DOI: 10.21276/sjet

Scholars Journal of Engineering and Technology (SJET)
Sch. J. Eng. Tech., 2017; 5(8):407-412

(An International Publisher for Academic and Scientific Resources)
www.saspublisher.com

ISSN 2321-435X (Online)
ISSN 2347-9523 (Print)

Research Article

A New Method to Forecast of Office Building Electricity Consumption: Genetic Algorithm - Neural Network

Yuanyuan Hao
School of Mathematics and Statistics, Qingdao University, Qingdao 266071, China

***Corresponding author**
Yuanyuan Hao
Email: 1193976187@qq.com

Abstract: Saving energy is becoming a trend in today's society, office as our research object, the power consumption is very huge. We can effectively predict and track the office electricity situation is more and more important. In this paper, we play the data advantage, use the neural network to do training, get the corresponding forecast value. In order to overcome the shortcomings of the slow convergence of the neural network and easy to fall into the local error, we combine the genetic algorithm to filter the weights and thresholds of the input layer and the hidden layer to speed up the training accuracy and efficiency. Through the actual data validation, we get the training error within 5.5%, the robustness is strong, and the genetic algorithm after the combination of time increased by 31%. Effectively proving the reliability of our model.
Keywords: energy saving, power consumption, neural network.

INTRODUCTION

The significance of the study

Today, with the rapid development of science, technology and society, people's living standards are increasing. The requirements about living conditions and working environment are getting higher and higher [1]. So the surrounding buildings continue to be update, whether residential or public buildings, its scale gradually expanded, the function gradually improved, following with the energy consumption is improved. On earth, the available resources are limited, so it is important for us to understand the consumption of resources in our daily lives. Different types of buildings are different in energy consumption. Accurately predict the energy consumption of one type of buildings, can help us tap the potential of building energy efficiency, reducing the cost of living, contributing to save the earth resources.

Research Status

According to China's current situation, housing construction scale is huge, but from the building enclosure and the composition of the house can be found within the situation about high energy consumption. Statistics show that [2], China's office buildings, educational institutions, shopping malls and other public buildings energy consumption is 10-20 times the civil buildings. It can be seen that the prediction of energy consumption in public buildings is the key to our work. According to the literature [2-4] research income, in China's Guangzhou, Hangzhou, Changsha and other places, in public buildings, shopping malls due to complex internal structure, partition function is different, whether in the whole air system or lighting system, the energy consumption is far more than other public buildings. In this regard, many of the literature to be studied. And the office buildings are related to the working environment, it has great energy-saving potential. As an object of this paper, we will study the energy consumption characteristics of office buildings in Shandong Province, and use the model to forecast them.

PRINCIPLE

K-means Clustering Algorithm

K-means clustering belongs to unsupervised learning [5]. There is no given y in the clustered sample, the purpose of clustering is to find each sample x the potential category y , and the sample x are make together of the same category y, in the clustering problem, we give the training sample is $\{x^{(1)},\cdots,x^{(m)}\}$, that $x^{(i)} \in R^n$ and there is no y.

K-means algorithm is to cluster the samples into y clusters (cluster), the specific algorithm described below [6]:

1. Randomly select a number of k clusters of centroid points $\mu_1, \mu_2, \cdots \mu_k \in R^n$;
2. Repeat the following procedure until it converges {

For each sample i, and calculate the class it should belong to $c^{(i)} := \arg\min_j \left\| x^{(i)} - \mu_j \right\|^2$.

For each class j, recalculate the centroid of this class

$$\mu_j := \frac{\sum_{i=1}^{m} 1\{c^{(i)} = j\} x^{(i)}}{\sum_{i=1}^{m} 1\{c^{(i)} = j\}}.\}$$

Where, K is the number of clusters we have given in advance, $C^{(i)}$ represents the class i and the class closest in the class k number of classes. The value of $C^{(i)}$ is one of 1 to k, the centroid u_j represents the guess of the center of the sample that belongs to the same class.

