Modeling Prediction and Research on Leaf Moisturizing Effect of Tobacco Redrying

Keqiang Li\textsuperscript{1}, Ruidong Li\textsuperscript{1,\ast}, Wenyong Huang\textsuperscript{1}, Zehui Yang\textsuperscript{1}, Xiaobing Zhang\textsuperscript{2} and Jianguo Liu\textsuperscript{2}

\textsuperscript{1}Yunnan Leaf Tobacco Redrying Co. Ltd, Kunming 650031, China
\textsuperscript{2}China Tobacco Zhejiang Industrial Co. Ltd, Huangzhou 310024, China

\textsuperscript{\ast}Corresponding author email: 371856443@qq.com

Abstract. This paper studies the influence of the process parameter setting of the hot-air leaf moisturizer on the quality index of the exit leaf during the secondary leaf conditioning of threshing and redrying, and establishes the corresponding prediction model. According to the characteristics of the characteristic process data of the secondary moisturizing, the prediction models of BP neural network and cyclic neural network are established. Call the current popular neural network writing framework TensorFlow's high-level API interface to build the neural network structure. Gradually optimize the activation function, optimizer, number of hidden layer neurons and other key parameters in the neural network structure to make the prediction results of the test set reach the best state. By inputting the parameters of the front steam nozzle pressure, the front-end water flow rate, the hot air temperature, the return air temperature, the temperature of the feed blade, and the water combination of the feed blade, two key tobacco leaf evaluation indicators are predicted for the outlet leaf moisture and temperature. Based on the mean square error, root mean square error, and average absolute error of the prediction results as quantitative indicators, the three indicators of BP neural network are 7.17, 2.68, 1.83, and the three indicators of cyclic neural network are 4.70, 2.16, 1.74. It is concluded that the cyclic neural network has the best prediction effect. It has certain reference value for the tobacco factory to adjust the process parameters of the secondary leaf conditioning according to different conditions.

Keywords: BP neural network, cyclic neural network, secondary moisturizing, tensorflow, outlet leaf temperature, outlet leaf moisture, mean square error.

1. Introduction
Threshing and redrying is the basis for the development of the tobacco industry\textsuperscript{[1]}. Threshing and redrying is an advanced raw tobacco processing technology, and the heating and humidification technology of tobacco leaves (including the secondary moisturizing stage) is the key link\textsuperscript{[2]}. In the moisturizing stage of tobacco redrying, Liu Zhengqin\textsuperscript{[3]} explored how to improve the traditional method, and proposed a method of heating and humidifying with pure steam. Experiments verified that the changes in various index parameters of the moisturizing leaf affected the tobacco after moisturizing. Quality has an impact. Long Minghai et al.\textsuperscript{[4]} took Hongda B1F and Yunyan 87XZF as the research objects and proved that the steam moisturizing method can well retain the aroma components of middle and high-grade tobacco leaves, and the soda-water mixed moisturizing method is beneficial to the degradation of macromolecular compounds. He Guangyu\textsuperscript{[5]} proposed an adaptive nonlinear PID control method based on the complex system characteristics of the two-run tube equipment with typical large, pure hysteresis, and strong interference. Through a digital simulation...
platform combined with an improved intelligent optimization algorithm, an adaptive nonlinear PID control method was proposed. Scholars such as Lian Changwei and Wang Fayong\cite{6,7} have proved through different field tests that the sensory quality and combustion performance of the obtained sheet smoke will change when the leaf moistening intensity, hot air humidification amount, and steam pressure at the outlet are different. Some scholars\cite{8-10} proposed adopting new equipment and control methods in the leaf conditioning stage, and using different processing methods and workflows can effectively improve the quality of leaf conditioning. Zhao Chaowen\cite{11} determined that the main processes of threshing and redrying were leaf moisturizing, threshing and destemming and redrying, and constructed an evaluation model of the influence of these processes on the quality of tobacco slices. Through this model, the comprehensive index score was calculated and standardized. Finally, the score $Q$ is obtained, the BP neural network is built with the score $Q$ as the output expectation, and the genetic algorithm is used to optimize the network, and an effective evaluation model is established. Many scholars have conducted research on the leaf moistening stage, but they are all based on experiments, and the conclusions of each experiment are not the same. They did not thoroughly explore the relationships existing in the data and establish general models. Deep learning is currently a hot spot in scientific research, and has been widely used in image recognition, natural language processing, automatic control systems, etc\cite{12-14}. Based on previous researches and the characteristics of the secondary leaf conditioning process data, this paper summarizes it as a black box optimization problem. Based on the adjustable process parameters of the hot-air leaf moisturizer and the characteristics of the tobacco leaves, a BP (Back Propagation) neural network prediction model is proposed to solve the instability and uncertainty of the quality indicators of exported tobacco at this stage. Reasonably adjust the process parameters to provide a theoretical reference.

