Revolutionizing Energy Management System with Machine Learning based Energy Demand Prediction and Theft Detection

Abilesh K S
Electrical and Electronics Engineering
Kumaraguru College of Technology
Coimbatore, India.
abilesh.20ee@kct.ac.in

Brian Maurish J
Electrical and Electronics Engineering
Kumaraguru College of Technology
Coimbatore, India.
brianmaurish.20ee@kct.ac.in

Naresh B
Electrical and Electronics Engineering
Kumaraguru College of Technology
Coimbatore, India.
naresh.20ee@kct.ac.in

Maithili P
Electrical and Electronics Engineering
Kumaraguru College of Technology
Coimbatore, India.
maithili.p.eee@kct.ac.in

***Abstract*—The project revolves around the comprehensive management of energy resources through real-time monitoring, theft detection, and demand prediction. It employs AC current and voltage sensors to monitor energy consumption, with data being timestamped and compiled into a dataset. Integrating Machine Learning (ML) techniques into energy management systems has emerged as a promising solution to address this challenge. The future energy demand is forecasted using Long Short-Term Memory (LSTM) neural networks, allowing for better resource planning. Additionally, theft detection is done through a Support Vector Machine (SVM) algorithm trained on historical data. The combination of these technologies is anticipated to revolutionize the energy sector by enabling proactive decision-making, fostering energy conservation, and minimizing financial losses due to energy theft. The gathered information is securely uploaded to the ThingSpeak cloud platform, ensuring data integrity and accessibility. To enhance user engagement and control, a web application is developed for visualizing energy usage, demand forecasts, and theft alerts. This project addresses the pressing need for efficient energy management, enabling users to make informed decisions and combat energy theft while contributing to sustainability and resource conservation.**

***Keywords*—Machine Learning, LSTM, SVM, demand forecast, theft Detection, data analysis, python, simulink**

I. INTRODUCTION

Energy forecasting has been a prominent issue because of the rising penetration of renewable energy resources (RES) in today's power system. Decision-makers and grid operators must know how much power RES will generate in the upcoming days and hours. In addition, the operation and planning of contemporary power systems depend heavily on the ability to forecast load demand and consumption. When there is less demand for power and excess power output from the RES, electrical energy storage is required [1]. It cannot, however, be stored because of the excessive cost, demanding maintenance, and short lifespan of energy storage. Utilities must constantly balance supply and demand as a result. These restrictions give rise to several intriguing features of energy forecasting, such as the requirement for exact precision and data collecting. Errors in forecasting result in an imbalance between supply and demand, which hurts operating costs, dependability, and efficiency.

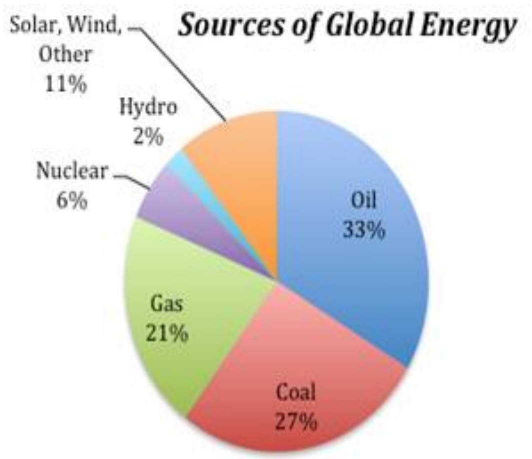

Fig. 1. Energy Production from Various Resources.

Models that are driven by data and make use of historical data on electricity output are viable options. In this research, data pre-processing techniques to evaluate the accuracy of the performance of the data-driven forecasting model by utilizing the available historical power data. The objective is to create a unified model for long-term forecasting with a step of short-term (hourly) precision and select an appropriate anomaly detection technology and data-driven methodology for energy production forecasting. Concerns about using energy resources sustainably are becoming more widespread. Fig. 1 shows the energy production from various resources. With population growth and industrialization, energy management has become a crucial issue. The project's goal is to solve these problems in this environment by developing an advanced system for demand forecasting, energy monitoring, and theft detection. Renewable energy sources can be incorporated into

979-8-3503-8434-5/24/$31.00©2024 IEEE

forecasting and management systems through data integration, specialized forecasting models, real-time monitoring, optimization algorithms, and grid integration. Data from renewable sources, such as solar irradiance and wind speeds, can be integrated into forecasting models to predict energy generation. Integration platforms aggregate data feeds from renewable sources, ensuring interoperability with existing infrastructure.

