Application of Clustering Algorithm by Data Mining in the Analysis of Smart Grid from the Perspective of Electric Power

Yaru Qi1,*, Junda Ren1, Ni Sun1, Yangqi Yu2
1Big Data Center of State Grid Corporation of China, Beijing 100052, China
2State Grid Info-telecom Great Power Science and Technology Co., Ltd., Fuzhou 350003, China

*Corresponding author e-mail: yaruqi@sgcc.com.cn

Abstract. In recent years, the research on the relationship between economic development and power consumption is also a focus of government departments at all levels. The development of electric power (EP) not only has an impact on power supply enterprises and EP industry, but also is closely related to the social and economic development of the whole region and residents’ life, that is, there must be some internal relationship between economic development and EP consumption. In many areas of our country, the situation of power consumption and economic development is not coordinated. Insufficient and untimely power supply will hinder economic growth, and excessive power supply will bring unnecessary waste of resources. This also reflects that the research on the relationship between economic development and power consumption is not in-depth, so the research on the relationship between power consumption and economic development has a certain theoretical significance. This paper studies the application of clustering algorithm based on data mining in the analysis of economic development characteristics (EDC) from the perspective of EP, uses K-means clustering algorithm to understand the relationship between EP development and economic development, understands its characteristic development application, studies the relationship between EP consumption and economy from the EP elastic coefficient method and output value unit consumption method, and uses K-means clustering algorithm to calculate, this paper uses the chart analysis method to analyze the correlation degree of power consumption structure and the relationship between different industrial GDP and power consumption. The results show that the total output of each industrial structure is in direct proportion to the power consumption. The total value of each industrial structure is increasing, and the power consumption is on the rise. For example, in 2016, the total value of the first industry was 186.81 trillion yuan, the second industry was 960.54 trillion yuan, the third industry was 515.96 trillion yuan, and the power consumption was 41.246 billion kwh, by 2020, the total value of the primary industry will be 220.66 trillion yuan, the secondary industry 1916.51 trillion yuan, the tertiary industry 866.22 trillion yuan, and its electricity consumption (EC) will be 57.697 billion kwh.
Keywords: Electric Power Perspective, Industrial Structure, Clustering Algorithm; Characteristics of Economic Development.

1. Introduction
As one of the important parameters in the barometer of national economic development, power big data carries economic identification signals, decision-making auxiliary value, policy guidance significance, etc., which are highly valued by the relevant departments of EP and economic decision-making departments [1-2]. Compared with the economic indicators such as the added value of the three major industries, trade deficit and unemployment rate, the power indicators are more real-time, objective and accurate [3-4]. In addition, there are two obvious advantages of using EP data to forecast economy: first, because EP products cannot be stored, it can avoid the lag of data statistics to the maximum extent; on the other hand, all aspects of modern economic system are obviously related to power products [5-6]. The change of EC can not only reflect the operation effect of economic policies, but also reflect the change trend of market economy [7-8]. Therefore, paying attention to and studying the relationship between power consumption and economic development, especially the impact of the growth or decline of certain quantitative factors between the relationship, can directly or indirectly grasp the production and operation status of all walks of life [9-10].

In the research on the application of clustering algorithm based on data mining in the analysis of EDC from the perspective of EP, many scholars at home and abroad have studied it and achieved certain results. Zhao x pointed out that data mining technology is an edge technology based on actual needs, which is mainly used to study the distribution characteristics of data objects and predict the change rules of data objects [11]. Deng J pointed out that the adjustment of the country's overall industrial structure has an impact on and constraints on the development of power economy, as well as opportunities and challenges for the development of the power industry. The market-oriented reform and development strategy of the power market are directly related to China's future energy strategic security and economic development situation, further research on power economic adjustment and power market development strategy will contribute to the sustainable development of China's power industry, and also play an important guiding and constructive role in national energy security and sustainable development [12].

This paper mainly studies the application of clustering algorithm based on data mining in the analysis of EDC from the perspective of EP. In this paper, based on data mining technology, we use k-means clustering algorithm to understand the relationship between power development and economic development, and understand its characteristics, development and application. This paper studies the relationship between EC and economy from the perspective of electricity elasticity coefficient method and output value unit consumption method, points out the development strategy of electricity market and the measures of economic development from the perspective of electricity, and proposes to grasp the market, emphasize the analysis and prediction of the market, cultivate the market, emphasize the potential development of the market, influence the market, and emphasize the product publicity and penetration of the market, emphasis on the power market possession, maintenance of the market, emphasize on the market after-sales service. This paper also uses K-means clustering algorithm to calculate and analyze the correlation degree of power consumption structure, and the relationship between different industrial GDP and power consumption.

