Research Article

Optimization of Printing and Dyeing Energy Consumption Based on Multimedia Machine Learning Algorithm

Xulan Zhang 1 and Yue Yu 2

1 China Knitting Industrial Association, Beijing 100020, China
2 Dacheng County Power Supply Branch of State Grid Jibei Power Co., Ltd., Langfang 065900, China

Correspondence should be addressed to Yue Yu; 20162103826@mails.imnu.edu.cn

Received 15 February 2022; Accepted 18 March 2022; Published 10 May 2022

Abstract

In order to solve the problem with the expansion of the industrial scale, the contradiction between energy resources and development of printing and dyeing enterprises must implement refined management and optimize the allocation of resources such as production technology, energy consumption types, and metering instruments. As a printing and dyeing textile industry with high energy consumption and high pollution, energy conservation and emission reduction have become difficult problem to be solved in this industry. The premise of optimizing resource allocation is to have an objective and scientific evaluation of the current energy resource allocation of printing and dyeing enterprises. In view of this, through the investigation of printing and dyeing enterprises, this paper puts forward the index system of enterprise energy consumption optimization evaluation. Based on the application of data warehouse and combined with historical data, a new energy consumption optimization evaluation method is proposed. This survey has basically understood the current situation of energy and water resources management of a printing and dyeing enterprise and pointed out the direction for the development of enterprise energy consumption optimization project in the next step. By means of multimedia informatization, with the help of Internet of Things sensing technology, building a new generation of energy management system, improving the refined management of energy measurement, and improving the energy assessment system, enterprises can achieve significant economic and social benefits in terms of energy conservation and emission reduction. On this basis, a low-power scheduling strategy for typical data center applications is designed and implemented. The algorithm uses the input data of different calculation parts, performs matching to determine the redundant part of the calculation process, and schedules the algorithm. Experimental data show that the mean square error of the limit tree is 0.0004248 and the mean square error of the decision tree is 0.01581; through calculation, the algorithm can achieve 23% and 17% energy saving.

1. Introduction

With the expansion of data scale, how to use the machine learning algorithm to analyze these data in real time or off-line and find the internal law of data has become an important way for various industries to improve the accuracy of decision-making [1]. For example, Google indexes and ranks 571 new websites that emerge every minute based on machine learning algorithms; Wal-Mart, the world’s largest retailer, needs to analyze and process more than 1 million transaction data per hour, so as to provide decision-making for its further business activities; Microsoft Research Asia, according to the air quality data provided by existing monitoring stations and other multiple data sources in the city, uses machine learning technology to fully analyze big data; it can be inferred in real time that urban air quality data contain fine particulate matter information [2, 3]. Therefore, machine learning has become an indispensable functional module in the data center. In this context, various machine learning systems running in data centers continue to emerge, such as GraphLab, Mahout, and MadLINQ. Among them, Mahout is a distributed machine learning library based on Hadoop, and it has been widely used by many companies (such as Yahoo, Twitter, and LinkedIn). In order to further optimize the consumption of printing and dyeing, we first conduct an in-depth characteristic analysis of typical machine learning algorithms for printing and dyeing [4].
Machine learning tasks are computationally intensive applications; therefore, its main energy consumption is reflected in data analysis and calculations. However, these algorithms are in the process of machine learning and need to continuously iteratively categorize and analyze data; there are a lot of redundant calculations in this process. These redundant calculations bring unnecessary energy consumption overhead. At the same time, in the data center, machine learning is mainly used for data analysis to find the internal laws of data, so there are no very strict requirements for the accuracy of users’ calculation results. In the recommendation system, you only need to obtain the items that users like, and the overly accurate calculation results are of little significance here [5, 6]. With the rapid development of the Internet of Things and artificial intelligence, some industrial enterprises try to achieve better energy-saving and emission reduction effects by collecting the data generated in the production process and analyzing it with some sensor equipment to a certain degree. Collect some production data in the production process, such as energy consumption and water consumption, by installing relevant sensor equipment of relevant equipment in the production process. Using the corresponding machine learning algorithm in a large number of production data, we can find some inherent laws and patterns in the production data, so as to achieve more accurate energy conservation and emission reduction.