2. The Establishment of BP Neural Network Model

2.1. Statement of Problem

This paper selects the data collected from the second leaf conditioning experiment conducted by a tobacco factory process laboratory within one year for research\cite{15}, the total number of samples is 100 groups, and some sample information is intercepted as shown in Table 1.

| No. | Front steam nozzle pressure (Bar) | Front end water flow (L/h) | Hot air temp. (°C) | Return air temp. (°C) | Feeding blade temp. (°C) | Feeding leaf moisture (%) | Outlet blade temp. (°C) | Outlet leaf moisture (%) |
|-----|---------------------------------|---------------------------|--------------------|----------------------|-------------------------|---------------------------|-------------------------|--------------------------|
| 1   | 0.00                            | 250.00                    | 80.00              | 50.00                | 29.30                   | 14.47                     | 45.33                   | 22.80                    |
| 2   | 0.10                            | 250.00                    | 94.00              | 60.00                | 29.60                   | 13.36                     | 53.00                   | 18.63                    |
| 3   | 0.20                            | 170.00                    | 106.00             | 67.00                | 29.70                   | 13.81                     | 60.00                   | 17.15                    |
| 4   | 0.35                            | 220.00                    | 118.00             | 78.00                | 28.90                   | 14.32                     | 61.00                   | 20.28                    |
| 5   | 0.46                            | 190.00                    | 135.00             | 82.00                | 29.70                   | 14.32                     | 70.33                   | 18.01                    |
| 6   | 0.00                            | 200.00                    | 80.00              | 50.00                | 29.90                   | 14.57                     | 41.67                   | 22.40                    |
| 7   | 0.10                            | 200.00                    | 92.00              | 57.00                | 29.20                   | 15.12                     | 50.00                   | 20.06                    |
| 8   | 0.20                            | 250.00                    | 106.00             | 67.00                | 28.60                   | 14.81                     | 58.67                   | 21.86                    |
| 9   | 0.33                            | 210.00                    | 117.00             | 73.00                | 29.40                   | 19.50                     | 62.33                   | 21.25                    |
| 10  | 0.46                            | 195.00                    | 132.00             | 80.00                | 29.50                   | 15.31                     | 68.33                   | 19.90                    |

It can be seen from the above table that each group of samples is to conduct process experiments on the quality indicators of export materials under different conditions. The first four parameters are equipment control parameters, the fifth and sixth parameters are inlet process parameters, and the last two are outlet quality indicators. For 100 groups of experiments, based on a large amount of research experience, it can be known that 80 groups of them can be used as the training set, and the remaining 20 groups of samples can be used as the test set as a reasonable allocation plan.
The main purpose of modeling is to form the relationship between the first six parameters and the latter two parameters. The first six parameters can be used to predict whether the tobacco leaf quality indicators meet the corresponding requirements after moistening. To solve the problem of the need to adjust parameters in production and life to deal with different situations. For example, in some areas, tobacco leaves have less moisture, and the corresponding parameters can be adjusted to increase the export leaf moisture to meet production needs.

2.2. Modeling of BP Neural Network

When modeling a BP neural network, we must first clarify a few points. First, there are several indicators (outputs) to be predicted. Secondly: There are several control parameters (inputs) to be adjusted. Finally, select the number of hidden layers and the number of neurons to determine the structure model of the neural network. For the structural design of the hidden layer, many researchers have done relevant research. The number of hidden layers and the number of neurons contained in each layer largely determines the memory capacity, training speed and generalization ability of the BP neural network. Too many nodes in the hidden layer will increase the amount of calculation and reduce the fault tolerance of the model; too few nodes in the hidden layer may cause problems such as increased local optimal conditions and decreased convergence accuracy [16,17]. From a theoretical analysis, the addition of the activation function allows the BP neural network with only one hidden layer to approximate any nonlinear system. Since the sample data is not large, this paper uses a three-layer BP neural network with only one hidden layer, six neuron input layers, and two neuron output layers.

