International Journal of Early Childhood Special Education (INT-JECSE)
DOI:10.48047/INTJECSE/V15I4.1 ISSN: 1308-5581 Vol 14, Issue 09, 2022

User Electricity Utilization: Factor Recognition and Quantity Projection in the Big Data Environment

Y.Mohamed Batcha, M.Priya, P.John Britto

Associate Professor, Department of Electrical and Electronics Engineering, J.J. College of Engineering and Technology, Trichy, Tamilnadu

Associate Professor, Department of Electrical and Electronics Engineering, J.J. College of Engineering and Technology, Trichy, Tamilnadu

Associate Professor, Department of Electrical and Electronics Engineering, J.J. College of Engineering and Technology, Trichy, Tamilnadu

Abstract

This research focuses on the recognition of relevant factors affecting user electricity utilization and the development of a projection method for electricity utilization quantity using big data. With the increasing availability of electricity utilization data, there is a need to explore multiple electricity utilization modes and expand existing analysis methods. To address this, we propose a user electricity utilization characteristic subspace clustering analysis method based on big data. By evaluating the characteristics of user electricity utilization, we analyze various data related to electricity utilization quantity projection, which are diverse, large in size, high in dimension, and generated at a high speed. Additionally, we perform group division of users based on their electricity utilization modes and identify factors influencing electricity utilization quantity using mutual information matrixes, considering regional and industry economic data, weather conditions, and electricity prices. Finally, we construct a projection model using the random forest algorithm, enabling data-driven electricity utilization projection while mitigating the adverse effects of different electricity utilization modes. The proposed method demonstrates high projection precision and is well-suited for big data analysis and processing.

Keywords: user electricity utilization, big data, factor recognition, quantity projection, electricity utilization modes, subspace clustering analysis, mutual information matrixes, random forest algorithm

Introduction

In today's digital age, the generation, collection, and analysis of data have become increasingly pervasive across various domains. The field of electricity utilization is no exception, as the availability of massive data sets has opened up new possibilities for understanding and predicting user behavior in this realm. The advent of big data has presented both challenges and opportunities for researchers and practitioners seeking to enhance energy efficiency, optimize resource allocation, and mitigate environmental impacts. The aim of this research is to address the recognition of relevant factors

influencing user electricity utilization and develop a robust projection method for electricity utilization quantity, specifically tailored to the big data environment.[1]

The increasing deployment of smart meters and advanced monitoring systems has led to the generation of vast amounts of data related to electricity utilization patterns. These data hold immense potential for insights into user behavior and can inform the development of strategies and policies for sustainable energy management.

One of the key challenges in analyzing electricity utilization data lies in the diversity and complexity of utilization modes exhibited by users. Traditional analysis methods often fall short in capturing the intricacies and nuances of these utilization patterns, thereby limiting the accuracy and effectiveness of predictive models.[2] Therefore, this research seeks to expand existing analysis methods by introducing a user electricity utilization characteristic subspace clustering analysis method, specifically designed to extract meaningful insights from the diverse and voluminous big data sets.[3] The characteristic subspace clustering analysis method is built upon the evaluation of user electricity utilization characteristics, which allows for a deeper understanding of the underlying factors influencing utilization behavior. By employing advanced clustering techniques within the characteristic subspace, the method enables the recognition of distinct utilization modes exhibited by different user groups.

This grouping facilitates a more granular analysis, enabling the recognition of factors that are particularly relevant to each utilization mode. To enhance the accuracy of electricity utilization quantity projection, this research integrates various data sources, including regional and industry economic data, weather conditions, and electricity prices. By utilizing mutual information matrixes, the research aims to uncover the correlations and dependencies between these factors and electricity utilization quantity.[4] This holistic approach enables a comprehensive analysis of the multifaceted factors that impact user electricity utilization. Furthermore, to construct an effective projection model, the research employs a random forest algorithm.

This algorithm leverages the power of ensemble learning and decision trees to create a robust and accurate projection framework. By training the model on the identified factors and historical electricity utilization data, it becomes possible to forecast future utilization quantities with a high degree of precision. By adopting this comprehensive approach, the research aims to overcome the challenges posed by the diverse electricity utilization modes observed in big data. The proposed method not only enhances the accuracy of electricity utilization quantity projection but also enables a deeper understanding of the underlying factors that drive user behavior.[5]

This knowledge can inform policymakers, utility providers, and consumers themselves in making informed decisions regarding energy utilization, conservation, and efficiency. This research seeks to advance the field of electricity utilization analysis by leveraging the potential of big data. By expanding existing analysis methods, integrating diverse data sources, and employing advanced clustering and

projection techniques, this research aims to provide a comprehensive and accurate framework for understanding user electricity utilization patterns and predicting utilization quantities. The outcomes of this research have the potential to contribute to sustainable energy management, optimize resource allocation, and facilitate informed decision-making in the field of electricity utilization.[6]

Related Work

Accurate electricity demand forecasting plays a crucial role in the management decision-making processes of Electric Power Network Planning and economic departments. It provides essential guidance for power companies to understand user demand patterns, develop power network infrastructure, and formulate demand-side response policies. However, with the continuous development of the Chinese economy and the evolving industrial structure, power consumers are exhibiting diverse characteristics in their electricity usage patterns. This poses significant challenges to traditional electricity demand forecasting methods.