In terms of energy conservation and emission reduction, the initial exploration of enterprises is generally to optimize the industrial process to a certain extent and to upgrade some key equipment, so as to achieve energy conservation and emission reduction to a certain extent and promote the recycling of resources. In printing and dyeing enterprises, there are many different ways to achieve energy-saving effect, including process improvement, waste water and waste gas waste heat recovery, equipment update and maintenance, and the use of green lighting. In recent years, with the development of environmental protection trend, the printing and dyeing industry began to study the new technologies of energy saving, consumption reduction, and emission reduction, advocating the concept of green printing and dyeing, and was committed to transforming from extensive to intensive economic development form [7]. Based on the above characteristics of machine learning algorithms, we designed and implemented an energy-saving mechanism for machine learning calculations in data centers. It is proposed to remove redundancy by matching input data, so as to achieve energy-saving methods. The core idea is through twice input matching, the similarity of its output is analyzed, and the energy-saving effect is achieved by reusing the calculation results and reasonable scheduling [8, 9]. The experimental data shows that on the basis of ensuring the accuracy of the algorithm, the algorithm can effectively reduce the redundant calculation of the data center machine learning algorithm, so as to achieve the effect of energy saving. The machine learning algorithm is shown in Figure 1. Starting from the data of the printing and dyeing process finalization process, this paper combines order-related information, process parameter-related information, energy-related multimedia information, etc., to construct energy consumption categories and energy consumption as predicted values, and conduct model training after certain data preprocessing. Through the application of the optimization algorithm, the energy consumption in the subsequent production process is predicted, and the predicted data are used to adjust the process parameters to a certain extent so as to achieve the effect of energy saving and emission reduction [10].

2. Literature Review

Zhou and others said that with the new trend of social development, under the new wave of Internet+ and big data, the deep mining processing of data has entered saturation. This requires people to re-emphasize machine learning algorithms and set up a reasonable algorithm; let the machine do an in-depth analysis of the data by itself. This can be to a large extent solve the work that cannot be handled manually [11]. Xu and others expressed the application of machine learning algorithms; it is widely used in all aspects and various fields of today’s society. For example, we use the Internet to search, use social platforms, and use e-commerce platforms [12]. Zhang and others obtained through data collection and used machine learning algorithms in the face of complicated data and information platforms on the Internet and the recommendation mechanism of shopping websites, which can analyze the user’s points of interest and make better pushes [13]. Yin and others said that the application of machine learning algorithms in the above-mentioned fields mainly reflected in the different data in the information that can be identified, for example, pictures, text, mixed graphics, audio, and so on; in addition, it can successfully convert voice to text output and identify the domain of the content of an article and the blocking of bad content [14]. Sosnicki stated that the advantages of this technology have been valued in many fields; from this, we can see the development prospects of this advantage [15]. Rao and others found through investigation that relatively speaking, the research of some machine learning algorithms focuses on the theoretical model of machine research. They attach great importance to the accuracy of the theoretical model and analyze various disadvantages faced after being put into use. In other words, in the field of machine learning algorithms, a number of mutually supporting algorithm linkage mechanisms have been introduced, and data classification creation mechanism, so with the help of multilayer models, carries out practical training to improve the feasibility of data processing tasks [16]. Su and others have found through research that the application of vector machines supported by machine learning algorithms is an important part of connecting other vector machines. Only by optimizing the scheme between vector machines can the optimal data classification for data processing be obtained. Based on the development of the era of big data, the speed of various information data processing has been accelerated. Based on the premise of development in the era of big data, various information data processing is speeded up [17]. Canejo and others said that from multiple perspectives of current social information data processing and traditional information data processing algorithms, some large-scale data are no longer applicable. In today’s society where data flow and information flow are exploding, in-depth data mining has gradually developed into a new trend [18]. Khan and others have found through surveys that the transformation of
enterprises and the progress of the industry require the improvement of machine learning algorithms to easily process huge amounts of information and data to achieve social development; thereby, the huge information data can be conveniently processed to realize the development of society [19]. Pasha and others proposed a machine learning algorithm through repeated experiments, which has great advantages in solving the current data mining problems. With the help of modern artificial intelligence technology and the modeling of various mathematical models, we can better use and solve the data problems [20]. Due to the relatively harsh production environment, the relevant sensor equipment is easy to damage under high temperature and high pressure conditions, resulting in some hesitation in intelligent exploration. There are also some obstacles in the upgrading of equipment. Considering the cost problem, most enterprises will not take too much risk to replace the equipment. Therefore, it is of great value to explore the feasibility of energy conservation and emission reduction from the perspective of data.

3. Methods

3.1. Data Preprocessing of the Setting Machine. This section mainly explains the improved neural network optimization method based on the gradient boosting tree model. According to the energy consumption model of the setting machine, the parameters of the model are adjusted to improve the generalization ability of the model. The specific steps are as follows.

Step 1. First, you need to set the initial value of the step size and the number of iterations. In general, initially choose a smaller step size for grid search to get the best number of iterations. Therefore, this paper just started by setting the step size to 0.1. Then, use the step size to perform grid search on the number of iterations and finally get a more suitable number of iterations.