In theory, there is an optimal number of neurons for each neural network model, but it is difficult to draw a conclusion directly through a specific formula. You can use the empirical formula for preliminary calculations, and then verify the accuracy of the model to get Optimal number of neurons.

\[ l = \sqrt{n + m + a} \]  

(1)

In the above formula, \( n \) is the number of neurons in the input layer, \( m \) is the number of neurons in the output layer, and \( a \) is a constant between 1 and 5. In this paper, \( n \) is 6 and \( m \) is 2, so the value range of \( l \) is 5-9. In the process of model training, set different numbers of hidden layer neurons and observe the number of neurons, the model training effect is the best, and the number of neurons is selected accordingly. The structure of the neural network is shown in Figure 1 below.

**Figure 1.** Neural network structure diagram.
After establishing the neural network model and determining the input and output data, write the neural network code and debug it. The first step: Read the data, use the first 80 sample data of the overall sample data as the training set, and the other 20 groups of data as the test set, and normalize the data and disorder the order. There are six adjustable process parameters in the secondary leaf conditioning process, with different dimensions and units. In order to eliminate this influence and speed up the convergence of the data, normalization is required. The purpose of data disorder is to eliminate the correlation between data and enhance the generalization ability of the model. The normalization method used in this article is a linear normalization method, and the calculation formula is as follows.

\[ x' = \frac{x - \min(x)}{\max(x) - \min(x)} \]  

(2)

Step 2: Build a neural network model, call the Sequential() function in Keras to build a three-layer neural network. Step 3: Set the hyperparameters in the model, call the compile() function, set the optimizer, learning rate, loss function, accuracy function, etc. Step 4: Train the neural network model, call the fit() function, specify the input features and corresponding labels of the training set, the number of cycles, and the input features and labels of the test set. Step 5: Save the parameters of the model, make predictions, and draw the image of the predicted value and the actual value, so as to better observe the prediction effect.

2.3. Parameter Analysis of BP Neural Network

In the BP neural network model, the number of neurons in the hidden layer has a significant impact on the accuracy of the entire model. The two factors are analyzed and discussed below to determine a reasonable plan to get the best prediction effect and model parameter.

To study the influence of a single factor on the overall model, the controlled variable method is adopted, and other parameters are kept uniform, and this factor undergoes different levels of changes. The main settings of several parameters are as follows: activation function (Activation='relu'), optimizer (Optimizer='rms- prop'), the size of each bunch of data (batch_size=20), and the data samples are made into a bunch of 20, sent to the neural network, to improve the computer's computing efficiency, the number of loop iterations (epochs=100), that is, the data samples should be fed to the neural network model in a loop, and the model is continuously trained to improve the accuracy. The regularization method chooses L2 regularization (kernel_regularizer=tf.keras.regularizers.l2()), regularization can effectively alleviate the overfitting problem of the model. Explain several important parameters, as described below.

The full name of RMSprop algorithm is Root Mean Square Prop, which is an optimization algorithm proposed by Geoffrey E. Hinton\cite{18,19}. Because the learning rate decay of the Adagrad algorithm is too radical, consider changing the calculation strategy of the second-order momentum, not accumulating all the gradients, and only focusing on the gradients in a certain window in the past. The exponential moving average is approximately the average over a period of time in the past, reflecting "local" parameter information, and using this method to calculate the second-order cumulative momentum.

\[ w_{t+1} = w_t - \eta_t \]  

(3)

\[ m_t = g_t, \quad V_t = \beta_1 \cdot V_{t-1} + (1 - \beta_1) \cdot g_t^2 \]  

(4)

\[ \eta_t = \frac{\alpha}{\sqrt{V_t}}, \quad m_t = \frac{\alpha}{\sqrt{V_t}} \cdot g_t \]  

(5)

