generation process of unknown objects. The experimental results show that this method performs excellent quantitative and qualitative evaluation. The theoretical basis of this method lies in constructing a processing model based on large-scale data sets and effectively compressing transform and project the high-dimensional feature space to a low-dimensional space. This strategy aims to improve processing performance by eliminating redundant information and preserving key data characteristics. Specifically, data-driven technology analyzes a large amount of raw data, identifies and extracts representative features, and then uses these features for subsequent data processing and analysis to improve efficiency and accuracy. In addition, quantitative and qualitative evaluation confirm the superiority and practicability of data-driven methods in dealing with complex, high-dimensional data. The input gate formula of the LSTM unit is shown in Equation (2).

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (2)$$

Where i_t represents the input gate output vector, W_i represents the weight matrix of the input gate, and b_i the bias term of the input gate. Data-driven technology combines multiple disciplines, such as signal processing, information fusion, and statistical machine learning. Its development benefits from the theoretical support of pattern recognition, machine learning, and data mining. In-depth exploration reveals the potential of neural networks, but building such models requires a lot of labeled data and parameter optimization. In this context, deep learning has become an innovative unsupervised learning strategy [24]. By mimicking the human brain's multi-level structure, deep learning can adaptively extract features from large unlabeled datasets, showcasing strong classification, generalization, and noise resistance abilities. It excels in fields like image, text, and language processing, becoming a key development in data mining and AI. Deep neural networks enhance data representation quality, particularly in large-scale data processing. This paper uses a data-driven deep learning approach for adaptive feature extraction of agricultural product online transaction data [25]. The candidate memory formulas for LSTM units are shown in Equation (3). Among them, C_t represents the candidate memory vector, *tanh* represents the hyperbolic tangent activation function, W_C represents the weight matrix of the candidate memory, and b_C represents the bias term of the candidate memory.

$$C_t = tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \quad (3)$$

2.2. Consumption Recommendation System

Data preprocessing, as an indispensable prerequisite step in the data mining process, covers a series of operations on the initial data, mainly including cleaning, integration, transformation, and simplification. The core purpose of this stage is to ensure that the data meets the technical requirements of mining activities and that its quality reaches the predetermined standard. Faced with the everyday challenges of noise, missing values, and inconsistency in raw data, the data preprocessing process can effectively improve data quality and ensure that the data is compatible with the application requirements of various mining algorithms. The memory update formula of the LSTM unit is shown in Equation (4).

$$C_t = f_t \cdot C_{t-1} + i_t \quad (4)$$

where C_t represents the memory cell state at the current time, C_{t-1} represents the memory cell state at the previous time, f_t represents the forgetting gate output vector, and i_t represents the input gate output vector. Whitening processing is central to natural image processing and unsupervised feature learning. Its main goal is to reduce the redundancy of input data, ensure that the correlation between features is minimized, and maintain the consistency of variance. This process involves multiple whitening techniques, including independent component analysis, principal component analysis, and zero-phase component analysis [20]. Principal Component Analysis (PCA) techniques focus on performing dimensionality reduction operations while preserving maximum variance information. In contrast, Zero-phase Component Analysis (ZCA) whitening to approximate the characteristics of the original data while retaining the original data dimensions through matrix inverse operation. The existing research shows that the whitening data can significantly enhance the convergence performance of the model. Based on this, the ZCA whitening strategy is chosen in this study, which aims to eliminate redundant data efficiently and thus accelerate the convergence process of the model. The output gate formula of the

LSTM unit is shown in Equation (5).

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \tag{5}$$

Where *ot* denotes the output gate output vector, Wo denotes the weight matrix of the output gate, bo denotes the bias term of the output gate, and $[h_{t-1}, x_t]$ denotes the connected vector. The progress of fuzzy theory benefits from the management and analysis of fuzzy associations in the real world [14]. The pioneering work in this research direction can be traced back to the first concept of Fuzzy Sets proposed by Feng [10] in the article "Fuzzy Sets" in 1965. Subsequently, Mamdani's contribution in 1974 was to apply fuzzy reasoning theory to the pressure and speed regulation of steam engines, which not only verified the practicality of this theory in industrial control, but also marked a key turning point for fuzzy theory to move towards practical application. Since then, fuzzy system has shown a rapid development and expansion trend in theoretical exploration and application practice, and its application scope covers automatic control, pattern recognition, cluster analysis and fraud detection. The hidden state update formula of the LSTM unit is shown in Equation (6). Where h_t represents the hidden state vector at the current time, o_t represents the output gate output vector, C_t represents the memory cell state at the current time, *tanh* hyperbolic tangent activation function.

$$h_t = o_t \cdot tanh(C_t) \tag{6}$$

3. LSTM Model Design

3.1. Sample Selection and Establishment of Evaluation Indicators

Deep learning, inspired by multi-level neural networks, excels in data pattern recognition and feature learning, especially in unstructured data like images, text, and language [27]. Despite its success, research on its application in agriculture-related e-commerce is still emerging. This study aims to address information asymmetry in online agricultural transactions by developing a model that predicts sales volume accurately, using deep learning algorithms. The "Crown Model" sales forecasting system was designed to leverage deep learning's self-learning and updating abilities, achieving precise classification and prediction of sales, thus supporting innovation and practical application in the field [28].

Figure 1 shows the flowchart for preprocessing and feature extraction of agricultural product sales data. The factors influencing agriculture-related e-commerce are complex and diverse, which should be highly valued. As a critical structural variable, the difference in logistics infrastructure among production regions significantly restricts the popularization and application efficiency of e-commerce [29]. On the other hand, the differentiation of the needs of different consumer groups should be considered. It directly determines the optional range and consumption patterns of commodities. In addition, the significant impact of holidays on sales volume reveals the seasonal law of consumer behavior, which is of great significance for forecasting market trends and formulating marketing strategies. Finally, the importance of promotion channels has become increasingly prominent in the information age. Effective information dissemination can enhance brand awareness and promote product market penetration. To sum up, these factors work together on the development process of agriculture-related e-commerce, forming an intertwined and dynamically changing influence network. The Embedding representation formula for agricultural product categories is shown in Equation (7). Where e_i represents the embedding vector and *ci* represents the predicted sales volume.

