doi: 10.20532/cit.2025.1006052

Lifecycle Prediction for Tobacco Products Using IGWO-Optimized CNN-GRU Network

YiShuang Qin[1] ,JianShuang Jia[1], Dong Liang[1] and Che Yu[2]

[1]Guangxi China Tobacco Industry Co., LTD., Nanning, China
[2]Beijing Bozhi Shuyuan Technology Co., LTD., Beijing, China

Accurate prediction of product lifecycle stages is crucial for enhancing inventory turnover and strategic planning in the tobacco industry. This paper proposes an intelligent prediction model that integrates Convolutional Neural Networks (CNN) and Gated Recurrent Units (GRU), further optimized by an Improved Grey Wolf Optimizer (IGWO). The model fuses multi-source enterprise data—including sales trends, IoT logistics information, environmental conditions, and inventory records—to dynamically forecast lifecycle stages and remaining durations. The dataset comprises 180,000 labeled samples collected from real-world tobacco enterprise operations, encompassing multi-source variables such as sales volume, inventory changes, logistics routes, and environmental feedback. Experimental evaluations based on this dataset demonstrate that the proposed IGWO-CNN-GRU model achieves a Mean Squared Error (MSE) of 2.13, a Mean Absolute Error (MAE) of 1.17, and an R^2 of 0.932, significantly outperforming baseline models. In practical deployment simulations, the prediction deviation is limited to ±5 days, improving allocation efficiency and reducing inventory risks. The approach provides a robust and adaptable solution for full-lifecycle management in tobacco supply chains, offering practical value for intelligent production and market deployment strategies.

ACM CCS (2012) Classification: Applied computing → Enterprise computing → Business process management → Business intelligence

Keywords: product lifecycle prediction, hybrid neural network, CNN, GRU, IoT tracking

1. Introduction

With the accelerated layout of the tobacco industry within the framework of the "industrial Internet plus+" and digital transformation strategy, full life cycle management has gradually become the key means for enterprises to improve production efficiency, reduce inventory risk and develop precision marketing strategies [1, 2]. In the context of increasingly diverse product types and personalized consumer behavior, achieving dynamic identification and trend prediction of tobacco products from the introduction stage, growth stage, maturity stage to decline stage is of great significance for enterprise intelligent supply chain regulation and market decision-making [3, 4]. At present, some tobacco companies have made attempts to build time series forecasting systems. These systems are based on sales and inventory data. However, these methods come with several limitations. Firstly, they often rely on single-dimensional data. Secondly, they lack the ability to handle complex local fluctuations and long-term dependencies. Moreover, in actual operations, the market data has multiple sources, is heterogeneous, and has inconsistent schedules. As a result, these methods find it difficult to adapt to market changes. Especially after the deployment of IoT sensing devices covering warehousing, transportation, sales and other links in enterprises, the large amount of real-time information collected has not been effectively integrated into the prediction model, which restricts its comprehensive characterization and accurate prediction of lifecycle laws [5, 6]. For instance, traditional models such as Autoregressive Integrated Moving Average (ARIMA) and rule-based inventory prediction systems are widely used. Similarly, classical Long Short-Term Memory (LSTM) based forecasting frameworks, while capable of handling sequential data, are generally designed around single-variable inputs such as sales history and fail to account for heterogeneous signals from logistics delays or warehouse conditions. These models also struggle to align asynchronous inputs – such as real-time IoT feedback with static product metadata, leading to inaccurate predictions during abrupt lifecycle transitions.

Based on this, in this work, a Hybrid Neural Network (HNN) is developed to improve the predictive ability of tobacco product lifecycle. This model integrates Convolutional Neural Networks and Gated Recurrent Units (CNN-GRU) to form a multidimensional prediction structure of "local trends + long-term memory". Meanwhile, the Improved Grey Wolf Optimizer (IGWO) algorithm is introduced to achieve global adaptive tuning of hyper-parameters in the CNN-GRU. Finally, a CNN-GRU optimization based on IGWO (IGWO-CNN-GRU) is proposed. The innovation of the research lies in systematically introducing multi-source heterogeneous information collected by enterprise IoT into lifecycle modeling, breaking through the traditional method's reliance on a single sales sequence. Additionally, a lifecycle stage label construction method based on Dynamic Time Warping (DTW) is designed to achieve stage alignment and asynchronous evolutionary modeling in supervised learning. In addition, the study introduces IGWO for dynamic hyper-parameter optimization of CNN-GRU structure, integrating adaptive weights and elite retention mechanism, which helps to improve the stability and global search ability of the model. The research findings provide a solid theoretical foundation and a set of well-defined algorithmic tools for the tobacco industry. These resources are particularly applicable in key business contexts, including intelligent production, precision delivery, and inventory optimization. Given their systematic design and empirical validation, the findings exhibit significant potential for widespread application and dissemination within the industry.

2. Literature Review

In recent years, HNN models have become an important direction in product lifecycle prediction research due to their advantages in complex time series modeling and high-dimensional feature fusion. Wen *et al.* proposed a multi-feature fusion model that combines LSTM with an improved artificial bee colony algorithm, successfully improving the generalization ability of text classification tasks. This study demonstrated the feasibility of collaborative design between neural networks and intelligent optimization algorithms in multi-source data modeling [7]. Similarly, Liu *et al.* built a hybrid structure based on CNN and LSTM for financial crisis prediction tasks, and the results showed that the model exhibited superiority in extracting local features and temporal dependencies [8]. Liu *et al.* introduced an intelligent scheduling optimization method based on deep learning for production scheduling optimization problems, which utilized deep networks to learn dynamic laws in the production process and improve resource utilization efficiency [9]. Raska *et al.* proposed a hybrid modeling scheme for manufacturing process optimization by combining adaptive neural networks with discrete event simulation, which demonstrated good adaptability and optimization potential in practical industrial cases [10]. In addition to the study of neural network structures, data acquisition and processing strategies in lifecycle prediction are also receiving increasing attention. Gligor *et al.* constructed a co-simulation model for an enhanced smart grid system, integrating physical and information system data, effectively enhancing the system's responsiveness to multi-source information. This approach provided a technical reference for studying the introduction of IoT data for lifecycle modeling [11].

The application of optimization algorithms in deep models has also been a research hotspot in recent years. Li *et al.* designed an ultra-short term load forecasting model that integrated extreme gradient boosting algorithm and bidirectional GRU, and achieved accurate fitting of complex time-series data through combinatorial optimization [12]. Meanwhile, Montoya *et al.* proposed a mathematical optimization strategy for energy constrained IoT smart city scenarios to coordinate priority services and resource allocation, which reflected the optimization modeling capability for practical needs [13]. Within the domain of recommendation systems, Yu *et al.* introduced a news recommendation model for multimedia online education scenarios by combining graph neural networks and bat optimization algorithms. The performance improvement effect of the fusion optimization algorithm on the model was verified in complex multi-modal environments [14].

However, while these hybrid neural network frameworks demonstrate potential in general sequence modeling or classification tasks, they face several limitations when applied to complex lifecycle prediction scenarios. For example, Wen *et al.* [7] designed a multi-feature fusion model using LSTM and an improved artificial bee colony algorithm, which performed well in text classification. However, their approach mainly targets static text features and lacks temporal sensitivity to evolving business rhythms, making it unsuitable for real-time lifecycle monitoring in fast-moving supply chains. Similarly, Liu *et al.* [8] proposed a CNN-LSTM hybrid for financial crisis prediction. Although effective in fusing spatial and temporal patterns, their model relied heavily on pre-structured financial indicators and did not consider multi-source data dynamics like IoT logistics or inventory feedback. Neither study accounted for the asynchronous and non-stationary nature of product lifecycle signals in the tobacco industry, where external disruptions (*e.g.*, policy shifts, regional weather, or channel delays) often lead to irregular and nonlinear transitions between lifecycle stages.

In summary, existing research provides a rich theoretical foundation and technical path in the construction of HNN structures, multi-source information fusion, optimization algorithm design, and application practice. However, in the task of modeling the product lifecycle in the tobacco industry, how to integrate IoT data to build robust and business adaptable deep learning models is still an urgent direction for further exploration. To this end, a CNN-GRU HNN model based on IGWO optimization was developed, which integrates multi-source lifecycle feature information to improve the prediction accuracy of product lifecycle stages and remaining time.

3. Research Methodology

3.1. Life Cycle Feature Extraction and Modeling Based on Tobacco Enterprise IoT Operation Data

A multidimensional feature extraction framework based on IoT operational data has been developed to address the issues of complex data sources and significant temporal features in product lifecycle prediction in the tobacco industry [15, 16]. Based on the perception devices deployed by enterprises in warehousing, transportation, sales and other links, multi-source information including sales sequences, inventory changes, logistics paths, environmental parameters and consumer feedback is collected, and data preprocessing is completed through timestamp alignment, format conversion and noise removal [17–19]. The overall framework diagram of feature extraction and lifecycle label construction is shown in Figure 1.