The main idea of regularization is to add an index that can describe the complexity of the model on the basis of the loss function. The loss function (loss) is used to express the accuracy of the model on the training set. In the optimization, the loss function is not directly used, but the complexity of the model R(w) is added. The goal of model optimization is loss(p)+q×R(w). Where q represents the proportion of the model's complex loss in the total loss, and p represents the weight w and bias b in the model.
There are two main functions describing the complexity of the model, one is L1 regularization, and the formula is as follows:

\[ R(w) = \|w\|_1 = \sum_i |w_i| \]

(6)

The other is L2 regularization, the formula is as follows:

\[ R(w) = \|w\|_2^2 = \sum_i |w_i|^2 \]

(7)

After setting the parameters, in order to observe the quality of the model training results, the corresponding observation indicators are formulated. The main indicators are the loss of the training set and the test set. The smaller the loss, the higher the accuracy of the model. In addition, in order to show the relationship between the prediction result and the true value more clearly, the predict() function is called to predict the sample data of the test set and compare it with the true value. The evaluation indicators are Mean Squared Error (MSE) and mean square Root Mean Squared Error (RMSE), Mean Absolute Error (MAE).

MSE is used to express a measure of the degree of difference between the estimator and the estimated value. The sample size in this article is fixed, and the square function of the distance between the point estimate \( \hat{\theta} \) and the actual value \( \theta \) is used, and the expectation of the function is calculated. The calculation formula is as follows:

\[ MSE(\hat{\theta}) = E(\hat{\theta} - \theta)^2 \]

(8)

RMSE is the square root of the ratio of the square of the deviation between the predicted value and the actual value to the number of observations \( n \), \( d_i^2 \) Represents the square of the deviation of a set of measured values from the true value, the formula is:

\[ RMSE = \sqrt{\frac{\sum d_i^2}{n}} \]

(9)

MAE is the average of the absolute values of the deviations of all single predicted values from the arithmetic mean. For a set of data \( \{x_1, \ldots, x_n\} \), the formula is:

\[ MAE = \frac{1}{n} \sum_{i=1}^{n} |x_i - m(x)| \]

(10)

Set the number of neurons in the hidden layer to a number from 5 to 9, and the results obtained after BP neural network training are shown in the following figures and table.

![Figure 2. Loss and prediction result graph for 5 neurons.](image)
Figure 3. Results of loss and prediction at 6 neurons.

Figure 4. Loss and prediction result graph for 7 neurons.

Figure 5. Graph of loss and prediction results at 8 neurons.

Figure 6. Results of loss and prediction at 9 neurons.
It can be seen from the above five figures that the difference in the number of neurons in the hidden layer has a significant impact on the prediction model. From the loss diagram on the left, it can be seen that the accuracy of the BP neural network prediction model basically meets expectations. The straight line is the loss rate of the training set, and the dashed line is the loss rate of the test set. Regardless of the number of neurons, the loss is reduced to less than 0.1 for the training set or the test set. When there are 5 neurons in the hidden layer, the first half of the loss map has a large difference between the training set and the test set. The main reason is that the number of neurons is small and the characteristics of the input data cannot be trained well. Adjust the weight matrix to an appropriate value. When there are 9 neurons in the hidden layer, the loss of the test set and the training set in the first half of the loss graph is very close, and the gap between the two curves in the second half increases significantly. The reason is that the number of neurons may be too large, which was created during the initial loop iteration. The model has a slight overfitting phenomenon. As a result, the accuracy of the second half of the forecast has decreased. Observe the comparison between the predicted results of the picture on the right and the real results. The upper part of the picture is the temperature forecast, and the lower part is the moisture forecast. The curves with triangles and plus signs are real data, and the other two curves are BP neural network prediction data. It can be seen from Figure 2 that the BP neural network model with 5 neurons in the hidden layer has an average performance in predicting the moisture content of the exported tobacco slices. It does not fluctuate much with the change of real data, and it is basically a straight line. As the number of neurons increases, the prediction effect is significantly improved. However, when the number exceeds 7, it is not possible to predict the values with large changes in the real temperature data. Therefore, it is concluded that when the number of hidden layer neurons is 7, the prediction effect is the best. See Table 2 for specific prediction and evaluation indicators.

