International Journal for Housing Science and Its Applications
Publish August 10, 2025. Volume 46, Issue 4 Pages 668-679

https://doi.org/10.70517/ijhsa46461

Quality Stability Prediction and Scheduling Optimization of Cigarette Production Process Empowered by Deep Learning Algorithms

Weihua Wang[1], Yunhai Gu[1], Fei Xia[1], Yang Lv[1,*] and Hengjiang Liu[1]
[1] Information Department, Hongta Group Chuxiong Cigarette Factory, Chuxiong, Yunnan, 675000, China
Corresponding authors: (e-mail: 13320568532@163.com).

Abstract Traditional quality management methods have problems such as insufficient prediction accuracy and slow response speed when facing complex production environments, which are difficult to meet the demand for refined management of modern production. In this study, the quality stability prediction model of cigarette production process based on Meta-DQN is constructed, which solves the deficiencies of traditional methods in small sample learning and environmental adaptability. 500 samples of working condition data such as temperature, humidity, airflow speed, etc. were collected through the production line sensors in cigarette factory H. The MQTT communication protocol was used for data transmission, and MinMaxScaler normalization was applied to ensure data consistency. The Meta-DQN prediction model is constructed by combining the Meta Reinforcement Learning MAML algorithm with the deep Q network, and the fast adaptation to different production tasks is achieved through the two-layer loop optimization mechanism. The experimental results show that the R^2 coefficients of determination of the training and test sets reach 0.992 and 0.957, respectively, and the model prediction accuracy is significantly improved. In the key parameters prediction validation, the average deviation of the predicted values of six process parameters from the real values is only 0.212, which is much lower than the standard setting of 0.565. Comparison experiments show that the Meta-DQN model can quickly converge in the initial operation stage of the equipment, effectively reducing the scrap rate, which is significantly better than that of the pure DQN algorithm. The method provides an efficient and intelligent solution for cigarette production quality management and has important engineering application value.

Index Terms Deep learning algorithm, cigarette production process, quality stability prediction, Meta-DQN, meta-reinforcement learning, production scheduling optimization

I. Introduction

With the development of information technology, the control of the quality of industrial products requires more and more, in order to meet the market demand, it is necessary to improve the quality of products. Quality is fundamental to the survival and development of all enterprises, and is also an important source of strength for enterprise development [1]. Continuous improvement of product quality can promote the development of enterprises, so that enterprises in the market competition in a favorable position [2]. For the production process has been using advanced automated control systems for cigarette machines, in order to improve the quality of the products produced by the cigarette machine as well as production efficiency, it is necessary to strictly control the production process of cigarette machines [3]-[5].

Traditional quality control methods, in the face of the complexity of the modern market and production environment gradually revealed bottlenecks. The traditional manual sampling method is limited by human resources and subjective judgment, and it is difficult to realize the comprehensive monitoring of the production process [6]. In addition, although rule-based automated systems can improve production efficiency, they lack sufficient adaptability to cope with rapid changes in the market and abnormalities in the production process, making the timeliness, accuracy and efficiency of quality control challenged [7]-[9]. Based on this, stability prediction and scheduling optimization in the cigarette production process is an innovative solution to meet the growing market demand and quality management challenges of the tobacco industry worldwide [10]-[12]. By introducing modern artificial intelligence technologies such as deep learning, cigarette manufacturers are able to realize real-time monitoring and intelligent adjustment of the whole production process [13], [14]. Deep learning algorithms can learn and discover potential patterns from a large amount of historical data to realize intelligent support for quality control [15], [16]. And deep learning technology, with its advantages in dealing with complex nonlinear problems, provides new ideas for dynamic adjustment of quality fluctuations in cigarette production and rapid learning and optimization of production process parameters [17]-[19].

Modern manufacturing industry is experiencing a profound change from the traditional production mode to the transformation of intelligent manufacturing, quality management as an important part of the core competitiveness of enterprises, its management level directly affects the market position and economic benefits of enterprises. Cigarette industry as an important branch of the traditional manufacturing industry, its production process involves silk, rolling, packaging and other complex processes, the quality control of each link has a decisive impact on the quality of the final product. In the current market environment where consumers' demand for product quality is constantly improving, how to realize accurate quality control of the production process has become a major challenge for cigarette companies. Traditional quality management methods mainly rely on manual empirical judgment and statistical process control, which has obvious deficiencies in dealing with complex production systems with multivariate coupling and obvious nonlinear characteristics, and it is difficult to realize accurate prediction and timely response to quality change trends. The rapid development of artificial intelligence technology, especially deep learning algorithms, provides new technical means to solve this problem. Deep learning has powerful feature extraction and pattern recognition capabilities, and is able to mine the hidden quality laws from a large amount of production data to achieve intelligent monitoring and prediction of complex production processes. However, traditional deep learning methods often suffer from poor adaptability and low learning efficiency when facing frequent changes in the production environment and limited sample data, which limits their application effect in actual industrial production.

Based on the above analysis, this study proposes a Meta-DQN model that combines meta-reinforcement learning with deep Q-networks for quality stability prediction in cigarette production process. The study first establishes a production data acquisition system based on sensor network to obtain key working condition parameters in the production process in real time; then builds a Meta-DQN prediction model and optimizes the initial parameters of the network through the MAML algorithm, so that the model can quickly adapt to different production tasks and environmental changes; finally verifies the prediction accuracy and practicability of the model through the actual production data to provide technical support for the cigarette enterprises to achieve intelligent quality management for cigarette enterprises to realize intelligent quality management to provide technical support.

II. Cigarette production process quality management theory foundation

The production site is a place for material conversion and configuration of people, machines and materials in the manufacturing process, which is the core of the cigarette enterprise. The good or bad management of the production site is directly related to the quality of the products and is the source of profit of the cigarette enterprise. Production site management optimization is the process of using advanced management theories to continuously improve site management methods and approaches, with the aim of improving product quality, increasing production efficiency, reducing production costs, and guaranteeing on-site production safety. Therefore, production site management optimization is the best way to solve the production site management problems of cigarette enterprises.