II. Proposed Approach

The following steps are proposed for the implementation of an Energy Management System with ML-based demand prediction and theft detection:

- Real-time monitoring with AC Sensors
- Data Upload to ThingSpeak cloud platform
- Data Timestamping and Compilation
- Integration of Machine Learning (ML) Techniques
- Demand Prediction using LSTM Neural Networks
- Theft Detection with Support Vector Machine (SVM) Algorithm

The primary problem this project seeks to address is the inefficient management of energy resources. Traditional methods of energy consumption monitoring often lack accuracy and real-time insights, leading to suboptimal resource allocation. Moreover, energy theft is a significant issue, resulting in substantial revenue losses for energy providers and an imbalance in the distribution of resources [2]. Fig. 2 shows the block diagram of the proposed system. The lack of reliable theft detection mechanisms further compounds these problems. Tackle these challenges, this project integrates advanced technology and data analytics to enable precise energy monitoring, proactive theft detection, and accurate demand prediction.

A. *Demand Prediction*

Demand prediction is a pivotal component of the project. Accurate forecasting of energy demand is essential for optimizing energy generation, distribution, and utilization. The project leverages Long Short-Term Memory (LSTM) neural networks, a type of recurrent neural network (RNN), to analyse historical energy consumption data and make predictions for future demand [3]. By employing LSTM models, which are well-suited for time series data, the system can provide users with valuable insights into when and how much energy will be required. This information can help energy providers plan more efficiently, reduce wastage, and improve resource allocation. In an LSTM neural network for forecasting future energy demand, specific features include historical energy consumption data, temporal patterns like seasonality and time of day, current consumption data, and historical trends. Model parameters such as the number of LSTM units, layers, dropout rates, and learning rates are adjusted during training for optimal performance.

Mean Absolute Error (MAE) quantifies the average absolute disparity between predicted and actual energy demand values, with lower MAE indicating greater accuracy. Forecast Bias gauges the model's tendency to consistently over or underestimate demand, ideally hovering near zero for balanced predictions. Visual Inspection involves comparing predicted values to actual ones via plots or charts, offering a qualitative assessment of the model's ability to capture trends, seasonality, and anomalies, providing valuable insights into its performance.

B. *Theft Detection*

Energy theft, whether through meter tampering or other illicit means, is a persistent problem in the energy industry. It leads to substantial revenue losses for energy providers and can even pose safety hazards. The project integrates a Support Vector Machine (SVM) algorithm for theft detection. This algorithm is trained on historical energy usage patterns to identify anomalies or irregularities in real-time data. When suspicious patterns are detected, the system can trigger alerts, allowing for immediate investigation and action. Various potential theft possibilities are trained in real-time to identify the theft with more accuracy. By effectively addressing energy theft, the project contributes to more equitable energy distribution and sustainability.

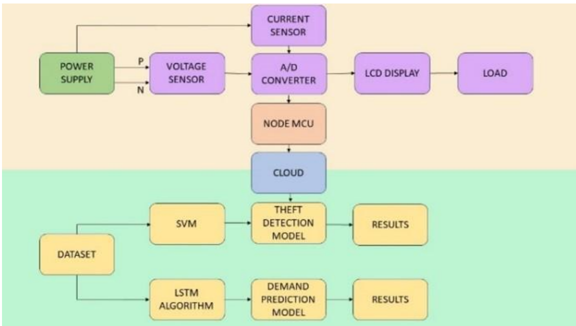

Fig. 2. Block Diagram of The Proposed System.