Step 2. Then, it is necessary to determine the maximum depth of the decision tree and the minimum number of samples required for re-division of internal nodes and obtain the optimal solution of the two parameters through grid search.

Step 3. Then, it is necessary to adjust the minimum number of samples required for the re-division of the internal nodes and the minimum number of samples of the leaf nodes at the same time and obtain the optimal solution of the combination of the two parameters through grid search.
Step 4. Perform a grid search on the maximum number of features for feature sampling.

Step 5. Determine the proportion of subsampling by grid search and then perform sample sampling.

Step 6. Finally, the fitting ability and generalization ability of the energy consumption model of the training machine are further enhanced by comprehensively using the method of halving the step size and doubling the maximum number of iterations.

According to the needs of different processes, the equipment involved will also have certain differences; next, we will briefly introduce the printing and dyeing process; some devices with higher energy consumption are used. The first is the singeing machine used in the singeing process; usually, after the grey fabric enters the printing and dyeing factory, it first needs to go through the processes of inspection, turning over, batching, distributing blanks, and sewing heads; the matching blank is to join the front and back sides of the two grey fabrics on the same side; then, the singeing process is carried out, and the main energy consumed in the singeing process is natural gas. The second is the dye vat used in the printing and dyeing process; the main energy consumed in the dyeing process is water. Finally, there is the setting machine equipment used in the printing and dyeing setting process. The main purpose of setting is to eliminate the stress and strain accumulated in the fabric. After the setting process, the surface of the fabric becomes smooth and wrinkle-free, and the size is stable, and it has good thermal stability. For polyester fiber, improve its anti-wrinkle and non-ironing properties. The shaping process is mainly divided into three steps. (1) Grey fabric setting: it is also called presetting used to remove unfavorable impurities and fixation. The yellowing produced during this process can be removed during bleaching. Have to be aware of is that this step is not good for dyeability and high-quality requirements for grey fabrics. (2) Semifinished product shaping: it is also called medium shaping; it is easy to wrinkle in the process before shaping. The purpose of this step is to make the cloth surface smooth and reduce stains. Have to be aware of is that in the process, the dye adsorption is reduced, before and after mercerizing and high requirements for pretreatment. (3) Finished product finalization: Immediate finalization, this process will be processed when the setting machine is running is as follows: if the abnormal value related to the setting machine rarely occurs, select the method of deleting the abnormal value for processing. If there are many abnormal values of the setting machine, it indicates that there may be a problem with the sensor installed on the sizing machine and need to replace the relevant sensors to collect data again. For intermediate cases, the average value correction method is generally selected to deal with abnormal values [23–25]. The basic principle of the average value correction method will be briefly introduced below.

Assuming that a certain period of time is \( T \), a certain runtime data value of the setting machine equipment at time \( t \) within \( T \) is \( x_t \), if at some \( t_i \), \( i = 1, 2, \ldots, n \), moment in between the value of a certain process parameter of the setting machine equipment is abnormal and can use the calculation result of the following formula to replace the abnormal monitoring value:

\[
x_t = \frac{\left( \sum_{i=t_i}^{t} x_i + \sum_{i=t_i+1}^{n} x_i \right)}{(n - 1)}
\]

In the formula, \((n - 1)\) is the period of time \( T \), the number of a certain technological parameter value of the setting machine equipment.

Missing value refers to the situation where the value corresponding to a certain attribute is empty. In the case of very few and many missing values, in the same way as outliers, for intermediate cases, choose multiple exponential smoothing methods to deal with and use three exponential smoothing methods to deal with missing values; the basic principle is as follows: for the missing value sequence of a certain process parameter of the setting machine equipment in a certain time period, according to the value of the process parameter of the setting machine equipment in the previous period of the missing value in this time period, on the basis of clarifying the length of the missing parameter sequence, the number of smoothing steps and data points is inserted; after clarifying the length of the missing parameter sequence, insert the number of smoothing steps and the number of data points [26, 27]. Use the following formula for smoothing:

\[
\begin{align*}
\hat{s}_t^1 &= a \hat{s}_t^1 + (1 - a) \hat{s}_{t-1}^1, \\
\hat{s}_t^n &= a \hat{s}_t^n + (1 - a) \hat{s}_{t-1}^n, \\
\hat{s}_t &= a \hat{s}_t^n + (1 - a) \hat{s}_{t-1}^n.
\end{align*}
\]
Among them, $s_t'$, $s_t''$, and $s_t^*$ are the values after primary, secondary, and smoothing, respectively; $a$ is the weight coefficient of smoothing processing. In general, set $a = 0.5$. So, the smoothed value of the missing value is as follows:

$$x_{t+m} = a_t + b_t m + \frac{1}{2}c_t m^2.$$  \hspace{1cm} (3)