$$e_i = Embedding(c_i) \tag{7}$$

$$r_{ui} = \mu + b_u + b_i + q_i^T \cdot p_u \tag{8}$$

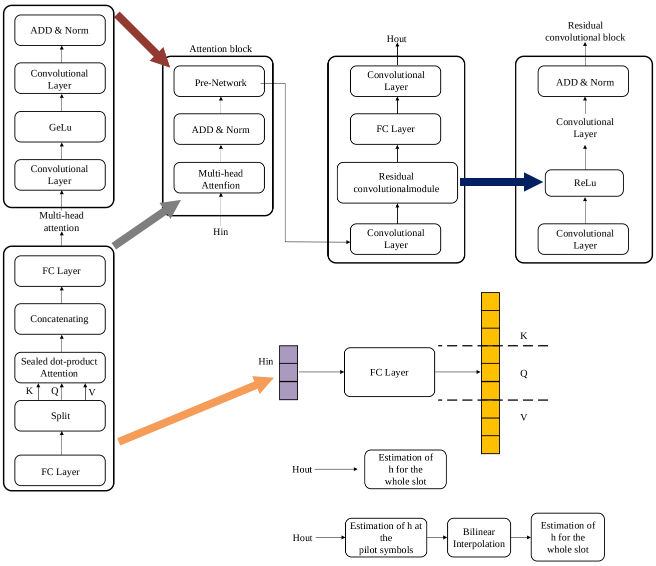


Figure 1. Flow chart of agricultural product sales data preprocessing and feature extraction.

The score prediction formula of the consumption recommendation system is shown in Equation (8). r_{ui} represents the predicted score, μ represents the global average score, b_u represents the preference bias, b_i represents the specific bias, and q_i, p_u represent the implicit vectors of goods and users. This paper focuses on the analysis of logistics strategies for eight selected destinations, which cover e-commerce active areas, including Beijing and Jiangsu Province, and then extend to the Beijing-Tianjin-Hebei and Jiangsu-Zhejiang-Shanghai Economic Zones, in addition to Liaoning Province and Guangdong Province, which are far away from their birthplaces [4]. Given the heterogeneity in type and dimension of the acquired data, to ensure the validity and precision of subsequent analysis, the initial data needs to be standardized to eliminate noisy elements and clearly define the categories of the input data. In order to avoid the bias caused by abnormal data distribution, the set value range is located in the [0, 1] interval, which helps to enhance the accuracy of the sales forecast and finally forms the representative numerical description of each attribute. The safety stock level formula for inventory management is shown in Equation (9). Among them, SS represents the safety inventory, Z represents the service level, LT represents the order lead time, and σ_d represents the standard deviation of demand.

$$SS = Z \times \sqrt{LT \times \sigma_d^2} \tag{9}$$

3.2. LSTM Classifier Selection

LSTM, a recurrent neural network designed to process time series data and capture long-term dependencies, is known for its superior performance. LSTM innovatively incorporates the gating mechanism compared to the traditional model. This design effectively solves gradient vanishing and explosion issues, making LSTM show significant advantages in sequence data classification tasks [5]. By learning the historical information of the input sequence, LSTM can accurately classify, thus addressing complex time series data challenges. It is widely applicable in speech recognition, natural language processing, and financial forecasting.

Figure 2 shows the flowchart for agricultural product sales forecasting. The advantages of the LSTM classifier are mainly reflected in its excellent memory ability and flexible adaptability to time series data [7]. With the help of an embedded gating mechanism, the model can intelligently and selectively retain or eliminate information, thus effectively capturing long-term dependencies. In financial market forecasting, LSTM can identify trends and patterns in time series and provide accurate forecast results. These application examples reveal the great potential of LSTM classifiers in dealing with complex and dynamic datasets.

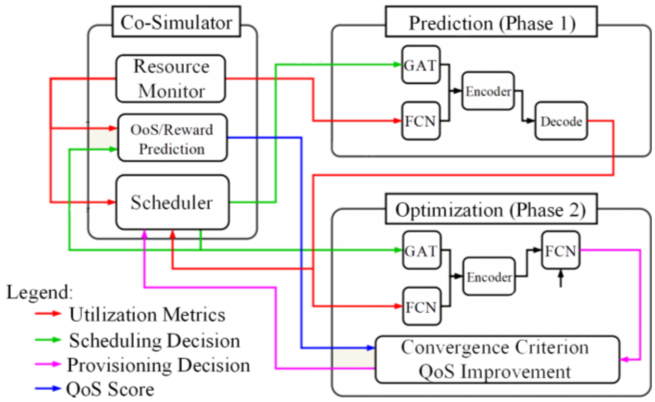
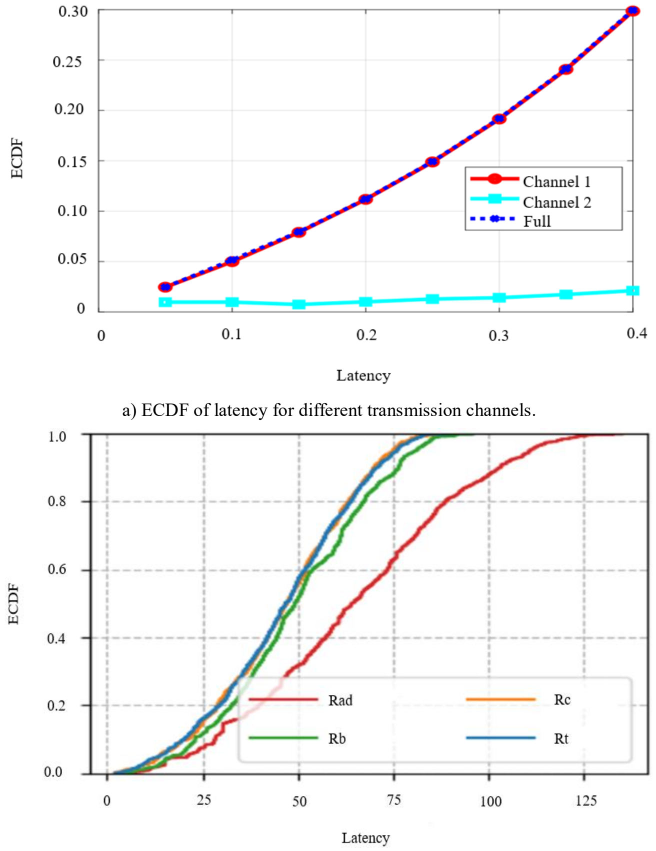


Figure 2. Agricultural product sales forecast flow chart.

In this paper, it is very reasonable to use LSTM model for time series prediction, because LSTM is good at capturing seasonal fluctuations and trend changes in sales data, and can accurately predict future sales, providing strong support for inventory management and supply chain optimization. The combination of LSTM and recommendation system will further enhance the value of the system. Through personalized recommendation, the system can provide customized agricultural products recommendation according to users' preferences, so as to improve the purchase conversion rate and user experience, ultimately optimize the sales strategy and enhance the competitiveness of the platform.



