Application of Data Mining Technology in Precise Investment Analysis of Distribution Networks

Wei Han*, Yunfei Song, Ye Wan, Kai Song

{Wei Han*: 18762549920@139.com}
{Yunfei Song: sugang_19872@163.com}
{Ye Wan: yangle_101@163.com}
{Kai Song: wangzhijian_1112@163.com }

STATE GRID HUAIAN POWER SUPPLY COMPANY,Huaian,223001,China

Abstract. In order to understand the economic benefits of precise investment in distribution networks in enterprises, an application research of data mining technology in precise investment analysis of distribution networks has been proposed. Firstly, using data mining and analysis methods, establish a precise investment strategy for the distribution network through decision-making evaluation and goal achievement of project investment, improvement of power grid operation level, and economic benefit evaluation. Secondly, in the application of data mining technology to precise investment strategies in distribution networks, data mining is used to identify the economic value contained in the distribution network data of power enterprises. By evaluating project investment and achieving goals, we ensure the use of data mining technology to improve precision investment in distribution networks, maximize the benefits of power enterprise power grid systems, avoid decision-making errors in distribution network economic projects, and provide guarantees for the healthy development of China's power industry.

Keywords: data mining; Power grid; Precision investment.

1 Introduction

As the Chinese economy enters a new stage, the transformation of energy consumption is gradually slowing down, making it difficult for power companies to improve their performance. Digital power grid. In order to better adapt to the changes in the internal and external environment of the energy industry and ensure the full role of data in electronic grid decision-making, it is urgent to use data mining and monitoring technologies for close analysis and monitoring. The important aspect of controlling investment decisions is to establish a quality analysis process for investment decisions, strengthen investment decision-making processes, flexibility, and all conditions, quickly and clearly understand changes in national macro policies, economic development, and economic reforms, make effective energy consumption investment decisions in a timely manner, improve the accuracy of power grid usage, help build a "digital national power grid", and support the development of energy grids [1-2].

MSIEID 2023, December 08-10, Guangzhou, People's Republic of China

DOI 10.4108/eai.8-12-2023.2344735

2 Project Investment Evaluation Methods

2.1 Predicting the future open capacity of the power grid

The prediction of the future open capacity of the power grid is the basis for evaluating the feasibility of the project. It can not only predict the rationality of project investment, but also reflect the power supply capacity of the power grid. If the future open capacity of the project after implementation is greater than 0, the project can be invested; If the future open capacity of the project is less than or equal to 0 after implementation, the project should stop investment.The future open capacity of lines in the power grid is jointly determined by the limited capacity of the line and the maximum load of the line. The future open capacity is represented by H. The calculation formula is as follows (1):

$$H = B \times 85\% - \max L_{mon} \tag{1}$$

In the formula: B represents the capacity of the line after investment; Imom is the monthly line load. If H is bigger than 0, the capacity after project implementation will meet future load requirements.

2.2 Analysis of Current Situation and Matching Degree

Analyze the current status of relevant indicators that can reflect the operation level, transfer capacity, power grid structure, equipment level, and power supply quality of the power grid, and identify weak points. For distribution network investment projects, using the method of calculating the effective investment rate and repeated investment rate of the project can fill the gap in the previous pre evaluation of investment benefits of distribution network projects and reduce the occurrence of the phenomenon of "beating the head to make decisions".Match the results obtained from the current situation analysis and future open capacity prediction with the project objectives. Set the future open capacity to have a "one vote veto" and record it as C. If the future open capacity is bigger than 0, C is 1; If the future available capacity is less than or equal to 0, then C is 0. Record the effective investment rate as T, and for individual projects, the effective investment rate is as follows (2):

$$\mathrm{T} = \mathrm{C} \times \frac{K_1}{K_2} \times 100\% \tag{2}$$

In the formula, k1 is the number of weak indicators that the project can improve; k2 is the total number of weak indicators.

