International Journal of Applied Operational Research
Vol. 13, No. 3, pp. 1-16, Summer 2025

Journal homepage: ijorlu.liau.ac.ir

Evaluating high-consumption and unusual subscribers in the smart gas meter network using machine learning and the Internet of Things in the cloud environment

K. Aghaei Badr, H. Mehrmanesh*, A. Fadavi Asghari

Received: 26 March 2025 ; **Accepted:** 12 June 2025

Abstract With the advancement of new technologies, the use of smart gas meters as a tool for managing energy consumption and optimizing energy resources is expanding. These meters can collect and analyze consumption data in real time with the help of Internet of Things (IoT) and machine learning. The aim of this research is to evaluate and identify high-consumption and abnormal subscribers in the smart gas meter network using machine learning algorithms and Internet of Things technology in a cloud environment. This research seeks to provide solutions to improve energy consumption management and reduce costs by identifying abnormal consumption patterns and providing optimization suggestions to subscribers. The importance of energy consumption management and the implementation of related policies have required governments to identify high-consumption subscribers and separate them from low-consumption subscribers. Accordingly, policies are being developed to fine or punish high-consumption subscribers based on their consumption and even reward low-consumption subscribers. This is possible more efficiently using a smart meter network in which data is transferred in real time on the Internet of Things network and stored in a cloud computing environment.

In this research, in line with this policy, an attempt has been made to design a model to identify and control high-consumption and irregular subscribers in the smart gas meter network. This model includes 5 variables: annual consumption, monthly consumption, consumption period, household size, and subscription type, which were implemented using 4 machine learning algorithms: random forest, decision tree, nearest neighbor, and XG boost.

The results show that the random forest algorithm was able to classify and identify high-use subscribers with 92% accuracy, followed by the XG boost algorithm with 91% accuracy, and then the nearest neighbor and decision tree algorithms with 90% accuracy.

The conclusion of this research shows that the use of machine learning algorithms and IoT technology in the smart gas meter network can help to accurately identify high-consumption and abnormal subscribers. This not only leads to energy consumption optimization and cost reduction, but also enables the implementation of effective policies for energy consumption management.

Keyword: High-Consumption Subscribers, Aberrant Subscribers, Meter Network, Gas, Machine Learning, Internet of Things, Cloud Computing.

*** Corresponding Author. (✉)**
E-mail: Has.mehrmanesh@iauctb.ac.ir (H. Mehrmanesh)

K. Aghaei Badr
Department of Information Technology Management, CT.C., Islamic Azad University, Tehran, Iran

H. Mehrmanesh
Department of Industrial Management, CT.C., Islamic Azad University, Tehran, Iran

A. Fadavi Asghari
Department of Industrial Management, CT.C., Islamic Azad University, Tehran, Iran

[Downloaded from ijaor.com on 2026-09-08]

[DOI: 10.71885/ijorlu-2025-2-706]

1 Introduction

Energy consumption is a significant issue that governments and local governments continually seek to control. This control is not limited to countries without energy resources. Still, countries with energy resources have also realized that one day their resources will end, and therefore, they must now seek to optimize consumption patterns [1]. With the advent of the fourth industrial revolution and emerging technologies like the Internet of Things and cloud computing, it is now possible to utilize tools such as smart meters. These meters are able to record consumption in real time and send it to the cloud computing environment, and store it in an unlimited space [2]. In this way, consumption information is always available to energy service providers, and therefore, its control and monitoring have also become possible [3].

Using smart meters, it is possible to find out how much each subscriber has consumed, and as a result, it is possible to distinguish high-consumption subscribers from low-consumption subscribers. Governments that seek to encourage low consumption and punish high consumption can use this smart meter tool to implement their policies [4]. In such a way that, by categorizing subscribers into low-consumption, high-consumption, and medium-consumption subscribers, they can implement the desired policies. This is done by entering a large volume of real-time data into machine learning algorithms and training data to identify high-consumption subscribers, and then based on the test data, the desired training can be evaluated and high-consumption subscribers can be distinguished from low-consumption subscribers with a percentage of error, and the necessary punitive policies can be applied to these subscribers, including disconnecting or reducing the energy load or even giving them a warning [5].

Considering the important issue of energy consumption and identifying high-consumption subscribers and separating them from low-consumption subscribers, this study attempts to design a model for this issue using smart meters connected to the Internet of Things and cloud computing that separates low-consumption subscribers from high-consumption subscribers. High-consumption subscribers include subscribers who generally use comfort devices such as jacuzzi or swimming pools and have high gas consumption. Considering that the high consumption of these subscribers can cause high energy waste and can also violate the rights of low-consumption subscribers, it is necessary to identify these subscribers by an automatic system, which is the goal of this study. In the present study, four machine learning algorithms, including decision tree, nearest neighbor random forest and XG boost are implemented so that the best algorithm among them is identified based on four criteria: accuracy, precision, f1score, recall, and any algorithm that obtains the highest values based on these four criteria is selected as the superior algorithm. By implementing this model, the present study attempts to present a new model with the lowest subscriber classification error rate for identifying high-consumption subscribers. This model can be implemented when implementing smart gas meters connected to the Internet of Things. The innovation of this research is the use of advanced machine learning algorithms and Internet of Things technology to identify high-consumption and unusual subscribers in the smart gas meter network, which helps improve energy consumption management. The contribution of this research is to provide practical solutions and effective policies for segmenting subscribers based on consumption patterns and optimizing energy resources at a macro level. The contribution of this research includes the following:

- Development of consumption management policies: Providing solutions to separate high-consumption subscribers from low-consumption subscribers in order to implement penalty and reward policies.

[Downloaded from ijaor.com on 2026-09-08]

[DOI: 10.71885/ijorlu-2025-2-706]

- Optimization of energy resources: Helping to improve energy consumption management and reduce costs by identifying consumption patterns and providing optimization suggestions.

The structure of the present paper is as follows: in the next section, a literature review is presented regarding research conducted in the field of smart gas meters, followed by an explanation of the research methodology, which includes the introduction of algorithms, followed by an analysis of the findings, and finally a conclusion.