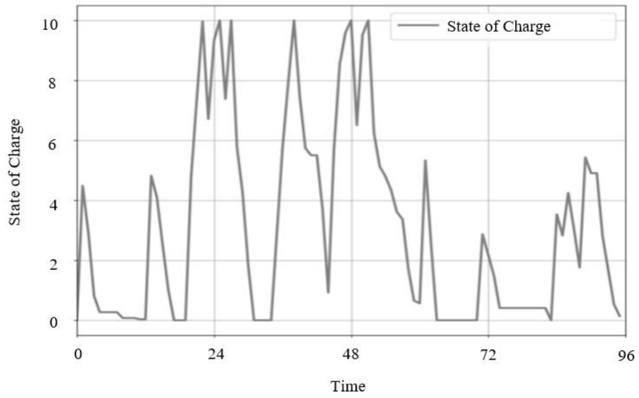
a) ECDF of latency for different transmission channels.

b) ECDF of latency across different regional distributions.

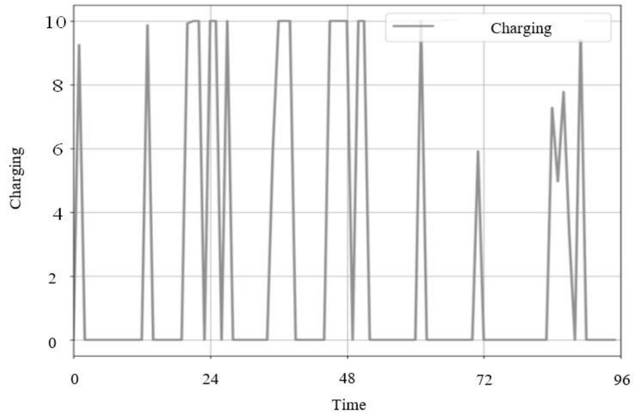
Figure 3. Time series chart of monthly sales of agricultural products.

The time series chart of monthly sales volume of agricultural products is shown in Figure 3. Selecting and optimizing LSTM classifiers involves many key dimensions: network architecture design, hyperparameter configuration, and refinement of training strategy. First, the choice of network architecture significantly impacts model performance [8]. In order to find the best hyperparameters, such as learning rate and batch size, we can use grid search and cross-validation methods. In addition, the proper application of regularization technology is an effective way to prevent overfitting and strengthen the model's generalization ability. In the training stage, adjusting the learning rate, selecting the appropriate optimization algorithm, and implementing strategies such as the early stop mechanism can accelerate the training process and improve the classification effect [9]. To sum up, the LSTM classifier can achieve higher classification accuracy and stability in specific application scenarios by comprehensively using the above optimization methods.

The comparison diagram of LSTM model training and validation losses is shown in Figure 4. As an extended application of logistic regression under multivariate classification tasks, the essence of a soft maximum classifier lies in normalizing the prediction output to ensure that the sum of all prediction probabilities reflects the probability of each classification being correctly identified intuitively [13]. Compared with other classification strategies, the superiority of the soft maximum classifier is not only limited to its excellent classification performance but also lies in its ability to provide specific probability values for each category, which is very important for subsequent analysis and decision-making.



a) Time series of battery state of charge.



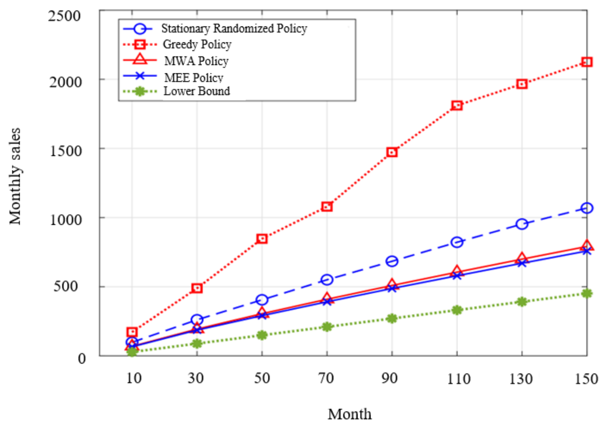
b) Time series of battery charging behavior.

Figure 4. Comparison of LSTM model training and verification losses.

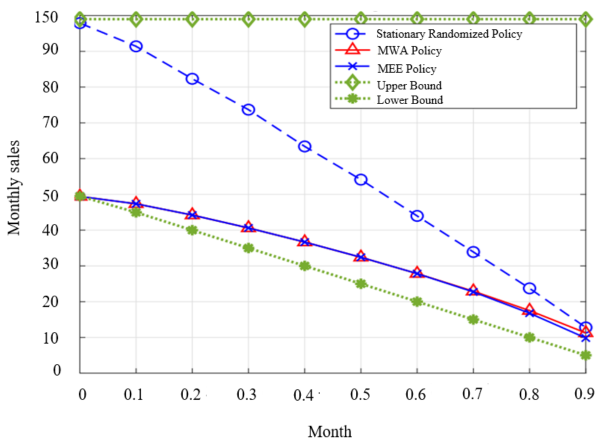
3.3. LSTM Model Structure

LSTM model is designed to solve long-term dependency challenge in time series data prediction. Its core structure includes the input, forget, and output gates [16]. The input gate is responsible for evaluating the input data under the current timestamp and deciding the details that should be injected into the memory unit. The forgetting gate performs the critical task of identifying and culling historical information in the memory unit irrelevant to the current prediction. The output gate determines the specific content of the output information to be generated according to the integration state of the memory unit. By precisely controlling the information flow path, LSTM effectively overcomes the gradient disappearance problem encountered by traditional recurrent neural networks when processing long-sequence data, thus demonstrating its powerful ability to capture long-term dependencies.

The sales forecast accuracy of different forecast time steps is shown in Figure 5. Its core function is to save long-term data, and the key to achieving this goal lies in the dynamic migration and management of data in the memory unit [18]. This process is driven by finely designed control mechanisms, ensuring adequate storage and timely information updates. Specifically, the input gate uses the Sigmoid activation function to accurately evaluate and decide which new information should be included in the memory unit to improve the model's ability to capture long-term dependencies. At the same time, the forgetting gate also uses the Sigmoid activation function to intelligently identify and eliminate information that is no longer valuable or outdated, thus avoiding unnecessary memory load. Finally, the output gate uses the tanh activation function, which generates output values according to the current state. This operation limits the range of output and promotes the effective transmission of information [19]. This ingenious design enables long-term and short-term memory networks to show significant advantages when dealing with complex time series data. It can not only efficiently retain critical information but also effectively filter out redundant or insignificant data, thereby greatly Improving the performance and efficiency of the model on long-term dependent tasks.



