Power Grid Investment Demand Forecasting Model Based on Data Mining

Yunhua Cong[1], Wen Xiang[2], Ying Wang[3], Feng Zuo[4], Yan Zhao[5], Bing Gu[6*]

[1]email: 282183442@qq.com, [2]email: 243789581@qq.com, [3]email: bo831129@126.com,

[4]email: fengheqi@126.com, [5]email: 502986464@qq.com, [6*]email: 459368167@qq.com

[1,2,3,4,5]State Grid Heilongjiang Electric Power Co., Ltd. Economic and Technological Research Institute, China

[2]School of Electrical and Information Engineering, Northeast Agricultural University, China

[3]School of Intelligent Science and Engineering, Harbin Engineering University, China

[6*]School of Economics and Management, Northeast Electric Power University, China

Abstract—Since the introduction of the "carbon peak" and "carbon neutral" action plans in 2020, the investment trend needs the joint sustained efforts of the supply side and the demand side. While the supply side adjusts the structure, the demand side also needs to make corresponding responses. Therefore, the future development of power enterprises should focus on the investment strategy and growth path within the key scope of the power grid. This paper identifies the key elements affecting power grid investment based on data mining, and constructs a power grid investment requirement forecasting model stemmed from intelligent mining algorithm, which provides a new method for power grid investment requirement calculating.

Keywords-Data mining; Power grid investment demand forecast; Nuclear principal component analysis; Support vector machine

1 Introduction

At present, the transformation investment of power grid enterprises is also facing many challenges in order to realize the investment transformation of "carbon peaking and carbon neutralization". Firstly, not only many external factors interfere with the investment of power grid enterprises, but also with the state's increasingly strict supervision on the investment of central enterprises and the pricing review of transmission-distribution price, resulting in the increasing pressure on the operation and investment of power grid enterprises at this stage [1-2].Secondly, when making investment, it is necessary to balance the relationship between long-term development and short-term demand, as well as the relationship between investment demand and investment capacity [3-4]. In contrast, more and more uncertain factors also interfere with power grid investment.Therefore, the corresponding value evaluation and prediction model is used to quantitatively analyze the value generated by power grid investment [5-6], so it is very necessary to form new businesses and new formats characterized by high investment value.

FFIT 2022, October 28-30, Shenzhen, People's Republic of China

DOI 10.4108/eai.28-10-2022.2328464

2 Mining key influencing factors of power grid investment demand

In order to better realize the accurate prediction of power grid investment requirement, constructing the input vector of power grid investment requirement prediction model is the fundamental purpose to select the impact of elements of power grid investment requirement and provide strong support for it. When selecting the impact of elements of power grid investment requirement, it is necessary to make a comprehensive, scientific and reasonable selection, and also consider the authenticity and availability of investment data. It is not only necessary to ensure that the selected factors are closely related to the power grid investment demand, but also to effectively reflect the changing trend of the power grid investment demand. According to the selection principles of influencing factors listed above, the specific selection factors are shown in Table 1:

Table 1. Influencing elements of power grid investment requirement

Category	Factors
Macroeconomic factors	Gross domestic product
	The industrial structure
	population
	Urbanization rate
	Household consumption index
	Investment in fixed assets
Power demand factors	Electricity consumption of the whole society
	Power grid sales
	Maximum load
	In the load
	Maximum peak-valley difference of load
Grid scale factors	220KV and above transmission line length
	Length of 110KV and below transmission line
	220KV and above transformer equipment capacity
	Capacity of 110KV and below substation equipment
Grid benefits	Unit power grid investment increase load supply
	Unit grid investment to increase electricity sales
	Power supply load per unit of grid assets
	Electricity sold per unit of grid assets
	Income from electricity sales of unit grid assets

3 Identification of key influencing factors based on data mining

3.1 Grey Relation Analysis

Grey relation analysis (GRA) is an analytical method that determines the correlation of various factors based on the similarity of their development trends. The specific operation procedures are as follows:

First, the subsequence and parent sequence are set, and dimensionless processing is performed for them, and the absolute value at time t is $\Delta_{0i}(t) = |x_0(t) - x_i(t)|, (t = 1,2,\cdots,n)$.