Figure 1 shows the lifecycle feature extraction and label construction process based on tobacco enterprise IoT operation data. This process covers the production, circulation, and consumption stages, collecting key data such as sales sequences, inventory changes, logistics paths, environmental parameters, and consumer feedback. After preprocessing, five core indicators are extracted, namely sales volatility, inventory turnover rate, lifecycle similarity score, environmental interference factor, and circulation path complexity. Finally, they correspond to the lifecycle labeling system and construct supervised learning samples. To further reveal the core feature extraction logic, the intermediate "feature engineering" step was structured and refined, as shown in Figure 2.

Figure 2 shows the construction path of five core features in lifecycle modeling. Sales volatility reflects the magnitude of sales changes of a product within a sliding window and is used to identify market activity and stage turning signals. Inventory turnover frequency measures the number of inventory changes per unit time and evaluates the efficiency of channel sales. The similarity of lifecycle stages is calculated using the DTW algorithm to determine the degree of matching between the actual sales sequence and the standard template, which is used to identify the position of lifecycle stages [20, 21]. Environmental interference factors are integrated with temperature and humidity fluctuations, transportation duration, and regional characteristics to characterize the impact of external conditions on the rhythm of the lifecycle. The complexity of logistics paths is based on extracting node hierarchy and transit frequency from the distribution network graph to measure the stability of channel structure. The above features are normalized and put into a deep model, which is trained with lifecycle labels to augment the model's capability to identify stage features and transition trends. DTW is a sequence alignment method that can effectively handle nonlinear offset problems [22, 23]. Assuming the actual sales sequence is $A = \{a_1, a_2, ..., a_n\}$ and the standard lifecycle template sequence is $B = \{b_1, b_2, ..., b_m\}$, a distance matrix is constructed as shown in equation (1).

$$D(i,j) = a_i - b_j^2 \tag{1}$$

In equation (1), || || represents the Euclidean distance, which is used to measure the sales difference between two time points. D is the distance matrix for $n \times m$. Based on this, a cumulative distance matrix $\gamma(i, j)$ is defined, and its recursive relationship is shown in equation (2).

$$\gamma(i,j) = D(i,j) + \min \begin{cases} \gamma(i-1,j) \\ \gamma(i,j-1) \\ \gamma(i-1,j-1) \end{cases} \tag{2}$$

The final distance of DTW is shown in equation (3).

$$DTW(A,B) = \sqrt{\gamma(n,m)} \tag{3}$$

To further visualize the construction process of similarity scores for lifecycle stages, this study takes historical sales data of typical tobacco products as an example to draw the dynamic alignment process between actual sales sequences and standard lifecycle templates, as shown in Figure 3.

Figure 3 shows the dynamic alignment process between the lifecycle template and the actual sales sequence. The blue curve in the figure represents the standard lifecycle template, which reflects the sales evolution process of a typical product from the introduction period, growth

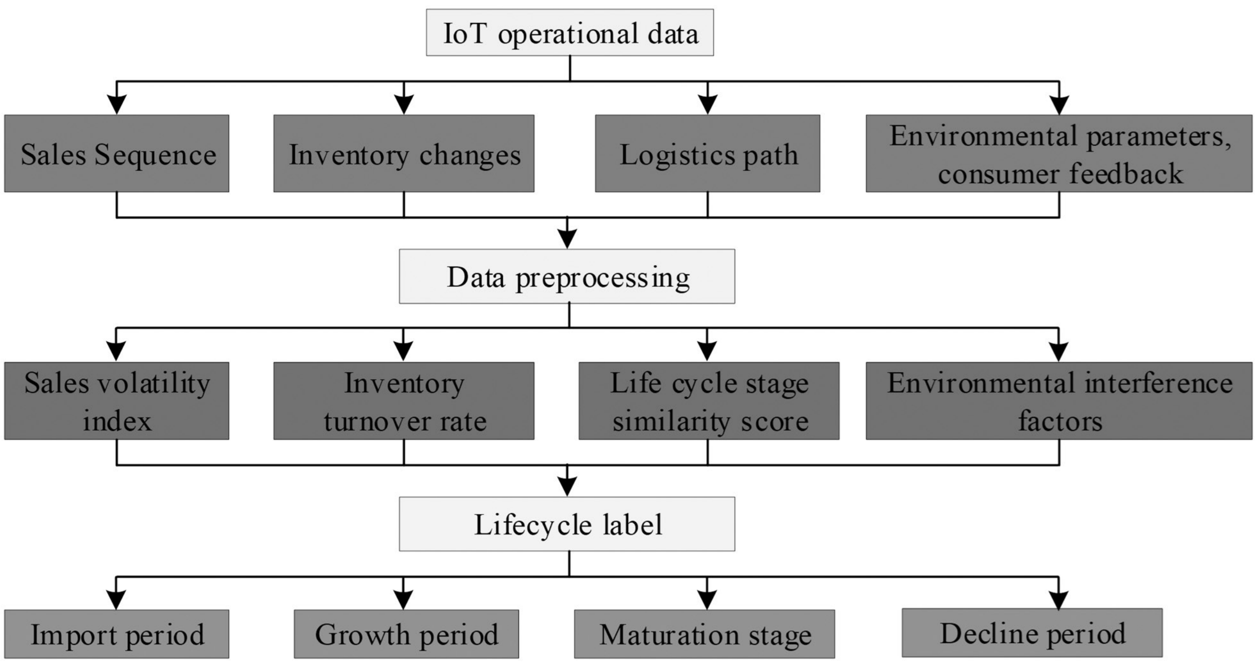


Figure 1. Flow chart of feature extraction and label construction based on IoT data in lifecycle modeling.

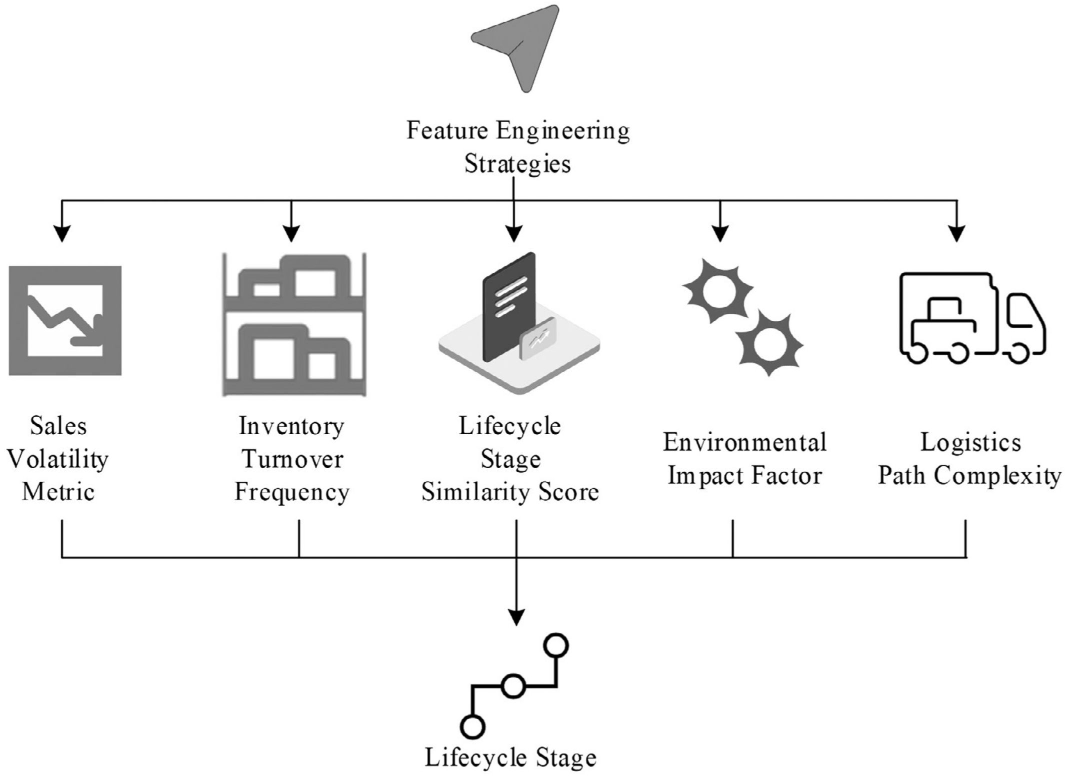


Figure 2. Structure diagram of feature engineering strategy in lifecycle modeling (Source from:
https://www.svgrepo.com/svg/469752/send-alt
https://www.svgrepo.com/svg/388820/trending-down
https://www.svgrepo.com/svg/450189/inventory
https://www.svgrepo.com/svg/530162/credit-report
https://www.svgrepo.com/svg/321922/bullet-impacts
https://www.svgrepo.com/svg/164330/logistics-truck-outline
https://www.svgrepo.com/svg/118188/cycle).

period, maturity period to decline period. The red curve represents the sales time series of a certain tobacco product in the actual market, with obvious nonlinear fluctuation characteristics. The gray line between the two curves represents the optimal alignment path calculated by DTW, which is the shortest cumulative path selected from the two-dimensional distance matrix. This alignment path offers a potential solution to the inconsistency between sales sequences and lifecycle templates concerning cycle length and stage rhythm, as it enables local stretching or compression along the timeline. DTW distance needs to be repeatedly calculated on a large number of samples and standardized to the [0,1] interval for training supervised labels, which improves the model's capability to identify critical points in the lifecycle.