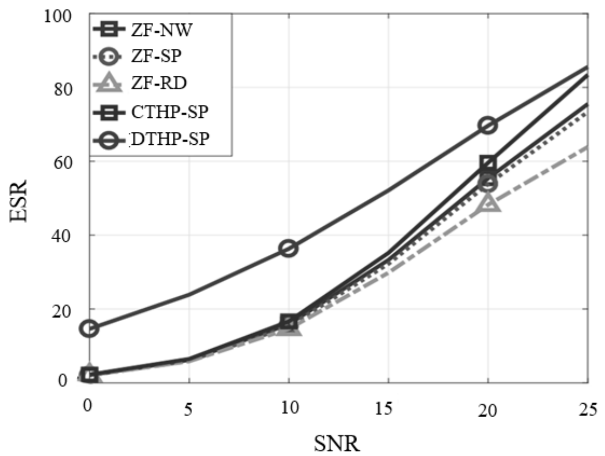
a) Monthly sales trends under different policy scenarios over time.



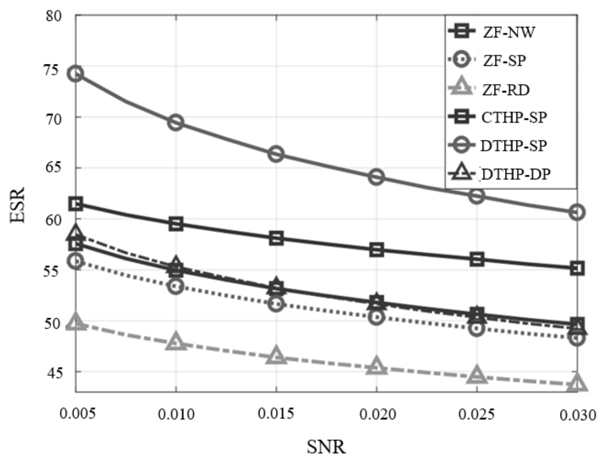
b) Forecast accuracy comparison at varying forecast time steps.

Figure 5. Sales forecast accuracy at different forecast time steps.

LSTM model shows a significant effect in the prediction of agricultural product sales, especially in the prediction of seasonal products. LSTM can effectively capture the long-term dependence and nonlinear characteristics in the data, which makes it particularly advantageous in dealing with seasonal fluctuations in agricultural product sales data. For example, the sales of some agricultural products such as fruits and vegetables will be strongly affected by seasonal changes. LSTM can accurately predict the sales volume in different seasons through the patterns and trends in historical data, so as to improve the prediction accuracy. However, the limitation of LSTM lies in its sensitivity to extreme outliers and emergencies. Although LSTM performs well in historical data, the performance of the model may be affected by sudden factors such as sudden weather changes and market fluctuations. In addition, the LSTM model has high requirements for the amount of data, and the prediction effect may be reduced in the case of incomplete or poor-quality data. Therefore, in practical application, it is necessary to combine other models or technologies to make up for these limitations, so as to further improve the accuracy and reliability of prediction.



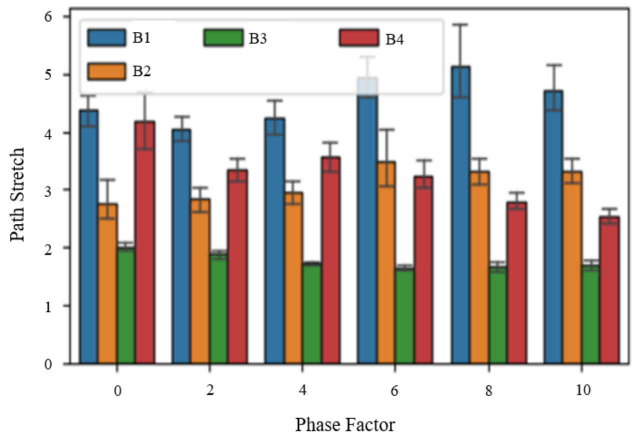
a) ESR performance comparison under varying SNR levels for different methods.



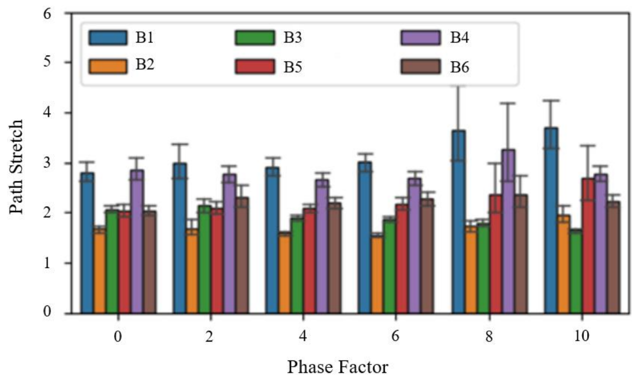
b) ESR sensitivity analysis across small-scale SNR values for different methods.

Figure 6. Chart of feature importance analysis.

The feature importance analysis diagram is shown in Figure 6. Based on the standard LSTM model, the researchers deeply explored diversified extension and variant strategies, aiming to improve the performance and applicability of the model significantly [21]. The bidirectional LSTM architecture effectively captures the bidirectional dependencies of sequences by integrating forward and backward LSTM networks, thus deepening the understanding of sequence context. Furthermore, by building a multi-level stacking structure, namely stacking LSTM, the model can learn more complex features and higher-level abstract representations, significantly enhancing its processing capabilities. At the same time, variant designs such as gated loop units simplify the structure of LSTM while maintaining its superiority in processing long sequence data. To sum up, these innovative extensions and variants jointly improve the flexibility and adaptability of the LSTM model in various practical application scenarios. The seasonal fluctuation chart of agricultural product sales is shown in Figure 7.



a) Path stretch variation across different phase factors for categories B1-B4.



b) Path stretch comparison with extended categories B1-B6 over phase factors.

Figure 7. Seasonal fluctuation chart of agricultural products sales.