Among them, $m$ represents the number of smoothing steps, set $m = 3$ here, $a_t$, $b_t$, and $c_t$ are the coefficients of smoothing value, and the calculation formula is as follows:

$$a_t = 3s_t' - 3s_t'' + s_t^*,$$

$$b_t = \frac{a}{2(1 - a)^2} \left[ (6 - 5a)s_t' - (10 - 8a)s_t'' + (4 - 3a)s_t^* \right],$$

$$c_t = \frac{a^2}{(1 - a)^2} \left( s_t' - 2s_t'' + s_t^* \right).$$  \hspace{1cm} (4)

(2) Characteristic code: the printing and dyeing order data sheet contains the characteristic fabric name; there are 10 types of fabrics in total; they are four-sided stretch, crepe satin, fragrant cloud yarn, silk satin, chiffon beads, stretch satin jacquard, burnt-out velvet, jinlun yarn, tree pattern forging, and composite silk plain weave. For category data, one-hot encoding will be used for processing. One-hot encoding has a good effect in dealing with multi-classification problems. The following will briefly introduce the processing method of one-hot encoding. If the fabric name data are marked according to the numbers 1, 2, 3, 4, 5, 6, 7, 8, 9, and 10, although there is no problem with accuracy, there will always be some misunderstandings. For example, 2 is twice as large as 1, or if it has no special meaning, 9 is greater than 0. In order to avoid such misunderstandings, they can be expressed in a more equal matrix form.

(3) Data merging is based on the process parameter data sheet of the setting machine; merge data on it; according to the rotational speeds of the 12 circulating fans, the average rotational speed of the circulating fans is obtained. In the same way, find the average speed of the 3 exhaust fans and the average temperature of the 11-section oven; the process time is obtained by subtracting the process end time from the process start time. At the same time, according to the order number, the basic information of the order and the process parameters of the setting machine and the energy consumption of the setting machine is integrated. The initial feature data set is obtained so far [28].

(4) Feature screening determines the target column in the feature data set or constructs the target column by certain rules and analyzes the correlation between each data item in the initial feature data set and the target column. By analyzing the correlation coefficient, the data items with a correlation coefficient greater than 0.5% are screened out as the final feature data set.

(5) Remove the unique attributes, usually some id attributes; these attributes can only be used in the process of data merging. In the description of the distribution law of the sample itself, it is of little significance, so it should be simply deleted at the end. As you can see from the basic order information table, the attribute order number is the only attribute, so it will be deleted after the data are sorted and merged.

Difficulties in the Preprocessing of Production Data. The production data of different printing and dyeing enterprises are quite different in terms of the data format and the specific data content, which also brings the difficulty to the data preprocessing to a certain extent.

(1) Sensor Data Acquisition Process. Due to the harsh environment such as high temperature and high pressure in the process of printing and dyeing, the requirements for all kinds of sensors are relatively high. In the actual use process, the sensor is easy to damage. In the process of sensor update, there are also certain differences in the data acquisition accuracy.

(2) Data Format and Standards. Different printing and dyeing enterprises have a large distance in the process of intelligent transformation, so there are also great differences in the data format and standards. The first is that for a certain device, the relevant data should be collected. Secondly, there is the specific significance of the feature term represented for different feature data.

3.2. Comprehensive Energy Consumption Model per Unit Output of the Setting Machine. The relevant data of the shaping machine used in this paper come from a printing and dyeing factory, and the data come from the No. 1 shaping machine equipment of the enterprise. The energy consumption during the finalization process for a large extent is related to the volume of orders; for example, if you need to shape the fabric with an order of 30,000 meters, the energy consumed must be far greater than the same fabric with an order volume of 3 kilometers. Therefore, an indicator that can measure the energy consumption of different orders is needed. By constructing a comprehensive energy consumption model per unit output of the setting machine, measure the energy consumption of different orders. The comprehensive energy consumption model per unit output of the setting machine is established, and the comprehensive energy consumption model per unit output of the setting machine is used to generate the predicted value of the energy consumption category of the setting machine. Combined with the energy consumption category, the preprocessed characteristic data set is selected to obtain the training characteristic data set. The process of constructing the
comprehensive energy consumption model per unit output of the setting machine is as follows: Extract energy-related data from the printing and dyeing sample data such as electricity consumption \( E \), gas consumption \( G \), water consumption \( W \), and order volume data meters \( M \). Through the formula of energy consumption per unit output of equipment, in order to calculate the energy consumption per unit output of a product, the formula is as follows:

\[
P = \frac{P_g}{N_{gh}}
\]  