2. The Application of Clustering Algorithm Based on Data Mining in the Analysis of EDC from the Perspective of EP

2.1. The Relationship Between EC and Economy
   (1) Electric elastic coefficient method
   As a common index, EC elasticity coefficient is a measure of the functional relationship between the annual growth of EC and the growth of social and economic GDP. The calculation method is EC elasticity coefficient = growth of EC / annual GDP growth of social economy. As an important evaluation standard in the production and life of national economy, EC elasticity coefficient is affected
by many factors, such as the change of economic structure, economic development, technological progress, supply and demand, etc. In general, if the power consumption elasticity coefficient is greater than 1, it means that the power industry is ahead of the economic development. Especially under the background of planned economy before the reform and opening up, the power industry, the government and other relevant departments usually regard the power elasticity coefficient as an important factor for the balance of power supply and demand. After the reform and opening up, with the comprehensive popularization of the market economy, the rapid development of science and technology, the large-scale application of various types of power equipment and electronic devices, and the wide popularization of energy-saving technology, the relationship between economic development and power demand has become more complex and unstable, and there may be a serious disconnection between them, which makes the power elasticity coefficient change and difficult to determine.

(2) Unit consumption method of output value

The output value unit consumption method is to determine the power demand according to the product output, output value and unit power consumption. It can be divided into output value unit consumption and product unit consumption. Generally, the output value unit consumption method is commonly used. The unit consumption method can be used to count the EC of primary, secondary and tertiary industries and residents. It also takes into account the changes of different industrial structures, technological improvement, electrification level and power saving and other factors. The key of unit consumption method in load forecasting is to determine the unit consumption of output value, which has the advantages of simple method and good short-term load forecasting effect. The disadvantage is that it needs a lot of detailed research work, and is relatively general, it is difficult to fully reflect the impact of modern economic, political, climate and other conditions.

2.2. Power Market Development Strategy

(1) Grasp the market and emphasize the analysis and prediction of the market

We should subdivide the sales market, obtain various market dynamic information through multiple channels, establish a multi-level market analysis and forecast management mechanism at the provincial, municipal and county levels, improve the depth of market analysis, strengthen the tracking analysis of important indicators and abnormal fluctuation indicators, and carry out in-depth analysis of supply and demand situation and load characteristics, sales market changes and market share, market system construction and operation, improve the accuracy of the forecast to provide support for the company's business decision-making. We should carry out continuous tracking and Analysis on various factors that have great influence on the power sales market, such as climate factors, macro-control policies, time of use pricing policies, business expansion and installation, small thermal power plants and self-provided power plants, etc., so as to grasp the characteristics of the regional market and judge the development trend, find the key points and breakthroughs in the development of the regional market, and formulate the competitive strategy of the regional market, determine marketing objectives.

(2) Cultivate the market and emphasize the potential development of the market

Industrial power consumption, especially chemical industry, steel and other high energy consumption industries, accounts for a large proportion, and these industries are greatly affected by the national industrial policy and macro-control, with uncertainty. Once the policy changes lead to fluctuations in power consumption, it is directly related to the stability of the power sales market. Residential electricity market is the second largest consumer group after industrial electricity, which has great growth potential and is the direction of power market cultivation. In cities, we should take the promotion of "household electrification" as an important measure to lead the sales market expansion, determine the work objectives, tasks and requirements of household electrification, set up special funds for the promotion of household electrification, formulate the promotion strategies and activity details of "household electrification", carry out various publicity activities, and strive to make the per capita living electricity level close to the national average level. In rural areas, we should seize the favorable opportunity of national new rural construction, do a good job in the popularization of household appliances, and vigorously expand the electricity market of rural residents.
(3) Influence the market, emphasize the product publicity to the market

Due to the characteristics of clean, efficient and easy to use, EP has always been a high and stable consumer group for a long time. However, with the gradual entry of natural gas and other energy sources and increasing impact, the power market is facing a new test. We should further formulate the work plan of training and publicity, strengthen publicity, and strengthen the influence of EP on the energy consumption market by compiling publicity pictures and manuals representing the corporate image, such as scientific power consumption, saving power consumption and market development, and adopting TV advertising and other publicity methods, so as to consolidate the existing consumer groups and constantly attract new electricity consumers.

(4) Penetration of the market, emphasizing the power market share

By taking economic, technical, administrative and legal measures, we should actively participate in the competition in the energy market, promote the equipment and technology of replacing oil with electricity, gas with electricity and coal with electricity, and make breakthroughs in electric vehicles, heat pumps, household electrification, electric kilns, rural power drainage and irrigation, so as to expand the use scope of electric energy in the terminal energy market preferential electricity price policies, such as electric boiler electricity price, guide end users to adopt new technologies and new products, reduce the use cost of electric energy, establish typical demonstration projects in various cities of the province, form market chain effect through orderly and standardized operation, and gradually expand market share.