Secondly, the correlation degree is calculated, $\xi_{0i}(t)$ is correlation coefficient, Δmax is the maximum of the absolute value distinction, Δmin is the minimum of the absolute value distinction and k is the resolution coefficient.

$$\xi_{0i}(t) = \frac{k\Delta max + \Delta min}{k\Delta max + \Delta_{0i}(t)} \tag{1}$$

Finally, the correlation degree was solved:

$$\gamma_{0i} = \frac{1}{n}\sum_{i=1}^{n}\zeta_{0i}(t) \tag{2}$$

3.2 Kernel Principal Components Analysis

Principal Components Analysis (PCA) is to project high-dimensional data to low-dimensional level by means of projection on the basis of ensuring the amount of information of original data, and then convert multiple indicator variables of relevant original data into a small number of unrelated comprehensive indicators. The eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_m$ with the cumulative contribution rate between 85%~95% correspond to the first, second, ⋯, n principal components one by one. However, the PCA only performs linear operations, and can not analyze the non-linear relationship in the data. The kernel principal component analysis (KPCA) can map the original data to the high-dimensional feature blank through nonlinear mapping by introducing different kernel functions, and then extract the principal components by using the traditional principal component analysis technique, so as to transform the nonlinear problems that cannot be solved by the traditional principal component analysis method into linear problems, so as to achieve the goal of dimension reduction. The specific operation steps are as follows: Set data sample$X = [x_1, x_2, \cdots, x_n], x_i = [x_{i1}, x_{i2}, \cdots, x_{ip}], i = 1,2, \cdots, n.p$is the number of samples.The covariance matrix of the sample is:

$$\frac{1}{n}\sum_{j=1}^{n} x_j^T x_j \tag{3}$$

Where n is the number of samples. Set nonlinear mapping $\Phi_x \rightarrow F, F$ consists of $\Phi(x_1)$, $\Phi(x_2)$, ⋯, $\Phi(x_n)$, At the same time, the kernel function is selected as:

$$K = \Phi(X)^T \Phi(X) = \left[k(x_i, x_j)\right]_{n\times n} \tag{4}$$

$$k(x_i, x_j) = \langle \Phi(x_i)^T, \Phi(x_j) \rangle = \Phi(x_i)^T \Phi(x_j) \tag{5}$$

Thus, the covariance matrix C in the feature space is solved. The formula of the covariance matrix is as follows:

$$C = \frac{1}{n}\sum_{j=1}^{n} \Phi(x_j)^T \Phi(x_j) \tag{6}$$

Thus, the solution equation of PCA in the feature space is:

$$\lambda V = CV \tag{7}$$

Where: λ is the characteristic value; V is the eigenvector, $V \in F\{0\}$.Since V belongs to the space generated by $\{\Phi(x_1), \Phi(x_2), \cdots, \Phi(x_n)\}$,so:

$$\lambda(\Phi(x_k) \cdot V) = \Phi(x_k) \cdot CV (k = 1, 2, \cdots, n) \tag{8}$$

And there is parameter $\alpha = \{\alpha_1, \alpha_2, \cdots, \alpha_n\}$, make V represented by $\Phi(x_k)(k = 1, 2, \cdots, n)$, namely:

$$V = \sum_{j=1}^{n} \alpha_i \Phi(x_i) \quad (9)$$

Combine formula (8) and formula (9) to obtain:

$$\sum_{j=1}^{n} \alpha_i [\Phi(x_k)\Phi(x_i)] \quad (10)$$
$$= \frac{1}{n} \sum_{i=1}^{n} \alpha_i \left[\Phi(x_k) \sum_{j=1}^{n} \Phi(x_j)\right] \left[\Phi(x_j)^T \Phi(x_i)\right]$$

Bring in the kernel function to obtain the equivalent form:

$$n\lambda\alpha = K\alpha \quad (11)$$

The eigenvalues and eigenvectors corresponding toK are obtained. Finally, the kernel principal components are extracted by PCA.

3.3 Mining key influencing factors based on GRA-KPCA algorithm

For the 20 elements impact power grid investment demand selected above, in order to ensure the accuracy of prediction, data mining technology is used, that is, the mining algorithm based on GRA-KPCA is used to excavate the central elements factors, so as to obtain the effective input vector of the prediction model.