(5)

Among them, \( P_g \) represents the comprehensive energy consumption of the equipment, in kilograms of standard coal, and the calculation formula is as follows:

\[
P_g = \sum_{i=1}^{n} (P_i \cdot E_i).
\]  

(6)

where \( E_i \) represents the production activity, the consumption of type \( i \) energy in kind; \( P_i \) represents the conversion factor of standard coal for type \( i \) energy; and \( P \) represents the comprehensive energy consumption per unit output of the equipment, in kilograms of standard coal per 100 meters. Through calculation, the comprehensive energy consumption per unit output of the setting machine uses these data to represent the energy consumption of the setting machine, after entering the relevant data; the formula is as follows:

\[
P = \frac{E \cdot P_x + G \cdot P_g + W \cdot P_w}{M/100}.
\]  

(7)

For features whose values are continuous variables, calculate the Pearson correlation coefficient between the feature column and the target column and keep the feature columns whose correlation coefficient value is greater than 0.5%. For some ordinal variables or even-interval feature data that do not meet the assumption of normal distribution, calculate the Spearman correlation coefficient between the feature column and the target column and keep the feature columns whose correlation coefficient value is greater than 0.5%. The retained features are listed as the feature column for model training \([29, 30]\). The Pearson correlation coefficient calculation formula between two variables is as follows:

\[
r_{xy} = \frac{\sum (X - \bar{X})(Y - \bar{Y})}{\sqrt{\sum_{i=1}^{n} (X_i - \bar{X})^2} \sqrt{\sum_{i=1}^{n} (Y_i - \bar{Y})^2}}.
\]  

(8)

The Spearman correlation coefficient is defined as the Pearson correlation coefficient between rank variables; the original data are based on its average descending position in the overall data and were assigned a corresponding level. Eliminate the data consumption and the average circulating fan speed by processing the data. The 10 features used to predict the energy consumption classification of printing and dyeing shaping machines include cloth name, gas meter temperature, ambient temperature, gas meter pressure of No. 1 shaping, humidity at the front end, tail humidity, process time-consuming, vehicle speed, average speed of exhaust fan, and average temperature of drying room. The whole process of model establishment and implementation is shown in Figure 2.

4. Results and Analysis

4.1. Experimental Environment. Hardware platform is as follows: 16 GB, Quad-Core Intel Core i5@2.3 GHz.

Software platform is as follows: macOS Catalina 10.15.2, PyCharm 2019.2.3, Jupyter Notebook.

In this section, two types of group experiments are done. The first type is K-means unsupervised experiments, as well as supervised gradient boosting tree experiments, random forests, and approximate entropy classification experiments. The second type is regression experiment, including gradient boosting tree model, general decision tree model, and limit tree model. The experimental environment is as follows: hardware platform is 16 GB, Quad-Core Intel Core i5@2.3 GHz and software platform is macOS Catalina 10.15.2, PyCharm 2019.2.3, Jupyter Notebook. For classification models, commonly used evaluation indicators include accuracy, precision, recall, and F1-score. In order to calculate these index data, you first need to understand several related values, as shown in Table 1.

With the four values in Table 1, we can calculate the accuracy, precision, recall, and F1-score values. The calculation formula and description are shown in Table 2.

For regression problems, Mean Square Error (MSE), Mean Absolute Error (MAE), and Mean Relative Error (MRE) are the main three algorithm evaluation indicators. Generally speaking, the smaller the MSE explains that the forecast data are better than the actual data, and the overall deviation is smaller. Same, the smaller the MAE and MRE, the better the prediction model. The RMSE is the square root of the sum of the square of the \( m \) ratio to the number of observations. It is used to measure the deviation between the observed value and the true value; the average absolute error is the average of the absolute error; it can better reflect the actual situation of the prediction value error; the standard deviation is the arithmetic square root which is used to measure the dispersion of a set of numbers itself. In Table 3, the calculation formula and description of each evaluation index are shown; the \( y \) that appears in the formula represents the actual data, \( y \)'represents predicted data, and \( L \) is the number of predicted data.

Through the gradient boosting tree model, regression analysis was performed on the data sample, the process parameters, and fabric types in the production process of the setting machine; data such as process time-consuming are used as input value; the calculated comprehensive energy consumption value is used as the output value, for training and performing prediction experiments on 1000 samples; on this basis, it is compared with the general decision tree model and the limit tree model. The contrast results of the regression model are shown in Figure 3.