2 Literature review

This section reviews the most important research in the field of study and the classification and identification of subscribers or theft in the smart meter network. The research is related to the last 2 years. Xia et al. [6] use theft detection methods in smart meters. Zhou et al. [7] use two unsupervised learning algorithms to detect abnormal inactivity in a household based on smart meter data. Hernandez et al. [1] detect anomalies in daily activities using smart meter data. Otuoze et al. [2] detect and confirm electricity theft in smart meter infrastructure using fuzzy inference system models and long short-term memory. Srivani et al. [8] conducted a study on smart meter billing. Fang et al. [9] conducted a study on anomaly detection based on big data analysis in smart meters. Kawosal et al. [10] focuses on improving theft detection using a data collection system and customer consumption patterns. Mbey et al. [11] detect energy theft in smart grids using a hybrid deep learning-based data analysis technique. Farooq et al. [12] work to detect cyberattacks on smart meters. Nkenyerye et al. [3] design a protocol for authentication in smart grids. This scheme includes an energy-efficient authentication scheme for smart meters with limited resources. Softan et al. [13] present a technological infrastructure for future urban development that includes smart meters and smart cities. Blazakis et al. [14] focus on evaluating the performance of smart meters and provide insights into energy management, dynamic pricing, and consumer behavior. Chaudhari et al. [15] analyze and predict residential electricity consumption using smart meter data. Farooq et al. [16] focuses on detecting energy theft in smart grids using quantum machine learning. Zhou et al. [7] predict residential energy consumption using long-term recurrent neural networks. Samiefard et al. [17] examine the role of smart meter data analytics in achieving sustainable development. Vaseei et al. [18] prioritize smart meters based on data control for grid resilience. According to above mentioned the research gap in this area is as follows:

1. **Lack of sufficient data:** Many smart gas meter networks still do not have sufficient data to accurately analyze the behavior of high-consumption and abnormal subscribers, which can lead to serious limitations in the accuracy of machine learning algorithms.
2. **Security and privacy challenges:** With the increasing use of IoT technology, there are concerns about data security and subscriber privacy, which necessitates the need to examine and solve these challenges.
3. **Diversity in consumption patterns:** Energy consumption patterns in different regions may vary greatly, and this diversity can lead to greater complexity in data analysis and identification of abnormal subscribers.
4. **Need for effective management policies:** There are still no effective management policies to segregate and manage subscribers based on consumption patterns, and this gap indicates the need for further research in this area.

[Downloaded from ijaor.com on 2026-09-08]

[DOI: 10.71885/ijorlu-2025-2-706]

3 Methods

In the present study, the aim is to identify high-consumption subscribers through smart gas meters in a cloud computing system using big data from data related to these meters. For this purpose, machine learning algorithms are used, which have the feature of classification, because in the present study, subscribers are classified into three categories based on input variables: low consumption, medium consumption, and high consumption, and the aim of the present study is to identify high-consumption subscribers. Data is collected from the smart gas meter dataset, which includes the following variables (Table 1):

Table 1 Input variables

Row	Variables	Signature	Variable type	Variable scale
1	Subscription type	X1	Input	Category
2	Annual consumption	X2	Input	Quantitative
3	Monthly consumption	X3	Input	Quantitative
4	Period consumption	X3	Input	Quantitative
5	Household size	X5	Input	Quantitative
6	Shared consumption level	Y1	Output	Category

There are 5 variables that determine the level of shared consumption, which are presented in the table below. These variables are entered into machine learning algorithms, and with their help, the level of shared consumption is determined. The input variables can be drawn as the following conceptual model in Fig. 1.

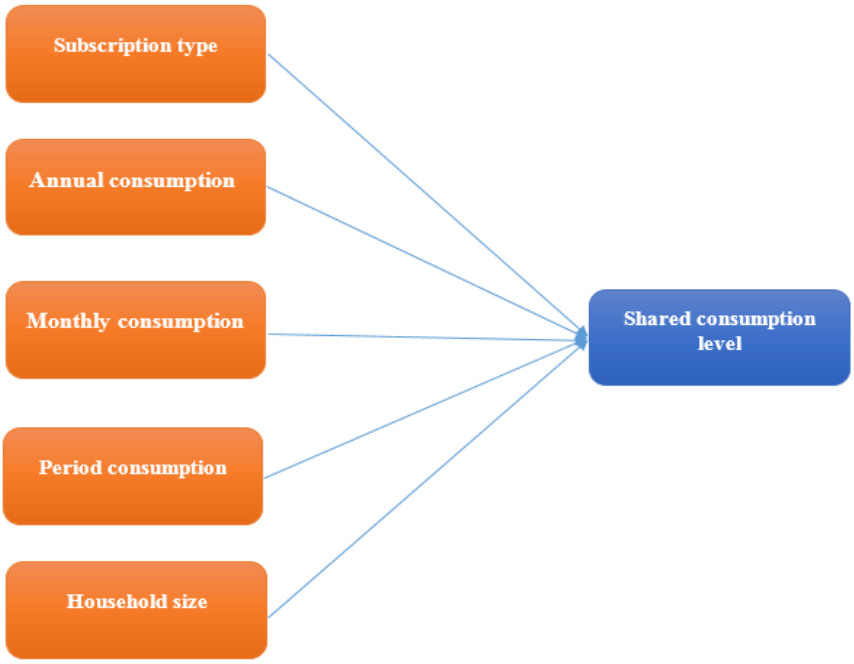

Fig. 1 Conceptual model

The implementation of the above conceptual model is done using four machine learning algorithms for classification. These algorithms are as follows:

1. Decision Tree Algorithm
2. Random Forest Algorithm

[Downloaded from ijaor.com on 2026-09-08]

[DOI: 10.71885/ijorlu-2025-2-706]

3. Nearest Neighbor Algorithm
4. XG Boost Algorithm

Decision Tree Algorithm

The decision tree algorithm classifies data through a hierarchical structure of conditional rules, where each node represents a decision based on a feature and the leaves indicate the final output. This model is simple, interpretable, and efficient for small datasets, but it often suffers from over fitting when applied to complex data.

Random Forest Algorithm

The random forest algorithm is an ensemble method that combines multiple decision trees trained on random subsets of data. The final prediction is obtained through majority voting in classification or averaging in regression tasks. This approach improves accuracy and reduces over fitting, although it is less interpretable and requires higher computational resources.

Nearest Neighbor Algorithm

The nearest neighbor algorithm (KNN) is a distance-based method that assigns a label to a new data point according to the majority class among its K closest neighbors. It does not require complex training and is easy to implement, but it can be computationally expensive with large datasets and is sensitive to the choice of K and data scaling.

XGBoost Algorithm

XGBoost is an advanced implementation of gradient boosting that builds models sequentially, with each step correcting the errors of the previous one. It is optimized for speed and accuracy, making it highly effective for large-scale and complex datasets. While it delivers strong predictive performance, its complexity makes parameter tuning more challenging.