4. Experimental Analysis of Consumption Recommendation System

4.1. LSTM Online Learning

Online learning extracts latent patterns from time series to enable real-time learning. Implementation methods include ensemble methods, kernel-based methods, Gaussian process methods, and gradient learning methods. This chapter focuses on gradient learning strategies suitable for LSTM cyclic networks. The loss function formula for time series prediction is shown in Equation (10). Among them, L represents the mean square error, N represents the number of samples, and y_t represents the real sales data.

$$L = \frac{1}{N}\sum_{t=1}^{N}(y_t - \hat{y}_t)^2 \tag{10}$$

$$X_t = \{x_{t-n}, x_{t-n+1}, \ldots, x_t\} \tag{11}$$

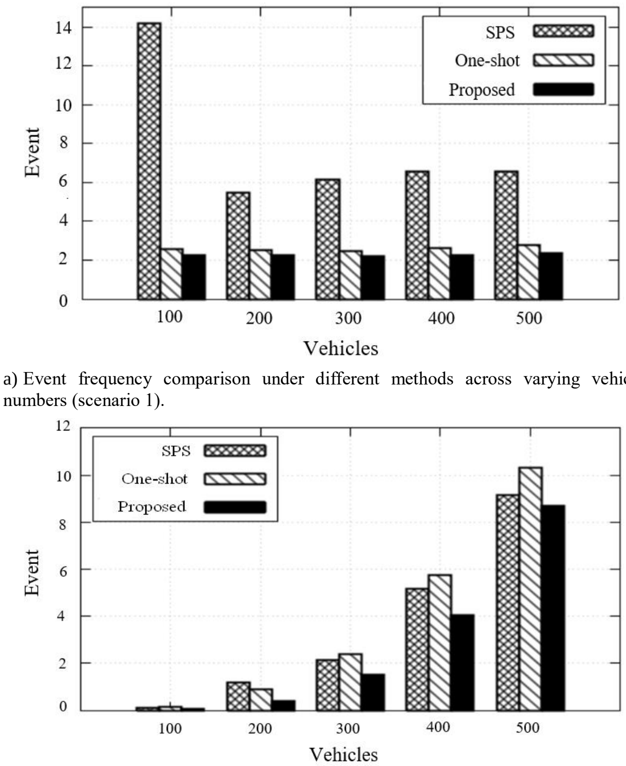
The input feature sequence formula of LSTM is shown in Equation (11). Where X_t denotes the input feature sequence at the current time and n denotes the sliding window size. In online learning, the ideal model needs to be able to automatically save valuable past information, identify predicted correlations, and use the latest data to update itself [22]. In this context, among all kinds of recurrent neural networks, long-term and short-term memory networks have shown significant advantages with their unique gating mechanism and memory unit design, making them an effective solution for tasks involving long-term dependencies in sequence data. Based on this characteristic, the LSTM network is recognized as a powerful tool to solve the streaming data problem. The data analysis table of sales forecast and consumption recommendation of e-commerce agricultural products based on LSTM model is shown in Table 1.

Table 1. Sales forecast and consumption recommendation data analysis table of e-commerce agricultural products based on LSTM model.

Category of agricultural products	Actual sales volume (unit: tons)	Forecast sales volume (unit: tons)	Prediction error (%)	Recommended score
Apple	120.5	118.7	1.49%	4.7
Orange	85.3	87.6	2.69%	4.3
Banana	100.2	98.9	1.30%	4.6
Strawberry	75.0	73.5	2.00%	4.5

Compared with the traditional neural network structure, the core characteristic of LSTM neurons is their ability to maintain and update the information of the previous time step at any point in time. This feature provides the LSTM neural network with a powerful memory function, effectively storing key features and capturing complex dependencies in long-term sequences. The fully connected layer plays a vital role when building an online learning system based on LSTM. Its main task is to transform the information extracted by LSTM units into numerical representations suitable for specific prediction tasks.

The distribution diagram of user purchase frequency is shown in Figure 8. Based on the training data, the LSTM neural network calculates the loss function value and gradient and optimizes the network weight accordingly to realize the iteration of the learning process. Different from the traditional LSTM model, the latter relies on offline learning and gradient descent algorithms to solidify model parameters, showing low adaptability and flexibility, especially when facing changes in water quality data [23]. On the contrary, the online learning feature of LSTM aims to capture and adapt to dynamic changes in the data stream through continuous training and weight updates, thereby significantly improving the efficiency and responsiveness of processing water quality data.



a) Event frequency comparison under different methods across varying vehicle numbers (scenario 1).

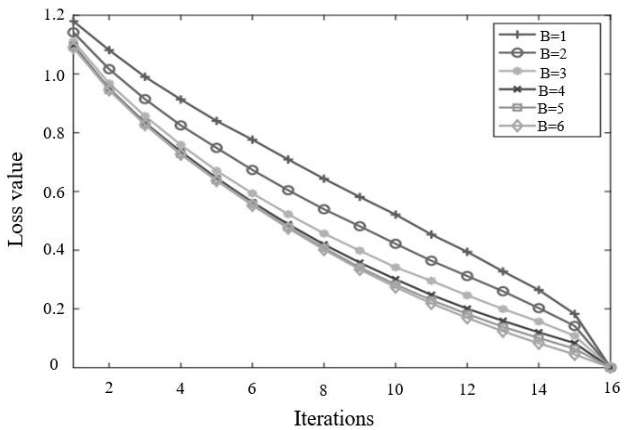
b) Event frequency comparison under different methods across varying vehicle numbers (scenario 2).

Figure 8. User purchase frequency distribution chart.

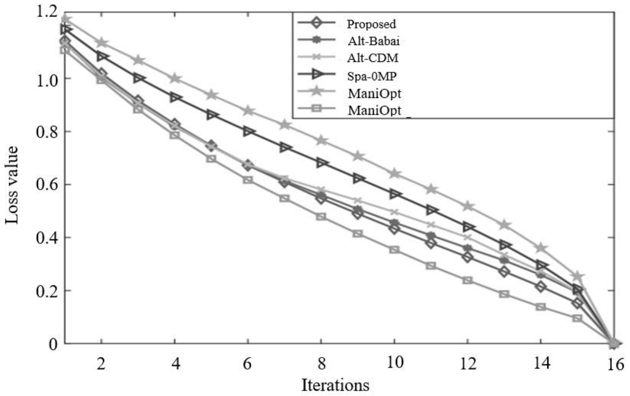
4.2. Experimental Analysis

The experimental data includes five groups, each containing six water quality indicators. This model adopts an end-to-end strategy and is designed specifically for predicting six indicators at a single sampling point, aiming to accurately adapt to the specific needs of water treatment facilities. This study discusses in detail the offline learning process, which involves dividing the dataset into training and testing parts to achieve model training. However, in practical applications, water quality data is generated in a continuous flow manner. Therefore, this study simulates this dynamic characteristic and uses online learning methods for prediction based on the constructed dataset.