II. A. Total Quality Management and PDCA Theory

II. A. 1) Total Quality Management Theory

Total Quality Management (TQM) is a way of doing business in which quality is at the core of an organization, based on the participation of all employees, with the goal of satisfying customers, all employees and the community as a whole, and achieving long-term business growth. Some researchers have suggested that TQM refers to carrying out production and providing services at the most economical level, while ensuring full compliance with customer needs, so that all parts of the organization form an organic system in developing quality, maintaining quality and improving quality [20].

Total quality management has three core features, namely, the participation of all employees in quality management, the whole quality management and the quality control of the whole product. All staff participation in quality management is that all staff, from senior managers to general staff, and even grass-roots staff, to participate in quality improvement, an important principle of TQM is to participate in the “improvement of the quality of the work of the core institutions. TQM is based on strict control in market research, product selection, research and testing, design, raw material procurement, production, inspection, storage, sales and other aspects. The production process of the company is analyzed from the procurement of raw materials, production and inspection, and corresponding improvement measures are proposed. The quality of the products, on the other hand, is tested and confirmed through sales and after-sales service.

II. A. 2) PDCA cycle theory

PDCA cycle, was proposed by the American expert Deming, PDCA cycle is widely used in all kinds of work. It mainly requires the work to be divided into four steps, including planning, implementation, checking and processing, and then summarize and learn from the successful experience in this improvement cycle, and improve the unsuccessful or fail to achieve the goal of the part into the next PDCA cycle to solve [21].

PDCA improvement cycle of a total of four stages, the first stage is P (plan), that is, to be based on the organization's goals or customer needs to develop work plans and related measures. The second stage is D (implementation), that is, the implementation of the work program that has been developed, this stage in accordance with the plan to carry out work to grasp the progress of the work plan. The third stage is C (check) stage, that is, according to the effect of the implementation of improvement measures and the development of the implementation of the plan, the implementation of the plan and the effect of checking, in this stage should be to the implementation of the process found in a timely manner to deal with the problem, including problem solving and implementation of the experience of the summary and so on. The fourth stage is the A (corrective) stage, which is to consolidate the results and summarize the experience based on the results and experience gained from carrying out the previous three stages of the work, and to import the unresolved problems in the next PDCA cycle, in which they will be resolved.

II. B. Cigarette production process quality improvement design

II. B. 1) Optimization of production quality management

On the basis of the theory of total quality management, the production process quality control of cigarette enterprises is the key to product quality. The core of production quality management lies in process management, process management can be effectively integrated and combined, can make the original plan to achieve the desired results. Total quality management theory-oriented cigarette enterprise production process as shown in Figure 1, can put the focus of our work on efficiency, in providing customers and other parties with confidence about the stability of organizational performance. Costs and cycle times are reduced through efficient use of resources. Various process analyses allow focused and prioritized ownership of improvement opportunities, clarification of responsibilities, and identification of metrics to measure the process to facilitate meeting customer requirements and increasing customer satisfaction.

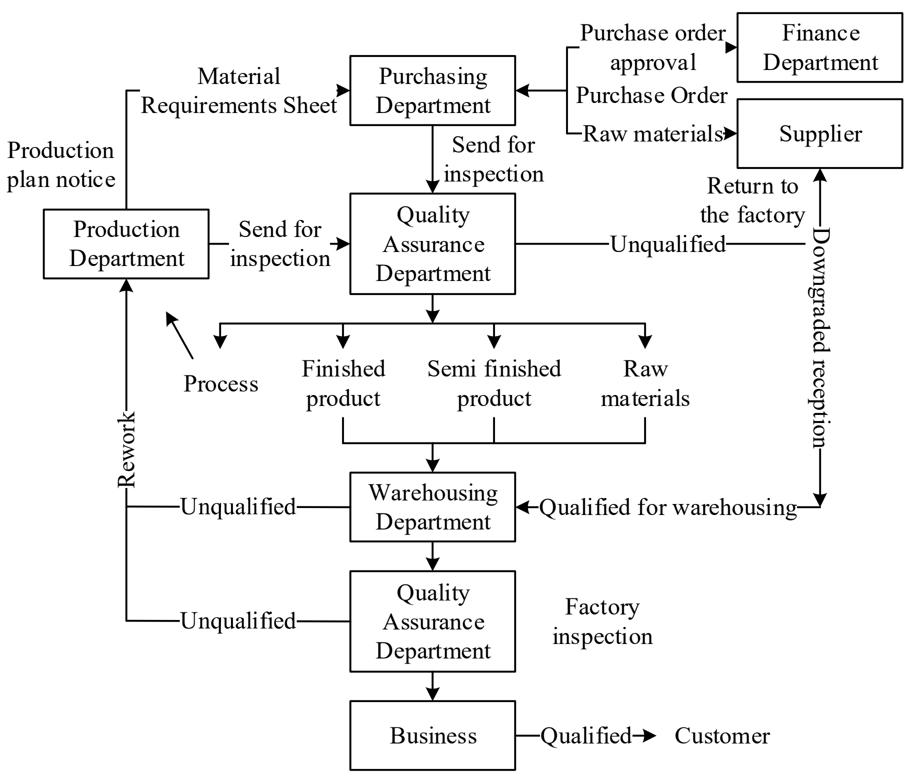


Figure 1: Flow chart of production process

The quality management of the production process of cigarette enterprises mainly involves the R&D Department, the Purchasing Department, the Quality Control Department, the Production Department, the Warehousing Department and other aspects. The use of process approach management for each organizational activity of the process of quality management program design. Production can best establish a set of standardized operating procedures, step by step, effectively improve product quality management.

