Effective Outlier Detection in Smart Home Energy Consumption Using Integrated Change Point Detection and Unsupervised Learning

Yamini Kodali, Kasaraneni Purna Prakash, Yellapragada Venkata Pavan Kumar

Corresponding Author: Yellapragada Venkata Pavan Kumar (Email ID: pavankumar.yv@vitap.ac.in)

INTRODUCTION & AIM

- As smart homes proliferate globally, smart meter energy consumption data has become vital for data-driven decisions, making data quality crucial for reliable analytics.

- However, smart meter data often contain anomalies such as outliers, missing values, and redundant entries, caused by communication delays, transmission errors, and device malfunctions.

- These anomalies can significantly compromise the accuracy of applications, including billing, contingency analysis, and energy forecasting.

- Among them, outliers are particularly detrimental, as they can distort statistical analysis, mislead machine learning models, and undermine overall system reliability.
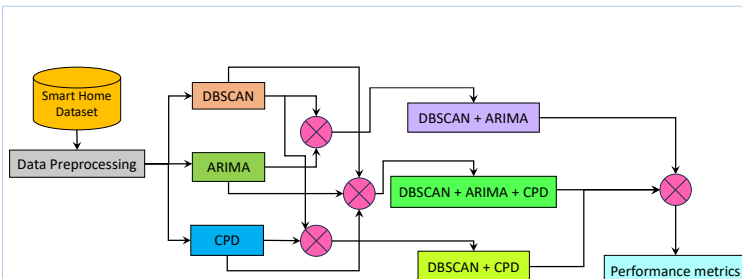
METHOD

Fig. 1. Conceptual workflow for Detection of Outliers

RESULTS & DISCUSSION

- A Sequential hybrid approaches ARIMA+DBSCAN, CPD+DBSCAN, and ARIMA+CPD+DBSCAN were computed along with individual models such as ARIMA, CPD, DBSCAN to detect outliers.

- The evaluation metrics were computed to know the performance of each model. The sequential hybrid methods performs better than individual methods.

- The results shows a sequential hybrid model interprets better performance with a precision of 0.96, sensivity of 0.94, F1-score of 0.95, and accuracy of 0.98.

Fig.2. Conceptual workflow for Detection of Outliers
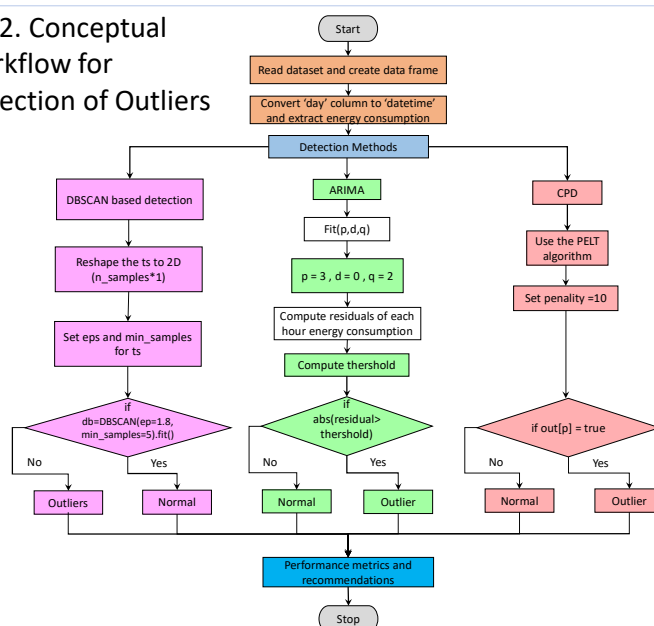
Model	Precision (%)	Recall (%)	F1-Score (%)	Accuracy (%)
ARIMA	0.91	0.89	0.90	0.92
CPD	0.85	0.72	0.78	0.90
DBSCAN	0.94	0.91	0.92	0.94
ARIMA+DBSCAN	0.93	0.90	0.91	0.96
CPD+DBSCAN	0.88	0.75	0.81	0.92
ARIMA+CPD+DBSCAN	0.96	0.94	0.95	0.98

Table 1. Evaluation Metrics

CONCLUSION AND FUTURE SCOPE

- Resultant sequential hybrid approach demonstrating the advantage of integrating statistical and clustering-based methods for robust outlier detection in smart home energy.

- Future research can focus on deep learning integration, real-time deployment, multivariate analysis.

KEY REFERENCES

1. X. Jia, P. Xun, W. Peng, B. Zhao, H. Li, and C. Shen, "Deep anomaly detection for time series: A survey," Computer Science Review, vol. 58, p. 100787, 2025, doi: https://doi.org/10.1016/j.cosrev.2025.100787

2. J. Paparrizos, P. Boniol, Q. Liu, and T. Palpanas, "Advances in Time-Series Anomaly Detection: Algorithms, Benchmarks, and Evaluation Measures," in Proceedings of the 31st ACM SIGKDD Conference on Knowledge Discovery and Data Mining V.2, Toronto ON Canada: ACM, Aug. 2025, pp. 6151–6161. doi: https://doi.org/10.1145/3711896.3736565