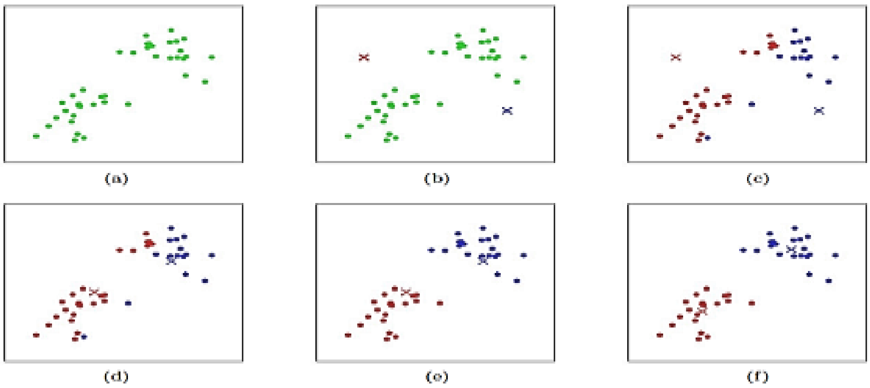

Fig-1: The Principle Structure of Cluster Analysis

BP neural network

BP neural network is a mathematical model which simulates the behavior of human neural network and carries on distributed parallel information processing algorithm. This network relies on the complexity of the system, by adjusting the relationship between a large number of nodes within the connection, so as to achieve the purpose of dealing with information. Its main working principle is the signal forward transmission, the error back propagation, in the forward transmission, the input signal from the input layer through the hidden layer processing, until the output layer Y. The basic principles can be further elaborated by the following figure.

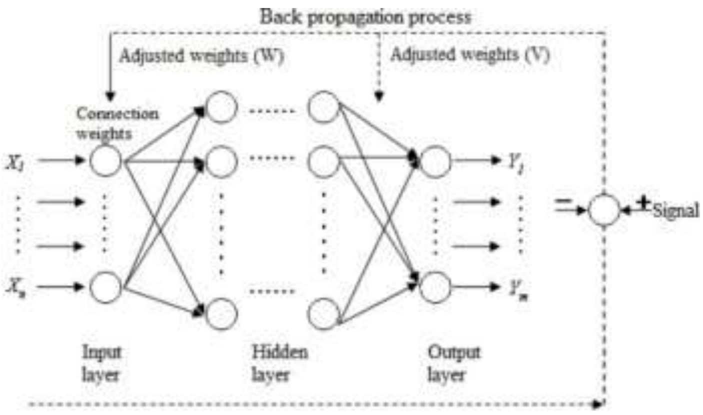

Fig-2: The Principle of BP Neural Network

The figure shows $X = (X_1, X_2 \cdots X_n)$ is the input layer, ω_{ij}, ω_{jk} is the weight coefficient,

a_j, b_k is threshold learning, rate is η and neuron activation function $\varphi(x)$, we chose the activation function $\varphi(x) = \frac{1}{1+e^{-x}}$.

Genetic algorithm

Genetic algorithm is an evolutionary algorithm that simulates the process of population evolution [7]. Genetic algorithms are usually implemented as a computer simulation. For an optimization problem, a certain number of candidate solutions (called individuals) can be abstractly expressed as chromosomes, allowing the population to evolve to better solutions. Traditionally, it is usually expressed in binary (that is, strings 0 and 1), but other representations can also be used. Evolution begins with a completely random individual population, followed by generation after generation. In each generation, the fitness of the whole population was evaluated, and a number of individuals (based on their fitness) were randomly selected from the current population, and a new life population was generated by natural selection and mutation. The population became the next iteration of the algorithm Current population [8].