II. B. 2) Combined with PDCA standardization

In the process of perfecting the enterprise cigarette production standardization, first of all, the enterprise standardization management system should be clarified, led by the Enterprise Management Department, as the management of the enterprise standardization of the main functional departments, is responsible for the management of the development of the enterprise's technical standards, work regulations, and the implementation of standard documents responsible for the supervision and inspection. In the process of improving the level of cigarette production standardization of cigarette factory enterprises, the overall idea can be in accordance with the PDCA way of working, in order to ensure the completeness and operability of the cigarette factory enterprise standards.

First, the preparation of draft standard operating instructions. The standard operating instructions are jointly prepared by enterprise quality management personnel, field managers and field operation technicians.

The second is to develop a standard process sequence table. Standard process flow is safe, fast and accurate to create a key foundation for quality products, through the development of standardized process sequence table, a clear order of operation, to prevent process errors lead to in-process quality problems.

Thirdly, establish the applicability of production operation instruction and process sequence table. Front-line personnel carry out on-site production according to the workflow specified in the established production standardized operating instructions and process sequence table, and problems occurring in the production process should also be reflected in the checklist in a timely manner.

Fourth, continuous improvement and refinement. For the situations appearing in the checklist of the third step, professional technical analysis meetings can be organized to strive for the best solutions. For the problems mentioned in the previous step, some experienced frontline operators suggest that special time products can be inspected separately, and the possible problems should be listed in detail. Through continuous improvement, the standardized operation manual is improved, and the guidance, operability and adaptability are greatly enhanced. Only in this way can standardized operations be promoted throughout the workshop.

To sum up, through the standardized operation improvement, not only can it enhance the manufacturing efficiency and product quality management level of the whole production site, but also create a strong guarantee for the rapid entry of new workers and technical upgrading, which in turn ensures the production continuity, efficiency and stability of the production site of the cigarette factory.

III. Cigarette production process quality stability prediction model

With the deepening of economic globalization, the competition in the cigarette market is intensifying. Whether cigarette enterprises can be in a favorable position in the competition, product quality has become a key link that can not be avoided. The concept of high-quality development has put forward new requirements for cigarette enterprises. Cigarette factory enterprises can only further enhance the market competitiveness of cigarette enterprises by improving the product quality management ability in the production process and reasonably introducing deep learning algorithms to assist in realizing the quality stability prediction of the cigarette production process.

III. A. Cigarette production data collection and processing

III. A. 1) Cigarette production data collection

In the production workshop of the cigarette factory, there are many different types of production machines, and the changes of each machine in the production process have a greater impact on the quality of cigarette production. This paper is based on a variety of sensors to collect data on the working conditions in the cigarette production process, and its specific process is shown in Figure 2.

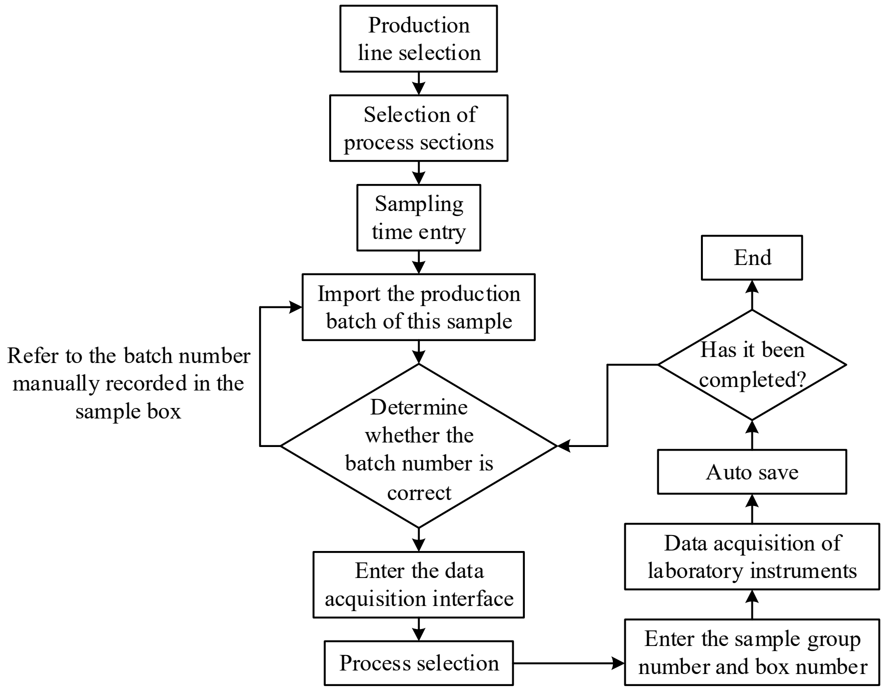


Figure 2: Data collection for cigarette production

The data acquisition module realizes the attribution of inspection items by batch. Usually, the production batch of a process section is unique, and each inspection item, such as moisture content, temperature, etc., is targeted at a specific process and a collection point. Therefore, the database table for batch attribution contains information such as batch number, production line, production brand, process section, process and inspection items, etc. The batch number is unique in the production process and can be used as the main identification code for the whole cigarette production process. Each inspection item is uniquely identified by batch number, production line, process section and process, and all inspection items of any batch can be associated, stored and managed through the batch, which is conducive to the analysis and comprehensive determination of cigarette production quality.

III. A. 2) Cigarette production data processing

This paper collects the temperature, humidity, and airflow speed characteristics of each working condition in the cigarette production process through various types of sensors in the production line of H Cigarette Factory, accesses the data collection microservice using the MQTT communication protocol, and selects 500 process samples, with a data collection frequency of once every 10 minutes.

Notify the processing micro-service through the message queue mechanism to normalize the received cigarette production data to ensure data consistency. The MinMaxScaler normalization method is selected to converge the data to between 0 and 1, and the calculation formula is as follows:

$$X^{\text{å}} = \left[X - \min(X)\right] / \left[\max(X) - \min(X)\right] \tag{1}$$

where $X^{\text{å}}$ is the normalized feature matrix, X is the feature matrix, and $\max(X), \min(X)$ are the maximum and minimum values of the feature matrix, respectively.