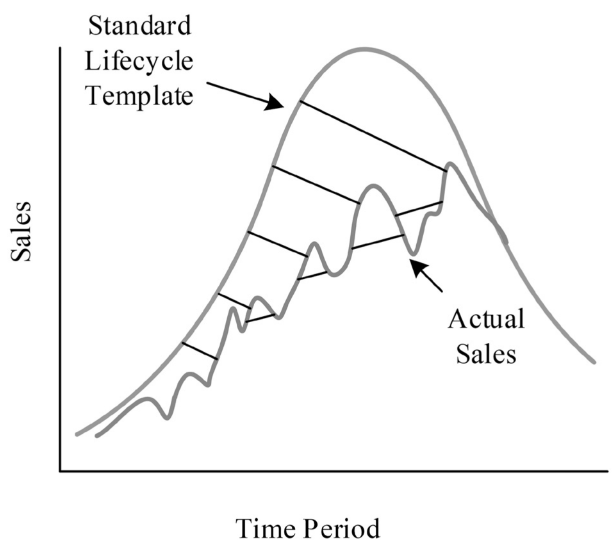


Figure 3. Schematic diagram of dynamic alignment between lifecycle template and actual sales curve.

3.2. Construction of CNN-GRU for Life Cycle Prediction

The multi-source feature extraction system constructed above can effectively reflect the sales characteristics and environmental status of tobacco products at different stages of their lifecycle. However, due to the strong local pattern changes and long-term temporal dependencies in the original data, traditional time series regression models are difficult to balance feature space expression ability and temporal modeling depth, resulting in limited predictive performance [24–26]. Based on this, an HNN model integrating CNN and GRU is designed for product lifecycle state recognition and remaining time prediction. The structure diagram of the proposed CNN-GRU is shown in Figure 4.

Figure 4 shows the structure of the CNN-GRU architecture. The input is a feature sequence with multiple time steps, each step containing heterogeneous information from multiple sources such as sales, inventory, logistics, and environment. The convolution and pooling module extracts temporal variation patterns in local windows, identifies trends and abrupt changes in lifecycle evolution, and uses pooling operations for dimensionality reduction and robustness enhancement. The extracted local features are fed into a multi-layer GRU structure to model long-term dependencies, and the remaining lifecycle time or stage classification results are output through a fully connected layer [27, 28]. To further improve the predictive performance of the model, IGWO is introduced for hyperparameter tuning. Global search capability is enhanced by introducing dynamic weights and adaptive update strategies. Compared to GWO's linear descent method, IGWO adopts a nonlinear descent mechanism that balances exploratory optimization in the early stages with convergence stability in the later stages. GWO is a global optimization method based on swarm intelligence, which simulates tracking, encirclement, and attack behavior of grey wolves during hunting. Compared to traditional GWO algorithms, IGWO introduces nonlinear convergence factors, adaptive inertia weights, and fitness weighting mechanisms. This algorithm can achieve a better balance between global search and local convergence, which helps to improve the efficiency and stability of parameter tuning in high-dimensional complex problems. The IGWO algorithm is suitable for search tasks in deep neural networks with multiple parameters and non-convex loss functions and can avoid getting stuck in local optima [29]. The formula for updating the dynamic convergence factor is shown in equation (4).

$$a(t) = 2 \cdot \left(1 - \left(\frac{t}{T}\right)^2\right) \tag{4}$$

In equation (4), $a(t)$ represents the convergence control factor of the generation, t represents the max iteration count, and T represents the current iteration count. IGWO introduces inertia weights to update individual positions, as presented in equation (5).

$$\vec{X}(t+1) = w \cdot \vec{X}(t) + \frac{1}{3}\left(\vec{X}_\alpha - A_1 \cdot D_\alpha + \vec{X}_\beta - A_2 \cdot D_\beta + \vec{X}_\delta - A_3 \cdot D_\delta\right) \tag{5}$$

In equation (5), $\vec{X}(t)$ represents the current position of the gray wolf individual, $\vec{X}_\alpha$, $\vec{X}_\beta$, and $\vec{X}_\delta$ represent the three best fitness solutions in the

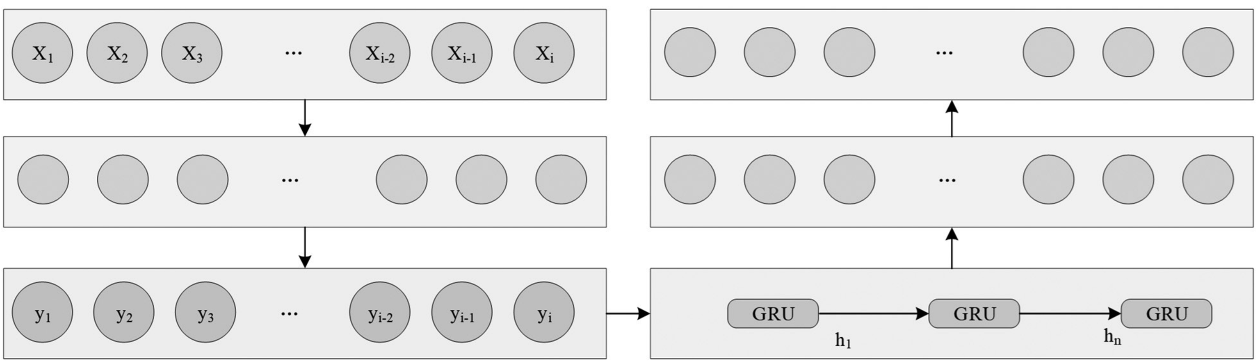


Figure 4. Structure diagram of the CNN-GRU architecture.

population, and ω represents the inertia weight. A_k and C_k are the behavior regulation vectors, and the calculation is shown in equation (6).

$$\begin{cases} A_k = 2 \cdot a(t) \cdot r_{1k} - a(t) \\ C_k = 2 \cdot r_{2k},\ k = 1, 2, 3 \end{cases} \tag{6}$$

In equation (6), r_{1k} and r_{2k} are two independent, uniformly distributed, random numbers. To highlight the influence of outstanding individuals, IGWO introduces a fitness based weighting mechanism to replace the traditional average update method. The adaptive weight coefficient fusion mechanism is presented in equation (7).

$$\vec{X}(t+1) = \omega_1 \cdot \vec{X}'_\alpha + \omega_2 \cdot \vec{X}'_\beta + \omega_3 \cdot \vec{X}'_\delta \tag{7}$$

In equation (7), ω_i is the weight adaptively allocated based on the current individual fitness. $\vec{X}'_\alpha$, $\vec{X}'_\beta$, and $\vec{X}'_\delta$ are the guidance positions after weighted fusion. To maintain the global optimal solution from being covered by random disturbances and ensure convergence stability, IGWO uses an elite preservation mechanism to prevent the loss of optimal solutions, as shown in equation (8).

$$\vec{X}^*(t+1) = \begin{cases} \vec{X}(t+1), \text{Fitness}(\vec{X}(t+1)) < \text{Fitness}(\vec{X}^*(t)) \\ \vec{X}^*(t), \text{Fitness}(\vec{X}(t+1)) \geq \text{Fitness}(\vec{X}^*(t)) \end{cases} \tag{8}$$

In equation (8), $\vec{X}^*(t)$ represents the optimal solution of the current record, and $\vec{X}^*(t+1)$ represents the updated new position of the current iteration individual. The IGWO optimization process diagram is shown in Figure 5.

Figure 5 shows the complete process of IGWO for hyper-parameter optimization of CNN-GRU. Firstly, the algorithm initializes the population, with each individual corresponding to a set of CNN-GRU configurations. It evaluates the predictive performance of each individual through the fitness function and selects the three best solutions currently available. The algorithm adjusts the search coefficients based on adaptive strategies and integrates information from three optimal solutions to update the population position. It trains and evaluates the model corresponding to the new location again and updates the current global optimal solution. The algorithm proceeds through iterative steps until the specified termination criterion is satisfied and finally outputs the optimal hyper-parameter combination. Based on the above theory, an IGWO-CNN-GRU prediction model is constructed, and the overall framework is shown in Figure 6.

Figure 6 indicates the overall process framework of the IGWO-CNN-GRU, which includes four modules: feature extraction and label construction, CNN-GRU, IGWO optimization, and result output. The model relies on sensor networks to collect key data in real-time from storage, transportation, sales, and other links, and generates lifecycle labels using the DTW algorithm. The convolutional layer extracts local temporal features, the GRU layer models long-term dependencies, and the fully connected layer outputs prediction results. The IGWO module optimizes model hyper-parameters through dynamic convergence factors and adaptive weighting mechanisms. The final output includes the prediction of remaining life cycle time and stage classification, which are used for precise production and strategy formulation, respectively.