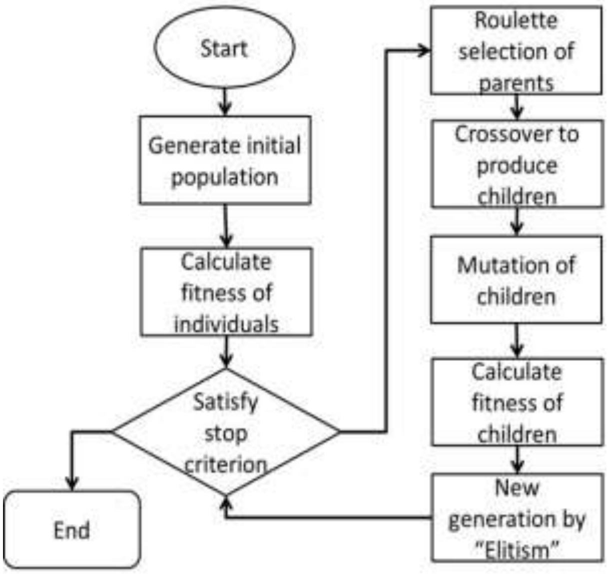

Fig-3: Genetic algorithm flow chart

ESTABLISH A FORECAST MODEL

According to the previous algorithm theory and practical problems, we choose the neural network to do the training of data. In order to overcome the shortcomings of neural network which is easy to fall into the local optimum, we propose a neural network algorithm based on genetic algorithm based on the characteristics of genetic algorithm.

Index selection

The energy consumption of an office building is mainly electricity, natural gas, artificial gas. Through the investigation and analysis of a large number of office buildings, we found that the impact of energy consumption factors is mainly building area, office space per unit of rent, total operating income, window design, lighting system design, the number of central air conditioning, latitude, Quantity, the number of ventilation equipment, the level of energy-saving equipment, occupancy rate. We choose the unit area of electricity (EPUA).

Because the data exists collinearity and repeatability, we choose to use K-mean to do clustering, and selecting the appropriate variables as an indicator. The following is a clustering result.

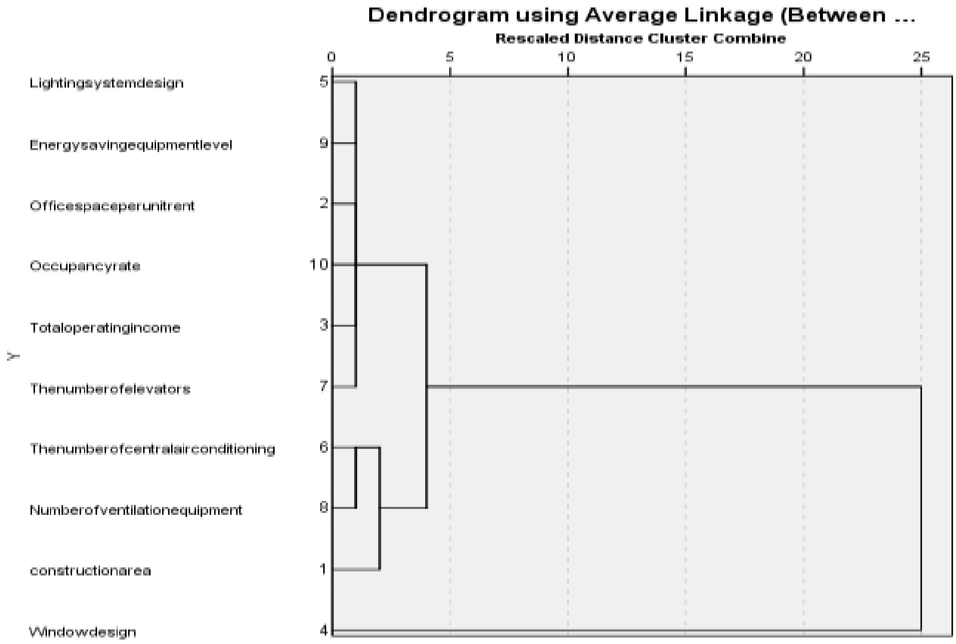

Fig-4 Clustering results

Data preprocessing

For the method of handling the vacancy value: ignore the record; remove the attribute; fill in the empty value manually; use the default value; use the attribute average; use the same sample average; predict the most likely value. Methods of processing noise data: sub-box; clustering; computer and manual inspection combined [9].

