



Comparative Analysis of Neural Network and Clustering Approaches for Time Series Anomaly Detection

Şule Şevval Karakaya⁽¹⁾, Vilda Purutçuoğlu^(1,2) and Ahmet Bursalı⁽³⁾

⁽¹⁾*Department of Biomedical Engineering Middle East Technical University Ankara, Türkiye
sule.karakaya@metu.edu.tr*

⁽²⁾*Department of Statistics Middle East Technical University Ankara, Türkiye
vpurutucu@metu.edu.tr*

⁽³⁾*Intelligent Application DC Dept, Huawei Turkey R&D Center İstanbul, Türkiye
ahmet.bursali@huawei.com*

Abstract

Outlier detection is widely used in many fields and may have different meanings depending on the purpose of analysis. Outliers may reduce data quality and model performance, but they may also indicate important warning signals. In electricity consumption data, this issue is especially important because abnormal values may distort forecasting and analysis results or reflect unusual consumption patterns that require attention. These values may be related to abnormal demand, sensor errors, user behavior changes, equipment malfunctions, and seasonal effects. Therefore, detecting outliers is important for energy analysis, forecasting, and decision support systems. In this study, the dataset includes electricity consumption measurements from four cities in Morocco: Laayoune, Boujdour, Marrakech, and Foug Eloued, with more than 67,000 smart meter observations recorded at 10-minute and 30-minute intervals.

This study comparatively examines anomaly detection methods under two main headings: neural network-based and clustering-based methods. Neural network-based methods are suitable for modelling nonlinear patterns, temporal dependencies, and complex consumption dynamics in time series data. Clustering-based methods are useful for identifying observations that form small, isolated, irregular, or structurally distinct groups without requiring strong distributional assumptions. The neural network-based methods used were GRU, LSTM, TCN, N-BEATS, and TimesNet. The clustering-based methods applied were K-Means with DTW distance, K-Medoids with DBA centers, HDBSCAN, and TICC. The study aims to determine which outliers are detected by different algorithms, whether small or isolated clusters can be interpreted as potential outlier groups, and how similar the detected patterns are across methods. After each method was applied, anomaly indices and counts were obtained for each region, and pairwise overlap was measured using Jaccard similarity.

The findings show that detected outliers vary depending on the algorithm used. Clustering-based methods showed higher overlap among themselves than neural network-based methods. The highest average overlap across all cities was observed between K-Means and K-Medoids, with a mean Jaccard similarity of 0.644. GRU and LSTM also showed relatively high similarity, with a mean value of 0.569. These results suggest that the detected anomaly sets are partly method-dependent. Overall, there is no single best anomaly detection method for this dataset, and using multiple methods together can provide a more reliable evaluation for electricity consumption analysis.

Acknowledgement: This study is supported by TÜBİTAK 1505 project (No: 5240032).

Keywords: Anomaly detection, neural networks, clustering methods, electricity consumption.