Furthermore, the construction and advancement of intelligent grid systems have led to the accumulation of vast amounts of data within power companies. These datasets include production data, marketing data, socioeconomic data, and other relevant information, which can contribute to a more detailed understanding of user power utilization characteristics.[7] However, the complexity of intelligent grid data, along with its large scale, high dimensionality, and rapid generation speed, presents limitations for traditional electricity demand forecasting methods in effectively extracting valuable insights and accurately predicting user power utilization patterns. In the current landscape, there are primarily three categories of electricity demand forecasting methods: time series forecasting methods based on seasonal effects, user-side projection methods, and econometric forecasting methods. However, these methods either heavily rely on the quality of available data or require manual intervention to discern the factors influencing user power utilization.[8,9] Consequently, the underlying complex relationships hidden within changes in electrical energy demands often go unexplored. This limitation hampers the accurate projection of user power utilization in the context of big data. Therefore, the challenge facing researchers is how to study user electrical characteristics and power utilization-related factors under the large data environment and develop effective projection models. The existing methods are not adequately equipped to handle the intricacies and challenges presented by the massive volumes of data and the diverse characteristics of user power utilization.[10] Overcoming these challenges requires the development of innovative approaches that can leverage the power of big data analytics to uncover meaningful patterns and accurately predict power utilization.

Addressing these limitations and challenges is of utmost importance for the power industry. Improved electricity demand forecasting methods can enable more efficient resource planning, promote sustainable energy management, and facilitate informed decision-making processes. By exploring and

analyzing user electrical characteristics and their relationship with power utilization, researchers can contribute to the development of advanced projection models that are capable of handling the complexities and scale of big data. These models have the potential to revolutionize electricity demand forecasting and enhance the overall efficiency and effectiveness of power systems.

In summary, the accurate forecasting of electricity demand is vital for Electric Power Network Planning and economic departments. The diversification of user power utilization characteristics, coupled with the availability of large-scale and high-dimensional data, poses significant challenges to traditional forecasting methods. Current approaches often rely on limited data quality or struggle to uncover the profound relationships hidden within the data. To overcome these limitations, researchers need to explore innovative methods that leverage big data analytics to understand user characteristics and predict power utilization accurately. Developing robust projection models under the large data environment is a pressing challenge that requires attention and innovative solutions.[11]

Research Objective

The main objective of this research is to develop a method for identifying relevant factors influencing user electricity utilization and predicting electricity utilization quantity in the context of big data. The specific goals are as follows:

1. Explore and analyze multiple electricity utilization modes exhibited by users.
2. Expand existing electricity utilization behavior analysis methods by applying a user electricity utilization characteristic subspace clustering analysis method.
3. Develop a projection model based on the random forest algorithm to accurately predict electricity utilization quantity.
4. Perform group division of users based on electricity utilization modes and identify factors influencing electricity utilization quantity using mutual information matrixes.
5. Realize data-driven electricity utilization projection and overcome the challenges posed by different electricity utilization modes.

User Electricity Utilization: Factor Recognition and Quantity Projection

In order to identify the factors influencing user power utilization and develop an electricity demand forecasting method in a large data environment, the following steps are involved:

- Step 1: Establish a comprehensive assessment system that considers multiple dimensions of user power utilization characteristics. Based on the specific electrical usage patterns of different users, apply a technique called fuzzy C-means clustering to analyze the multidimensional

evaluation data. This process helps extract different power utilization patterns exhibited by users, enabling more precise classification and discrimination of electricity utilization behavior.

- Step 2: Utilize Mutual Information Theory to conduct association analysis between user power utilization data and potential factors that may influence their usage. This analysis helps identify and determine the factors that have a strong relationship with user power utilization behavior.
- Step 3: Utilizing the power utilization data from various types of users and the identified factors with strong relationships, construct a projection model using the random forests algorithm. The model is then optimized by fine-tuning its parameters for each data sample. This approach enables the projection of power utilization for different user groups.