The voltage and current values drained by the load in the household are measured by the AC voltage and current sensor modules which are sent to the NodeMCU unit by converting the analog value into digital. AC current and voltage sensors are calibrated to ensure accurate measurements and then installed at key points within the electrical system. They are wired to capture readings, which are digitized by data acquisition devices and transmitted to the central monitoring system. Integrated into the energy monitoring system, these readings are timestamped, stored, and analysed in real-time for various functions such as demand forecasting and theft detection. Regular validation checks and maintenance ensure sensor accuracy and reliability, allowing for effective energy management. The microcontroller sends the values to the cloud and local display unit. ThingSpeak is used as a cloud interfacing unit for our project which continuously stores the data from the NodeMCU. A real-time graph for the data will be plotted and the data could be extracted in the form of an Excel sheet [4].

To mitigate communication failures between sensors and the central monitoring system, several strategies can be deployed. Redundant communication channels, such as Wi-Fi, cellular networks, or wired connections, offer backup options to ensure continuous data transmission despite disruptions. Data buffering and caching mechanisms in sensors store information locally during outages, guaranteeing no loss of data when connectivity is regained. Additionally, implementing remote monitoring and diagnostic capabilities enables administrators to proactively assess the health of sensors and communication links, facilitating timely maintenance and issue resolution. The data will be fetched to the machine learning models such as LSTM which is made to run locally through Thonny Python IDE for energy demand prediction and SVM through Google collab for theft detection. All the results from the models will be made available on the customized website for the user. The data in the cloud are well protected through the authentication key and login page for the web page, which prevents hackers from cyber-attacks.

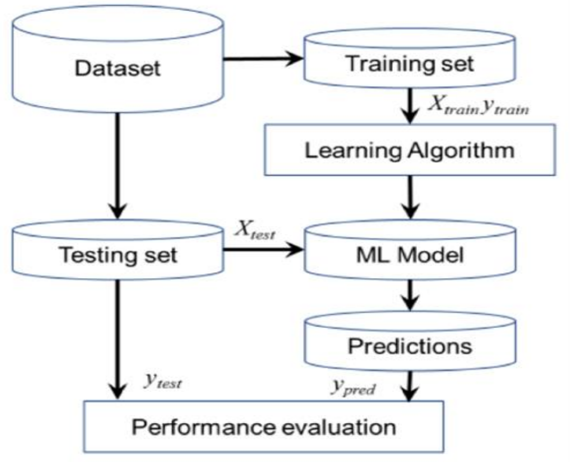

Fig. 3. Machine Learning Process Flowchart.

The necessity of this project is measured by the growing importance of sustainable energy management. With the depletion of natural resources and environmental concerns, it is imperative to use available energy resources efficiently. Accurate demand prediction ensures that energy generation and distribution is aligned with actual requirements, reducing waste and environmental impact. Furthermore, theft detection is essential to curbing illegal energy consumption, thereby ensuring fair access and revenue protection. The project's integration with cloud-based platforms enhances accessibility and scalability, making it an invaluable tool for both energy providers and consumers. The advantages of this project are manifold. First, it empowers energy providers with real-time insights into energy consumption, enabling them to make data-driven decisions for resource allocation. Second, the integration of demand prediction ensures that energy is produced and distributed efficiently, minimizing wastage. Third, theft detection contributes to revenue protection and resource conservation.

The cloud-based data storage and web application enable easy access, monitoring, and control for users, enhancing transparency and accountability in energy management. Fig. 3 shows the machine learning process flowchart. The project offers a holistic solution to the pressing challenges of energy resource management, serving the interests of both providers and consumers while promoting sustainability and responsible energy usage [5]. LSTMs excel in recognizing subtle relationships and adjusting to changing demand patterns, which is crucial in dynamic business environments. Additionally, their ability to selectively remember and forget information makes them effective in capturing both short-term fluctuations and long-term trends in demand. Applying Support Vector Machines (SVM) for energy theft detection presents several advantages. SVMs are particularly well-suited for binary classification tasks, making them effective in distinguishing between normal energy consumption patterns and anomalous behaviours associated with theft.