(5) Maintain the market and emphasize the after-sales service to the market

With the expansion of the market scale, the number and scope of customers using electricity are also increasing. If the follow-up services cannot be linked up, it will lead to the disappearance of some consumer groups. Therefore, it is particularly important to maintain the market and do a good job of after-sales service. According to the users and service requirements of the electricity sales market, the first thing is to ensure the product quality, that is, the electricity provided to customers should have qualified voltage, frequency and power supply reliability. Secondly, different types of electricity customer groups should be distinguished to provide differentiated services to meet the consumption needs of different levels. Thirdly, we should deepen the customer relationship, provide value-added services for customers, let customers have a sense of surprise, so as to become loyal users of the enterprise.

2.3. K-means Clustering Algorithm

K-means is a clustering technology based on prototype and partition. It tries to find that the number of clusters K specified by the user is represented by the centroid, which is used to define the prototype, where the centroid is the mean value of a group of points, usually used for objects in n-dimensional continuous space. Let N denote the number of observations, \(x_i\) denote the ith observation, K denote the number of categories, \(C_i\) denote the set of serial numbers of observations belonging to the 1st category, and C (i) denote the serial numbers of categories to which observation i belongs. The commonly used distance measure in K-means clustering method is minowski distance:

\[
d(x, y) = \left[ \sum_{r=1}^{p} |x_r - y_r|^m \right]^{1/m}
\]

By default, M = 2, that is, the Euclidean distance is adopted by default, and each observation object is reassigned to the class with the smallest distance from the center of the class:

\[
C_i = \arg \min_{i \leq j \leq k} d(x_i, y_i)
\]
Where argmin means to find the value of parameter (l) to minimize the function $d(x_i, v_j)$ and recalculate the category center:

$$v_j = \arg \min_{\forall C_i} \sum_{i \in C_i} d(x_i, v)$$  \hspace{1cm} (3)

Continue to cycle until the center of all categories changes little or reaches the maximum number of cycles specified in advance.

3. Experimental Study

3.1. Subjects
This paper mainly studies the application of clustering algorithm based on data mining in the analysis of EDC from the perspective of EP. By studying the relationship between EP development and economic development, this paper studies the relationship between EP consumption and economic development, focusing on the structure of EP consumption, and studies the development characteristics between them through the economic development and EP consumption of various industrial structures.

3.2. Experimental Process Steps
In this paper, based on data mining technology, we use K-means clustering algorithm to understand the relationship between power development and economic development, and understand its characteristics, development and application. This paper studies the relationship between EC and economy from the perspective of electricity elasticity coefficient method and output value unit consumption method, points out the development strategy of electricity market and the measures of economic development from the perspective of electricity, and proposes to grasp the market, emphasize the analysis and prediction of the market, cultivate the market, emphasize the potential development of the market, influence the market, and emphasize the product publicity and penetration of the market, Emphasis on the power market possession, maintenance of the market, emphasis on the market after-sales service. This paper also uses K-means clustering algorithm to calculate and analyze the correlation degree of power consumption structure, and the relationship between different industrial GDP and power consumption.

4. Experimental Research and Analysis of Clustering Algorithm Based on Data Mining in the Analysis of EDC from the Perspective of EP

4.1. Correlation Analysis of Power Consumption Structure
This paper mainly studies the application of data mining-based clustering algorithm in the analysis of EDC from the perspective of power. In order to analyze the impact of power consumption structure on economic development and understand the proportion of power consumption structure, this paper uses clustering algorithm to collect and integrate the power consumption structure through data mining technology, and then calculate its relevance. The results are shown in Table 1.

| Correlation degree | Comprehensive ranking |
|--------------------|-----------------------|
| primary industry   | 0.46                   | 4                     |
| the secondary industry | 0.84             | 1                     |
| the tertiary industry | 0.47              | 3                     |
| Residential electricity | 0.48              | 2                     |
As can be seen from Figure 1, the weighted ranking of the impact of each structure on the power consumption structure from large to small is the secondary industry, with the correlation degree of 0.84, residential power consumption, the correlation degree of 0.48, the tertiary industry, the correlation degree of 0.47 and the primary industry, with the correlation degree of 0.46.