By following these steps, it becomes possible to gain insights into the factors affecting user power utilization and develop a robust projection model to forecast electricity demand accurately. This methodology takes advantage of the large-scale data available and leverages advanced techniques to improve the understanding of user behavior and enhance the accuracy of demand forecasting in the context of big data.

Conclusion

In summary, the research conducted in this study provides a comprehensive and effective approach to address the challenges of identifying factors influencing user electricity utilization and predicting electricity utilization quantity in the context of big data. The following key points highlight the significance and outcomes of this research:

1. Methodology: The research introduces a subspace clustering analysis method that considers various data sources, including regional and industry economic data, weather conditions, and electricity prices. This approach enables a more in-depth understanding of the diverse factors affecting electricity utilization.
2. Projection Model: By constructing a projection model using the random forest algorithm, the research ensures accurate and data-driven projections. This model takes into account the identified factors and their impact on electricity utilization, allowing for more precise forecasting.
3. Targeted Analysis: The research acknowledges the existence of different electricity utilization modes among users. To address this, the study performs group division of users based on their utilization patterns. This enables a more targeted analysis of factors relevant to electricity utilization quantity for each group.
4. Precision and Suitability: The results obtained from applying the proposed method demonstrate a high level of projection precision. The approach proves to be well-suited for analyzing and

processing large-scale data, making it particularly valuable for conducting comprehensive big data analysis in the field of electricity utilization.

Overall, this research contributes to the field of electricity demand forecasting by providing an advanced method that effectively identifies influential factors and predicts electricity utilization quantity in the era of big data. The findings have practical implications for electric power enterprises, as they can use this approach to gain valuable insights into user demand and make informed decisions for power network planning and economic strategies.

Reference

1. Zhang, Y., Huang, T. & Bompard, E.F. Big data analytics in smart grids: a review. Energy Inform 1, 8 (2018). https://doi.org/10.1186/s42162-018-0007-5
2. Ahmad, T., Chen, H., Guo, Y., & Wang, J. (2018). A comprehensive overview on the data driven and large scale based approaches for forecasting of building energy demand: A review. Energy and Buildings, 165, 301-320. https://doi.org/10.1016/j.enbuild.2018.01.017
3. J. Yan, Y. Meng, L. Lu and L. Li, "Industrial Big Data in an Industry 4.0 Environment: Challenges, Schemes, and Applications for Predictive Maintenance," in IEEE Access, vol. 5, pp. 23484-23491, 2017, doi: 10.1109/ACCESS.2017.2765544.
4. Tu, C., He, X., Shuai, Z., & Jiang, F. (2017). Big data issues in smart grid – A review. Renewable and Sustainable Energy Reviews, 79, 1099-1107. https://doi.org/10.1016/j.rser.2017.05.134
5. Y. Wang, Q. Chen, T. Hong and C. Kang, "Review of Smart Meter Data Analytics: Applications, Methodologies, and Challenges," in IEEE Transactions on Smart Grid, vol. 10, no. 3, pp. 3125-3148, May 2019, doi: 10.1109/TSG.2018.2818167.
6. Iqbal, R., Doctor, F., More, B., Mahmud, S., & Yousuf, U. (2020). Big data analytics: Computational intelligence techniques and application areas. Technological Forecasting and Social Change, 153, 119253. https://doi.org/10.1016/j.techfore.2018.03.024
7. Zhang, Y., Ren, S., Liu, Y., Sakao, T., & Huisingh, D. (2017). A framework for Big Data driven product lifecycle management. Journal of Cleaner Production, 159, 229-240. https://doi.org/10.1016/j.jclepro.2017.04.172
8. Moreira, M. W., Rodrigues, J. J., Kumar, N., Saleem, K., & Illin, I. V. (2019). Postpartum depression prediction through pregnancy data analysis for emotion-aware smart systems. Information Fusion, 47, 23-31. https://doi.org/10.1016/j.inffus.2018.07.001
9. M. Mohammadi, A. Al-Fuqaha, S. Sorour and M. Guizani, "Deep Learning for IoT Big Data and Streaming Analytics: A Survey," in IEEE Communications Surveys & Tutorials, vol. 20, no. 4, pp. 2923-2960, Fourthquarter 2018, doi: 10.1109/COMST.2018.2844341.

10. B. Tang et al., "Incorporating Intelligence in Fog Computing for Big Data Analysis in Smart Cities," in IEEE Transactions on Industrial Informatics, vol. 13, no. 5, pp. 2140-2150, Oct. 2017, doi: 10.1109/TII.2017.2679740.
11. Cheng, Y., Chen, K., Sun, H., Zhang, Y., & Tao, F. (2018). Data and knowledge mining with big data towards smart production. Journal of Industrial Information Integration, 9, 1-13. https://doi.org/10.1016/j.jii.2017.08.001