If the project construction goals do not match the indicators in weak links, that is, the effective investment rate of the project has not reached 100%, then calculate the duplicate investment rate of the project, record the duplicate investment rate as 0, and the calculation formula is as follows (3):

$$0 = \frac{L}{n} \times 100\% \tag{3}$$

2.3 Assessment of power grid operation level

The operational level mainly compares and analyzes the expected operational level, transfer capacity, power grid structure, equipment level, and power supply quality before and after the construction of power grid construction projects. The evaluation indicators for operational level should start from the entire power grid construction project and its scope of influence. All aspects, details, etc. involved should be considered and reflected in the pre evaluation indicator system. At the same time, it is necessary to purposefully select meaningful and important indicators as the focus of investigation, so that the pre evaluation indicator system has strong purpose and pertinence. Based on the above requirements, the establishment of a pre evaluation index system for the operation level of the power grid is shown in Table 1. This index system is applied to both the analysis indicators for current situation and matching analysis[3-4].

Table 1. Pre-assessment power grid operation index system

Primary index	Secondary index
Operation level	Heavy overload line ratio Heavy overload distribution transformer ratio line assembly distribution transformer capacity Line ratio with capacity greater than 12 MVA for assembling and distributing transformers. Average load rate of distribution transformer
Transfer capacity	Line "N-1" pass rate Passing rate of substation full stop verification
electric power network composition	Utilization ratio of 10 kV outlet interval 10 kV power supply radius greater than 3 km Line ratio Typical connection ratio of 10 kV
Equipment level	Cable rate of line Distribution transformer automation coverage
Power supply quality	Voltage qualification rate Power supply reliability

2.4 Economic Benefit Evaluation

The economic benefits are mainly reflected in the expected annual return rate, which is obtained by calculating the ratio of the expected annual return of the project to the investment amount of the project. The expected investment income of distribution network projects is mainly reflected in two aspects: firstly, increasing power supply income by improving the power supply capacity of the power grid; The second is to reduce line losses by optimizing the power grid structure to reduce power supply costs. Therefore, the indicators in Table 2 are used as the basis for evaluating economic benefits[5].

Table 2. Annual Rate of Return Assessment Indicators

Evaluation target	indicators for performance check
Annual yield	Annual increase in power supply revenue Annual loss reduction power gain Project investment amount
Investment return rate	Annual net profit of the project
Investment payback period	Time required for the project
net present value	Discounted future cash flows to present value
Internal rate of return	The discount rate at which the net present value of the project is zero

The calculation formula is (4):

$$R = (M_1 + M_2) / c \times 100\% \tag{4}$$

In the formula, M1 represents the income from increased electricity sales during the evaluation period; M2 represents the return on reduced electricity consumption during the evaluation period; c represents the amount invested in the project.

2.5 Post investment evaluation methods for projects

Analyze the degree of achievement of various goals in power grid construction projects, compare them with pre evaluation data, review the implementation status of engineering construction goals, macro goals, etc., and identify the reasons for not achieving or achieving low levels of achievement. The evaluation formula for the achievement of project goals is:

$$Z = \frac{m}{k_2} \times 100\% \tag{5}$$

In the formula, m represents the actual number of goals achieved by the project.

The operational level mainly refers to the analysis of the achievement of the goals of the power grid construction project, as well as the power grid operation level, transfer capacity, power grid structure, equipment level, and power supply quality before and after the project is completed, in order to evaluate the degree of achievement of the self construction goals and macro goals of the power grid construction project. The content structure and weight of the indicator system for post evaluation of power grid construction projects shall be consistent with the content structure and weight of the pre evaluation indicator system, and the results of the post evaluation indicator system and weight shall be consistent with those of the pre evaluation. By collecting indicator data for each time period before and after project construction, calculate the improvement or decrease of indicators after project construction, and weighted average the improvement or decrease of indicators with corresponding weights to calculate the degree of improvement in the operation level before and after project implementation[6].

3 Precise Investment Strategies for Distribution Networks under Data Mining

Apply the Delphi method to estimate the weight vector of the primary indicators in the indicator system $(0.23\,,0.25\,,0.15\,,0.17\,,0.2)^{\mathrm{T}}$, which correspond to the weights of operating level, transfer capacity, power grid structure, equipment level, and power supply quality.Based on the distribution network operation level indicator system of a certain city's power enterprise and the actual data collected from each grid, the indicator values of the grid at different periods are determined. The raw data composed of evaluation indicator values is scored and standardized, and the obtained data is shown in Table 3 (Table 3 only shows the standardized data for each quarter of a certain grid from 2020 to 2022)[7].