(1) Data consistency processing

In view of the collected data, because some office public data is incomplete, they cannot give specific data, only gives the data in which interval, in order to facilitate the subsequent calculation of the data, to the data consistency. So we choose the maximum method to process the data.

(2) Eliminate the dimension of the data

We collect the data, such as occupancy rate, construction area, etc., their units are not unified, in order to facilitate comparison, we choose the standardized method to eliminate the scale.

Set the office building energy consumption and the impact of the original data is $x_{ij}(i=1,2,\cdots,n,\ j=1,2,\cdots,m)$, the new standardized index value is:

$$x_{ij}^{*}=\frac{x_{ij}-\overline{x}_j}{s_j}\in[0,1] \qquad (1)$$

Where, $\overline{x}_j=\frac{1}{n}\sum_{i=1}^{n}x_{ij}$ is average value that the $j-th$ indicator relates to n rated objects, $s_j=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_{ij}-\overline{x}_j)^2}$ is mean square error that the $j-th$ indicator relates to n rated objects.

Establish a forecasting model

In view of the genetic algorithm is a probabilistic adaptive iterative optimization process, which is with good global search performance and not easy to fall into the local minimal. Even if the definition of the fitness function is not continuous or irregular, it can have large probability to find the overall optimal solution, and parallel processing, so it has characteristics that searches does not depend on the gradient information, which can be used to optimize the BP neural network.

The genetic algorithm is used to optimize the initial weight threshold of BP neural network, search for a large range replaces the general selection of the initial weight, and then apply the BP algorithm to fine tune the network in this solution space, thus search for the optimal solution or approximate optimal solution. It is not only achieved the complementary advantages of both, but also does it play a wide range of neural network mapping capabilities and genetic algorithm global search capabilities, which speeds up the network learning speed and improve the learning process of the approximation ability and generalization ability [10].

The mathematical description of the improved genetic algorithm for BP neural network is as follows:

$$\begin{cases} E_1(v,w,\zeta,\eta) = \frac{1}{2}\sum_{i=1}^{M_1}\sum_{t=1}^{m}\left[y_i(t)-\hat{y}_i(t)\right]^2 \\ s,t,w \in R^{m\times n}, v \in R^{p\times n}, \zeta \in R^p, \eta \in R^n \end{cases} \quad (2)$$

Where: E_1 is the total error of the network training sample, $y_i(t)$ is the ideal output signal, $\hat{y}_i(t)$ is the actual output of the network:

$$\hat{y}_i(t) = f\left\{\sum_{i=1}^{m} v_{it} g\left[\sum_{j=1}^{p} w_{ij}x_j(t)-\zeta_i\right]\right\}-\eta_t \quad (3)$$

Where: w_{ij} is the connection weight of the input layer node to the hidden layer node, v_{it} is the connection weight of the hidden layer node to the output layer node, ζ_i is the output threshold of the process neurons, η_t is the output neuron threshold, g is the excitation function of hidden neurons, f is the excitation function of the output neuron.

Set $E_2 = \frac{1}{M-M_1}\sum_{k=M_1}^{M}\sum_{t=1}^{m}\left[y_i(t)-\hat{y}_i(t)\right]^2$ is the average mean square error of the sample, indicating the reliability of the network output data. In order to make the design of the network has a good generalization ability, which should make E_2 less than a given error ε, while meets the reliability of network output.

The weight correction is expressed as:

$$\Delta w_{ij}(n) = -\eta g(n) + \alpha\Delta w_{ij}(n-1)$$
(4)

Where: η is the adaptive learning rate, $g(n)$ is the gradient of the current error function to the weight, α is the momentum factor, n is the number of iterations.

