TradeRES
New Markets Design & Models for 100% Renewable Power Systems

TradeRES Research Bulletin

Clustering-Based Filtering of Big Data to Improve Forecasting Effectiveness and Efficiency

Tiago Pinto [1], Tânia Rocha[1], Arsénio Reis[1], Zita Vale[2]
[1] *Universidade de Trás-os-Montes e Alto Douro and INESC-TEC, Porto, Portugal.*
[2] *Research Group on Intelligent Engineering and Computing for Advanced Innovation and Development (GECAD), Intelligent Systems Associated Laboratory (LASI), Polytechnic of Porto (P.PORTO), P-4200-072 Porto, Portugal.*

Full paper: https://doi.org/10.1007/978-3-031-20215-5_12

Summary

The rise of Big Data introduces new challenges for decision-making, as vast and diverse data sources collected at varying times and with considerable uncertainty make the decision-making process a harsher task. Existing methods often struggle to address these new problems. This paper presents a novel data selection methodology that filters big volumes of data, so that only the most correlated information is used in the decision-making process in each given context. By applying a clustering algorithm, creates sub-groups of data according to their correlation, used to feed a forecasting process that uses the relevant data for each situation, while discarding data that is not expected to contribute to improving the forecasting results. This approach enables faster, more computationally efficient, and effective forecasting. A case study on electricity market data demonstrates that this method. Results show that the data selection increases the forecasting effectiveness of forecasting methods, as well as the computational efficiency of the forecasts, by using less yet more adequate data.

Highlights

- Big Data's complexity challenges traditional decision-making methods.
- A new data selection method filters relevant data clusters to enhance forecasting.
- Case study results show improved forecasting accuracy and efficiency.