**Table 2. The prediction results of the BP model with different numbers of neurons.**

| Number of hidden layer neurons(n) | MSE   | RMSE | MAE  |
|-----------------------------------|-------|------|------|
| 5                                 | 15.54 | 3.94 | 2.82 |
| 6                                 | 9.65  | 3.11 | 2.09 |
| 7                                 | 7.17  | 2.68 | 1.83 |
| 8                                 | 14.15 | 3.76 | 2.76 |
| 9                                 | 11.49 | 3.39 | 2.51 |
| 10                                | 9.99  | 3.16 | 2.40 |
| 11                                | 12.76 | 3.57 | 2.68 |

Observe the above table, because when the number of neurons in the hidden layer is 9, all indicators have decreased. In order to prevent the occurrence of a local optimum, two sets of experiments were carried out downward. It can be seen that when the number of neurons is 7, MSE, RMSE, and MAE all perform better than other cases. It can be determined that when the number of neurons in the hidden layer is 7, it is the optimal BP neural network structure. It is verified from the side that there is an optimal number of neurons in each neural network. Therefore, the final BP neural network structure selected in this article is a three-layer BP neural network structure with 6 neurons in the input layer, 7 neurons in the hidden layer, and 2 neurons in the output layer.

**3. Recurrent Neural Network Modeling**

The experimental data of the second leaf moisturizing characteristic process has been trained and predicted by the BP neural network and good results have been achieved. It is impossible to improve the accuracy rate by adjusting the parameters, so we should start from the data itself, observe its characteristics and propose a more reasonable and more realistic modeling method. The second set of leaf moisturizing test process in threshing and redrying plant is as follows: adjusting the first set of parameters, taking samples after the hot air moisturizer has worked for a period of time; adjusting the second set of parameters, and then taking samples after hot air moisturizing the leaves for a period of time…Repeating the above steps until all the experiments are completed. The experimental procedures of the threshing and redrying plant are continuous and there are no shutdown, cleaning or other processes. As a result, the previous group of tests will have a certain impact on the latter group of tests.
The main reason is that the tobacco leaves of the last test are left in the hot air moisturizer, and the change process of the first group of parameters and the second group of parameters is not a sudden change process, there will be a small amount of buffer time in the middle. This leads to an unique phenomenon that each adjacent group of experiments influencing each other and the non-adjacent experimental groups basically having no influence. Recurrent neural network is also proposed under a similar background. The main problem it solves is to consider the context when forecasting, and establish the connection between adjacent data. RNN will use the information stored in the memory of the previous step as a part of predicting the result of the next step. More complex cyclic neural networks such as LSTM and GRU will consider long-term dependence issues, not only establishing the connection between adjacent memories, but also will establish connections between memories which are far apart. Then, starting from the data itself and the principle of modeling, RNN is better than BP neural network in predicting the data of the characteristic process of secondary moisturizing. Therefore, based on this point of view, a recurrent neural network is established to verify the correctness of the hypothesis.

3.1. Recurrent Neural Network Design

In essence, the recurrent neural network is not much different from the BP neural network, except that the neurons are replaced by memories. The number of memories will affect the calculation effect of the entire RNN, the number and speed of reading information. According to the empirical formula, the amount of memories are about 1 to 1.5 times the number of training sets. The model training conducted this time includes a training set of 80 samples and a test set of 20 samples, so the number of memories is approximately between 80 and 120. The number of recurrent cores contained in the recurrent calculation layer in the RNN largely determines the training speed, the accuracy of the recurrent neural network and the generalization ability of the model. Too many cores will increase the amount of calculation and reduce the fault tolerance of the model; rarely cores may cause an increase in local optimal conditions and obvious oscillations in the accuracy of prediction results. In the case of a small amount of data, based on previous experience, select 2 loop cores as the main structure of the RNN. In order to reduce the over-fitting of the data, the Dropout layer is added to the RNN structure design\(^{20,21}\). In 2012, in order to reduce the phenomenon of over-fitting, Professor Hinton proposed the concept of Dropout. To put it simply, in the training of the neural network, the training unit is removed according to a certain probability, so that the parameters are not updated in this round of training, and the overall structure of the neural network will become as shown in Figure 7.