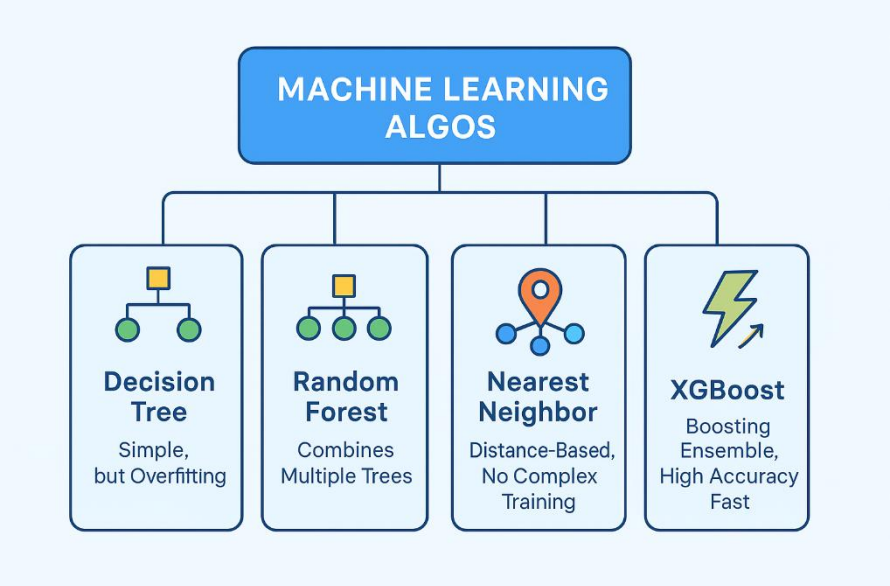

Fig. 2 Machine learning algorithms flowchart

These algorithms are implemented using Python, and their comparison is based on criteria related to classification algorithms. These criteria are introduced below.

1. Accuracy
2. Precision
3. Recall
4. F1score

[Downloaded from ijaor.com on 2026-09-08]

[DOI: 10.71885/ijorlu-2025-2-706]

The higher the score obtained from the algorithms regarding these criteria, the higher the efficiency of the respective algorithm. Accuracy indicates the number of correctly classified samples compared to the total sample data. Its calculation formula is as follows:

$$accuracy = \frac{TN + TP}{TN + FP + TP + FN} \tag{1}$$

In the above equation, TN is the total number of true negatives, TP is the total number of true positives, FP is the total number of false negatives, and FN is the total number of false negatives. Precision indicates the positive predictive value in classifying the data sample. Its formula is as follows:

$$Precision = \frac{TP}{TP + FN} \tag{2}$$

The next criterion is recall, which is defined as sensitivity or true positive rate. Its formula is as follows.

$$Recall = \frac{FP}{FP + TN} \tag{3}$$

And finally, the final criterion for evaluating the efficiency of machine learning algorithms in the F1Score classification is that it calculates both precision and recall together, and is as follows.

$$F1score = 2 * \frac{Precision * Recall}{Precision + Recall} \tag{4}$$

In the above formula, precision is multiplied by recall in the numerator, but in the denominator, these two criteria are added together and multiplied by 2, which results in the score value, and the higher it is, the better the performance of an algorithm. The data processing algorithms in the Clementine software were implemented on a personal computer with CPU Intel dual core 2.20 GHZ, RAM 4 GB and operating system Windows 10 configuration.

4 Results

In order to model the data, since the answer is not clear, which of the classification models should be used, first determine the desired answer using the clustering method. The K-Means method is used for clustering. In order to obtain the best number of clusters in this method, we use the Davies-Bouldin index comparison for 2 to 12 clusters (the maximum number of categories is based on 12 questions). The Davies-Bouldin index is actually the ratio of the distance of observations within clusters to the distance of clusters from each other. Therefore, it is obvious that lower values of this index indicate better clustering. For this, we first create the desired answer based on questions q_1 to q_12, which represent future categories for prediction, and then use it to create classification models and predict the type of performance for new people in the future. The results of obtaining the best number of clusters based on the mentioned questions are as follows in Table 2.

[Downloaded from ijaor.com on 2026-09-08]

[DOI: 10.71885/ijorlu-2025-2-706]

Table 2. Davies-Bouldin index

Clustering k	Davies-Bouldin Index
2	-0.989584123
3	-1.729566753
4	-1.8329627
5	-1.712606639
6	-1.72694021
7	-1.776220166
8	-1.69659013
9	-1.627926684
10	-1.649545961
11	-1.663281633
12	-1.741171018

Based on the above findings, the number of clusters is two, the most appropriate number for clustering based on questions q_1 to q_12 on the available dataset. (The lowest index indicates the best clustering). The number of clusters as the best number for clustering brings the following information on the dataset:

Clustering k = 2
Performance:
-----Avg. within centroid distance: -0.836
-----Avg. within centroid distance_cluster_0: -0.295
-----Avg. within centroid distance_cluster_1: -1.859
***Davies Bouldin: -0.990
Cluster Model:
Cluster 0: 4064 items
Cluster 1: 2151 items
Total number of items: 6215

Then, the evaluation of the presented model is discussed. As mentioned, four algorithms were used to evaluate the model, and the algorithms were evaluated and compared based on four criteria: accuracy, precision, recall, and f1score. The results of this comparison are presented in Table 3.

Table 3. Comparison of algorithms and model evaluation

Algorithms	Accuracy	Precision	Recall	F1-score
Decision tree	0.9067	0.8974	0.9211	0.9091
Random forest	0.9267	0.9333	0.9211	0.9272
K-nearest neighborhood	0.9000	0.9296	0.8684	0.8980
XGboost	0.9133	0.9200	0.9079	0.9139

[Downloaded from ijaor.com on 2026-09-08]

[DOI: 10.71885/ijorlu-2025-2-706]

The above table compares the algorithms in terms of the four criteria mentioned. At a glance, it can be seen that the random forest algorithm performs better than the other algorithms because it has the highest value in terms of each criterion, but for more details, the above results are presented in the form of the following graphs.