In this paper, the prediction effect of LSTM model is effectively proved by using clear evaluation indexes, such as Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE) and R^2 (coefficient of determination). RMSE is used to measure the difference between the predicted value and the actual value. The smaller the value, the higher the prediction accuracy of the model. MAPE shows the relative size of the prediction error, and the lower MAPE value means that the model has better prediction accuracy. R^2 reflects the ability of the model to explain data variation. The closer the R^2 value is to 1, the better the model can fit the actual data. Through the verification of these indicators, the research shows that LSTM model shows good accuracy and stability in the prediction of agricultural product sales, which further proves the effectiveness of the model in the e-commerce environment.



a) Loss value variation over iterations for different parameter settings (B=1 to B=6).



b) Loss value comparison over iterations among different recommendation algorithms.

Figure 9. Relationship between click-through rate of recommendation system and user purchase conversion rate.

The relationship between the click through rate of the recommendation system and the conversion rate of user purchases is shown in Figure 9. During the experiment, we constructed LSTM offline and online learning architectures based on the preprocessed water quality data of each sampling point. Set the window size of the offline learning module to 1 for the LSTM model and its related experimental results as a comparison standard. Taking point 1 as an example, we design a sequence model containing only one LSTM unit and use the grid search method to determine the number of hidden layer nodes to be 18. The purpose is to optimize the prediction efficiency of water quality data through this online learning mechanism. By comparing the indexes of these two models in the experimental stage, we can make an all-round evaluation to select the best learning strategy and improve the accuracy and efficiency of water quality data processing. The implicit state transfer formula of the multi-layer LSTM model is shown in Equation (12). Where h_t represents the hidden state and X_t represents the input feature sequence at the current time.

The relationship between the click through rate of the recommendation system and the conversion rate of user purchases is shown in Figure 9. During the experiment, we constructed LSTM offline and online learning architectures based on the preprocessed water quality data of each sampling point. Set the window size of the offline learning module to 1 for the LSTM model and its related experimental results as a comparison standard. Taking point 1 as an example, we design a sequence model containing only one LSTM unit and use the grid search method to determine the number of hidden layer nodes to be 18. The purpose is to optimize the prediction efficiency of water quality data through this online learning mechanism. By comparing the indexes of these two models in the experimental stage, we can make an all-round evaluation to select the best learning strategy and improve the accuracy and efficiency of water quality data processing. The implicit state transfer formula of the multi-layer LSTM model is shown in (12). Where h_t represents the hidden state and X_t represents the input feature sequence at the current time.

$$h_t = LSTM\,(h_t, X_t) \tag{12}$$

The final sales volume forecast formula is shown in Equation (13), where y_t represents the predicted sales volume at the current time, W_y represents the weight matrix of the output layer, b_y represents the bias term of the output layer, and h_t represents the hidden state. This chapter deeply analyzes the online learning mechanism of long and short-term memory networks and constructs the corresponding theoretical framework and practical process. In order to verify the effectiveness of online learning, an experiment is designed to compare and analyze the time series prediction effects of LSTM online learning strategy and offline learning strategy in a single water quality monitoring point. The results show that the model based on online learning shows more significant superiority in prediction accuracy. In particular, given the possible missing initial data in the original water quality time series data, the challenges encountered when using LSTM for online learning modeling suggest that we must pay attention to the time series dependence in the time series prediction process and adopt data filling strategies promptly to ensure that model performance is not affected.

$$y_t = W_y \cdot h_t + b_y \tag{13}$$

4.3. Application of LSTM Online Learning in E-Commerce Agricultural Product Demand Forecasting

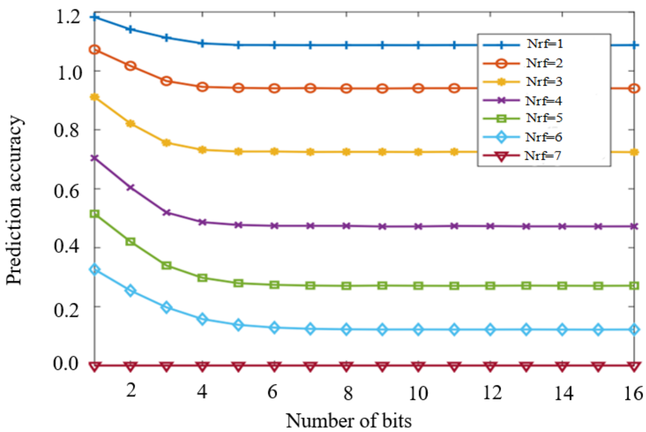
LSTM, as an outstanding time series prediction tool, shows significant advantages in a wide range of application fields [17]. Especially in the field of e-commerce, especially in the agricultural products market, accurate demand forecasting is imperative for merchants to achieve efficient inventory management, flexible price strategy formulation, and improved customer satisfaction. Given its unique ability to deal with long-range dependencies in time series data, the LSTM online learning model becomes an ideal choice for agricultural product demand forecasting in e-commerce. With the ability to update model parameters in real-time, the LSTM online learning mechanism can adjust the forecasting model promptly according to the dynamic changes of the latest sales data, thereby significantly improving the forecasting accuracy and response speed. The evaluation index formula is shown in Equation (14). Among them, RMSE represents the root mean square error, N represents the number of samples, and y_t represents the real sales volume data.

$$RMSE = \sqrt{\frac{1}{N}\sum_{t=1}^{N}(y_t - \hat{y}_t)^2} \tag{14}$$

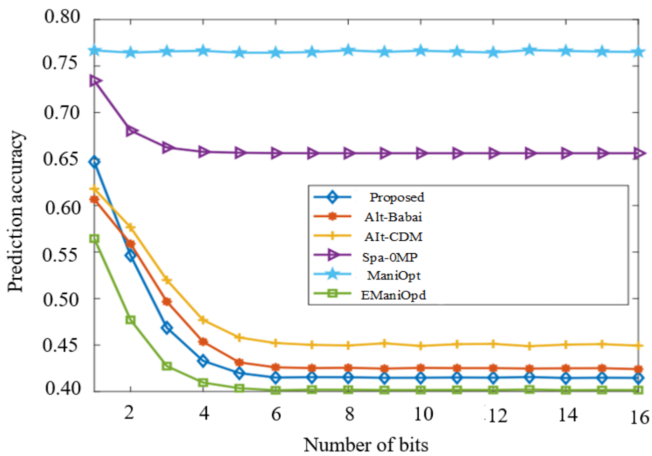
The average absolute percentage error formula is shown in Equation (15). MAPE represents the average absolute percentage error, N represents the number of samples, and y_t represents the real sales volume data. In building the LSTM online learning model, the key in the initial stage is to focus on collecting and preprocessing historical sales data. This step includes but is not limited to excluding outliers, implementing standardized operations, and handling missing values. Then, the construction of the model and the reasonable setting of hyperparameters become the core contents, explicitly involving the determination of the number of hidden layer nodes, the selection of the learning rate, and the setting of the length of the input sequence. Compared with traditional offline learning methods, the significant advantage of LSTM online learning is its ability to adapt to new data by gradually updating network weights, thus avoiding redundant training of the entire data set and effectively improving computational efficiency. In addition, optimizing the model structure and adjusting the training strategy aims to make it more closely adapt to the demand fluctuation characteristics in the e-commerce environment.