![Figure 7. Neural network running diagram with the Dropout layer added.](image-url)

As shown in the figure above, after adding the Dropout layer, some training units do not update the parameters. Among them, the unit that is not trained is to generate vector 0 and vector 1 with a certain probability through the Bernoulli function, vector 0 is the unit that does not participate in training. Because the final result contains the temperature of the outlet leaves and the moisture of the outlet leaves, a fully connected layer containing two neurons is added at the end. When using Sequential() in Keras to build a network, the structure of a recurrent neural network are mainly composed of a simple
RNN layer with 100 memories, a Dropout layer, a simple RNN layer with 120 memories and a fully connected layer with 2 neurons.

3.2. Parameter Selection and Calculation of Recurrent Neural Network
The recurrent neural network model is based on the context to make predictions, so the division of the sample data test set and training set is slightly different from that of the BP neural network. The BP neural network uses the first 80 sets of samples as the training set, and the last 20 sets of samples as the test set. The RNN test set has 4 samples per interval. One sample is taken into the test set, and the remaining samples are used as the training set, then the test set contains samples numbered 5, 10, 15, 20.... According to the dividing principle, the size of each piece of data should be as small as possible (batch_size=3). Through the experience of adjusting the parameters of the BP neural network, other parameters are set as: optimizer (Optimizer='RMSprop'), loss function (Loss='Mean Squared Error'), callback function monitoring value (monitor='Validation Loss'). Loop iterations required further discussion. Three sets of model training experiments were carried out, with 100, 200, and 500 loop iterations as adjustable parameters. The results are as follows.

Figure 8. Figure of RNN model loss and prediction results with 100 iterations.

Figure 9. Figure of RNN model loss and prediction results with 200 iterations.

Figure 10. Figure of RNN model loss and prediction results with 500 iterations.
It can be seen from the figure on the left that the increase or decrease of the number of iterations does not significantly affect the loss of the test set. The loss mainly oscillates up and down with the horizontal axis 0.06 as the center line which also shows that the increase in the number of iterations does not significantly improve its accuracy. Observing the figure on the right, it is found that when the number of iterations of the data loop is too many or too few, it will cause the problem of insufficient fitting ability of the model. When the number of loop iterations is 200, the final test indicators MSE, RMSE, and MAE are much smaller than the other two cases. It is reasonable to set the parameter to 200, and the specific prediction results are shown in Table 3.

**Table 3.** RNN model prediction results for different loop iterations.

| Number of loop iterations | MSE    | RMSE   | MAE    | Time consuming (s) |
|---------------------------|--------|--------|--------|--------------------|
| 100                       | 12.65  | 3.56   | 2.79   | 12.37              |
| 200                       | 4.70   | 2.16   | 1.74   | 19.01              |
| 500                       | 14.58  | 3.82   | 2.68   | 42.59              |

4. **Forecasting Models of Comparative Analysis and Discussion**

This paper models, analyzes and predicts the characteristic process data of the secondary moisturizing leaves through three methods of BP neural network and recurrent neural network. The predicted values of moisture and temperature of tobacco leaves at the outlet of the last 20 groups of samples by three different methods are shown in Table 4.

**Table 4.** The prediction results of BP algorithm, RNN algorithm.

| Number | Temperature of Outlet Leaves (°C) | Moisture of Outlet Leaves (%) |
|--------|----------------------------------|--------------------------------|
|        | Actual Value BP RNN Actual Value BP RNN |
| 1      | 43 44.78 45.66 20.85 19.94 20.07 |
| 2      | 55 49.81 52.59 22.97 21.31 21.35 |
| 3      | 60 60.13 58.87 21.62 20.02 18.94 |
| 4      | 59 63.80 62.43 20.70 20.10 19.40 |
| 5      | 73 68.82 66.72 20.15 19.56 18.63 |
| 6      | 44 44.84 45.71 18.97 20.54 20.46 |
| 7      | 52 53.16 55.22 23.10 20.62 20.17 |
| 8      | 58 60.40 58.71 20.01 20.27 19.33 |
| 9      | 61 63.01 59.74 21.50 21.31 20.31 |
| 10     | 65 67.22 66.97 22.07 19.82 18.96 |
| 11     | 45 48.28 45.94 19.87 21.19 21.04 |
| 12     | 55 53.68 53.75 22.00 21.20 20.00 |
| 13     | 60 61.14 58.99 22.18 20.85 19.87 |
| 14     | 63 63.12 61.86 21.02 21.12 20.27 |
| 15     | 68 66.52 65.80 21.26 20.85 20.06 |
| 16     | 42 52.71 45.91 17.75 19.47 18.11 |
| 17     | 60 55.26 54.73 21.94 21.45 20.86 |
| 18     | 63 59.85 58.19 21.86 22.03 21.15 |
| 19     | 63 63.61 61.81 21.42 21.55 20.62 |
| 20     | 70 67.44 67.43 19.81 20.35 19.49 |