Firstly, the data of power grid investment demand and its influencing factors are standardized, and the gray correlation degree between each influencing element and the time series of power grid investment demand is calculated by using gra. The gray correlation degree is sorted according to the size, and the influencing elements with a gray degree of association greater than 0.85 are selected for the first time as the primary influencing factors of power grid investment demand.

Secondly, the main influencing factors preliminary selected by GRA will be used as the input vector, x to F conversion by mapping Φ, Get the kernel matrix K, then determine the first L eigenvalues and their eigenvectors whose aggregate variance dedication rate is greater than 95%.

Finally, PCA was used to extract the nuclear principal components Y, it is additionally the import vector of the power investment requirement forecasting model, and its normalizing computed outcome is Z.

4 Construction of power grid investment requirement forecasting model based on Intelligent Mining Algorithm

4.1 Support Vector Machine

The support vector machine (SVM) model projects the information in the import space into the high-dimensional feature blank through nonlinear reflecting, and then performs linear regression. Further, the problem to be solved is reduced to a convex quadratic programming problem with linear constraints.

Define the training set (x_i, y_i) with number n and the nonlinear mapping as $\psi(x)$:

$$\{(\mathrm{x_i}, \mathrm{y_i}) | \mathrm{i} = 1,2,\cdots,\mathrm{n}\} \quad (12)$$

$$\psi(x) = \left(\varphi(x_1), \varphi(x_2), \cdots, \varphi(x_n)\right) \quad (13)$$

The two accomplish the reflecting from the specimen import blank R^d to the high-dimensional feature R^k through the linear regression function, namely:

$$f(x_1) = \omega^T \varphi(x_t) + b \quad (14)$$

Where ω is the weight coefficient vector in the high-dimensional blank, and b is the offset of the model. The model needs to find the optimal ω, b in order to minimize the structural risk. At the same time, during the calculation of model parameters, SVM generally applies the criterion of construction hazard minimization, namely:

$$\min J = \frac{1}{2}\|\omega\|^2 + c \cdot \sum_{i=1}^{n}(\zeta_i + \zeta_l^*)$$
$$\text{s.t.}\begin{cases} y_i - \omega^T\varphi(x_i) - b \le \varepsilon + \zeta_i \\ \omega^T\varphi(x_i) + \}b - y_i \le \varepsilon + \zeta_i \\ \zeta_i, \zeta_l^* \ge 0 (i = 1,2,\cdots,n) \end{cases} \quad (15)$$

Where, c is the regularization modulus in the majorization procedure, and ζ_i、ζ_l^* are the slacken factors to readjust the construction risk. Next, the Lagrange multiplier is introduced:

$$L\begin{pmatrix} \omega \\ \xi_l \\ \xi_i^* \\ \alpha \\ \alpha^* \\ c \\ \beta \\ \beta^* \end{pmatrix} = \begin{matrix} \frac{1}{2}\|\omega\|^2 + c \cdot \sum_{i=1}^{n}(\zeta_i + \zeta_l^*) - \\ \sum_{i=1}^{n} \alpha_l[\omega^T\varphi(x_i) + b - y_i + \varepsilon + \zeta_i] - \\ \sum_{i=1}^{n} \alpha_l^* [y_i - \omega^T\varphi(x_i) - b + \varepsilon + \zeta_l^*] - \\ \sum_{i=1}^{n}(\beta_i\zeta_i + \beta_i^*\zeta_l^*) \end{matrix} \quad (16)$$

According to Karush-Kunhn-Tucker algorithm:

$$\begin{cases} \frac{\alpha L}{\alpha \omega} = 0 \rightarrow \omega = \sum_{i=1}^{n}(\alpha_l - \alpha_l^*)\varphi(x_l) \\ \frac{\alpha L}{\alpha b} = 0 \rightarrow \sum_{i=1}^{n}(\alpha_l - \alpha_l^*) = 0 \\ \frac{\alpha l}{\alpha \zeta_l} = 0 \rightarrow \omega = c - \alpha_l - \beta_l = 0 \\ \frac{\alpha L}{\alpha \zeta_l^*} = 0 \rightarrow \omega = c - \alpha_l^* - \beta_l^* = 0 \end{cases} \quad (17)$$