III. Simulation Model And Integration

MATLAB Simulink offers a robust platform for electrical circuit simulation, featuring a dedicated library of blocks representing key components like resistors, capacitors, inductors, sources, and semiconductor devices. Users leverage these blocks to construct detailed circuit diagrams, enabling dynamic system simulation under diverse conditions. Simulink supports various solvers to accurately capture transient responses, frequency characteristics, and steady-state behaviour. Fig. 4 shows the Simulink model of the proposed system. The integration of custom MATLAB code enhances modelling flexibility, making it a valuable tool for engineers and researchers. Applications span from electronic circuit design and optimization to the analysis of control systems within electrical devices, showcasing Simulink's versatility in modelling and analysing intricate electrical phenomena.

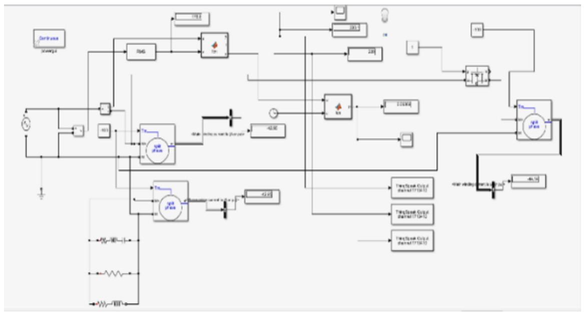

Fig. 4. Simulink Model of The Proposed System.

The simulation model of the proposed work, designed in MATLAB SIMULINK is shown in the above figure. The model includes the simulated block components of the MATLAB SIMULINK. The simulation model includes,

- AC supply fed to five types of loads.
- Scope for measuring the current across the load.
- Additional AC motor circuit controlled by a MOSFET switch for theft circuit.
- A switch to turn ON and OFF the MOSFET switch.
- MATLAB function block to convert the voltage and current into power.
- ThingSpeak block attached to export the readings to the cloud.

An AC supply is fed to five types of loads which consist of two split-type induction type motors with their loaded condition, an R load which may be of iron boxes or heaters, an RL load of any other type of motor, and an RLC load. A current and voltage sensor monitors the RMS current and voltage parameters of these loads continuously. Suitable formulas are applied for the MATLAB function blocks to determine the power and the energy drawn by the load. This system forms the household block. All the readings are uploaded to the ThingSpeak cloud by using the ThingSpeak module present in MATLAB [6].

Security measures implemented to safeguard the integrity of data uploaded to the ThingSpeak cloud platform include encryption of data transmission using protocols like HTTPS, role-based access control for restricted data access, strong authentication methods such as username/password or API keys, data validation to prevent injection attacks, regular data backups, logging and monitoring of user activities, and compliance with industry standards and regulations. Another block is set up to simulate the theft of electricity from the household. The theft circuit consists of a split-phase induction motor which directly connected to the main supply of the household. The theft circuit can be activated or deactivated by using a MOSFET switch. The theft circuit is purposefully kept turned Off for collecting and modeling data for a non-theft period. The theft circuit is kept on for collecting and modeling data for the theft period.

IV. Results And Discussion

The Simulation results by MATLAB and the results are shown below.

A. *Data Collection from The ThingSpeak*

The data obtained through the simulation of the model is uploaded to the cloud using the ThingSpeak module using the respective Channel ID and API key. Three fields are set up to obtain the Current, Voltage, and Energy consumption respectively in real time.

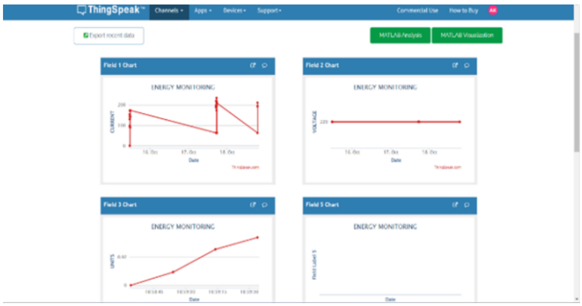

Fig. 5. ThingSpeak Result.

The data are uploaded to the cloud using a suitable microcontroller with Wi-Fi capability. The data uploaded to the cloud could be extracted to fetch it into the Machine Learning models for energy prediction and theft detection. Fig. 5 shows the ThingSpeak result.