In this paper, we choose random sampling to select 80% of the total 400 as the training set, and the remaining 20% of the total 100 as the test set.

In this paper, the coefficient of determination, root mean square error, average absolute root mean square error are used to assess the quality of cigarette process, and the relevant formula is:

$$R^2 = 1 - \sum_{i=1}^{N}(y_i - y^*)^2 / \sum_{i=1}^{N}(y_i - \tilde{y}_i)^2 \tag{2}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \tilde{y}_i)^2} \tag{3}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N} | y_i - \tilde{y}_i |^2 \tag{4}$$

where y_i is the actual score value and $\tilde{y}_i$ is the predicted sample score value.

III. B. Meta-DQN based stability prediction

III. B. 1) Deep Q-network model

The DQN algorithm is trained and updated based on a Markov decision process, which consists of a state space, an action space, a reward function and a state transfer probability. In this process, the intelligent body interacts with the environment, observes the current state, takes an action, and receives a reward.The main structure of the DQN algorithm is two neural networks and a pool of experience, one is the goal network and the other is the action network. In the DQN algorithm, the intelligent uses the action network to select an action and also uses the goal network to compute the goal Q-value. The target Q-value is used to update the weights of the action network to improve the accuracy of the Q-value estimation. The main reason for employing the target network is that if the same neural network is used to compute both the Q-value and the target Q-value, it will result in the estimation of the Q-value and the computation process of the target Q-value interacting with each other, which in turn may lead to algorithm divergence. By using a target network, this influence can be effectively reduced and the convergence and stability of the algorithm can be improved [22].

(1) At each time step t, the intelligent body interacts with the environment to obtain a quaternion $(s, a, r, s^{'})$ of states, actions, rewards, and next states. These quaternions are stored in an experience playback pool for subsequent training. The empirical playback pool avoids correlation and bias in the data and increases the efficiency of data utilization.

(2) A batch of data is randomly selected from the experience playback pool for training. Each quaternion $(s, a, r, s^{'})$ has a corresponding target Q value, which is calculated by the target network. The formula for the target Q-value is as follows:

$$Q(s, a) = r + \gamma * \max(Q^{'}(s^{'}, a^{'})) \tag{5}$$

where r is the reward, γ is the discount factor, $Q^{'}$ is the target network, and $\max(Q^{'}(s^{'}, a^{'}))$ is the maximum Q value of all actions feasible under the next state $s^{'}$ of the maximum Q value, and the target Q value represents the cumulative payoff of performing the action a in state s.

(3) Calculate the loss function, which is calculated using the mean square error loss function. The formula for calculating the loss function is as follows:

$$L(\theta) = 1 / N * \Sigma(Q(s, a; \theta) - y)^2 \tag{6}$$

where θ is the parameter of the action network, N is the number of samples, $Q(s, a; \theta)$ is the value of Q computed by the action network, and y is the value of the target Q.

(4) Update the parameters of the action network, and by deriving the loss function, the gradient formula can be obtained as:

$$\nabla \theta L(\theta) = 1 / N * \Sigma(\nabla \theta Q(s, a; \theta) * (Q(s, a; \theta) - y)) \tag{7}$$

The parameters of the action network are updated using an optimization algorithm to minimize the loss function. The updated parameters of the action network will be used for the next time the batch data is extracted for training.

(5) Update the parameters of the target network, periodically update the parameters of the target network, and copy the parameters of the action network into the target network. This ensures that the parameters of the target network are more stable, reduces the variance of the Q-value estimation, and improves the stability of the algorithm.

III. B. 2) Meta-reinforcement learning optimization

The model-agnostic meta-learning (MAML) approach is a learning framework that provides meta-learners that are used to train a variety of base learners.The basic idea of MAML is to find better initial parameters of the neural network, such that with good initial parameters, the model is able to learn a new task quickly with fewer gradient descent steps. The goal of small-sample meta-learning is to train a model that can quickly adapt to different tasks with only a small amount of sampled data and a small number of iterative training sessions. To accomplish such a goal, the model is trained and learned during the meta-learning process, and the model that continues to be trained on top of the meta-model requires only a very small amount of sampled data and the number of training sessions to be able to adapt quickly to new tasks.

Define the model f as a mapping from the set of states x to the set of actions a. During the meta-learning process, the goal of model training is to quickly adapt to a large number of different tasks.The framework of MAML is able to be applied to all kinds of deep learning problems from supervised to unsupervised learning. Therefore, the learning tasks are conceptualized as follows:

$$\mathrm{T} = \{\mathrm{L}(x_1,a_1,...,x_H,a_H), q(x_1), q(x_{t+1} \mid x_t,a_t), H\} \tag{8}$$

L is the loss function, the distribution $q(x)$ of the initial state set, the transfer distribution $q(x_{t+1} \mid x_t,a_t)$, and the length of the entire task H. In general independent identically distributed supervised learning, $H=1$. The model generates samples of length H and produces the output a_t every elapsed time t, with a loss function $\mathrm{L}(x_1,a_1,...,x_H,a_H) \to R$. This loss function is able to serve as a feedback for that task, defined in the form of a misclassification loss or cost function in a Markov decision process, depending on the specific training task.

In meta-learning, the distribution on the task $p(\mathrm{T})$ is considered and the goal of the model is to adapt to such a distribution. For meta-training, tasks T_i are extracted from $p(\mathrm{T})$, the model is trained using the extracted samples, and the sample tasks T_i are used to feed back the corresponding loss $\mathrm{L}_{\mathrm{T}i}$, and use another batch of sample tasks T_i''s for error testing. At the end of meta-training, brand new tasks are drawn from $p(\mathrm{T})$ samples, and the performance of meta-learning is weighed by the performance of the model after learning from K samples.