At this point, the kernel function of the stacking function used in the optimization process is defined as:

$$K(x_i, x_j) = \varphi(x_i)^T \varphi(x_j) \quad (18)$$

At this point, formula (15) can be converted into the following form:

$$maxW(\alpha,\alpha^*) = -\frac{1}{2}\sum_{i,j=1}^{n}(\alpha_i-\alpha_i^*)\cdot(\alpha_j-\alpha_j^*)\cdot K(x_i,x_j)+\sum_{i=1}^{n}(\alpha_i-\alpha_i^*)\,y_i - \sum_{i=1}^{n}(\alpha_i-\alpha_i^*)\varepsilon$$

$$s.t.\begin{cases}\sum_{i=1}^{n}(\alpha_i-\alpha_i^*)=0\\ 0\le \alpha_i,\alpha_i^*\le c\end{cases} \quad (19)$$

Lastly, the forecast formula of the model is as follows:

$$f(x)=\sum_{i=1}^{n}(\alpha_i-\alpha_i^*)\,K(x_i,x_j)+b \quad (20)$$

Where $K(x_i,x_j)$ is the radial basis function in the following form:

$$K(x_i,x_j)=exp\left(\frac{-\|x_i-x_j\|^2}{2\sigma^2}\right) \quad (21)$$

4.2 Grey Wolf Optimization

Gray wolf optimization (GWO) is an algorithm that imitate the population hierarchy and hunting behavior of wild gray wolves in nature. Taking wolves in nature as an example, wolves can be divided into four levels. The responsibilities of each level are shown in Table 2.

Table 2. Division of gray wolf population responsibilities

level	Duties and responsibilities
α	Directing the population, making hunting decisions
β	Dominate other wolves and assist in decision-makingα
δ	Obey and execute the decisions of higher authority, can be controlledω
ω	Follow orders and assist with the hunt

Assume that there is a gray wolf population with an individual number of M and a search space size of K. The identifier of gray wolf is i, their position is described as $x_i=(x_{i1},x_{i2},\cdots,x_{ik})$, and the process of wolves surrounding their prey is :

$$\mathrm{D}=\left|\mathrm{E}\cdot \mathrm{X_p(t)}-\mathrm{x(t)}\right| \quad (22)$$

$$\mathrm{x(t+1)}=\mathrm{x_p(t)}-\mathrm{B}\times \mathrm{D} \quad (23)$$

Where, D stands for the proximity between the prey and the wolf, $\mathrm{x(t)}$ is the location of the gray wolf after being changed and optimized, and B、E is the parameter vector of the model. The calculation formula for B、E is as follows :

$$\mathrm{B}=2\mathrm{a}\mathrm{r_1}-\mathrm{a} \quad (24)$$

$$\mathrm{E}=2\mathrm{r_2} \quad (25)$$

When the level α、β、δ in the wolf group is close to the prey, the location of the prey can be estimated. The gray wolf group updates the location iteratively according to the α、β、δ. In order to ensure the data accuracy, the trends change qualitatively operator and nonlinear convergence factor are leaded into the conventional gray wolf algorithm to change and transform the position of the gray wolf population combined with α、β、δ. The location update method is as follows:

$$X(t+1)=\frac{X_\alpha(t)+X_\beta(t)+X_\delta(t)}{3} \tag{26}$$

Accordingly, the convergence factor in formula (24) is improved:

$$\alpha=2-\left(e^{\frac{1}{t_{max}}}-1\right)\cdot\frac{2}{(e-1)} \tag{27}$$

At the same time, the differential evolution algorithm is used to improve the ameliorated gray wolf algorithm. The algorithm embodies three segments: mutation, intersection and choice. First, the population is redefined:

$$x_{ij}(0)=x_{ij}^L+rand(0,1)\left(x_{ij}^U-x_{ij}^L\right) \tag{28}$$

Mutate N initial populations in the random search space:

$$V_i(t+1)=x_{rl}(t)+F[x_{r2}(t)-x_{r3}(t)] \tag{29}$$

Then carry out cross activity to improve the abundance of the animal community:

$$U_{i,j}(t+1)=\begin{cases}V_{i,j}(t+1), rand(0,1)\le CR\\ X_{i,j}(t), otherwise\end{cases} \tag{30}$$