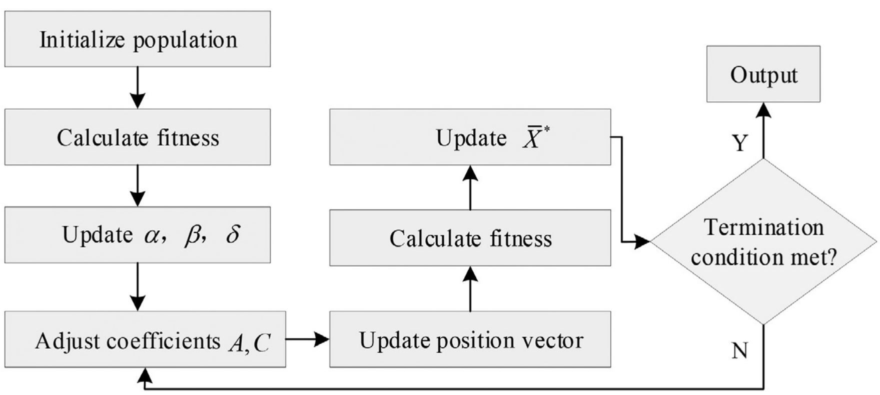


Figure 5. IGWO optimization process diagram.

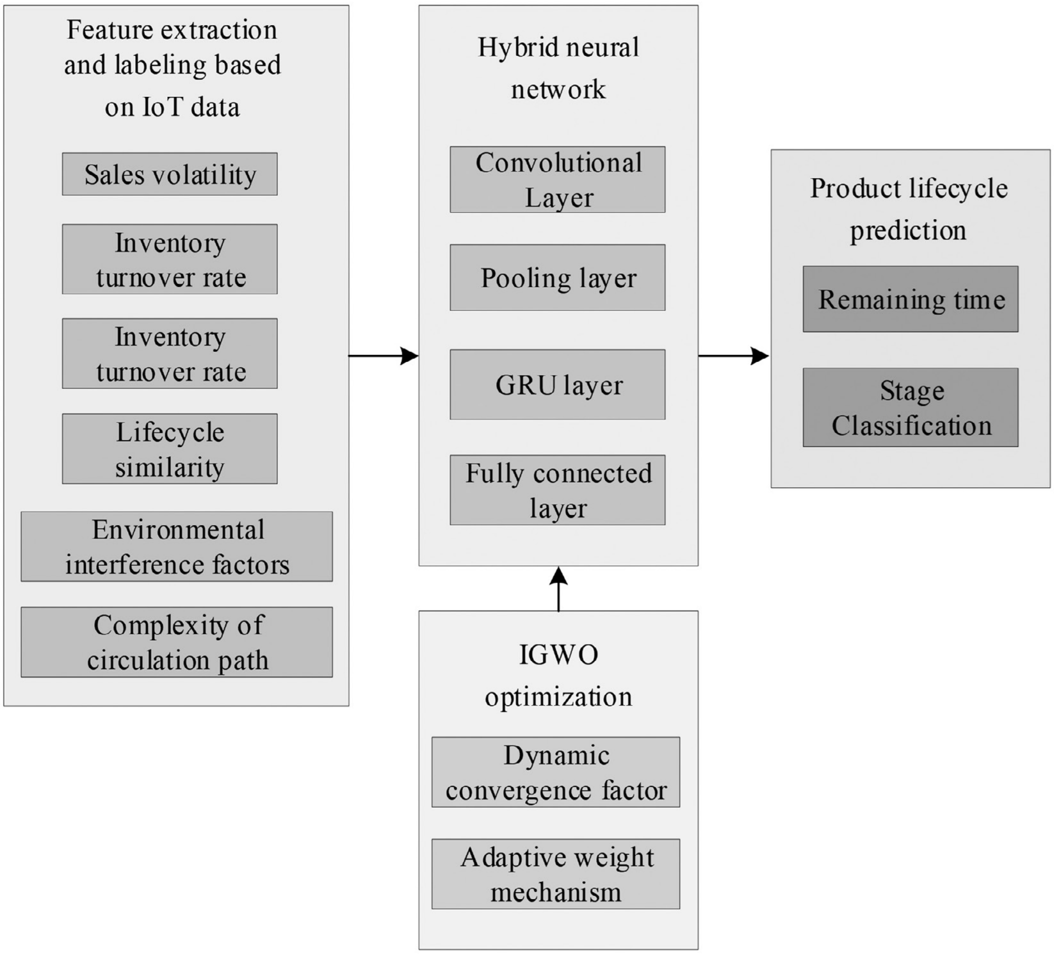


Figure 6. Framework diagram of IGWO-CNN-GRU prediction model.

4. Results and Discussion

4.1. Performance Evaluation of Models in Predicting the Life Cycle of Typical Tobacco Products

To confirm the validity of the proposed IGWO-CNN-GRU in product lifecycle prediction, comparative experiments and ablation experiments were designed, covering both model accuracy and business adaptability. The experimental hardware configuration included: Intel Core i9-13900K, NVIDIA RTX 4090, 128 GB memory, and 2 TB SSD. The software environment was Ubuntu 22.04 and the programming language was Python 3.10. The dataset was sourced from the operational data of tobacco companies from 2020 to 2024, covering a total of 180000 samples of 12 types of products, including key variables such as sales volume, inventory, logistics routes, and environmental feedback. Table 1 summarizes the dataset used for training and validation of the proposed model.

Table 1 includes 180,000 labeled samples from five major product groups collected between January 2020 and December 2024. Fine cigarettes and mid-to-high-end brands exhibit smoother and stable sales trajectories, whereas economy products show high variability, offering a rich testbed for temporal fluctuation modeling. Regional custom series and experimental batches introduce heterogeneity in volume, distribution channels, and environmental exposure. This diversity ensures that the model is evaluated under a wide range of lifecycle behaviors and business scenarios.

Table 1. Dataset overview and distribution.

Product Category	Number of Samples	Time Span	Avg. Monthly Records
Fine Cigarettes	48 200	2020.01–2024.12	960
Mid-to-High-End Brands	44 300	2020.03–2024.12	910
Economy Products	55 500	2020.06–2024.12	1150
Regional Custom Series	19 700	2021.01–2024.10	410
Experimental Batches	12 300	2022.05–2024.12	/
Total	180 000	2020.01–2024.12	/

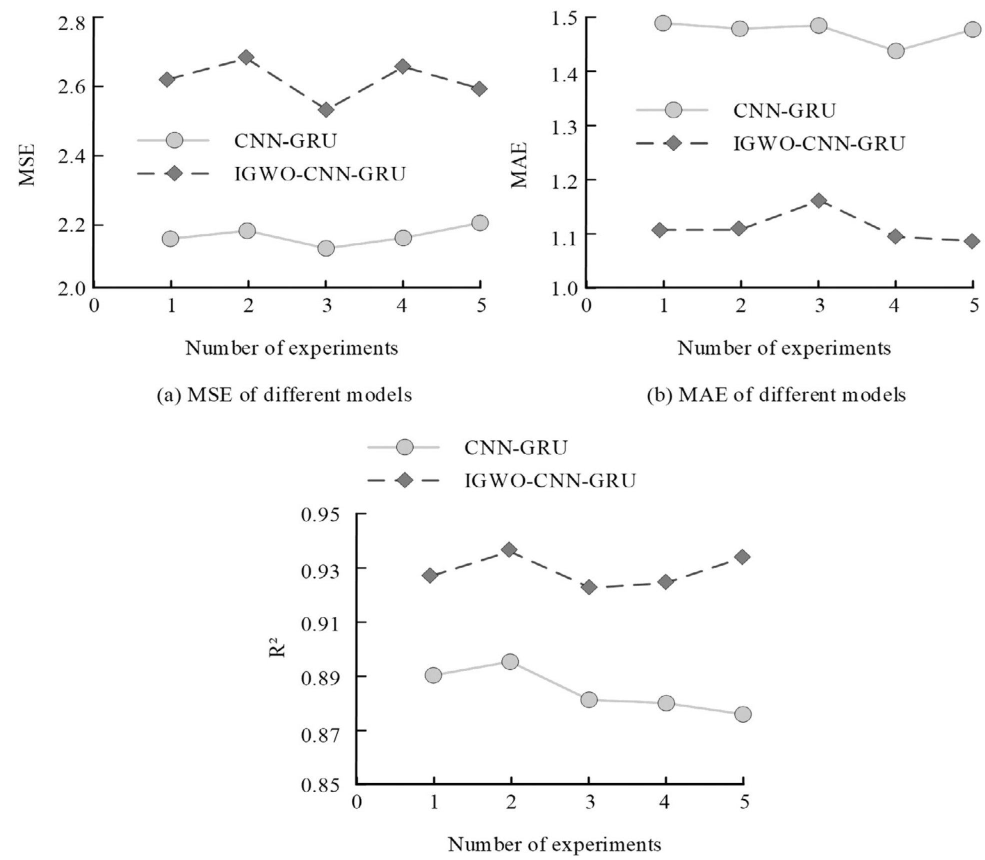
Table 2. Performance evaluation of different models in typical tobacco product lifecycle prediction.