In order to observe and compare the results of the three methods more intuitively, subtract the true value from the predicted value of the three methods and take the absolute value, which is recorded as the absolute error. Calculating the error distribution of each method. In the multiple linear regression, the results are two linear equations, but when the error results are calculated, the errors of the outlet leaf temperature and outlet leaf moisture prediction indicators are counted together. The final results are shown in the table 5.
Table 5. Comparison of absolute error of predicted value of BP algorithm, RNN algorithm.

| Absolute error range (i) | Multiple linear regression | BP neural network | Recurrent neural network |
|--------------------------|-----------------------------|-------------------|-------------------------|
| 0<i<0.5                 | 12                          | 9                 | 3                       |
| 0.5<i<1                 | 6                           | 6                 | 9                       |
| 1<i<1.5                 | 4                           | 6                 | 9                       |
| 1.5<i<2                 | 4                           | 5                 | 8                       |
| 2<i<3                   | 8                           | 8                 | 5                       |
| 3<i<4                   | 3                           | 2                 | 3                       |
| 4<i<5                   | 2                           | 3                 | 1                       |
| 5<i                     | 1                           | 1                 | 2                       |

Combining the above two tables, it can be concluded that the prediction effect of RNN is better than that of BP neural network. The prediction data with absolute error less than 2 can be considered as small error. BP neural network prediction results have 26 prediction data with small errors, and RNN has 29 predictions, indicating that the prediction of recurrent neural network is more stable than the previous method, and the prediction error is smaller.

If the absolute error is greater than 3, it can be considered that the error is relatively large. In this regard, the prediction data of the two methods have little difference. The mean square error (MSE) is the sum of the predicted value minus the square of the true value, and the mean value is calculated. To a certain extent, it can directly reflect the pros and cons of the prediction effect. The mean square error of the two methods is 7.17 and 4.70. It can be seen that the performance of the BP neural network prediction method is poor, mainly because the amount of data is not large and the characteristics of the method cannot be used effectively. The prediction result of the recurrent neural network is the best. The mutual influence of adjacent test groups in the secondary leaves moistening process provides the basis for RNN modeling and prediction. Because the various memories are related to each other, improving the accuracy of the model is not particularly dependent on the amount of data. Accurate predictions can be made with only a small amount of data. Therefore, the method of recurrent neural network should be used when modeling and predicting the characteristic process data of the secondary leaves moisturizing.

5. Conclusion

In this paper, by extracting the characteristics of the secondary leaf moisturizing process data, establishing the BP neural network and the Recurrent neural network prediction model, the influence of the process parameter setting of the hot air moisturizing machine on the quality index of the exported leaves is studied. The results show that when the BP network predicts the secondary leaf conditioning process data, the number of neurons selected affects the training effect of the network, resulting in over-fitting or under-fitting. Secondly, the small data size also limits the prediction results of the BP network. The prediction result of the RNN is better among them. The mutual influence of adjacent test groups in the second leaf moistening process provides the basis for RNN modeling and prediction. Because the various memories are related to each other, it is not particularly dependent on the amount of data. To improve the accuracy of the model, only a small amount of data is needed to make accurate predictions. Therefore, the method of RNN should be used when modeling and predicting the characteristic process data of the secondary moisturizing. In the future, multiple sources and large amounts of data will be introduced for this process link, and the neural network model will be optimized to make the prediction results more accurate, so as to guide the processing technology.