In accordance with greedy algorithm, select the preferable individuals after advance as the new generation of individuals:

$$X_{i,j}(t+1)=\begin{cases}U_i(t+1), f[U_i(t)]\le f[X_i(t)]\\ X_j(t), otherwise\end{cases} \tag{31}$$

The general idea of building the power grid investment demand prediction model is shown in Figure 1:

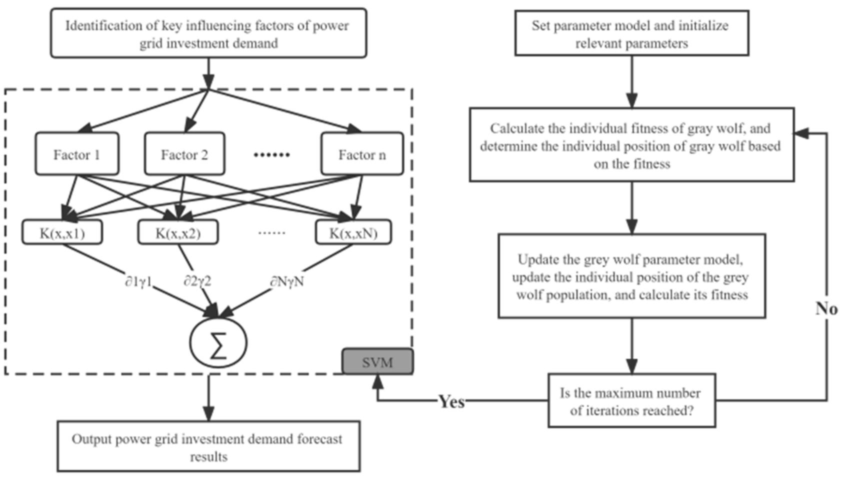

Figure 1. Application process of power grid investment demand forecasting model

5 Conclusions

Based on preliminary selection of influencing factors, this paper uses the influencing factor mining algorithm based on GRA-KPCA to confirm the key influencing factors of power grid investment demand, and obtains the effective import vector of prediction model. The constructed prediction model adopts SVM as the main prediction algorithm, and combines GWO algorithm to optimize the parameters of SVM, thereby improving the accuracy of the prediction model, providing a novel technique for the investment demand divination of power grid enterprises.

ACKNOWLEDGMENTS

The project of State Grid Heilongjiang Electric Power Co., Ltd. "Research on the theory, model and application of transmission and distribution investment demand prediction and investment optimization under the background of carbon emission reduction".

Note: Northeast Electric Power University and State Grid Heilongjiang Electric Power Co., Ltd. Economic and Technological Research Institute are in no particular order, and they are the co-first affiliation of this paper.

References

[1] Jang,H.Y;Du,E.S;Jin,C;Xiao,J.Y;Zhang,N.(2021)Multi time scale energy storage capacity optimization planning for multinational interconnected power systems with high proportion of clean energy grid connection.Chinese Journal of electrical engineering,41 (06):2101-2115.

[2] Huang,W;Liu,Q;Yang,S.W;Xiong,W.P;Liu,Z.F.(2017)Security situation awareness method based on power supply capability of active distribution system.Power automation equipment,37 (08): 74-80.

[3] Xiao,X;Wang,Z;Zhang,HS;(2020)Operational risk assessment of "China Pakistan Economic Corridor" power investment project. International economic cooperation, 06: 138-147.

[4] Wang,S.X;Liang,D;Ge,L.J;(2016)Key technologies of situation awareness and situation guidance for intelligent distribution network.Power system automation,40 (12): 2-8.

[5] GE L, LI Y, XIAN Y, et al.(2020) A FA-GWO-GRNN method forshort-term photovoltaic output prediction.2020 IEEE Power & Energy Society General Meeting (PESGM).

[6] Zhao,H.S;Ma,L.B;(2019) Big data compression of intelligent distribution network based on tensor Tucker decomposition. Chinese Journal of electrical engineering, 39 (16): 4744-47524976.