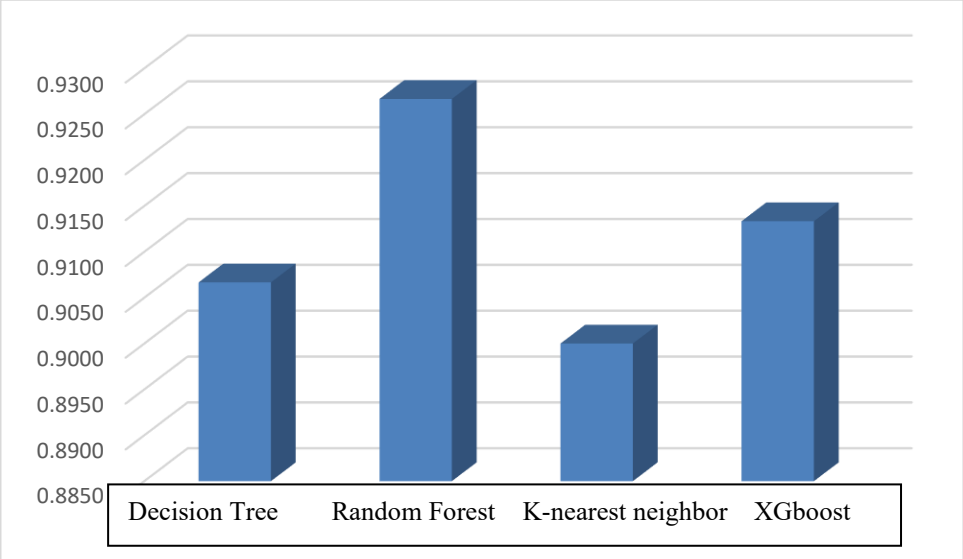

Fig. 3 Comparison of algorithms and evaluation of the model in terms of accuracy criteria

Based on Figure 3, it can be seen that the accuracy, which indicates the accuracy in classifying high-use subscribers, is higher for the random forest algorithm than for other algorithms, and therefore, this algorithm has performed better than other algorithms in this criterion.

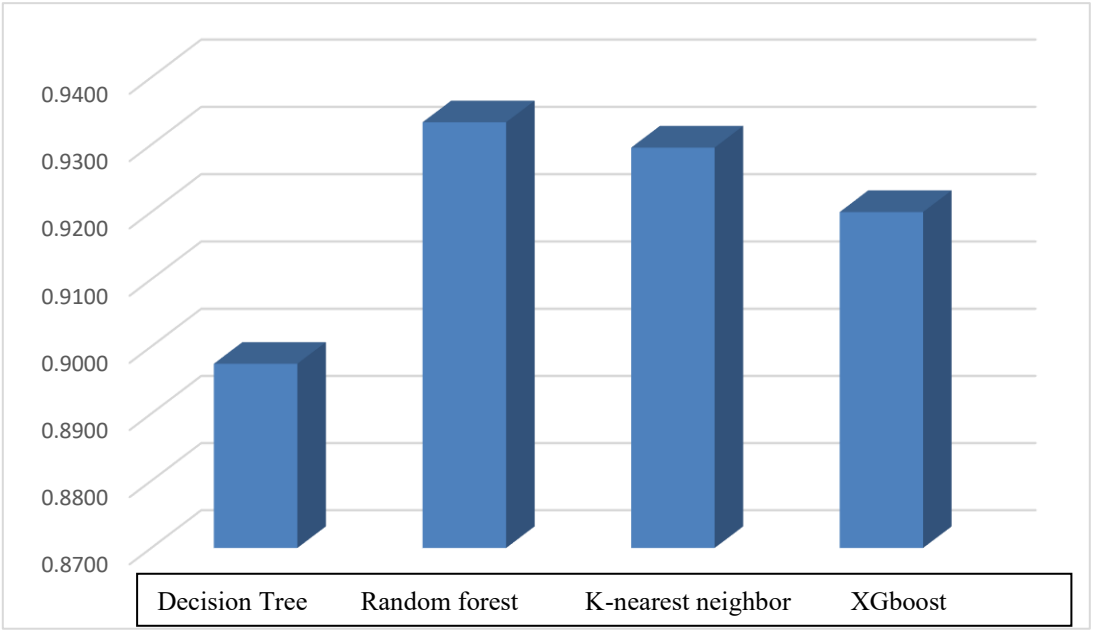

Fig. 4 Comparison of algorithms and evaluation of the model in terms of Precision criterion

[Downloaded from ijaor.com on 2026-09-08]

[DOI: 10.71885/ijorlu-2025-2-706]

In Figure 4, the results of comparing the algorithms in terms of the precision criterion indicate the superiority of the random forest algorithm, followed by the nearest neighbor algorithm, which is slightly behind the random forest algorithm in terms of this criterion.

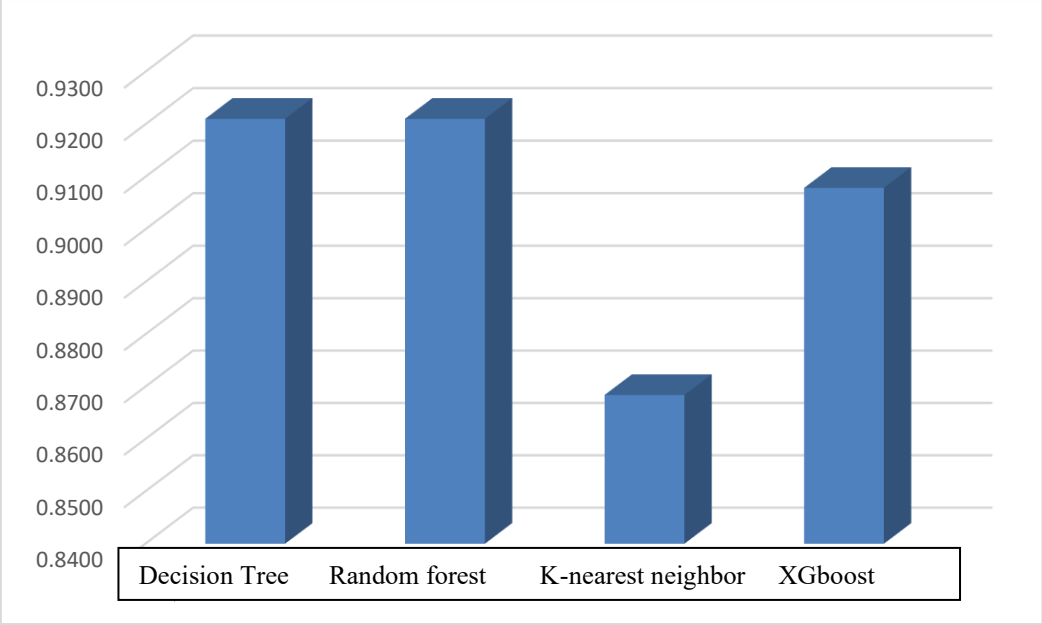

Fig. 5 Comparison of algorithms and evaluation of the model in terms of the Recall criterion

In Figure 5, it can be seen that the random forest algorithm and the decision tree algorithm obtained the same Recall value, and therefore, in terms of this criterion, both of these algorithms are recognized as superior algorithms. The XG Boost algorithm is after them, and finally, the Nearest Neighbor algorithm is located.