III. B. 3) Meta-DQN models

DQN is a discrete control-oriented algorithm, but there are still drawbacks of overfitting to the environment and long training time of the algorithm, in order to improve the performance of the algorithm as well as quickly learn the optimal strategy, the idea of meta-learning is introduced into the DQN algorithm, and the Meta-DQN model is proposed for the prediction of the quality stability of the cigarette production process, and its specific framework is shown in Fig. 3. The Meta-DQN algorithm is based on the MAML algorithm, which was initially proposed to solve the learning problem of small samples. The Meta-DQN algorithm refers to the idea of MAML algorithm, which was proposed to solve the learning problem of small samples at the beginning, and this algorithm is also suitable for reinforcement learning, which can solve the ineffective problem of the initial learning of DQN algorithm and make its learning more efficient and accurate.

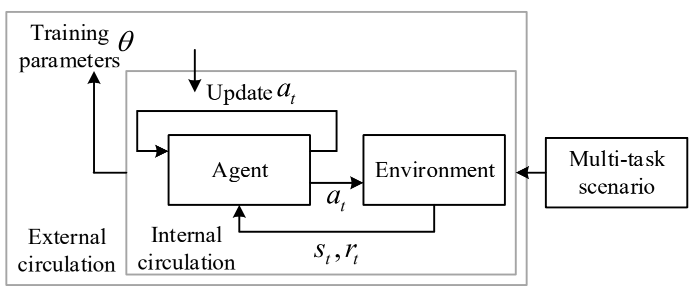


Figure 3: Meta-DQN model architecture

The advantage of the MAML algorithm is that it can be easily combined with reinforcement learning, neural networks and other algorithmic structures to learn, and various loss functions can be used to optimize and refine the algorithmic model, which improves the algorithmic accuracy, and avoids the difficult problem of parameter expansion and structural change of the model, and greatly improves the algorithmic learning efficiency, compared to the updating function and learning rule methods previously used for refining the model. For the DQN algorithm, the reward function is set by the task, the evaluation of the strategy is carried out by setting Target Networks, and the parameter gradient is obtained by the derivation of the difference between the current network and the output of the target network.

According to the algorithmic diagram, it can be seen that in each reinforcement learning problem there exists an initial state distribution $P(T_i)$, applying meta-learning to reinforcement learning, the loss function of each task corresponds to the reward function of reinforcement learning, and the model f_θ of learning for the strategy of reinforcement learning π, the loss function of the task τ_i and the model f_θ is:

$$L_{T_i}(f_\theta) = -E_{x_t,a_t - f_\theta, a_{\tau_i}}\left[\sum_{t=1}^{H} R_i(x_t,a_t)\right] = \sum_{n=1}^{N} l^n(f_{e_i}) \tag{9}$$

After the loss function is determined, the network parameters θ are updated using the policy gradient algorithm, and each additional update step during adaptive learning of the network function requires a new sample from the current policy $f_{\theta_i'}$ for training, and then the model parameters are updated again using gradient descent in the inner loop, i.e.:

$$\theta_i' = \theta - \alpha \nabla_\theta L_{T_i}(f_\theta) \tag{10}$$

The structure of the algorithm is divided into two stages, the first stage uses the MAML algorithm to obtain a set of initialized parameters applicable to a variety of similar tasks, so that the model has an initial strategy, and then the use of deep reinforcement learning targeted learning training for the new environment task, so that the model can quickly learn the optimal strategy to accelerate the convergence speed of the model.

The initialized model parameters θ are used in the outer loop to update the parameters for a certain learning task after one learning to get the parameters applicable to this new task as:

$$\theta^n = \theta - \beta \nabla_\theta \sum l_{T_i}(f_{\theta_i}) \tag{11}$$

where $\nabla_\theta L(f_\theta)$ is calculated as:

$$\nabla_\theta L(f_\theta) = \nabla_\theta \sum_{n=1}^{N} l^n(f_{\theta_i}) = \sum_{n=1}^{N} \nabla_\theta l^n(f_{\theta_i}) \tag{12}$$

The Agent's dynamic model of the environment varies with the model parameters θ, and its posterior knowledge distribution varies with the sampling trajectory $D=(x_1,a_1,...,x_H)$, action a_t, and state s_{t+1} updates. In order to improve the exploration strength of the Agent, it is important to make the actions taken at different times differentiated, i.e., the posterior distribution $p(\theta \mid a_{t+1},s_{t+1})$ is different from the prior distribution $q(\theta \mid a_t,s_t)$, and the difference of the prior and posterior distributions can be expressed by KL The dispersion is then expressed as:

$$D_{kl}(p(\theta \mid a_{t+1},s_{t+1}) \| q(\theta \mid a_t,s_t) = \sum_{i=1}^{N} p(\theta \mid a_{t+1},s_{t+1}) \log \frac{p(\theta \mid a_{t+1},s_{t+1})}{q(\theta \mid a_t,s_t)} \tag{13}$$

First introduce the cross entropy formula for discrete random variables as:

$$H(p,q) = -\sum_{i=1}^{n} p(x_i)\log(q(x_i)) \tag{14}$$

The specific KL scatter inference process is as follows:

$$\begin{aligned} D_{kl}(p \| q) &= \sum_{i=1}^{n} p(x_i)\log(p(x_i)) - \sum_{i=1}^{n} p(x_i)\log(q(x_i)) \\ &= -H(X) + \left[-\sum_{i=1}^{n} p(x_i)\log(q(x_i))\right] \\ &= \left[-\sum_{i=1}^{n} p(x_i)\log(q(x_i))\right] - H(x) \end{aligned} \tag{15}$$

To encourage the Agent to explore more, $D_{kl}(p(\theta \mid a_{t+1},s_{t+1}) \| q(\theta \mid a_t,s_t))$ should be maximized as an internal reward for the learning process, so the Agent's reward function can be set as:

$$R_{t+1} = r(s_1,a_1) + \eta D_{kj}(p(\theta \mid a_{t+1},s_{t+1}) \| q(\theta \mid a_t,s_t)) \tag{16}$$

where η is the learning factor.