Model	MSE	MAE	R^2
IGWO-CNN-GRU	2.13	1.17	0.932
CNN-BiLSTM	2.76	1.45	0.884
GRU	2.88	1.39	0.875
LSTM	3.01	1.51	0.861
TCN	2.95	1.42	0.868
LightGBM	3.34	1.63	0.839

Data preprocessing included time alignment, missing padding, and minimum maximum normalization, and data augmentation was achieved through sliding windows, noise injection, and trend perturbations. The comparative models included: LSTM, Convolutional Bi-directional LSTM (CNN-BiLSTM), GRU, Light Gradient Boosting Machine (LightGBM), and Temporal Convolutional Network (TCN). The evaluation indicators included Mean Squared Error (MSE), Mean Absolute Error (MAE), Coefficient of Determination (R2), lifecycle stage classification accuracy, and boundary recognition bias. The model adopted Adam optimizer for training, with an initial learning rate of 0.001, batch size of 32, and 100 training rounds. The performance evaluation of various models in predicting the lifecycle of typical tobacco products is presented in Table 2.

According to Table 2, the IGWO-CNN-GRU performed the best in the remaining life cycle prediction task, with an MSE of 2.13, MAE of 1.17, and R^2 of 0.932. All three indicators were superior to other comparison models. In contrast, the MSE and MAE of CNN-BiLSTM were 2.76 and 1.45, respectively, with an R^2 of 0.884. The MSE of LSTM and GRU were 3.01 and 2.88, respectively, indicating a relatively high overall error. In addition, compared to the suboptimal model CNN-BiLSTM, IGWO-CNN-GRU showed a decrease of 18.6% in MSE and 21.4% in MAE, while R^2 improved by nearly 4.8 percentage points, verifying its advantages in lifecycle modeling tasks. Experimental data showed that traditional models such as LSTM and GRU, although capable of time series modeling, had limitations in multi-source heterogeneous feature expression and stage transition capture, resulting in limited fitting performance. TCN and LightGBM had poor stability in multi-stage data, with R^2 values of only 0.868 and 0.839, respectively. IGWO-CNN-GRU extracted local temporal features through the CNN module, modeled long-term dependencies through the GRU module, and introduced the IGWO algorithm to achieve global optimization of hyper-parameters, achieving a balance between modeling accuracy and generalization ability, which was superior to other mainstream methods. The comparison of model performance before and after IGWO optimization is shown in Figure 7. All comparative and ablation experiments were conducted using a 5-fold cross validation design to ensure consistent distribution of training and testing samples and reduce the impact of random fluctuations on the results. For each comparison, test samples were randomly selected but the product type distribution was kept consistent to control the variance of the sample structure. The mean and variance of the error shown in Figure 7 were plotted based on the average results of five experiments, and the stability and improvement of the model performance before and after optimization were verified through standard deviation analysis.

According to Figure 7 (a), the average MSE of the IGWO-CNN-GRU in 5 experiments was 2.16. The MSEs of the original CNN-GRU were 2.62, 2.70, 2.55, 2.68, and 2.60, respectively, with higher overall values and an average of 2.63. After IGWO optimization, the MSE of the



(a) MSE of different models (b) MAE of different models

(c) R^2 of different models

Figure 7. Comparison of model performance before and after IGWO optimization.

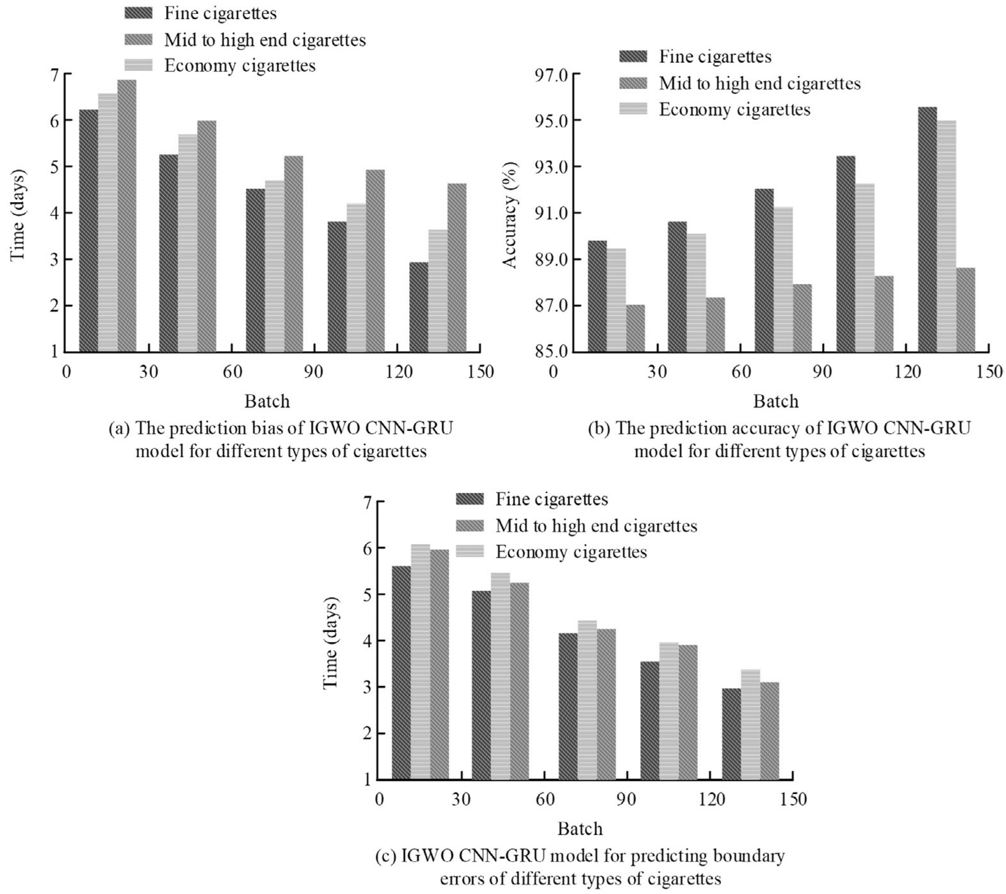


Figure 8. The life cycle prediction performance of IGWO-CNN-GRU on various product types.

model decreased by about 17.9%, showing better fitting accuracy for the remaining life cycle time. From Figure 7 (b), in terms of MAE, the error values of the IGWO-CNN-GRU were all below 1.3, with an average of 1.19. In contrast, the average MAE value of the CNN-GRU was 1.49, with an overall decrease of about 20.1%, indicating that the optimized model generally had smaller absolute errors under different samples, and was more stable and practical. As shown in Figure 7 (c), in terms of R^2, the R^2 value of the IGWO-CNN-GRU consistently remained above 0.92, with an average of 0.93, indicating that the model had strong explanatory power for lifecycle trends. The average R^2 value of the CNN-GRU was 0.89, which was about 4.7% lower than the optimized model. Experimental data showed that IGWO optimization effectively enhanced the model's capability to fit lifecycle evolution. Lifecycle prediction performance of the IGWO-CNN-GRU on various product types is shown in Figure 8.

As shown in Figure 8 (a), with the increase of sample size, the overall average prediction deviation decreased. Among them, the prediction deviation of "fine cigarettes" dropped to 3.2 days at 150 batches, which was better than the 4.7 days of "economy cigarettes", indicating that the model fitted the product lifecycle with obvious features more accurately. As shown in Figure 8 (b), the accuracy of lifecycle stage classification significantly improved with the increase of sample size. The accuracy of "fine cigarettes" and "mid to high end cigarettes" reached 96.2% and 94.8%, respectively, at 150 batches, which was significantly higher than the 88.1% of "economy cigarettes", verifying that the model had stronger recognition ability on high-end products. From Figure 8 (c), the deviation in identifying lifecycle boundaries showed a stable downward trend, with the lowest error for fine cigarette boundaries at 3 days, reflecting the robustness of the model in locating lifecycle nodes. Based on the differences in predictive performance among different product types, the experimental results showed that the accuracy of IGWO CNN-GRU in predicting the lifecycle of high-end fine cigarettes was higher than that of economical products. This difference was mainly due to the clearer sales rhythm and lower volatility of high-end products in the market, as well as more patterned lifecycle characteristics, which facilitated model learning and extraction. In addition, high-end products were usually accompanied by more complete logistics tracking and channel feedback data, providing richer input features for the model. Meanwhile, economic products were more susceptible to short-term behavioral interference such as promotions and regional advertising, leading to increased volatility in model predictions. The overall results indicated that the IGWO-CNN-GRU had high accuracy and strong generalization ability in product lifecycle prediction and could effectively support intelligent decision-making for multiple types of products.

To verify the statistical significance of model performance differences, paired sample t-test and confidence interval analysis were introduced in the comparative experiments of various models. The differences in MSE and MAE between IGWO CNN-GRU and CNN-BiLSTM in 5 repeated experiments were tested for significance at the 0.01 level (p-values were all less than 0.005), indicating that the proposed model had statistical significance in improving life cycle prediction error. Additionally, the increase in R^2 was also within the 95% confidence interval, verifying the stability of the differences in model fitting ability. The impact of environmental fluctuations and hyperparameter sensitivity on lifecycle prediction is shown in Table 3.