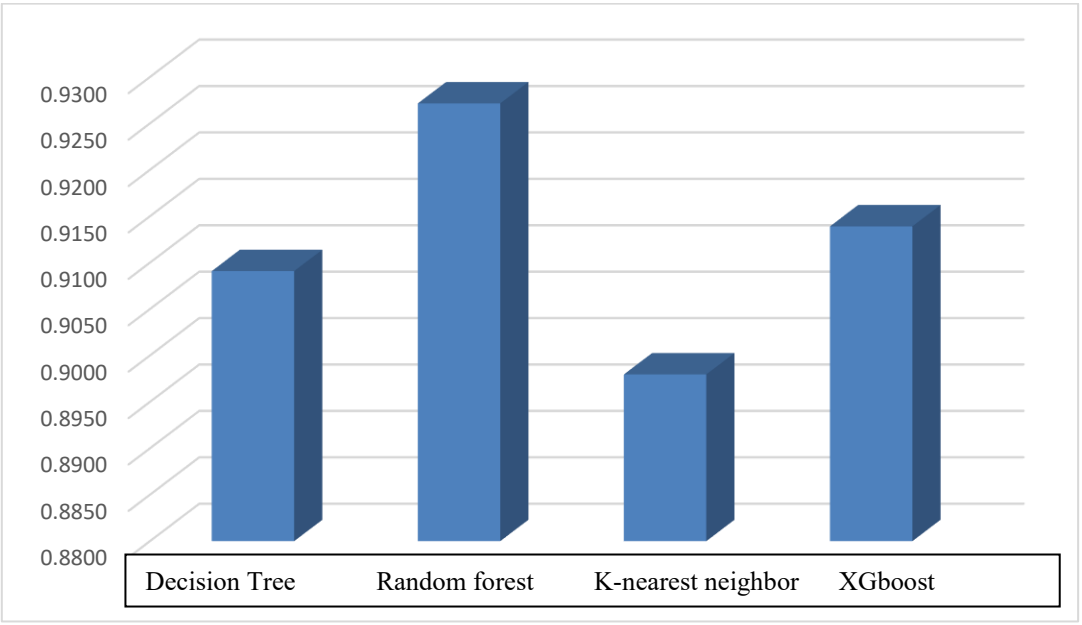

Fig. 6 Comparison of algorithms and model evaluation in terms of f1score criterion

[Downloaded from ijaor.com on 2026-09-08]

[DOI: 10.71885/ijorlu-2025-2-706]

In Figure 6, the algorithms are compared in terms of the f1score criterion, the explanation of which is provided in the methodology section along with its formula. In Figure 6, it can be seen that the random forest algorithm has the highest f1score value and is therefore considered the superior algorithm in terms of this criterion.

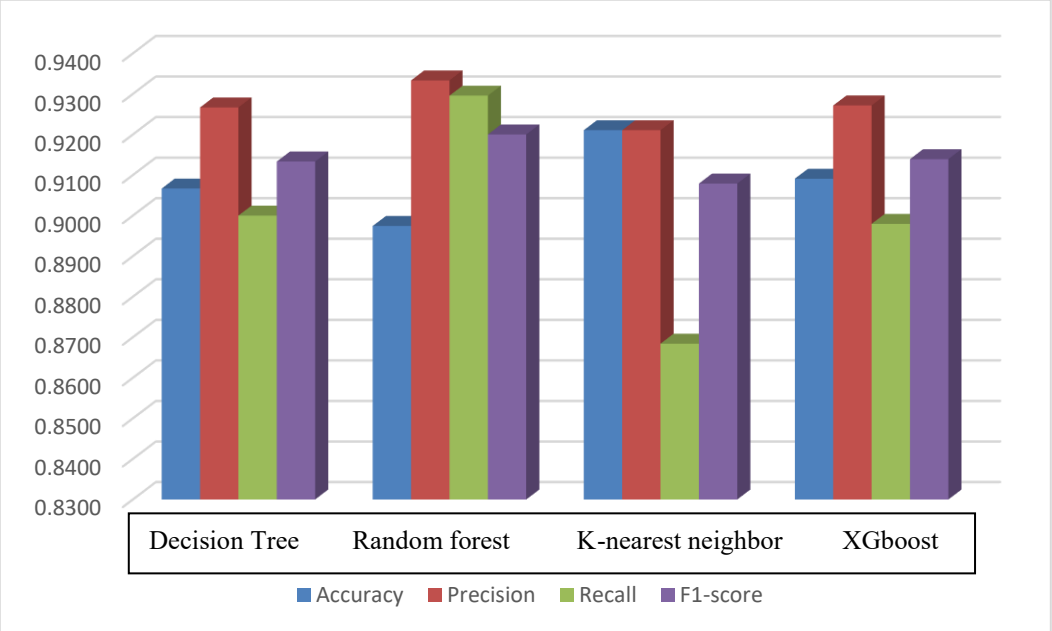

Fig. 7 Comparison of algorithms and overall model evaluation

In Figure 7, the algorithms are compared in a general overview, which clearly shows that the random forest algorithm has a significant advantage over other algorithms, and therefore, the superior algorithm in the present study is the random forest algorithm, which has the highest value in terms of all four selection criteria, according to Figure 7. Of course, it should be noted that the distance between other algorithms and this algorithm is not much, but because the goal is to select the best algorithm, the random forest algorithm is selected as the superior algorithm. Next, the confusion matrix is presented, and the best algorithm is displayed in terms of efficiency. The confusion matrix is shown in Figure 8.

[Downloaded from ijaor.com on 2026-09-08]

[DOI: 10.71885/ijorlu-2025-2-706]