IV. Cigarette production process quality stability prediction application

In recent years, with the Internet + Industry 4.0 in full swing, the cigarette industry is more and more aware of the importance of production quality management in the enterprise lean, improve the information technology on the allocation efficiency of industrial production resources, can promote the supply side structural reform at a higher level. Through the data to measure the level of production quality management, and then use the data to force the cigarette production quality management.

IV. A. Validation of integrated quality prediction results

IV. A. 1) Meta-DQN model fitting

The training set and test set error is an important index for evaluating the effect of Meta-DQN model on the quality stability prediction of cigarette production process. Based on the cigarette production data given in the previous section, the training set and test set are trained, and the relationship between the error and the number of iterations is shown in Fig. 4. As can be seen from the figure, when the number of iterations exceeds 3000, the data in the training set and test set reach the optimal result, and cannot continue to converge, which indicates that the structure of the constructed Meta-DQN model used for the prediction of quality stability in the cigarette production process is reasonable.

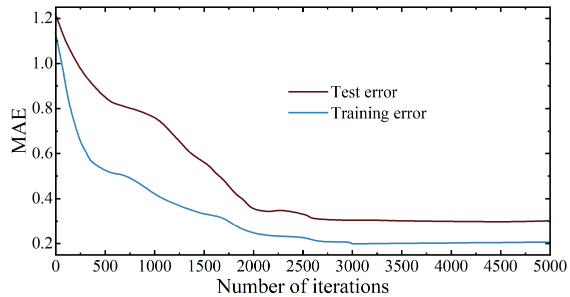
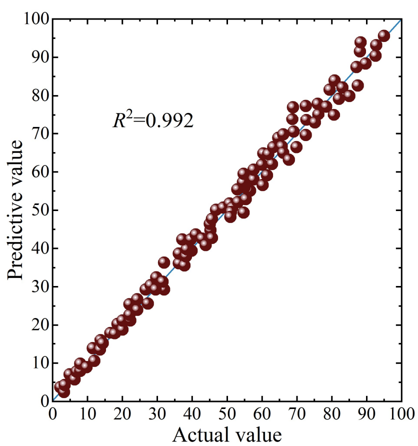
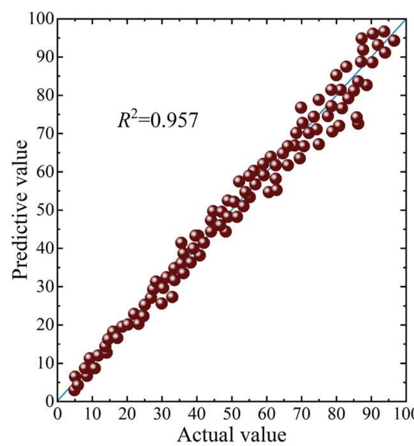


Figure 4: Error curve of Meta-DQN model

On this basis, this paper uses the R2 coefficient to analyze the fitting of the Meta-DQN model, and obtains the fitting relationship between the predicted value and the actual value of the quality stability of the cigarette production process as shown in Fig. 5, in which Fig. 5(a)~(b) shows the fitting results of the training set and test set samples, respectively. From Fig. 5(a), it can be seen that the predicted and actual values of the samples in the training set and test set are distributed near a straight line, and the R2 coefficient of determination of the actual value and the predicted value reaches 0.992 after testing, which indicates that the predicted value and the actual value of the Meta-DQN model for the integrated quality of the cigarette production process are relatively close to each other, and it is good for the prediction of the quality of the cigarette production process. From Fig. 5(b), it can be seen that the predicted and actual values of the test set samples are also distributed near the straight line, and the R2 coefficient of determination of the actual and predicted values reaches 0.957 after the test, which indicates that the intrinsic relationship between the key indexes and the quality of the cigarette production process has strong correlation ability, and it is capable of predicting the comprehensive quality of the cigarette production process.



(a) Training set



(b) Test set

Figure 5: The relationship between the actual value and the predicted value

IV. A. 2) Verification of the quality of cigarette production

Aiming at the data in the cigarette production process as a test sample, the main key parameters are selected as loose retort-discharge moisture content, loose retort-discharge temperature, first-stage charging-discharge moisture content, first-stage charging-discharge temperature, second-stage charging-discharge moisture content, and second-stage charging-discharge temperature, which are denoted by P1~P6. The cigarette filament making process standards are denoted as standard L1 and standard L2, which correspond to the above six key parameters. The cigarette production quality data based on standard L1 is used as a training sample to construct a network model to analyze the relationship between adjustable parameters and key parameters. The cigarette production quality data based on standard L2 is used as a test sample to validate the network model and determine the usability of the network model. Cigarette production quality data based on standard L2 production is used as a test sample to compare and analyze the predicted and true values of each key parameter in the test sample to verify the accuracy and adaptability of the network model, and the specific results are shown in Table 1.

The results are shown in Table 1. It can be seen that the deviation between the predicted values and the real values of each key parameter of the Meta-DQN model for the cigarette production process is within 1.0, with an average deviation of 0.212, which is lower than the average deviation of the center value of the setting interval of the corresponding key parameter of the criterion L2 and the real value of 0.565, indicating that the Meta-DQN model has a higher accuracy in the prediction of the stability of the quality of the cigarette production process. It shows that the Meta-DQN model has high accuracy in predicting the quality stability of the cigarette production process, and it can better describe the relationship between the process parameters and the quality indexes of the cigarette production process.