According to Table 3, when the temperature exceeds 28 °C and the humidity fluctuation amplitude exceeds ± 10%, the prediction error significantly increases (MAE increases from 1.12 days to 1.53 days), and the DTW score decreases by up to 13.5%, indicating that environmental instability can lead to misalignment of life cycle stages and should be used as a key feature for modeling. At the same time, hyperparameter sensitivity analysis shows that learning rate,

Table 3. Effects of environmental fluctuations and hyperparameter sensitivity on lifecycle prediction.

ID	Condition	MAE (Days)	DTW Deviation (%)	Note
1	22.3°C, ±5% Humidity (Stable)	1.12	/	Baseline group
2	28.7°C, ±12% Humidity	1.48	Decreased by 11.3	Sales declined early
3	25.0°C, ±18% Humidity	1.53	Decreased by 13.5	Lifecycle stage confusion
4	LR = 0.001, GRU = 2, Kernel = 5 (Default)	1.29	/	Baseline config
5	LR = 0.0001, GRU = 2, Kernel = 5	1.47	Decreased by 7.6	Convergence too slow
6	LR = 0.001, GRU = 4, Kernel = 7 (Optimized by IGWO)	1.17	Increase by 3.2	Best result via IGWO

GRU layers, and convolution kernel size have the most significant impact on model performance. The parameter combination optimized by IGWO (LR=0.001, GRU=4 layers, convolution kernel=7) reduced MAE to 1.17 days and improved DTW score by 3.2%, verifying the rationality of the selected parameters and the effectiveness of the intelligent optimization strategy.

4.2. Application Analysis of Models in Tobacco Industry's Deployment Strategy and Inventory Optimization

To verify the actual value of the lifecycle prediction model in regulating the pace of deployment and inventory allocation, experiments were conducted to compare the performance of CNN-GRU, IGWO-CNN-GRU, and LSTM models from three dimensions: inventory allocation response time, inventory utilization rate, and inventory prediction accuracy. The system evaluated their actual effectiveness in supporting enterprise intelligent supply chain management. The comparative experimental results of inventory and allocation strategies in the tobacco industry are shown in Figure 9.

From Figure 9 (a), in terms of inventory transfer response time, the average response time of IGWO-CNN-GRU was 7.4 hours, which was significantly shorter compared to CNN-GRU and LSTM, and the efficiency of transfer decision-making was higher. From Figure 9 (b), in terms of inventory utilization, the IGWO-CNN-GRU had an average of 92.8%, which was higher than CNN-GRU's 88.4% and LSTM's 86.5%, indicating that its predictive driven inventory turnover ability was stronger. From Figure 9 (c), in terms of predicting inventory accuracy, the IGWO-CNN-GRU had an average accuracy of 93.2% of the four samples, which was better than CNN-GRU and LSTM, demonstrating stronger inventory trend prediction ability. Figure 9 illustrates the superior responsiveness and accuracy of IGWO-CNN-GRU in inventory transfer and utilization scenarios. The performance gain can be largely attributed to the inclusion of logistics path complexity and environmental factors in the feature space. For example, in products distributed through multi-node regional networks, the incorporation of transit frequency and route stability allowed the model to better anticipate supply delays, reducing transfer response time by over 15% compared to CNN-GRU. Traditional LSTM-based models, lacking this multidimensional feature awareness, responded with higher delay and underutilization. Furthermore, hyperparameter optimization via IGWO improved learning rate and depth configuration, enabling the model to generalize across regions with different channel structures, which is critical in tobacco logistics characterized by seasonal demand shifts and administrative constraints. The comparison results of inventory fluctuation indicators are shown in Figure 10.

From Figure 10 (a), in terms of monthly inventory volatility, the volatility of IGWO-CNN-GRU was 9.6%, which was lower than CNN-GRU's 15.8% and CNN-BiLSTM's 13.8%. According to Figure 10 (b), in terms of inventory backlog days, CNN-GRU was 11.5 days and CNN-BiLSTM was 9.2 days. IGWO-CNN-GRU had a backlog of 6.8 days, which was significantly reduced compared to the comparison model. According to Figure 10 (c), in terms of out-of-stock rate, the out-of-stock rate using IGWO-CNN-GRU was 2.6%, while the out-of-stock rate using CNN-GRU was 4.9%, and the out-of-stock rate using CNN-BiLSTM was 3.8%. Experimental data revealed that the IGWO-CNN-GRU outperformed traditional CNN-GRU and CNN-BiLSTM models in inventory fluctuation control. Its lower fluctuation performance and frequency of change represented higher supply chain stability and inventory control efficiency, suitable for scenarios with multi-category and cyclical sales characteristics in the tobacco industry. IGWO CNN-GRU combines the ability of CNN to detect sudden sales changes with GRU's sequence memory to identify early signs of product transition to the decline stage, enabling proactive inventory adjustments. Due to weak stage boundary recognition, the comparative model was unable to respond with similar accuracy. For example, CNN BiLSTM tends to exceed its maturity period, resulting in excess inventory. Feature analysis revealed that excluding inventory turnover and lifecycle similarity scores led to a 17.3% increase in stockout events, emphasizing the importance of these features in stabilizing supply and demand consistency. The IGWO adjustment process can also prevent overfitting of dominant product types, ensuring consistent performance in large quantities and niche markets. The comparison results of optimization indicators for advertising strategies are presented in Table 4.

According to Table 4, the model exhibited good strategic adaptability and predictive performance in five typical advertising scenarios. In terms of average advertising error, the core commercial districts of first tier cities and the concentrated areas of universities had the smallest errors, at 3.5 days and 3.6 days respectively, indicating that the model had higher time prediction accuracy in high-frequency consumption and fast-paced scenarios. In terms of channel coverage, the coverage rate of transportation hubs in third tier cities was the highest, at 93.1%, indicating that the model had a stronger ability to regulate complex circulation paths than other regions. The matching degree of sales rhythm was higher than 85% in all regions, with transportation hubs and university areas exceeding 88%, indicating that the model could effectively coordinate market rhythm to adjust the delivery rhythm. In terms of the decrease in unsold rates, transportation hubs in third-tier cities also performed the best, with a decrease of 24.3%. The overall results confirmed the model's dual optimization ability for advertising efficiency and unsold risk in dynamic markets.

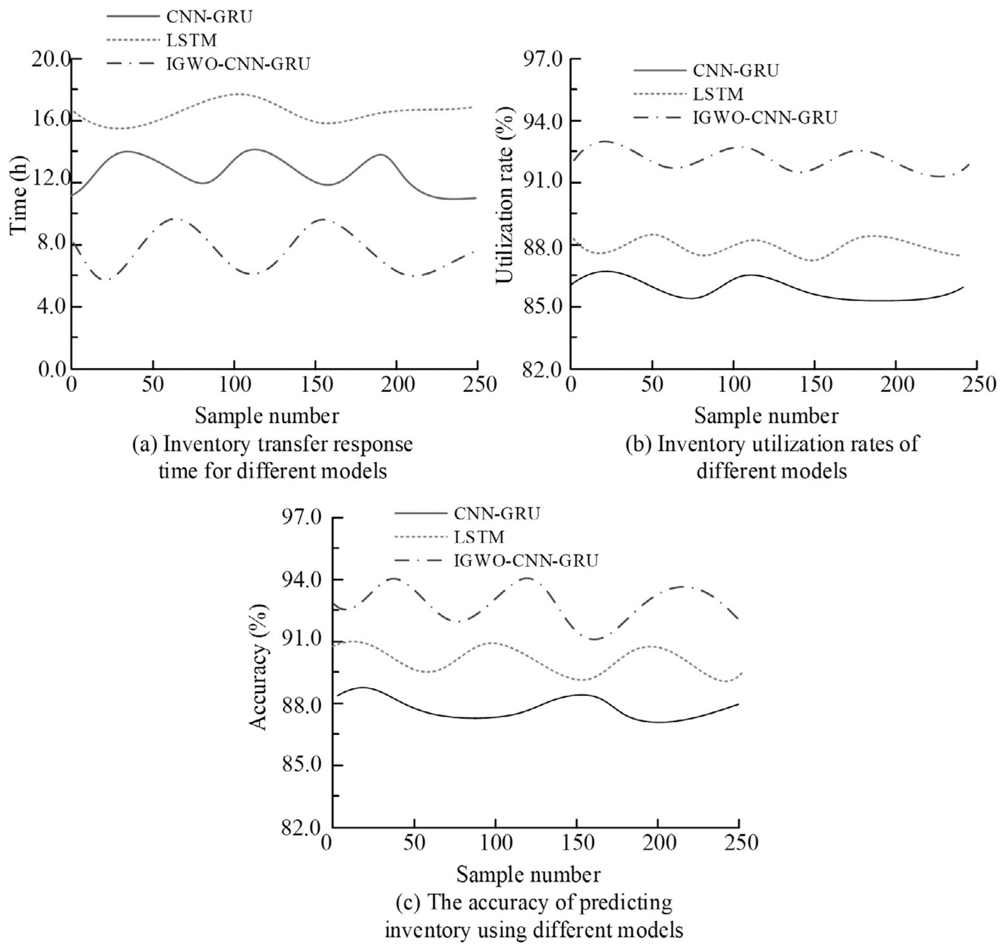


Figure 9. Experimental results of inventory and allocation strategies comparison in the tobacco industry.