$$MAPE = \frac{1}{N}\sum_{t=1}^{N}\left|\frac{y_t - \hat{y}_t}{y_t}\right| \times 100\% \tag{15}$$

The sales changes of agricultural products under different promotion strategies are shown in Figure 10. In e-commerce, the LSTM online learning model is applied to the agricultural product demand forecasting experiment, which reveals that compared with the traditional offline learning method, the LSTM online learning model has obvious superiority in forecasting accuracy and response timeliness. The research indicates that the LSTM online learning mechanism can sense and respond to the changing trend of sales data in real-time, thus allowing for highly accurate demand forecasting. The core advantage of this method lies in its real-time update feature, which enables enterprises to quickly adjust inventory management and supply chain strategies, effectively avoiding the risk of product surplus or shortage. To sum up, LSTM online learning provides an efficient and dynamic solution for agricultural product demand forecasting in an e-commerce environment, highlighting its broad application potential and value prospects.



a) Prediction accuracy under varying bit numbers across different feature sets.



b) Comparison of prediction accuracy among different models under varying bit numbers.

Figure 10. Changes in sales of agricultural products under different promotion strategies.

5. Conclusions

We discuss the research of using long-term and short-term memory network model to predict agricultural product sales on e-commerce platforms and optimize the user's consumption experience with recommendation system. The LSTM model shows significant advantages in processing time series data. When predicting the sales data of agricultural products of an e-commerce platform, LSTM model can effectively capture the time-dependence and seasonal characteristics of the sales data. We used the platform's sales data from 2019 to 2022 for model training and validation, and the data sample size reached 12,000. After several rounds of experimental parameter adjustment, the final selected model parameters are 50 hidden layer nodes, 0.001 learning rate and 100 training rounds. In terms of prediction accuracy, the LSTM model shows high accuracy, with a root mean square error of 38.4, an average absolute error of 29.6, and an R^2 value of 0.87. This shows that the LSTM model can better fit the actual sales data and has strong forecasting ability.

In the construction of recommendation system, this paper combines users' historical purchase behavior and the forecast results of agricultural product sales, and provides users with personalized consumption recommendation through collaborative filtering algorithm. To test the effectiveness of the recommendation system, we conducted recommendation experiments with 2000 active users on the platform. The experimental results show that the click-through rate of the recommendation system is 4.5%, which is 1.3 percentage points higher than that of the traditional rule-based recommendation system. The recommendation accuracy rate reaches 68.2%, while the F1 value is 0.64, which shows that the system can effectively capture users' interests and provide recommendations that meet users' needs.

This paper also analyzes the sales forecast of different categories of agricultural products, and finds that LSTM model has better forecasting effect on seasonal agricultural products than non-seasonal agricultural products. In the sales volume forecast of fruit products, the R^2 value of LSTM model reached 0.92, while in the forecast of rice products, the R^2 value was 0.79. This shows that the LSTM model can flexibly adjust and optimize the prediction effect when dealing with agricultural products with different characteristics.

The agricultural product sales forecasting system based on LSTM model shows high application value in e-commerce environment, and can provide scientific decision support for platform managers. At the same time, combined with the consumption optimization function of the recommendation system, it can significantly improve users' shopping experience and increase the user stickiness and sales conversion rate of the platform. Future research can further explore how to combine other deep learning models, such as convolutional neural networks or variational autoencoders, to improve prediction accuracy and optimize the algorithm of the recommendation system to provide more accurate and intelligent consumption recommendations.

References

[1] Ai Q., Tian H., Wang H., Lang Q., Huang X., Jiang X., and Jing Q., "Comparative Analysis of ARIMA and LSTM Model-Based Anomaly Detection for Unannotated Structural Health Monitoring Data in an Immersed Tunnel," *CMES-Computer Modeling in Engineering and Sciences*, vol. 139, no. 2, pp. 1797-1827, 2024. DOI:10.32604/cmes.2023.045251.

[2] Al-Selwi S., Hassan M., Abdulkadir S., Muneer, A., Sumiea E., Alqushaibi A., and Ragab M., "RNN-LSTM: From Applications to Modeling Techniques and Beyond-Systematic Review," *Journal of King Saud University-Computer and Information Sciences*, vol. 36, no. 5, pp. 102068, 2024. DOI:10.1016/j.jksuci.2024.102068

[3] Alzghoul A., Backe B., Löfstrand M., Byström A., and Liljedahl B., "Comparing a Knowledge-based and a Data-Driven Method in Querying Data Streams for System Fault Detection: A Hydraulic Drive System Application," *Computers in Industry*, vol. 65, no. 8, pp. 1126-1135, 2014. DOI:org/10.1016/j.compind.2014.06.003

[4] Asaithambi S., Ravi L., Devarajan M., Almazyad A., Xiong G., and Mohamed A., "Enhancing Enterprises Trust Mechanism through Integrating Blockchain Technology into E-Commerce Platform for SMEs," *Egyptian Informatics Journal*, vol. 25, pp. 100444, 2024. DOI:10.1016/j.eij.2024.100444

[5] Cang Y. and Wang D., "A Comparative Study on the Online Shopping Willingness of Fresh Agricultural Products Between Experienced Consumers and Potential Consumers," *Sustainable Computing: Informatics and Systems*, vol. 30, pp. 100493, 2021. https://doi.org/10.1016/j.suscom.2020.100493

[6] Casanova R. and Conde A., "Enhancement of LSTM Models based on Data Pre-Processing and Optimization of Bayesian Hyperparameters for Day-Ahead Photovoltaic Generation Prediction," *Computers and Electrical Engineering*, vol. 116, pp. 109162, 2024. DOI:10.1016/j.compeleceng.2024.109162

[7] Daza A., Rueda N., Sánchez M., Espíritu W., and Quiñones M., "Sentiment Analysis on E-Commerce Product Reviews Using Machine Learning and Deep Learning Algorithms: A Bibliometric Analysis, Systematic Literature Review, Challenges and Future Works," *International Journal of Information Management Data Insights*, vol. 4, no. 2, pp. 100267, 2024. https://doi.org/10.1016/j.jjimei.2024.100267