Table 1: Comparison between the real values and values predicted

Index	L2	Measured result		Prediction result		L2 Error	M-P
		Means	SD	Means	SD		
P1/%	17.2±1.9	16.581	0.815	16.815	1.521	0.619	0.234
P2/℃	56.9±3.4	56.492	1.072	56.238	1.437	0.408	0.254
P3/%	20.2±1.3	20.075	0.518	20.029	0.599	0.125	0.046
P4/℃	48.1±3.5	47.364	7.232	47.563	0.456	0.736	0.199
P5/%	21.6±1.7	21.027	0.751	21.279	1.796	0.573	0.252
P6/℃	53.8±2.8	52.869	0.939	53.154	1.748	0.931	0.285
Means	-	-	-	-	-	0.565	0.212

IV. B. Stability analysis of cigarette production process

IV. B. 1) Optimization of mass stability prediction

In the cigarette production process, the production equipment often appear brand and production batch switching, as well as equipment start shift initial production state changes, at this time the DQN deep learning model because of the lack of sample size, in the case of small samples, its parameter prediction deviation is large, in the industrial control process can not meet the requirements of the product manufacturing process, easy to produce a short period of time the high rate of cigarette rejects, affecting the production cost of the situation. Based on this, this paper introduces the MAML algorithm to optimize the DQN model, and proposes the Meta-DQN model for optimizing the prediction of quality stability in the cigarette production process.

The MAML algorithm is used as a predecessor model optimization, so as to solve the cigarette production process parameter adjustment and optimization method due to the large initial deviation caused by the DQN for the small-sample learning process, and we designed a comparative experiment. The performance of DQN algorithm alone and Meta-DQN model in the initial operation stage of the equipment. Figure 6 shows the results of the comparison of the optimization of the quality stability prediction of the cigarette production process.

The data plots show that after the introduction of the MAML approach, the system is able to quickly adapt to the new production conditions and achieve better performance on the new tasks with a small number of fine-tuning steps. This proves the effectiveness of MAML in parameter tuning and optimization in the cigarette production process. The DQN model incorporating MAML meta-learning is better able to perform stable weight control output during the initial operation stage of the equipment, and can effectively adjust the relevant parameters in the SRM weight control system to achieve stable cigarette weight and reduce the scrap rate. While relying only on the DQN learning algorithm in the initial operation stage of the device due to the learning content of the small sample, the greedy strategy is used in the selection of the execution of the action, and its convergence is slower, resulting in a large fluctuation in the weight control during the period, which greatly exceeds the standard of industrial production cost control. At the same time, the data also show the adaptability of the model in the face of multiple parameter changes and the application effect in actual production, further verifying the feasibility of the method. It should be noted that the introduction of MAML also needs to take into account the uncertainty in the actual production environment. Whether the parameters learned in the training process of the model can be well generalized to the actual production needs to be verified and tested. In addition, to better cope with the uncertainty in production, MAML can be combined with other methods, such as reinforcement learning or model predictive control, to improve the robustness of the system.

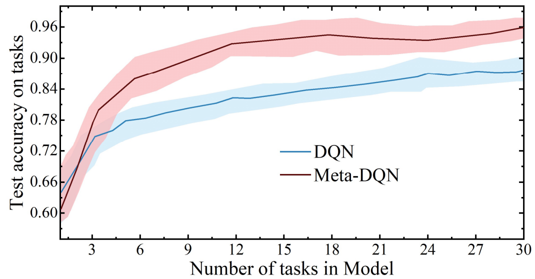


Figure 6: Prediction and optimization of quality stability

IV. B. 2) Stability of chemical composition content

In the process of cigarette production and processing, a sampling point is set in a process under the production process flow, and the next process after that is rolling and wrapping, so it can be assumed that the sample obtained at the sampling point is the finished cigarette, and it is analyzed by using the principal component analysis to obtain the principal component score map of the finished cigarette. Based on the principal component score plot, it can be seen that the tobacco samples at the sampling point and the finished cigarette samples are in the same principal component space, so the finished cigarette model can be used directly to predict the conventional chemical composition of the samples at the sampling point.

Three different batches of tobacco samples were selected from the sampling points for chemical composition analysis, and the statistical data on the conventional chemical composition of the samples between different batches at the sampling points were obtained as shown in Table 2. As can be seen from the table, the mean values of total sugar, reducing sugar, chloride, total phytochemicals, total nitrogen and potassium of the samples between the three batches at the sampling point were 24.90%, 23.08%, 0.57%, 2.32%, 2.10% and 2.32%, respectively. The standard deviations of the main chemical components in the tobacco were also close to each other, which indicated that the quality consistency and stability of the cigarette products among different batches under the same sampling point were relatively good.

Table 2: Stability of chemical composition content.

-	-	First batch	Second batch	Third batch
Total sugar	Average value (%)	24.89	24.92	24.88
	Standard deviation	0.65	0.67	0.64
Reducing sugar	Average value (%)	23.11	23.05	23.07
	Standard deviation	0.61	0.62	0.65
Chlorine	Average value (%)	0.57	0.57	0.58
	Standard deviation	0.06	0.05	0.05
Total alkaloids	Average value (%)	2.31	2.32	2.32
	Standard deviation	0.07	0.07	0.07
Total nitrogen	Average value (%)	2.08	2.12	2.10
	Standard deviation	0.05	0.04	0.04
Potassium	Average value (%)	2.26	2.24	2.45
	Standard deviation	0.09	0.09	0.09

V. Conclusion

Through the application of deep learning algorithms in the prediction of quality stability in the cigarette production process, the Meta-DQN model shows excellent results in several key performance indicators. The model reaches the optimal convergence state after more than 3000 iterations of training, and the fitting effect of the training set and test set is good, which proves the rationality and effectiveness of the model structure. In the key process parameters prediction validation, the prediction accuracy of six parameters, including water content and temperature of loose moisture return material, water content and temperature of primary charging material, and water content and temperature of secondary charging material, is significantly improved, and the average prediction deviation is controlled within 0.212, which is a significant improvement compared with the deviation level of 0.565 in the traditional method. The stability analysis of chemical composition shows that the average contents of total sugar, reducing sugar, chlorine, total phytochemical alkali, total nitrogen and potassium of the tobacco samples between different batches are stable at 24.90%, 23.08%, 0.57%, 2.32%, 2.10% and 2.32%, respectively, and the standard deviation of each component is kept at a low level, indicating that the consistency of the quality of the production process is good. meta-DQN model By introducing the meta-learning mechanism, the limitations of the traditional DQN algorithm in small sample learning and environment adaptation are effectively solved, and the parameter optimization and quality control can be realized quickly in the initial operation stage of the equipment, which significantly reduces the scrap rate and production cost. This study provides a feasible technical path for intelligent quality management in the cigarette industry, which is of great significance for improving production efficiency and product quality.