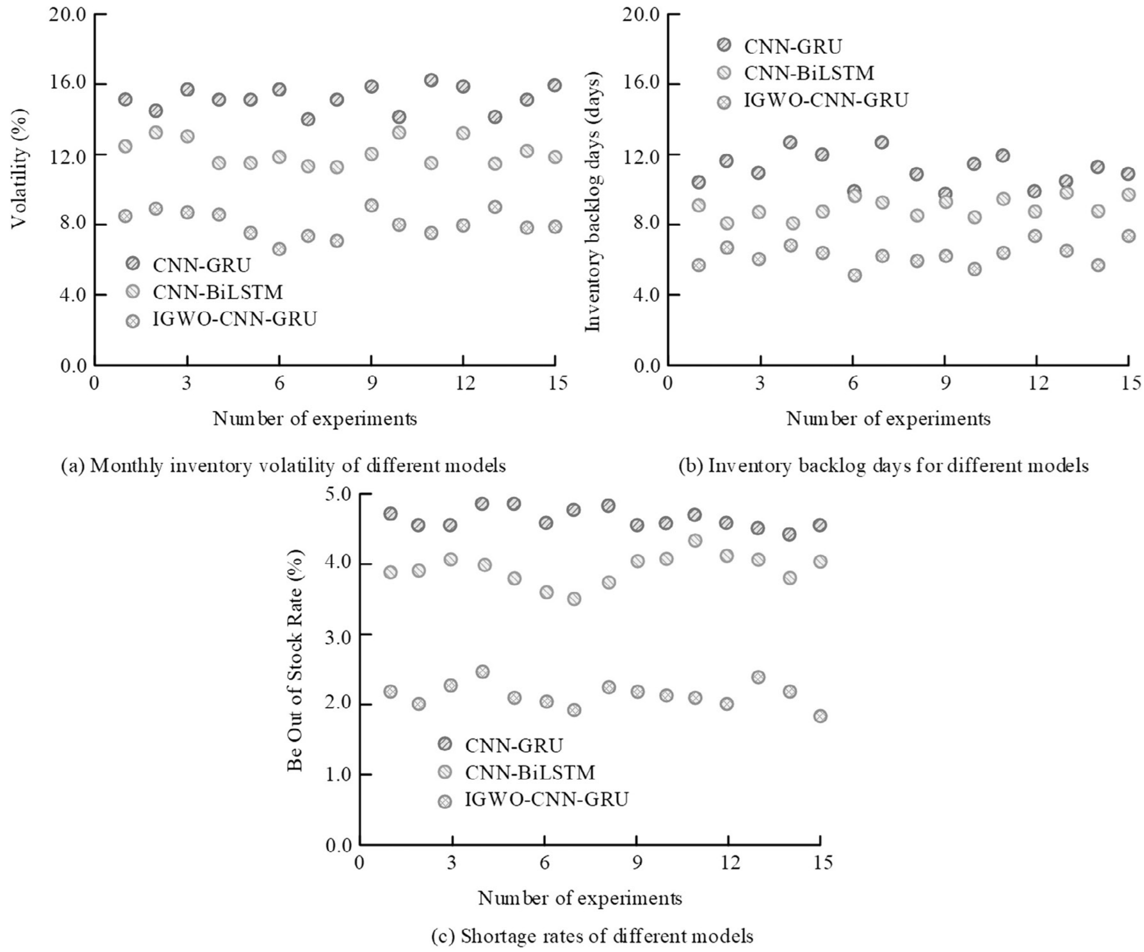


Figure 10. Comparison results of inventory fluctuation indicators.

Table 4. Comparison results of optimization indicators for advertising strategies.

Placement scenario	Average advertising error (days)	Channel advertising coverage rate (%)	Sales rhythm matching degree (%)	Decrease in unsold rate (%)
Core business districts in first tier cities	3.5	91.2	87.4	22.1
Residential concentration areas in second tier cities	4.2	88.5	85.9	19.8
Transportation hub in third tier cities	3.8	93.1	89.2	24.3
Remote county market	4.0	89.7	86.7	20.5
Concentrated areas of universities	3.6	90.6	88.1	21.7

5. Conclusion

In the context of promoting intelligent management in the tobacco industry, accurately predicting the product lifecycle has become the key to improving inventory turnover efficiency and formulating refined deployment strategies. Aiming at the shortcomings of existing methods in multi-source heterogeneous data fusion and time series modeling, an HNN model integrating CNN and GRU was proposed, and IGWO was introduced for hyper-parameter optimization. The experiment outcomes revealed that the model possessed notable superiority in predicting the remaining time of the lifecycle. IGWO improved the fitting ability and stability of the CNN-GRU, with an average MSE reduction of 17.9% and a MAE reduction of 20.1% in multiple rounds of experiments. In tests of different product types, the accuracy of classifying the lifecycle of "fine cigarettes" was 96.2%, and the error in identifying lifecycle boundaries was controlled within 3 days, verifying the high recognition ability of the model for lifecycle turning points. In terms of application experiments, the IGWO-CNN-GRU outperformed the comparison model in regard to inventory allocation response time of 7.4 hours, inventory utilization rate of 92.8%, and predicted inventory accuracy of 93.2%. The experiment outcomes revealed that the IGWO-CNN-GRU performed excellently in predicting the remaining time of the lifecycle. Compared with mainstream models such as GRU, LSTM, and CNN-BiLSTM, the IGWO-CNN-GRU showed significant improvements in prediction error and fitting degree, indicating that the model could more accurately reflect the evolutionary trend of tobacco products at various stages of the lifecycle. However, current research still has certain limitations. The construction of lifecycle labels depended on the alignment of sales rhythm and stage templates, which introduced temporal consistency fluctuations—particularly for products with strong seasonality, sporadic sales spikes, or abrupt marketing interventions. In such cases, lifecycle label deviation may occur, leading to potential misclassification. For instance, limited-edition promotional products released during festivals often show short-lived sales bursts that differ significantly from standard stage curves. Additionally, the model was trained on tobacco industry-specific datasets, and its generalization capability to cross-industry or highly diversified inventory structures has not been systematically validated. Broader evaluation is needed across various enterprise profiles and supply chain conditions. Another key limitation arises in cold-start scenarios, where newly launched products lack historical data. In these cases, the model's performance deteriorates due to insufficient pattern references.

Future work will explore metadata-driven initialization, semi-supervised learning using product embeddings, and adaptive clustering to mitigate early-stage prediction gaps. Moreover, the research will further introduce graph neural networks to capture structural correlations among products, distribution channels, and region-based consumption behaviors. This approach may improve lifecycle inference under sparse or noisy input conditions. In terms of practical implementation, several challenges must also be addressed. Real-time deployment within tobacco enterprise systems may be affected by data latency, particularly in IoT integration across fragmented platforms or under low-bandwidth logistics infrastructure. Sensor data loss, asynchronous updates, or irregular reporting from warehousing and transportation nodes could lead to partial feature drop-out, undermining model accuracy. Additionally, despite the model's lightweight neural structure, computational limitations at edge nodes, such as in production-line controllers or warehouse gateways, may restrict inference speed or parallel deployment. Future work will consider model compression, quantization, and edge-device optimization, ensuring efficient inference with limited hardware resources.

Declaration of Competing Interests

The authors declare no conflict of interest.

Funding

This research was funded by "Research on intelligent selection and operation model of cigarette customers based on three portraits" project.

Data availability

Data used in this study is proprietary.