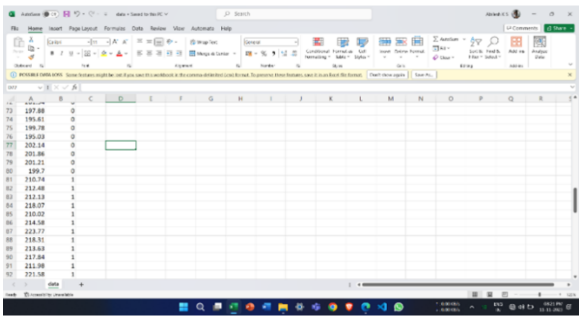

Fig. 6. Data collection on Excel sheet.

Fig. 6 shows the data collected on an Excel sheet. The values obtained from the simulation are targeted as 1 to indicate that the theft circuit is intentionally kept On and the values obtained from the simulation are targeted as 0 to that the theft circuit is intentionally kept Off [7]. These historical data on energy consumption will be analysed by the SVM model to determine any abnormalities in energy consumption. Before compiling timestamped energy consumption data into a dataset for machine learning, several preprocessing steps are conducted. This involves cleaning the data by managing missing values and outliers, normalizing features to ensure uniform scales, and performing feature engineering to extract relevant information. Categorical variables are encoded, and temporal aggregation may be applied to reduce dimensionality and capture trends. Finally, the dataset is split into training and testing sets for model evaluation.

B. *Evaluation of The SVM Model for Theft Detection*

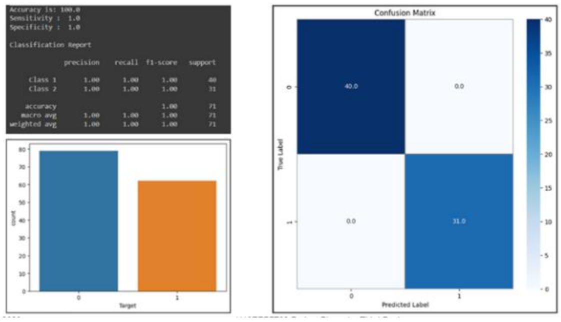

Fig. 7. Model Evaluation of SVM.

Model evaluation involves using various metrics and techniques to gauge how well a trained machine-learning model generalizes to new, unseen data. Fig. 7 shows the Model Evaluation of SVM. Common evaluation metrics include accuracy, precision, recall, F1 score, mean squared error, and many others, depending on the nature of the problem (classification, regression, etc.). Cross-validation is another technique used to assess a model's performance by splitting the dataset into multiple subsets for training and testing. The overarching goal is to ensure that the machine learning model can make accurate predictions on new, previously unseen data and is not overfitting or underfitting the training data [8].

Historical data is utilized in training the Support Vector Machine (SVM) algorithm for theft detection by first identifying relevant features through feature engineering, labelling data to indicate instances of theft, and then training the SVM model to classify new instances of energy usage. Measures to ensure the model's reliability include data quality assurance, balancing the dataset, regular model updating with new data, employing cross-validation techniques, and tuning decision thresholds. These steps collectively enhance the model's accuracy and effectiveness in detecting energy theft, thereby contributing to more proactive energy management within the system.

C. *Demand Forecast Result*

Demand forecasting is developed in Thonny Python using NumPy, Pandas, MatPlotLib, TensorFlow, and SKlearn libraries. By suitable selection of the training and test dataset, the model is trained with an epoch value of 50. Higher the value of epoch results in higher accuracy of the model. The graph is plotted using the plot function available.

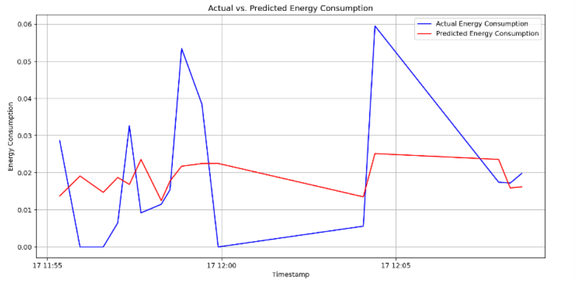

Fig. 8. Actual vs. Predicted Energy Consumption.