As can be seen from Figure 3, in terms of predicting trends. The gradient boosting tree model is better than the decision tree model and the limit tree model in predicting the trend. In terms of predicting specific values, the gradient
boosting tree model is better overall. In the later stage of the energy consumption forecasting process of the enterprise, compared with the previous model, compared with the actual energy consumption, the gradient boosting tree model is better overall. By setting the gradient boosting tree model as a fitness function, the initial value of the learning rate is 0.1, and the number of iterations is 1000, and the optimization curves for different fabric types are obtained as shown in Figure 4.

Substituting the bestT, bestS, bestW, bestC, and bestP obtained by optimization into the model, compare its actual energy consumption per unit output with the optimized energy consumption per unit output, as shown in Figure 5; it is a comparison of energy consumption before and after optimization of Xiangyun yarn.

It can be found from the comparison before and after the energy consumption optimization of Xiangyun yarn that energy consumption has been reduced to a certain extent; in
Table 3: Evaluation index of the regression model.

| Algorithm evaluation index | Calculation formula | Description |
|----------------------------|---------------------|-------------|
| Mean square error          | $MSE = \frac{1}{L} \sum_{i=1}^{L} (y_i - y'_i)^2$ | The mean of the sum of squares of the difference between actual data and predicted data; this value is used to measure the "average error" of the forecast data. |
| Mean absolute error        | $MRE = \frac{1}{L} \sum_{i=1}^{L} |y_i - y'_i|/y_i$ | The mean value of the sum of the absolute values of the difference between actual data and predicted data; the main advantage of this evaluation index is that it can overcome absolute and relative errors, the situation where positive and negative cancel each other out. |
| Average relative error     | $MAE = \frac{1}{L} \sum_{i=1}^{L} |y_i - y'_i|$ | The difference between the actual data and the predicted data, and the mean value of the ratio of the absolute value of the actual data, the evaluation index is mainly used to reflect the mean value of the percentage of the absolute error to the true value. |

Figure 3: Comparison results of regression models.

Figure 4: Optimization curve.
order to a certain extent, the effect of energy saving and emission reduction has been achieved. In terms of optimization models, traditional genetic algorithms, differential evolution algorithms, and particle swarm optimization are used, compared with the proposed improvement based on the gradient boosting tree model and evolutionary algorithm for comparison. The comparison result is shown in Figure 6, which is the optimized result of Xiangyun yarn.

5. Conclusion

In recent years, people’s pursuit of quality of life has become higher and higher, especially in terms of ecological environment. In addition, the country vigorously strengthens the construction of ecological civilization, during the inspection of the general secretary in Anji, Zhejiang and puts forward the idea that green water and green mountains are golden mountains and silver mountains. As a printing and dyeing textile industry with high energy consumption and high pollution, energy saving and emission reduction have become urgent problem to be solved in the industry. Initially, the optimization of the technological process and the optimization of the scheduling of the workshop achieved a certain effect to a certain extent. However, with the rise and maturity of Internet of things and big data technology, it is possible for us to explore energy-saving and emission reduction methods of printing and dyeing enterprises from the perspective of data. The printing and dyeing processes contain a large amount of production data, including process parameters and energy consumption data, and these large amounts of data also contain some unknown laws. Obtained through experimental results, the average absolute error of the limit tree is 0.01391, and the average relative error is 0.072461, the average absolute error of the decision tree is 0.01979, and the average relative error is 0.062579. Based on the data of printing and dyeing process, combined with order related information, process parameter related information, and energy-related multimedia information, the energy consumption category and energy consumption are constructed as the predicted value. After certain data preprocessing, the energy consumption of the subsequent production process is predicted by using the optimization algorithm for model training, and the process parameters are adjusted to a certain extent by using the prediction data, so as to achieve the effect of energy conservation and emission reduction. Energy consumption classification and regression method of printing and dyeing shaping process is studied. Starting from the data of the finalization process of the printing and dyeing process, combined with order-related information, process parameter-related information, energy consumption-related multimedia information, etc., the gradient lifting tree model is used for the classification and regression prediction of the shaping machine data, and the effect is compared with some other methods through comparative experiments to prove the effectiveness of the method.