References

[1] X. Zhang *et al.*, "A Novel Industrial Big Data Fusion Method Based on Q-learning and Cascade Classifier", *Computer Science and Information Systems*, vol. 21, no. 4, pp. 1629–1649, 2024. http://dx.doi.org/10.2298/CSIS240314051Z

[2] M. Hamdi *et al.*, "Deep Recurrent Neural Networks Distributed on a Hadoop/Spark Cluster for Fall Detection", *International Journal of Computers Communications & Control*, vol. 19, no. 3, p. 6428, 2024. http://dx.doi.org/10.15837/ijccc.2024.3.6428

[3] H. Sun, "Optimizing Manufacturing Scheduling with Genetic Algorithm and LSTM Neural Networks", *International Journal of Simulation Modelling*, vol. 22, no. 3, pp. 508–519, 2023. http://dx.doi.org/10.2507/IJSIMM22-3-CO13

[4] N. El Faydy and L. El Abbadi, "Interpretive Structural Modelling of Critical Success Factor for Lean Product Lifecycle Management in Industry 4.0", *International Journal of Production Management and Engineering*, vol. 11, no. 1, pp. 65–72, 2023. http://dx.doi.org/10.4995/ijpme.2023.18840

[5] Z. Li and M. Yao, "Optimized Test Case Design for Automotive Safety Driving Virtual Training System Using Formal Activity Diagrams", *Tehnicki vjesnik – Technical Gazette*, vol. 31, no. 6, pp. 2014–2021, 2024. http://dx.doi.org/10.17559/TV-20240126001293

[6] S. Bahaa *et al.*, "DataOps Lifecycle with a Case Study in Healthcare", *International Journal of Advanced Computer Science and Applications*, vol. 14, no. 1, pp. 136–144, 2023. http://dx.doi.org/10.14569/IJACSA.2023.0140115

[7] T. Wen, "A Multi-feature Fusion Model Based on Long and Short Term Memory Network and Improved Artificial Bee Colony Algorithm for English Text Classification", *Computer Science and Information Systems*, vol. 21, no. 4, pp. 1607–1627, 2024. http://dx.doi.org/10.2298/CSIS240314050W

[8] Z. Liu *et al.*, "A Hybrid CNN and LSTM Based Model for Financial Crisis Prediction", *Tehnicki vjesnik – Technical Gazette*, vol. 31, no. 1, pp. 185–192, 2024. http://dx.doi.org/10.17559/TV-20230307000415

[9] A. Y. Liu *et al.*, "Deep Learning for Intelligent Production Scheduling Optimization", *International Journal of Simulation Modelling*, vol. 23, no. 1, pp. 172–183, 2023. http://dx.doi.org/10.2507/IJSIMM23-1-CO4

[10] P. Raska *et al.*, "Using Adaptive Neural Networks for Optimising Discrete Event Simulation", *International Journal of Simulation Modelling*, vol. 23, no. 2, pp. 227–238, 2024. http://dx.doi.org/10.2507/IJSIMM23-2-678

[11] A. Gligor *et al.*, "Augmented Cyber-Physical Model for Real-Time Smart-Grid Co-Simulation", *International Journal of Computers Communications & Control*, vol. 20, no. 1, p. 6914, 2025. http://dx.doi.org/10.15837/ijccc.2025.1.6914

[12] S. Li *et al.*, "Factoring Electrochemical and Full-Lifecycle Aging Modes of Battery Participating in Energy and Transportation Systems", *IEEE Transactions on Smart Grid*, vol. 15, no. 5, pp. 4932–4945, 2024. http://dx.doi.org/10.1109/TSG.2024.3402548

[13] G. A. Montoya *et al.*, "A Mathematical Optimization Approach for Prioritized Services in IoT Networks for Energy-Constrained Smart Cities", *International Journal of Computers Communications & Control*, vol. 20, no. 1, p. 6912, 2025. http://dx.doi.org/10.15837/ijccc.2025.1.6912

[14] J. Yu *et al.*, "News Recommendation Model Based on Encoder Graph Neural Network and BAT Optimization in Online Social Multimedia Art Education", *Computer Science and Information Systems*, vol. 21, no. 3, pp. 989–1012, 2024. http://dx.doi.org/10.2298/CSIS231225025Y

[15] V. N. D. Almeida *et al.*, "Knowledge Transfer in Multi-objective Multi-agent Reinforcement Learning via Generalized Policy Improvement", *Computer Science and Information Systems*, vol. 21, no. 1, pp. 335–362, 2024. http://dx.doi.org/10.2298/CSIS221210071A

[16] M. Motiejauskas and G. Dzemyda, "EfficientNet Convolutional Neural Network with GramMatrices Modules for Predicting Sadness Emotion", *International Journal of Computers Communications & Control*, vol. 19, no. 5, p. 6697, 2024. http://dx.doi.org/10.15837/ijccc.2024.5.6697

[17] L. N. Han *et al.*, "Balancing Material Supply-Demand with ARIMA and Neural Networks", *International Journal of Simulation Modelling*, vol. 22, no. 4, pp. 712–722, 2023. http://dx.doi.org/10.2507/IJSIMM22-4-CO18

[18] X. Yang *et al.*, "Vehicle Driving Behavior Recognition and Optimization Strategies Based on Cloud Computing and SSA-BP Algorithm", *Studies in Informatics and Control*, vol. 33, no. 3, pp. 17–28, 2024. http://dx.doi.org/10.24846/v33i3y202402

[19] S. P. and S. P. Devi, "A Sporadic Classification and Regression-Based Approach to Intermittent Demand Forecasting in Smart Supply Chain", *Tehnicki vjesnik – Technical Gazette*, vol. 31, no. 5, pp. 1544–1552, 2024.
http://dx.doi.org/10.17559/TV-20240305001374

[20] M. S. Unluturk *et al.*, "Convolutional Neural Network for Cotton Yield Estimation", *Studies in Informatics and Control*, vol. 33, no. 2, pp. 109–117, 2024.
http://dx.doi.org/10.24846/v33i2y202410

[21] J. F. Tsai *et al.*, "Performance Evaluation of Logistics Service Providers Under Uncertain Environment Using a rDANP-U Model", *Studies in Informatics and Control*, vol. 33, no. 3, pp. 61–71, 2024.
http://dx.doi.org/10.24846/v33i3y202406

[22] T. Xia *et al.*, "An Electric Vehicle Charging Control System Using LSTM Encoding-GRU Decoding", *Tehnicki vjesnik – Technical Gazette*, vol. 31, no. 5, pp. 1530–1535, 2024.
http://dx.doi.org/10.17559/TV-20230708000789

[23] S. Canchila *et al.*, "Natural Language Processing: An Overview of Models, Transformers and Applied Practices", *Computer Science and Information Systems*, vol. 21, no. 3, pp. 1097–1145, 2024.
http://dx.doi.org/10.2298/CSIS230217031C

[24] M. Pekarcikova *et al.*, "Inventory Management Supported by Tecnomatix Plant Simulation Tool", *International Journal of Simulation Modelling*, vol. 23, no. 2, pp. 251–262, 2024.
http://dx.doi.org/10.2507/IJSIMM23-2-682

[25] A. S. Akopov, "A Clustering-Based Hybrid Particle Swarm Optimization Algorithm for Solving a Multisectoral Agent-Based Model", *Studies in Informatics and Control*, vol. 33, no. 2, pp. 83–95, 2024.
http://dx.doi.org/10.24846/v33i2y202408

[26] H. Zhang *et al.*, "A Dynamic Credit Evaluation Approach Using Sensitivity-Optimized Weights for Supply Chain Finance", *Tehnicki vjesnik – Technical Gazette*, vol. 30, no. 6, pp. 1951–1958, 2023.
http://dx.doi.org/10.17559/TV-20230801000841

[27] K. Cho *et al.*, "On the Properties of Neural Machine Translation: Encoder-Decoder Approaches", in *Proceedings of the 8th Workshop on Syntax, Semantics and Structure in Statistical Translation (SSST-8), Doha, Qatar*, 2014, pp. 103–111.
http://dx.doi.org/10.3115/v1/W14-4012

[28] Y. LeCun *et al.*, "Gradient-Based Learning Applied to Document Recognition", *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.
http://dx.doi.org/10.1109/5.726791

[29] S. Mirjalili *et al.*, "Grey Wolf Optimizer", *Advances in Engineering Software*, vol. 69, pp. 46–61, 2014.
http://dx.doi.org/10.1016/j.advengsoft.2013.12.007

Received: June 2025
Revised: July 2025
Accepted: August 2025

Contact addresses:
YiShuang Qin*
Guangxi China Tobacco Industry Co., LTD.
Nanning
China
e-mail: 18677093066@163.com
*Corresponding author

JianShuang Jia
Guangxi China Tobacco Industry Co., LTD.
Nanning
China
e-mail: xxqin517@163.com

Dong Liang
Guangxi China Tobacco Industry Co., LTD.
Nanning
China
e-mail: aloysiousliang@gmail.com

Che Yu
Beijing Bozhi Shuyuan Technology Co., LTD.
Beijing
China
e-mail: db_sours@163.com

YiShuang Qin obtained a master's degree from Business Administration Institute at Guangxi University School of Business in 2012. Currently, she is employed at Guangxi China Tobacco Industrial Co., Ltd. Her main research directions include brands, products and markets.

JianShuang Jia graduated from China Central Radio and Television University, Hubei Provincial Federation of Trade Unions Branch. Currently, he is employed at Guangxi China Tobacco Industrial Co., Ltd. He works in manufacturing. His main research direction are markets.

Dong Liang graduated from Beihang University (Beijing University of Aeronautics and Astronautics) with a master's degree in 2013. Currently, he is employed at Guangxi China Tobacco Industrial Co., Ltd., where he specializes in information technology and artificial intelligence-related initiatives. His primary research focuses on artificial intelligence and data mining.

Che Yu obtained a doctor's degree from the University of Science and Technology Beijing in 1999. He is currently working at Beijing Bozhi Shuyuan Technology Co., Ltd. His main research direction is big data model algorithm research and brand marketing.