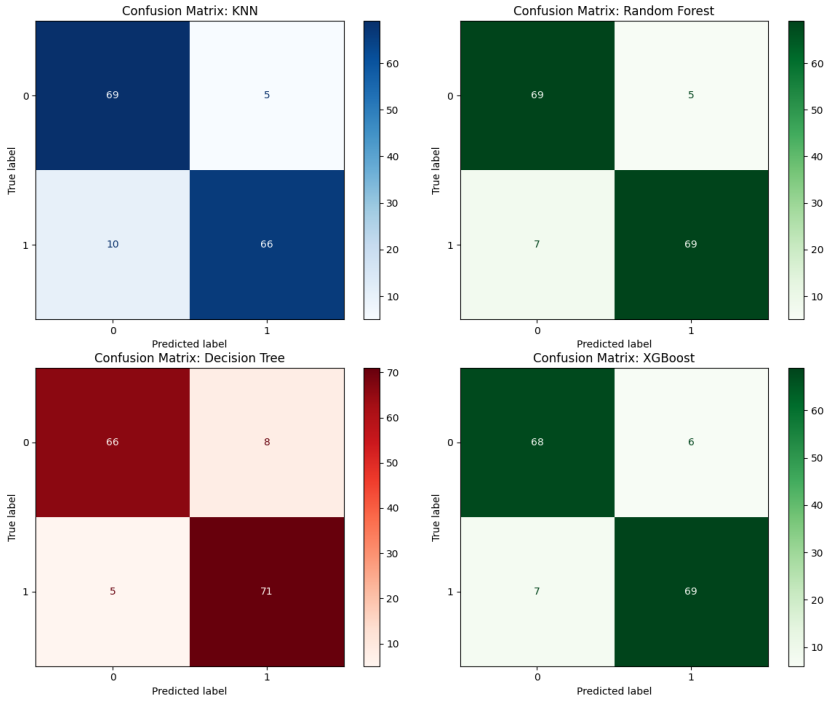

Fig. 8 Confusion matrix

Regarding the confusion matrix, it should be noted that this matrix is specific to the classification operation and not estimation, and given that the goal of the present study was to classify high-consumption subscribers in terms of consumption, this matrix was presented. Based on this matrix, any algorithm whose main diameter has higher values and whose sub-diameter has values close to zero is the superior algorithm in terms of efficiency. The result shows that although the random forest algorithm has lower values in terms of the main diameter compared to the decision tree algorithm, in the sub-diameter, these values tend more toward zero, and therefore, the random forest algorithm is considered a more suitable algorithm for classifying and detecting high-consumption and abnormal subscribers in the smart gas meter network. Next, the algorithms are compared in terms of the ROC curve. Each of the algorithms whose curve is higher indicates a higher predictive power and classification.

The Receiver Operating Characteristic (ROC) curve is a widely used tool for evaluating the performance of classification models. It illustrates the trade-off between the True Positive Rate (TPR) and the False Positive Rate (FPR) at different decision thresholds. By plotting TPR on the vertical axis and FPR on the horizontal axis, the ROC curve demonstrates the model's ability to distinguish between positive and negative classes. A key metric derived from this curve is the Area under the Curve (AUC), which provides a single measure of performance. An AUC value of 0.5 indicates random guessing, while values above 0.7 reflect acceptable performance, and values greater than 0.9 demonstrate excellent discriminatory power.

In summary, the ROC curve, along with the AUC measure, offers a robust approach for comparing and interpreting classification models across various threshold levels, making it an essential technique in modern predictive analytics. The ROC graph is presented in Figure 9.

[Downloaded from ijaor.com on 2026-09-08]

[DOI: 10.71885/ijorlu-2025-2-706]

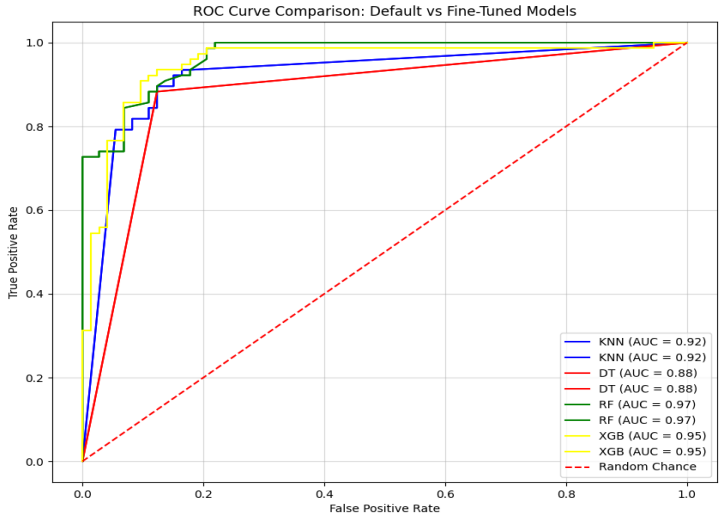

Fig. 9 Comparison of algorithms in terms of ROC curve

In the ROC chart above, as can be seen, the green curve, which is related to the random forest algorithm, is higher than the other curves, followed by the yellow curve, which is related to the XG Boost algorithm. Therefore, it can be said that based on the ROC curve, the performance of the random forest algorithm is in a better position than other algorithms. After determining the superior algorithm, it should be noted that the features are ranked in terms of importance and the effect they have on the output variable, i.e., high-consumption subscribers. In fact, based on the ranking of the features, it can be seen which feature can have the most impact on the classification of high-consumption, low-consumption, and medium-consumption subscribers in the smart gas meter network. The results are presented in the form of output graphs from the Python software.

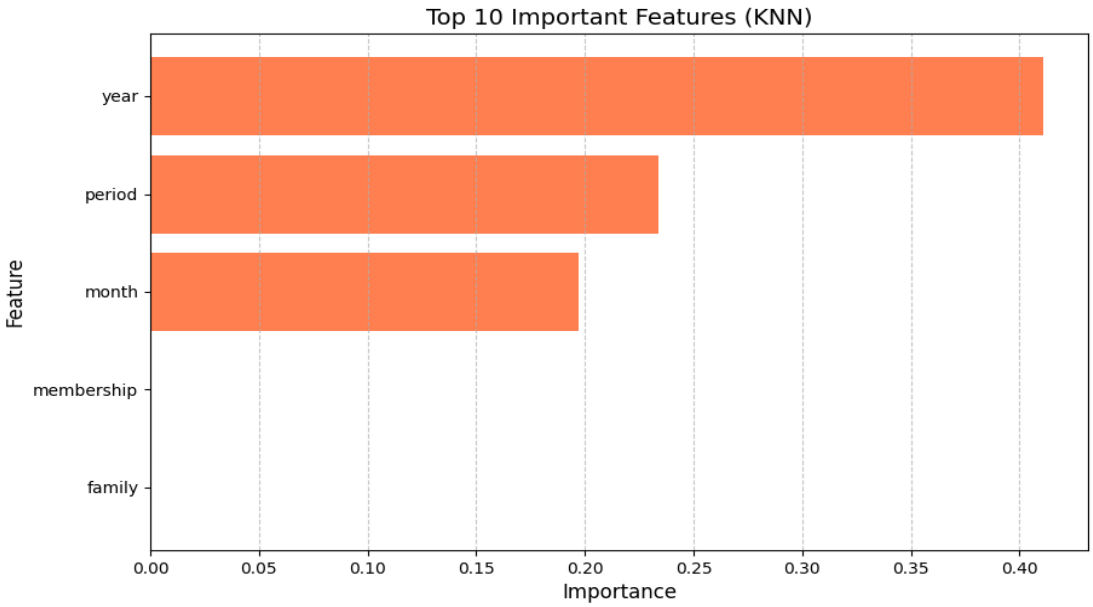

Fig. 10 Feature ranking based on the nearest neighbor algorithm

[Downloaded from ijaor.com on 2026-09-08]

[DOI: 10.71885/ijorlu-2025-2-706]

According to the nearest neighbor algorithm, annual consumption is more influential than other variables, followed by period and month consumption (Figure 10).