References

[1] Karthigaikumar, P. (2021). Industrial quality prediction system through data mining algorithm. Journal of Electronics and Informatics, 3(2), 126-137.

[2] Ren, L., Meng, Z., Wang, X., Lu, R., & Yang, L. T. (2020). A wide-deep-sequence model-based quality prediction method in industrial process analysis. IEEE Transactions on Neural Networks and Learning Systems, 31(9), 3721-3731.

[3] Anyan, X., Shanlian, L., Meizhou, D., Wenzhao, L., Liang, G., & Guangfu, H. (2017). The design and application of evaluation method for quality stability of cigarette production process. Food and Machinery, 33(2), 183-188.

[4] Cao, J., Kong, X., Ji, Q., & Zhang, M. (2018). Research on evaluation method of cigarette manufacturing process capability. International journal of computing science and mathematics, 9(6), 621-627.

[5] Ji, Q., Li, M., Xu, G., & Zhang, W. (2024). Quality Improvement of Cigarette Manufacturing Process Driven by Deep Learning: Application of Wasserstein Distance Algorithm. IEEE Access.

[6] Zhi, D., Xie, L., Xia, Y., & Shi, J. (2023). Cigarette Production Line Flow Automatic Control System Based on PLC Technology and Embedded Software. Scalable Computing: Practice and Experience, 24(4), 1185-1190.

[7] Gang, S., Xiaoming, W., Junfeng, W., Shufang, L., Zhuowen, L., & Wang, X. (2021). Analysis of the information management system in the manufacturing process of cigarette enterprises using fuzzy AHP. Journal of Intelligent & Fuzzy Systems, 40(4), 8257-8267.

[8] Lin, J., & Guo, H. (2020, August). Application of Machine Learning in Cigarette Quality Evaluation Method. In Journal of Physics: Conference Series (Vol. 1607, No. 1, p. 012011). IOP Publishing.

[9] Liu, H., Yuan, G., Yang, L., Liu, K., & Zhou, H. (2022). An appearance defect detection method for cigarettes based on C-CenterNet. Electronics, 11(14), 2182.

[10] Bi, S., Zhang, B., Mu, L., Ding, X., & Wang, J. (2020). Optimization of tobacco drying process control based on reinforcement learning. Drying Technology, 38(10), 1291-1299.

[11] Gao, J., Xie, W., Shi, D., Wu, J., & Wang, R. (2024). Synchronized optimization of the logistics system of a tobacco high-bay warehouse under production task fluctuations. Engineering Research Express, 6(4), 045580.

[12] Wang, W., Zhao, F., Wang, M., Sun, Q., Gao, H., Jian, W., ... & Luo, S. (2025). Multi-objective optimisation of cigarette production planning and inventory management. International Journal of Computing Science and Mathematics, 21(1), 64-76.

[13] Qu, R., Yuan, G., Liu, J., & Zhou, H. (2021, October). Detection of cigarette appearance defects based on improved SSD model. In Proceedings of the 2021 5th International Conference on Electronic Information Technology and Computer Engineering (pp. 1148-1153).

[14] Wang, X., Wang, X., Wu, J., Zheng, K., Pang, Y., & Gang, S. (2021). Research on quality traceability of cigarette by combining PDCA quality cycle with information strategy based on fuzzy classification. Journal of Intelligent & Fuzzy Systems, 40(4), 8217-8226.

[15] Wang, W., Zhang, L., Wang, J., & Long, Z. (2022). Locating Defects and Image Preprocessing: Deep Learning in Automated Tobacco Production. Journal of Sensors, 2022(1), 6797207.

[16] Yin, J., Wang, J., Jiang, J., Xu, J., Zhao, L., Hu, A., ... & Cai, M. (2024). Quality prediction of air-cured cigar tobacco leaf using region-based neural networks combined with visible and near-infrared hyperspectral imaging. Scientific Reports, 14(1), 31206.

[17] Yuhang, S. (2025). Deep Learning Algorithm-Based Appearance Detection of Cigarette Strip Packs. International Journal of High Speed Electronics and Systems, 2540262.

[18] Chen, F., Liao, M., Deng, C., & Su, L. (2024). Research on digital online inspection of product quality in tobacco production enterprises. International Journal of Innovation and Sustainable Development, 18(5-6), 629-647.

[19] Zhuoran, X. I. N. G., Songshuang, D. I. N. G., Kai, Z. H. A. N. G., Ming, M. A., Wenlong, G. U. O., Xudong, L. I. U., & Xiangdong, S. H. I. (2025). Research Progress of Deep Learning and Computer Vision in Tobacco Leaf Production. Journal of Agricultural Science & Technology (1008-0864), 27(1).

[20] Ahinful Abraham Ato,Opoku Mensah Abigail,Koomson Samuel,Cobblah Collins,Takyi Godfred & Kwarteng Abdul Hamid. (2025). Achieving banking industry innovation performance using total quality management: an empirical study. The TQM Journal,37(5),1292-1319.

[21] Hui Chen & Baiquan Chen. (2024). Research on the Teaching Mode of Hydraulics based on PDCA Cycle Theory. Academic Journal of Management and Social Sciences,7(2),44-48.

[22] Talip Çay,Emre Ölmez,Nermin Tanık & Cemil Altın. (2024). EEG Based Cigarette Addiction Detection with Deep Learning. Traitement du Signal,41(3).