4.2. EC and GDP

Macroeconomic development is the leading factor affecting power demand. Therefore, by studying the relationship between GDP and power consumption, this paper analyzes the correlation between GDP and power consumption, and summarizes the functional relationship between them. Different power demands of different industries also lead to different power consumption of different industrial structures. In order to understand the relationship between the total value of industrial structures and power consumption, this paper uses data mining technology and clustering algorithm to collect and sort out the data of GDP and power consumption of different industries in China in recent years. The results are shown in Table 2.

| Year | Total value of primary industry | Total value of secondary industry | Total value of tertiary industry | EC  |
|------|--------------------------------|----------------------------------|---------------------------------|-----|
| 2016 | 186.81                         | 960.54                           | 515.96                          | 412.46 |
| 2017 | 189.04                         | 1163.55                          | 600.58                          | 462.46 |
| 2018 | 207.01                         | 1377.64                          | 692.47                          | 496.49 |
| 2019 | 212.78                         | 1613.22                          | 782.51                          | 515.25 |
| 2020 | 220.66                         | 1916.51                          | 866.22                          | 576.97 |
Figure 2. Analysis of EC and GDP

It can be seen from Figure 2 that the total output of each industrial structure is in direct proportion to the power consumption. The total value of each industrial structure is increasing, and the power consumption is on the rise. For example, in 2016, the total value of the first industry is 186.81 trillion yuan, the second industry is 960.54 trillion yuan, the third industry is 515.96 trillion yuan, and the power consumption is 41.246 billion kwh. By 2020, the total value of the first industry is 186.81 trillion yuan, the second industry is 960.54 trillion yuan, and the third industry is 515.96 trillion yuan, the total value of the primary industry is 220.66 trillion yuan, that of the secondary industry is 1916.51 trillion yuan, that of the tertiary industry is 866.22 trillion yuan, and its EC is 57.697 billion kwh.

5. Conclusion
This paper studies the application of clustering algorithm based on data mining in the analysis of EDC from the perspective of EP, and understands the relationship between EP consumption and economic development, which is conducive to the optimization and adjustment of economic policy and EP planning, and has certain practical significance. This paper studies the relationship between EP development and economic development, understands its characteristics, development and application, studies the relationship between EP consumption and economy from EP elasticity coefficient method and output value unit consumption method, uses K-means clustering algorithm to calculate, and uses chart analysis method to analyze the correlation degree of EP consumption structure, as well as the relationship between different industrial GDP and EP consumption. Paying attention to and studying the relationship between power consumption and economic development can directly or indirectly grasp the production and operation status of all walks of life, which is of great significance to China's economic development.

References
[1] Medium and Long-term Electric Power Development Considering Operating Characteristics of High Proportion of Renewable Energy [J]. Automation of Electric Power Systems, 2017, 41 (21): 9 - 16.
[2] Ali Y, Chichirova N D . Electric power system of Yemen, its structure and characteristics [J]. Power Engineering Research Equipment Technology, 2019, 21 (3-4): 38 - 43.
[3] Kang, Goo E. Analysis of Electrical Characteristics of Power MOSFET with Floating Island [J]. Journal of Nanoscience and Nanotechnology, 2016, 16 (12): 12971 - 12974.
[4] Zhou P, Jin R Y, Fan L W. Reliability and economic evaluation of power system with renewables: A review[J]. Renewable and Sustainable Energy Reviews, 2016, 5 (8): 537 - 547.
[5] Tokarski T. Evaluation of direct current electric power systems of the aircraft based on
characteristics of a transient state [J]. Diagnostyka, 2018, 20 (1): 81 - 91.

[6] Hohs D, Schuller D, Stein R, et al. Microstructure Characteristics of Electrical Steel for Electrical Power Converters [J]. Praktische Metallographie, 2017, 54 (9): 615 - 635.

[7] Ali N D, Zakri A A. Identifying Characteristic of Power Quality Problems on Solar Electric Power Generation [J]. International Journal of Electrical Energy and Power System Engineering, 2018, 1 (2): 6 - 8.

[8] Jiehao L I, Wang J, Wang S, et al. Study on Contact Resistance Characteristics of Electrical Connectors of Power Battery Module [J]. Journal of Mechanical Engineering, 2020, 56 (11): 80.

[9] Zhou S, Chen Z, Zheng C, et al. Electrical Characteristics of Power Lithium-Ion Batteries for Full Ocean Depth Submersible [J]. Shanghai Jiaotong Daxue Xuebao/Journal of Shanghai Jiaotong University, 2019, 53 (1): 49 - 54.

[10] Choi Y, Song S, Kim J, et al. Thermal and Electrical Characteristics of Hollow Former with Fault Current Limiting Function for Superconducting Power Cable [J]. IEEE Transactions on Power Delivery, 2021, PP (99): 1 - 1.

[11] Zhao X, Cai Q, Ma C, et al. Economic evaluation of environmental externalities in China's coal-fired power generation [J]. Energy Policy, 2017, 102 (5): 307 - 317.

[12] Deng J, Wang M, Zhou Y, et al. Partial discharge characteristics of uniform gap in oil-impregnated paper insulation under switching impulse voltage [J]. IEEE Transactions on Dielectrics & Electrical Insulation, 2017, 23 (6): 3584 - 3592.