The LSTM model, designed to forecast future energy consumption by leveraging historical data, has yielded results that are thoughtfully presented in a graphical format to enhance user comprehension. Fig. 8 shows the Actual versus the predicted energy consumption. The visual representation allows for a clear and intuitive understanding of the model's predictions [9]. Through a graphical interface, users can easily interpret trends, patterns, and fluctuations in the predicted energy consumption over time. Facilitate a comprehensive evaluation, the predicted values are periodically compared against the actual energy consumption data. This side-by-side comparison provides users with a valuable benchmark for assessing the accuracy and reliability of the LSTM model. By visually aligning the predicted and actual consumption values on the same graph, users can readily identify any disparities and gain insights into the model's performance.

The graphical representation not only serves as a tool for post-analysis but also enables real-time tracking of how well the model aligns with actual energy consumption patterns. This visualization fosters transparency and accountability in the forecasting process, empowering users to make informed decisions based on the model's predictions. The graphical representation of LSTM model results serves as a crucial element in making complex energy consumption data more accessible and user-friendly [10]. It transforms raw predictions into a visually comprehensible format, facilitating a more intuitive and effective interpretation of the model's forecasting capabilities. The system continuously collects real-time data on energy consumption, processes it to extract features, and updates machine learning models like LSTM networks for demand forecasting. The performance of the model is evaluated and refined using feedback loops, while dynamic threshold adjustments ensure adaptability to changing patterns. Automated alerts notify stakeholders of significant deviations, facilitating timely interventions. This iterative process ensures that the models remain accurate and effective in predicting energy consumption and detecting anomalies.

D. Website Development for Users

A user-friendly website has been developed to provide an interactive platform for live remote monitoring of household status, system theft detection, and future energy consumption prediction. The website's front end is crafted using a combination of HTML, CSS, and JavaScript, ensuring a visually appealing and responsive design.

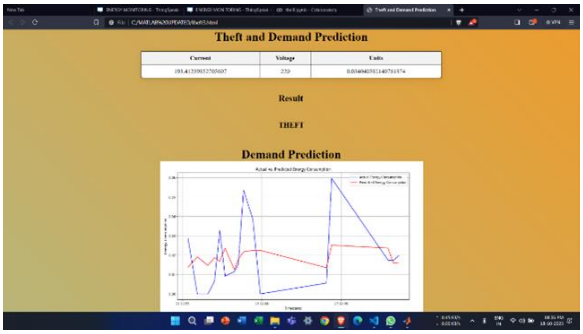

Fig. 9. Website for Users.

Fig. 9 shows an interactive website designed for users. Through a seamless user interface, individuals can access real-time updates on the status of various household parameters, allowing for immediate awareness and response to any irregularities. The incorporation of JavaScript enhances the website's functionality, enabling dynamic and asynchronous updates without requiring page reloads. Users can easily navigate through different sections of the website, accessing detailed information on the current state of their household systems and energy consumption patterns. User authentication and authorization within the web application are ensured through methods like username/password authentication or two-factor authentication, verifying user identities against stored data. Authorization mechanisms control user access based on roles and permissions, limiting users to only relevant data and functionalities. Encryption techniques safeguard sensitive data during transit and storage, while secure session management prevents unauthorized access to user accounts. Adherence to secure coding practices, regular updates, and thorough logging and monitoring further bolster the application's security, safeguarding against potential threats and data breaches.

One of the notable features of this website is the inclusion of a theft detection system, which actively monitors and alerts users about any suspicious activities related to energy consumption. This proactive approach empowers users to take preventive measures promptly, contributing to enhanced security and resource efficiency. The web application uses intuitive and interactive graphs to visualize energy usage, demand forecasts, and theft alerts. Line charts display historical consumption patterns and real-time trends while overlaying predicted values for future trends. Theft alerts are highlighted with color-coded indicators or dedicated panels, pinpointing suspicious activities. Users can adjust time frames, zoom in on data points, and set threshold alerts for abnormal patterns, facilitating effective energy resource management.