The specific steps of the improved BP neural network based on genetic algorithm are as follows [11]:

1. The initial BP neural network structure is established, the network related parameters and functions are set, and the initial population is generated by coding. The optimal weight threshold is obtained by optimizing the population, selection, crossover, mutation and so on.
2. Optimizing the BP neural network with the optimal weights and thresholds, and get the best BP neural network structure to meet the requirements.

The design flow chart is shown in the figure.

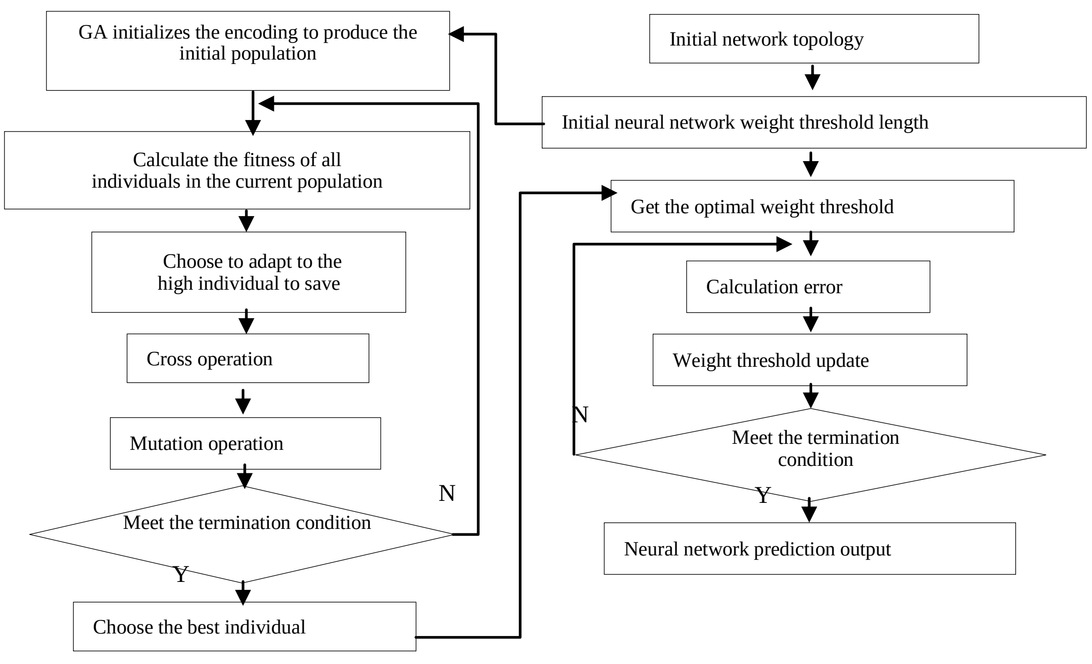

Fig-5: Model structure diagram

EXAMPLE AND VERIFICATION

We randomly selected two office buildings from all the offices in 17 cities. Respectively, office buildings C in Linyi and offices D in Yantai. According to prepared training in our selected indicators. Two cities 9 months of forecast data and the actual data and the error were as follows. The actual data and the forecast data and the error of the two cities in 9 months are as follows.