Acknowledgments

This work is supported by China National Tobacco Corporation's major science and technology project "Research and application of "characterized and modular" threshing and redrying characteristics that highlight the tobacco leaf style in the producing area" (No.110201901031 (FK-02)), China National Tobacco Corporation Yunnan Branch Technology The planned key project "Research on the uniformity control of threshing and redrying slices" (No. 2018530000241024).
References

[1] Mingde Hua. Study on some problems in the development of threshing and redrying industry in China[J]. Acta Tab. Sin, 1998(02):3-5.

[2] Yuekun Wang, Yukun Hua, Hongwei Li et al. Overview of the research progress of threshing and redrying[J]. Acta Agric.Jiangxi, 2013,25(12):57-61.

[3] Zhengqin Liu. Research on the new processing method of threshing, redrying and moisturizing leaves[C]. Guangxi: Constructing a new system Stimulate new vitality. 2020:10.

[4] Minghai Long, Xiaolong Zhang, Xianguo Wang et al. The effect of threshing and redrying and moistening methods on the quality of tobacco leaves[J]. Hubei Agricultural Sciences, 2016,55(01):108-111+116.

[5] Guangwei He. Research on Auto Disturbance Rejection Control Algorithm of Threshing and Redrying Hot Air Moisturizer[D]. Hunan:Hunan University, 2018.

[6] Changwei Lian, Tengjian Zhang, Minrui Lu et al. The effect of different leaf moisturizing intensities on the quality and energy consumption of threshing and redrying in-process products[J]. Journal of Anhui Agricultural Sciences, 2020,48(15):194-196.

[7] Fayong Wang, Shaohui Niu, Chao Xu et al. Screening of leaf moisturizing parameter model based on tobacco slice structure and aroma preservation processing[J]. Jiangsu Agric.Sci, 2020,48(09):223-227.

[8] Cheng Chen, Ming Zhou, Xingchong Yang et al. Discussion on Improving the Feeding Mode of Leaf Conditioner [J]. Modern Agricultural Machinery, 2019(06):45-46.

[9] Chaoming Wang. The design and application of a new type of hot air moisturizing leaf circulation system[J]. Plant Engineering Consultants, 2019(05):48-50.

[10] Yanling Liu, Zeli Wang, Qiang zhou et al. Establishment of a water temperature control model for the moisturizing section of threshing and redrying leaves[J]. Anhui Agricultural Science Bulletin, 2017,23(22):109-111.

[11] Chaowen Zhao. Research on the influence of threshing and redrying technology on the processing quality of tobacco leaves[D]. Kunming University of Science and Technology, 2020.

[12] Yujia Duan, Ting Ju. Evaluation of the effectiveness of code review opinions based on deep learning[J]. Electronic Science and Technology, 2020,33(01):39-45.

[13] Junhua Cheng, Guohui Zeng, Jin Liu. Research on classification method of complex background image based on deep learning [J]. Electronic Science and Technology, 2020,33(12):59-66.

[14] Chao Yan, Zhanquan Sun, Engan Tian et al. Research progress of medical image segmentation technology based on deep learning[J]. Electronic Science and Technology, 2021,34(02):7-11.

[15] Yunlan Tang. Cigarette product parameter optimization design method and application research integrating neural network and multi-objective evolutionary algorithm[D]. Hunan:National University of Defense Technology, 2008.

[16] Bing Cai. The design method of BP neural network hidden layer structure[J]. Journal of Tonghua Teachers College, 2007,28(2):18-19.

[17] Jiani Fan, Zhenlei Wang, Feng Qian. Research Progress of BP Artificial Neural Network Hidden Layer Structure Design [J]. Control Engineering of China, 2005 (S1):109-113.

[18] Zou F, Shen L, Jie Z et al. A Sufficient Condition for Convergences of Adam and RMSProp[C]. California: Conference on Computer Vision and Pattern Recognition (CVPR). 2019.

[19] Xu D, Zhang S, Zhang H et al. Convergence of the RMSProp deep learning method with penalty for nonconvex optimization[J]. Neural Networks, 2021, vol 139:17-23.

[20] Yuanyuan Chen, Zhang Yi. Adaptive sparse dropout: Learning the certainty and uncertainty in deep neural networks[J]. Neurocomputing, 2021, 450: 354-361.

[21] Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, et al. Dropout: a simple way to prevent neural networks from overfitting[J].The Journal of Machine Learning Research, 2014,15(1): 1929–1958.