[8] Fecke W., Danne M., and Musshoff O., "E-Commerce in Agriculture-The Case of Crop Protection Product Purchases in a Discrete Choice Experiment," *Computers and Electronics in Agriculture*, vol. 151, pp. 126-135, 2018. https://doi.org/10.1016/j.compag.2018.05.032

[9] Feng G. and Zhang M., "The Coupling Coordination Development of Rural E-Commerce and Rural Revitalization: A Case Study of 10 Rural Revitalization Demonstration Counties in Guizhou," *Procedia Computer Science*, vol. 199, pp. 407-414, 2022. https://doi.org/10.1016/j.procs.2022.01.197

[10] Feng G., Analysis and Synthesis of Fuzzy Control

Systems, *CRC Press*, 2018. https://www.taylorfrancis.com/books/mono/10.1201/EBK1420092646/analysis-synthesis-fuzzy-control-systems-gang-feng

[11] Garg S. and Krishnamurthi R., "A CNN Encoder Decoder LSTM Model for Sustainable Wind Power Predictive Analytics," *Sustainable Computing: Informatics and Systems*, vol. 38, pp. 100869, 2023. DOI:10.1016/j.suscom.2023.100869

[12] Gupta B., Gaurav A., Attar R., Arya V., Alhomoud A., and Chui K., "LSTM Based Neural Network Model for Anomaly Event Detection in Care-Independent Smart Homes," *CMES-Computer Modeling in Engineering and Sciences*, vol. 140, no. 3, pp. 2689-2706, 2024. DOI:10.32604/cmes.2024.050825.

[13] Kabaja B., Wojnarowska M., and Varese E., "Environmental Labels in E-Commerce," *Procedia Computer Science*, vol. 225, pp. 1053-1061, 2023. https://doi.org/10.1016/j.procs.2023.10.093

[14] Kessy A., Lewin A., and Strimmer K., "Optimal Whitening and Decorrelation," *arXiv Preprint*, vol. arXiv:1512.00809v4, pp. 1-14, 2015. https://doi.org/10.48550/arXiv.1512.00809

[15] Khand K. and Senay G., "Evaluation of Streamflow Predictions from LSTM Models in water- and Energy-Limited Regions in the United States," *Machine Learning with Applications*, vol. 16, pp. 100551, 2024. DOI:10.1016/j.mlwa.2024.100551

[16] Li J. and Li X., "Coopetition in Social Commerce: What Influences Livestreaming Knowledge Sharing in Agricultural Clusters?," *Electronic Commerce Research and Applications*, vol. 64, pp. 101383, 2024. https://doi.org/10.1016/j.elerap.2024.101383

[17] Li Y. and Zhang H., "EEG Signal Recognition of VR Education Game Players based on Hybrid Improved Wavelet Threshold and LSTM," *The International Arab Journal of Information Technology*, vol. 22, no. 1, pp. 170-181, 2025. https://doi.org/10.34028/iajit/22/1/13

[18] Lin J., Li L., Luo X., and Benitez J., "How do Agribusinesses Thrive Through Complexity? The Pivotal Role of E-Commerce Capability and Business Agility," *Decision Support Systems*, vol. 135, pp. 113342, 2020. https://doi.org/10.1016/j.dss.2020.113342

[19] Lu Y., He Y., and Ke Y., "The Influence of E-Commerce Live Streaming Affordance on Consumer's Gift-Giving and Purchase Intention," *Data Science and Management*, vol. 6, no. 1, pp. 13-20, 2023. https://doi.org/10.1016/j.dsm.2022.10.002

[20] Migenda N., Möller R., and Schenck W., "Adaptive Dimensionality Reduction for Neural Network-Based Online Principal Component Analysis," *PLOS ONE*, vol. 16, no. 3, pp. 1-32, 2021. https://doi.org/10.1371/journal.pone.0248896

[21] Ren J., Li H., Zhang M., and Wu C., "Massive-Scale Graph Mining for E-Commerce Cold Chain Analysis and Optimization," *Future Generation Computer Systems*, vol. 125, pp. 526–531, 2021. https://doi.org/10.1016/j.future.2021.06.057

[22] Roumeliotis K., Tselikas N., and Nasiopoulos D., "LLMs in E-Commerce: A Comparative Analysis of GPT and LLaMA models in Product Review Evaluation," *Natural Language Processing Journal*, vol. 6, pp. 100056, 2024. https://doi.org/10.1016/j.nlp.2024.100056

[23] Schwering D., Sonntag W., and Kühl S., "Agricultural E-Commerce: Attitude Segmentation of Farmers," *Computers and Electronics in Agriculture*, vol. 197, pp. 106942, 2022. https://doi.org/10.1016/j.compag.2022.106942

[24] Wang Y. and Yu J., "A Co-LNP Academic Award Prediction Mechanism based on the CEw-LSTM Model," *Procedia Computer Science*, vol. 222, pp. 468-477, 2023. DOI:10.1016/j.procs.2023.08.185

[25] Wubet Y. and Lian K., "How Can We Detect News Surrounding Community Safety Crisis Incidents in the Internet? Experiments Using Attention-based Bi-LSTM models," *International Journal of Information Management Data Insights*, vol. 4, no. 1, pp. 100227, 2024. DOI:10.1016/j.jjimei.2024.100227

[26] Zhang J., Fast Qualification of Solder Reliability in Solid-state Lighting System, Delft University of Technology, 2015. https://resolver.tudelft.nl/bf95c658-555a-45ee-a07e-394dfddd74ac

[27] Zhou J., Yang C., Wang X., and Cao S., "A Soft Sensor Modeling Framework Embedded with Domain Knowledge Based on Spatio-Temporal Deep LSTM for Process Industry," *Engineering Applications of Artificial Intelligence*, vol. 126, pp. 106847, 2023. https://doi.org/10.1016/j.engappai.2023.106847

[28] Zhou X., Sheil B., Suryasentana S., and Shi P., "Multi-Fidelity Fusion for Soil Classification via LSTM and Multi-Head Self-Attention CNN model," *Advanced Engineering Informatics*, vol. 62, pp. 102655, 2024. DOI: 110.1016/j.aei.2024.102655.

[29] Zhu X., Chen G., Ni C., Lu X., and Guo J., "Hybrid CNN-LSTM Model Driven Image Segmentation and Roughness Prediction for Tool Condition Assessment with Heterogeneous Data," *Robotics and Computer-Integrated Manufacturing*, vol. 90, pp. 102796, 2024. https://doi.org/10.1016/j.rcim.2024.102796

Jianwei Wu he obtained a Bachelor's degree in Department of Information Management at Zhengzhou University in 2005. Since September 2005, he has been working at Zhengzhou College of Finance and Economics. His current research interests include E-Commerce and Rural E-Commerce.