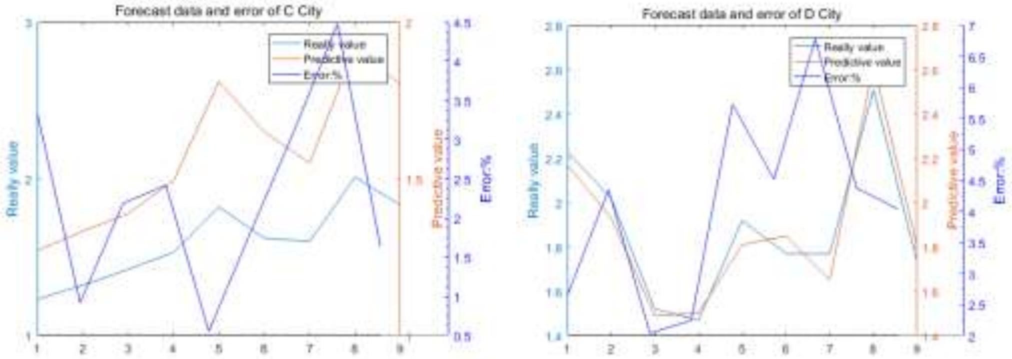

Fig-6: Predictive results and errors

According to the data in the figure, the average error for each group is 2.27% and 4.08%, respectively. In order to test the robustness of the structure, we have a number of groups of study and training, in the number of hidden layer 8 under the premise, the average error is less than 5.5%.

Under the premise of introducing the genetic algorithm, our algorithm is compared with the simple neural network. The final result shows that the average time is 3.86s and the speed is improved by 31%. Indicating that our model is efficient.

CONCLUSION

In order to optimize the management and to save energy as the starting point, this paper realizes the advance forecast of the power consumption, and prepares the follow-up energy-saving equipment improvement and scientific management [12].

From the simplicity and reliability, the data is input layer, after several training and training, and finally established a stable prediction model [13]. To achieve the mapping between the data and to achieve a low sensitivity, the actual operation is important. The final experimental results show that the average error rate of the new algorithm is 4.53%, and the learning time is obviously shortened by 31%. It is proved that our algorithm is effective for predicting the electricity consumption, which can improve the efficiency of management and save a lot of cost.

REFERENCES

1. Zhang T, Siebers PO, Aickelin U. Modelling electricity consumption in office buildings: An agent based approach. Energy and Buildings. 2011; 43(10): 2882-2892.
2. Yibi J. Based on multi-level energy consumption model of public building visual energy consumption monitoring platform key technology. Zhejiang University, 2012.
3. Xiaoping D, Ying L. Energy consumption status and energy saving analysis of public buildings in Hangzhou. Building Science. 2010; (8): 36-39.
4. Xiangxiang Z, Jun R, Shurong Y. Survey and Measurement of Energy Consumption in Office Buildings in Guangzhou. Architecture and Urban Physical Environment in the Process of Urbanization: Proceedings of the 10th National Conference on Building Physics.
5. Hartigan JA, Wong MA. Algorithm AS 136: A k-means clustering algorithm. Journal of the Royal Statistical Society. Series C (Applied Statistics). 1979; 28(1): 100-108.
6. Kanungo T, Mount DM, Netanyahu NS. An efficient k-means clustering algorithm: Analysis and implementation. IEEE transactions on pattern analysis and machine intelligence. 2002; 24(7): 881-892.
7. Whitley D. A genetic algorithm tutorial. Statistics and computing; 1994; 4(2): 65-85.
8. Houck CR, Joines J, Kay MG. A genetic algorithm for function optimization: a Matlab implementation. Ncsu-ie tr. 1995; 95(09).
9. Wu CL, Chau KW, Fan C. Prediction of rainfall time series using modular artificial neural networks coupled with data-preprocessing techniques. Journal of Hydrology. 2010; 389(1): 146-167.
10. Fu Z, Mo J. Springback prediction of high-strength sheet metal under air bending forming and tool design based on GA–BPNN. The International Journal of Advanced Manufacturing Technology, 2011; 53(5): 473-483.
11. Montana D J, Davis L. Training Feedforward Neural Networks Using Genetic Algorithms. IJCAI. 1989; 89: 762-767.
12. Jones P, Lockwood A. The management of hotel operations. Cengage Learning EMEA, 2002.
13. Cano EL, Moguerza JM, Ermolieva T. Energy efficiency and risk management in public buildings: strategic model for robust planning. Computational Management Science. 2014; 11(1-

2): 25-44.