According to the formula of the legitimate weight method, calculate the legitimate value and legitimate weight of each evaluation index in the second level indicator layer for the first level indicator layer. Taking power supply quality and its second level indicator power supply reliability and voltage qualification rate as examples, calculate the legitimate value and legitimate weight.What are the standardized data weights for power supply reliability:$P_{11} = \frac{0.3}{0.3+0.3+0.4} = 0.3$; $P_{21} = \frac{0.3}{0.3+0.3+0.4} = 0.3$; $P_{31} = \frac{0.4}{0.3+0.3+0.4} = 0.4$.

Reliability rate of power supply:$e_1 = -\frac{1}{ln3}\sum_{i=1}^{3} P_{il}ln(P_{il}) = 0.9912$.

Similarly, the legitimate value $\mathrm{e}_2 = 0.9952$ of voltage qualification rate.

The difference coefficient $\mathrm{g}_1 = 0.0088$ of power supply reliability; The difference coefficient $\mathrm{g}_2 = 0.0048$ of voltage qualification rate[8].

The legitimate right of power supply reliability:$W_1 = \frac{0.0088}{0.0088+0.0048} = 0.64706$.The qualified rate of the voltage $\mathrm{W}_2 = 0.35294$.The evaluation index system of the operation level of the distribution network and the rights of the secondary indicators are shown in Table 3.

Table 3. Results after the standardization of a certain grid index

Primary indicators	Secondary indicators	2020	2021	2022
Operating level	Heavy overload line ratio	1.0	1.0	1.0
	Heavy overload distribution transformer ratio	0.9	1.0	1.0
	Line installation and distribution transformer capacity/MVA	0.5	0.4	0.2
	Ratio of lines with installed transformer capacity greater than 12 MVA	0.4	0.6	0.6
	Average load rate of distribution	0.9	0.9	0.9

	transformer			
Transfer capacity	Line "N-1" pass rate	0.1	0.4	0.4
	Pass rate of substation full stop verification	0.0	0.0	0.0
Power grid structure	10 kV outgoing line interval utilization rate	0.8	0.9	1.0
	Ratio of lines with a 10 kV power supply radius greater than 3 km	0.1	0.1	0.2
	10 kV typical connection rate	0.6	0.1	0.1
Equipment level	Line Cabling Rate	0.0	0.0	0.0
	Distribution automation coverage rate	1.0	1.0	1.0
Power supply quality	Power supply reliability	0.3	0.3	0.4
	Voltage qualification rate	0.8	1.0	1.0

The Table 4 index system is used to conduct the pre-evaluation of the three transformation projects of an electric power enterprise in a city, and the time point of the evaluation is before the project starts. The details of these 3 items are shown in Table 5.

Table 4. Evaluation index system of distribution network operation level and the rights of secondary indicators

Primary indicators	Secondary indicators	Information legitimacy	Legitimate power
Operating level(0.23)	Heavy overload line ratio	0.99	0.07204
	Heavy overload distribution transformer ratio	0.99	0.08144
	Line installation and distribution transformer capacity/MVA	0.9408	0.51536
	Ratio of lines with installed transformer capacity greater than 12 MVA	0.9755	0.21362
	Average load rate of distribution transformer	0.9865	0.11753
Transfer capacity(0.25)	Line "N-1" pass rate	0.9661	0.07652
	Pass rate of substation full stop verification	0.5913	0.92348

Power grid structure(0.15)	10 kV outgoing line interval utilization rate	0.9672	0.27630
	Ratio of lines with a 10 kV power supply radius greater than 3 km	0.9472	0.44417
	10 kV typical connection rate	0.9668	0.27953
Equipment level(0.17)	Line Cabling Rate	0.8809	0.65228
	Distribution automation coverage rate	0.9365	0.34772
Power supply quality(0.2)	Power supply reliability	0.9912	0.64706
	Voltage qualification rate	0.9952	0.35294