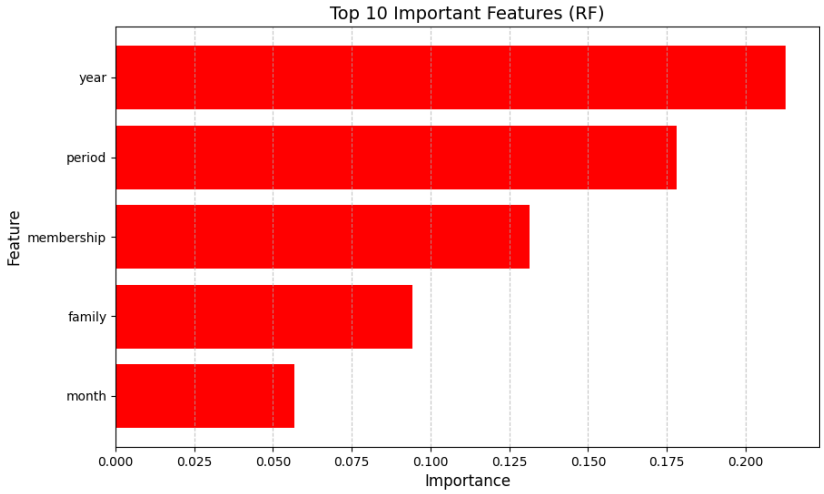

Fig. 11 Feature ranking based on the random forest algorithm

In Figure 11, it can be seen that the random forest algorithm ranks the importance of the features in the form of annual consumption, period consumption, subscription type, household dimension, and monthly consumption, respectively. In other words, according to this algorithm, the most important feature is annual consumption, which is similar to the nearest neighbor algorithm. Given that the superior algorithm in the present study is the random forest algorithm, the results obtained from the ranking by this algorithm are considered the criterion for the conclusion in the present study.

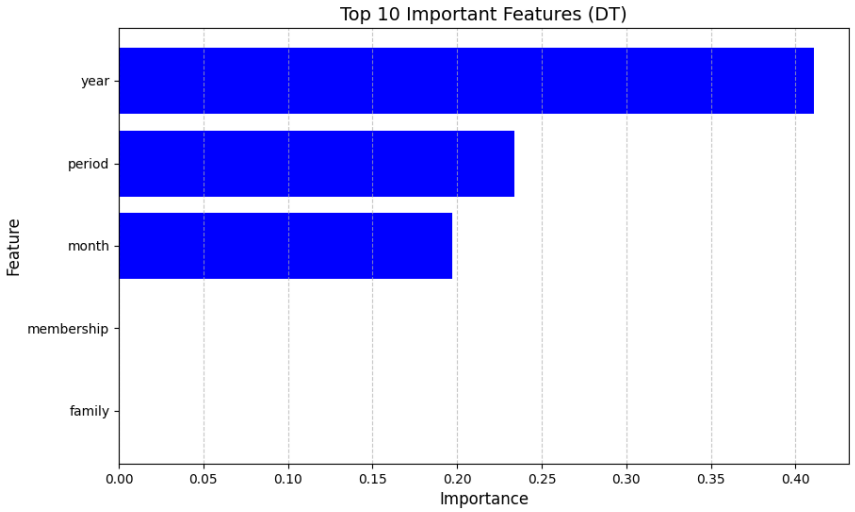

Fig. 12 Ranking of features based on the decision tree algorithm

In Figure 12, it can be seen that the decision tree algorithm performs the ranking similarly to the nearest neighbor algorithm and still emphasizes annual consumption as the most influential characteristic, followed by period consumption and monthly consumption.

[Downloaded from ijaor.com on 2026-09-08]

[DOI: 10.71885/ijorlu-2025-2-706]

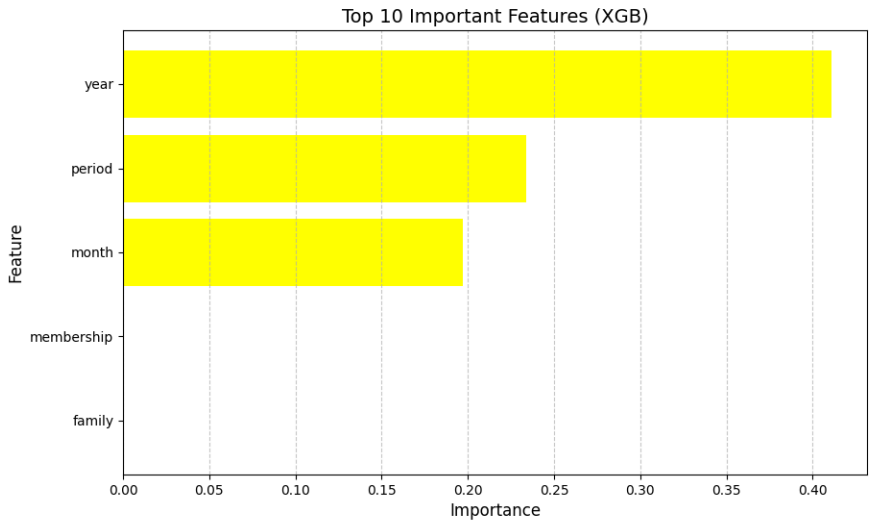

Fig. 13 Feature ranking based on the XG Boost algorithm

Although the XG Boost algorithm showed a strong performance in the present study, it provided weaker results compared to the random forest algorithm. Based on Figure 13, it can be seen that this algorithm also considers the ranking of features separately, including annual consumption, period consumption, and monthly consumption. It is noteworthy that the reason why the random forest algorithm ranks the features completely and ranks all 5 features is the complete performance of this algorithm, which distinguishes it from other algorithms in terms of predicting and classifying high-consumption subscribers.