This paper mainly explores the energy consumption optimization of printing and dyeing enterprises, and there are some reasons posed by industry restrictions. First of all, it is the current situation of most printing and dyeing industry; the table, line, and other equipment are easy to damage, data are inaccurate, and refinement is not in place. Some enterprises, especially small enterprises, still use the equipment in the state of ten or twenty years ago. Considering the cost problem, they do not introduce too many new equipment. Many orders are still made with old equipment. It leads to high energy consumption and more serious pollution. As a printing and dyeing enterprise with high energy consumption and high pollution, only by sharing some relevant data can scientific research really play a role. Only by truly paying attention to relevant scientific research can we explore the value as soon as possible to better improve the output and quality of the daily production of enterprises.
This paper mainly explores the data of energy consumption optimization of printing and dyeing enterprises, in which there are some problems caused by industry restrictions. The first is the current situation of most printing and dyeing industries. The tables, lines, and other equipment on the machine are easy to be damaged, the data are inaccurate, and the refinement is not in place. Some enterprises, especially small enterprises, still use the equipment in the state of more than ten or twenty years ago. Considering the cost, they do not introduce too many new equipment. A large number of orders are still produced with old equipment, resulting in high energy consumption and serious pollution. The second is the problem of data. As it involves enterprise production, the confidentiality of data is also very important. In terms of data security and confidentiality, most small enterprises cannot make a balance and can only stay in the more original production mode. The enterprises that have just begun to explore production data have problems such as inconsistent data format, which leads to certain limitations in exploring secondary energy conservation and emission reduction from data.

**Data Availability**

The data used to support the findings of this study are available from the corresponding author upon request.

**Conflicts of Interest**

The authors declare that there are no conflicts of interest.

**References**

[1] Z. Zhou, R. Xu, D. Wu, Z. Zhu, and H. Wang, "Illumination correction of dyeing products based on grey-edge and kernel extreme learning machine," *Optik*, vol. 127, no. 19, pp. 7978–7985, 2016.

[2] C. Satish, "Jaya spider monkey optimization-driven deep convolutional lstm for the prediction of COVID-19," *Bio-Algorithms and Med-Systems*, vol. 16, no. 4, pp. 145–151, 2020.

[3] L. H. Zhang, "On optimizing the sum of the Rayleigh quotient and the generalized Rayleigh quotient on the unit sphere," *Computational Optimization and Applications*, vol. 54, no. 1, pp. 111–139, 2013.

[4] C. Liu, H. Lin, S. Gong, Y. Ji, and Q. Liu, "Learning topic of dynamic scene using belief propagation and weighted visual words approach," *Soft Computing*, vol. 19, no. 1, pp. 71–84, 2015.

[5] S. Narula, K. Shameer, A. M. Salem Omar, J. T. Dudley, and P. P. Sengupta, "Machine-learning algorithms to automate morphological and functional assessments in 2d echocardiography," *Journal of the American College of Cardiology*, vol. 68, no. 21, pp. 2287–2295, 2016.

[6] P. Charonytkakis, M. Plakia, I. Tsamardinos, and M. Papadopoulou, "On user-centric modular qoe prediction for voip based on machine-learning algorithms," *IEEE Transactions on Mobile Computing*, vol. 15, no. 6, pp. 1443–1456, 2016.

[7] P. Plawiak, T. Sosnicki, M. Niedzwiecki, Z. Tabor, and K. Rzecki, "Hand body language gesture recognition based on signals from specialized glove and machine learning algorithms," *IEEE Transactions on Industrial Informatics*, vol. 12, no. 3, pp. 1104–1113, 2016.

[8] H. Y. Tsao, P. Y. Chan, and C. C. Y. Su, "Predicting diabetic retinopathy and identifying interpretable biomedical features using machine learning algorithms," *BMC Bioinformatics*, vol. 19, no. 59, pp. 283–121, 2018.

[9] J. A. Atkinson, G. Lobet, M. Noll, P. E. Meyer, M. Griffiths, and D. M. Wells, "Combining semi-automated image analysis techniques with machine learning algorithms to accelerate large scale genetic studies," *GigaScience*, vol. 6, pp. 1–7, 2017.

[10] V. Prasad, T. S. Rao, and M. S. P. Babu, "Thyroid disease diagnosis via hybrid architecture composing rough data sets theory and machine learning algorithms," *Soft Computing*, vol. 20, no. 3, pp. 1179–1189, 2016.

[11] F. Azmat, Y. Chen, and N. Stocks, "Analysis of spectrum occupancy using machine learning algorithms," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 9, pp. 6853–6860, 2016.

[12] G. Bartsch, A. P. Mitra, S. A. Mitra et al., "Use of artificial intelligence and machine learning algorithms with gene expression profiling to predict recurrent nonmuscle invasive urothelial carcinoma of the bladder," *The Journal of Urology*, vol. 195, no. 2, pp. 493–498, 2016.

[13] C. Stoica, J. Camejo, A. Banciu et al., "Water quality of danube delta systems: ecological status and prediction using machine-learning algorithms," *Water Science and Technology*, vol. 73, no. 10, pp. 2413–2421, 2016.