Table 5. Shows the 3 distribution network projects used for the evaluation

Planning reseau	Category	Project	Total line capacity / kVA	Estimated maximum load / kW	Amount of investment / first	Weakness	Expected improvement in key indicators/%
A	reform	1	11195	7252	2286549	For the public transformer overload and low voltage problems, improve the reliability of power supply	Heavy overload distribution transformer 4.57 decreased to 4 The voltage pass rate rose from 99.92 to 99.93 Power supply reliability of 99.93 rose to 99.99
A	reform	2	12280	8214	4399022	Improve the power supply reliability	Power supply reliability of 99.93 rose to 99.99
B	low pressure	3	19825	8540	4792327	For the public transformer overload, load imbalance, low voltage problems	4.95 decreased to 3.6 Voltage qualified rate rose from 99.93 to 99.96 Power supply reliability 99.90 increased to 99.94 The average load rate of the distribution transformer decreased from 49.18 to 42.8

First, by calculating the future open capacity of the investment line C:

C_1=11195×80%-7252＞0,C_2=12280×80%-8214＞0,C_3=19825×80%-8540＞0,

Obtain all lines with future open capacity greater than 0.

Evaluate the implementation status of three projects using the same method as the pre evaluation. The goal achievement rate for Project 1 is 100%, the goal achievement rate for Project 2 is 33.3%, and the goal achievement rate for Project 3 is 100%.Apply the indicator system and weights shown in Table 4 as the post evaluation indicator system and weights for power grid operation level, calculate the improvement degree of operation level of three projects before and after commencement, and compare the investment efficiency of the three projects[9]. The operating level before project 1 started was 0.359, and the level after project completion was 0.422, an increase of 0.063; The grid operation level before project 2 started was 0.664, and the level after project completion was 0.698, an increase of 0.034; The grid operation level before project 3 started was 0.355, and the level after completion was 0.62, an increase of 0.265. Although Project 1 and Project 2 belong to the same grid, Project 1 has brought significant improvement in the operation level of the power grid. Therefore, it is recommended to invest in Project 1 for Grid A[10].

4 Conclusion

From this, it can be seen that in the current precise investment strategy of power enterprises' distribution networks, the use of data mining technology and evaluation model analysis can effectively reduce or avoid errors in project investment decisions. The application of data mining technology in the precise investment strategy of distribution network is an important link in the precise management of distribution network projects. Constructing an investment evaluation model can effectively position the economic benefits of the distribution network after investment, provide guidance for the precise investment of power enterprises in the distribution network, and achieve the optimal construction of China's power grid system.

References

[1] Chen, H. , & Lian, Q. . (2021). Poverty/investment slow distribution effect analysis based on hopfield neural network. Future Generation Computer Systems, 122(11)45.

[2] Brehm, C. L. A. . (2021). Nestedness of eco-industrial networks: exploring linkage distribution to promote sustainable industrial growth. Journal of Industrial Ecology, 25(1)36.

[3] Yang, H. , & Zhang, W. . (2022). Data mining in college student education management information system. International Journal of Embedded Systems,36(3), 15.

[4] Zhou, W. , & Yang, T. . (2021). Application analysis of data mining technology in ideological and political education management. Journal of Physics: Conference Series, 1915(4), 042040 (7pp).

[5] Shima T A .(2023).The evaluation of wastewater treatment plant performance: a data mining approach.Journal of Engineering, Design and Technology(6),1785-1802.

[6] Wang, L. . (2021). Dynamic financial economic fluctuation model based on non-normal distribution. Journal of intelligent & fuzzy systems: Applications in Engineering and Technology, 40(2)74.

[7] Adegbola, A. A. , Ozigis, I. I. , & Muhammad, I. D. . (2021). Conceptual design of gas distribution pipeline network for estates in nigeria. Nigerian Journal of Technology, 4(1), 25-36.

[8] Wang, S. Y. , Wang, S. S. , Yu, X. , Xu, H. , Li, Y. Y. , & Jiang, H. Y. , et al. (2023). Design and analysis of a hybrid magnetic shielding system: application for the magnetic non-destructive testing of circuits. Superconductor Science and Technology, 36(3), 035001-.

[9] Liu, H. . (2021). Big data precision marketing and consumer behavior analysis based on fuzzy clustering and pca model. Journal of Intelligent & Fuzzy Systems: Applications in Engineering and Technology,36(4), 40.

[10] Olga L H .(2023).Sentiment, we-talk and engagement on social media: insights from Twitter data mining on the US presidential elections 2020.Internet Research(6),2058-2085.