5 Conclusion

In this research, the identification and control of high-consumption and abnormal subscribers in the smart gas meter network were addressed using machine learning algorithms and the Internet of Things in a cloud computing environment. Using a dataset of gas subscribers in the smart meter network, learning was performed by the algorithms, and based on the learning, subscribers were divided into three categories of high-consumption subscribers, with medium consumption and low consumption, and the machine learning algorithms based on classification were able to differentiate subscribers based on 5 input variables, annual consumption, monthly consumption, consumption period, household size and type of consumption subscription. With an accuracy of 0.92, the random forest algorithm was able to classify gas-consuming subscribers, while other algorithms showed slightly lower accuracy of about 0.9 and 0.91. In terms of ROC curves and confusion matrix, the random forest algorithm also performed better than the others, which shows that the random forest algorithm is able to classify high-consumption subscribers with an accuracy of over 90 percent, and in fact, the 5 input variables introduced, including annual, monthly, period, household size, and type of subscription, are able to explain the consumption of gas subscribers in the smart meter network by up to 92 percent. Therefore, it can be said that the model has high power in predicting and classifying high-consumption subscribers. On the other hand, based on the feature ranking, it was found that annual consumption can explain the consumption more than other features, followed by period consumption and type of subscription. The present study attempted to design a model based on big data from the Internet of Things and with 5 input variables, which showed an

[Downloaded from ijaor.com on 2026-09-08]

[DOI: 10.71885/ijorlu-2025-2-706]

accuracy of over 92 percent. Future research can implement the model of the present study in other smart meter networks, such as electricity and water, or evaluate the model of the present study with new variables. Developing predictive models based on deep learning to predict subscriber behavior and early identification of abnormal subscribers can help improve energy consumption management. Also, research on designing and developing effective management policies based on data analysis results to optimize energy consumption and reduce costs for subscribers and utility companies can be considered for future research by researchers. The results of the study can provide industry managers with insight into how to identify high- and low-consumption customers and implement incentive and penalty policies regarding consumption, which are now being considered as macro-energy management policies around the world.

References

1. Hernández, Á., Nieto, R., de Diego-Otón, L., Pérez-Rubio, M. C., Villadangos-Carrizo, J. M., Pizarro, D., & Ureña, J. (2024). Detection of anomalies in daily activities using data from smart meters. *Sensors*, *24*(2), 515.
2. Otuoze, A. O., Mustafa, M. W., Sultana, U., Abiodun, E. A., Jimada-Ojuolape, B., Ibrahim, O., ... & Abdullateef, A. I. (2024). Detection and confirmation of electricity thefts in Advanced Metering Infrastructure by Long Short-Term Memory and fuzzy inference system models. *Nigerian Journal of Technological Development*, *21*(1), 112-130.
3. Nkenyereye, L., Thakare, A., Khataniar, P., Imandi, R., & BN, P. K. (2025). Lightweight Authentication Protocol for Smart Grids: An Energy-Efficient Authentication Scheme for Resource-Limited Smart Meters. *Mathematics*, *13*(4), 580.
4. Ibrahim, Q., & Qassab, M. (2025). Smart City and Smart Metering: A Technological Infrastructure for Future Urban Development.
5. Koukouvinos, K. G., Koukouvinos, G. K., Chalkiadakis, P., Kaminaris, S. D., Orfanos, V. A., & Rimpas, D. (2025). Evaluating the performance of smart meters: insights into energy management, dynamic pricing and consumer behavior. *Applied Sciences*, *15*(2), 960.
6. Xia, X., Xiao, Y., Liang, W., & Cui, J. (2022). Detection methods in smart meters for electricity thefts: A survey. *Proceedings of the IEEE*, *110*(2), 273-319.
7. Zhou, Z., Liu, Y., Hu, T., & Wang, C. (2023). Two unsupervised learning algorithms for detecting abnormal inactivity within a household based on smart meter data. *Expert Systems with Applications*, *230*, 120565.
8. Severiche-Maury, Z., Uc-Rios, C. E., Arrubla-Hoyos, W., Cama-Pinto, D., Holgado-Terriza, J. A., Damas-Hermoso, M., & Cama-Pinto, A. (2025). Forecasting Residential Energy Consumption with the Use of Long Short-Term Memory Recurrent Neural Networks. *Energies*, *18*(5), 1247.
9. Fang, J., Liu, F., Su, L., & Fang, X. (2024). Research on Abnormity Detection based on Big Data Analysis of Smart Meter. *WSEAS Transactions on Information Science and Applications*, *21*, 348-360.
10. Kawoosa, A. I., Prashar, D., Anantha Raman, G. R., Bijalwan, A., Haq, M. A., Aleisa, M., & Alenizi, A. (2024). Improving electricity theft detection using electricity information collection system and customers' consumption patterns. *Energy Exploration & Exploitation*, *42*(5), 1684-1714.
11. Mbey, C. F., Bikai, J., Yem Souhe, F. G., Foba Kakeu, V. J., & Boum, A. T. (2024). Electricity Theft Detection in a Smart Grid Using Hybrid Deep Learning-Based Data Analysis Technique. *Journal of Electrical and Computer Engineering*, *2024*(1), 6225510.
12. Farooq, A., Shahid, K., & Olsen, R. L. (2024). Securing the green grid: A data anomaly detection method for mitigating cyberattacks on smart meter measurements. *International Journal of Critical Infrastructure Protection*, *46*, 100694.
13. Softah, W., Tafakori, L., & Song, H. (2025). Analyzing and predicting residential electricity consumption using smart meter data: A copula-based approach. *Energy and Buildings*, *332*, 115432.
14. Blazakis, K., Schetakis, N., Badr, M. M., Aghamalyan, D., Stavrakakis, K., & Stavrakakis, G. (2025). Power theft detection in smart grids using quantum machine learning. *IEEE Access*.
15. Chaudhari, A. Y., Mulay, P., & Chavan, S. (2025). The role of smart electricity meter data analysis in driving sustainable development. *MethodsX*, *14*, 103196.
16. Farooq, A., Shahid, K., & Olsen, R. L. (2025). Prioritization of smart meters based on data monitoring for enhanced grid resilience. *Computer Communications*, *234*, 108082.

[Downloaded from ijaor.com on 2026-09-08]

[DOI: 10.71885/ijorlu-2025-2-706]

17. Samieifard, M., Abolghasemian, M., & Pourghader Chobar, A. (2024). The impact of innovation, performance, and e-commerce development in the online shop on online marketing: A case study in the industry. *Interdisciplinary Journal of Management Studies*, *18*(1), 1-17.
18. Vaseei, M., Agha, M. N. J., Abolghasemian, M., & Chobar, A. P. (2024). Investigating the role of transformative technologies and smart processes on sustainable business. In *Building smart and sustainable businesses with transformative technologies* (pp. 38-51). IGI Global Scientific Publishing.

[Downloaded from ijaor.com on 2026-09-08]

[DOI: 10.71885/ijorlu-2025-2-706]