[14] A. Awaysheh, J. Wilcke, F. Elvinger, L. Rees, W. Fan, and K. L. Zimmerman, "Evaluation of supervised machine-learning algorithms to distinguish between inflammatory bowel disease and alimentary lymphoma in cats," *Journal of Veterinary Diagnostic Investigation*, vol. 28, no. 6, pp. 679–687, 2016.

[15] J. Liu, Y. Ye, C. Shen, Y. Wang, and R. Erdelyi, "A new tool for cme arrival time prediction using machine learning algorithms: cat-puma," *The Astrophysical Journal*, vol. 855, no. 2, p. 109, 2018.

[16] M. A. H. Khan, B. Thomson, R. Debnath, A. Motayed, and M. V. Rao, "Nanowire-based sensor array for detection of cross-sensitive gases using pca and machine learning algorithms," *IEEE Sensors Journal*, vol. 20, no. 11, pp. 6020–6028, 2020.

[17] A. R. Verma, B. Gupta, and C. Bhandari, "A comparative study of ecg beats variability classification based on different machine learning algorithms," *Augmented Human Research*, vol. 5, no. 1, pp. 16–18, 2020.

[18] I. Mandal and S. Pal, "Modelling human health vulnerability using different machine learning algorithms in stone quarrying and crushing areas of Dwarka river Basin, Eastern India," *Advances in Space Research*, vol. 66, no. 6, pp. 1351–1371, 2020.

[19] M. Roimi, A. Neuberger, A. Shrot, M. Paul, and Y. Geffen, "Early diagnosis of bloodstream infections in the intensive care unit using machine-learning algorithms," *Intensive Care Medicine*, vol. 46, no. 3, pp. 454–462, 2020.

[20] A. R. Olivera, V. Roesler, C. Iochpe et al., "Comparison of machine-learning algorithms to build a predictive model for detecting undiagnosed diabetes - ela-brasil: accuracy study," *Sao Paulo Medical Journal*, vol. 135, no. 3, pp. 234–246, 2017.

[21] R. Rama Devi and M. Abualkibash, "Intrusion detection system classification using different machine learning algorithms on kdd-99 and nsl-kdd datasets—a review paper," *International Journal of Computer Science and Information Technology*, vol. 11, no. 03, pp. 65–80, 2019.
[22] B. W. Carter, “Machine learning algorithms utilizing functional respiratory imaging may predict COPD exacerbations,” Academic Radiology, vol. 26, no. 9, pp. 1200-1201, 2019.
[23] I. Mandal, “Machine learning algorithms for the creation of clinical healthcare enterprise systems,” Enterprise Information Systems, vol. 11, no. 6-10, pp. 1-27, 2016.
[24] J. Sherman, “Db exploring how machine-learning algorithms might assist decision-making,” The Pentagon, vol. 33, no. 25, pp. 12-13, 2017.
[25] B. Ashuri and J. Wang, “Predicting enr construction cost index using machine-learning algorithms,” International Journal of Construction Education and Research, vol. 13, no. 1, pp. 47–63, 2017.
[26] H. I. Shousha, A. H. Awad, D. A. Omran, M. M. Elnegouly, and M. Mabrouk, “Data mining and machine learning algorithms using il28b genotype and biochemical markers best predicted advanced liver fibrosis in chronic hepatitis c,” Japanese Journal of Infectious Diseases, vol. 71, no. 1, pp. 51–57, 2018.
[27] M. Fatima and M. Pasha, “Survey of machine learning algorithms for disease diagnostic,” Journal of Intelligent Learning Systems and Applications, vol. 09, no. 01, pp. 1–16, 2017.
[28] A. A. Jamali, R. Ferdousi, S. Razzaghi, J. Li, R. Safdari, and E. Ebrahimie, “Drugminer: comparative analysis of machine learning algorithms for prediction of potential druggable proteins,” Drug Discovery Today, vol. 21, no. 5, pp. 718–724, 2016.
[29] F. Sun, G. B. Huang, Q. M. Jonathan Wu, S. Song, and D. C. Wunsch II, “Efficient and rapid machine learning algorithms for big data and dynamic varying systems,” IEEE Transactions on Systems, Man, and Cybernetics: Systems, vol. 47, no. 10, pp. 2625-2626, 2017.
[30] T. Lee, K. B. Lee, and C. O. Kim, “Performance of machine learning algorithms for class-imbalanced process fault detection problems,” IEEE Transactions on Semiconductor Manufacturing, vol. 29, no. 4, pp. 436–445, 2016.