V. CONCLUSION

In conclusion, this project presents an integrated system that effectively addresses the critical aspects of energy management, including real-time monitoring, demand prediction, and theft detection. By combining hardware components and advanced data analytics techniques, it offers a comprehensive solution to energy resource optimization and security challenges, ensuring efficient utilization of energy resources and safeguarding against unauthorized consumption. This project's contribution to sustainability, efficiency, and accountability makes it a valuable tool for energy providers and consumers in an increasingly energy-conscious world. Future work includes enhancing predictive analytics by refining machine learning models with advanced techniques like ensemble methods or deep learning, improving demand forecasting and anomaly detection accuracy. Additionally, the project aims to integrate renewable energy sources like solar or wind power into monitoring and management systems to optimize grid integration and overall resource efficiency. User interface enhancements will be made through iterative improvements based on user feedback and usability studies, while scalability and deployment will focus on expanding the project to larger systems and diverse geographical regions, ensuring reliability and compatibility with various energy infrastructures and regulatory environments.

REFERENCES

[1] Nitin K Mucheli, Umakanta Nanda,D Nayak,P K Rout,S K Swain,S K Das,S M Biswal,"Smart Power Theft Detection System",2019 Devices for Integrated Circuit (DevIC)

[2] Rohit Andore,S.S. Kulkarni,A. G Thosar,"Energy Meter and Power Theft Monitoring System",2023 IEEE International Students Conference on Electrical Electronics and Computer Science (SCEECS)

[3] Sumit Mohanty,M. Mohamed Iqbal,Parvathy Thampi M.S.,"Controlling and Monitoring of Power Theft using Internet of Things",2021 International Conference on Design Innovations for 3Cs Compute Communicate Control (ICDI3C)

[4] Sanujit Sahoo,Daniel Nikovski,Toru Muso,Kaoru Tsuru,"Electricity theft detection using smart meter data",2015 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT)

[5] Taimur Shahzad Gill,Durr E Shehwar,Hira Memon,Sobia Khanam,Ali Ahmed,Urooj Shaukat,Abdul Mateen,Syed Sajjad Haider Zaidi,"IoT Based Smart Power Quality Monitoring and Electricity Theft Detection System",2021 16th International Conference on Emerging Technologies (ICET)

[6] D. Syed, H. Abu-Rub, S. S. Refaat and L. Xie, "Detection of Energy Theft in Smart Grids using Electricity Consumption Patterns," 2020 IEEE International Conference on Big Data (Big Data), Atlanta, GA, USA, 2020, pp. 4059-4064, doi: 10.1109/BigData50022.2020.9378190.

[7] Mel Keytingan M. Shapi, Nor Azuana Ramli, Lilik J. Awalin, Energy consumption prediction by using machine learning for smart building: Case study in Malaysia, Developments in the Built Environment, Volume 5, 2021, 100037, ISSN 2666-1659, https://doi.org/10.1016/j.dibe.2020.100037. (https://www.sciencedirect.com/science/article/pii/S266616592030034X)

[8] Mehmet Güçyetmez, Husham Sakeen Farhan, Enhancing smart grids with a new IOT and cloud-based smart meter to predict the energy consumption with time series, Alexandria Engineering Journal, Volume 79, 2023, Pages 44-55, ISSN 1110-0168, https://doi.org/10.1016/j.aej.2023.07.071. (https://www.sciencedirect.com/science/article/pii/S1110016823006622)

[9] Ejaz Ul Haq, Can Pei, Ruihong Zhang, Huang Jianjun, Fiaz Ahmad, Electricity-theft detection for smart grid security using smart meter data: A deep-CNN based approach, Energy Reports, Volume 9, Supplement 1, 2023, Pages 634-643, ISSN 2352-4847, https://doi.org/10.1016/j.egyr.2022.11.072. (https://www.sciencedirect.com/science/article/pii/S2352484722024581)

[10] M. J. Jeffin, G. M. Madhu, A. Rao, G. Singh and C. Vyjayanthi, "Internet of Things Enabled Power Theft Detection and Smart Meter Monitoring System," 2020 International Conference on Communication and Signal Processing (ICCSP), Chennai, India, 2020, pp. 0262-0267, doi: 10.1109/ICCSP48568.2020.9182144.