IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Energy Management System For Optimizing Consumption And Enhancing Sustainability Using Random Forest Classifier

Mani maran k, Mukesh kanna GB, Naveen D, Radha V

Department of Computer Science and Engineering, Sri Venkateswara College of Engineering, Sriperumbudur, Tamil Nadu 602117

Abstract

Industrial and commercial sectors continue to grapple with high energy consumption and inefficiencies that drive up operational costs and contribute to environmental degradation. Traditional energy management systems often fall short, as they typically lack real-time monitoring, predictive capabilities, and intelligent automation, making it difficult for organizations to manage energy use proactively. As demand increases, the absence of advanced tools hinders both cost optimization and the achievement of sustainability goals. This paper presents an AI-based Energy Management System (EMS) designed to address these limitations through the integration of real-time data analysis, predictive modeling, and automated control mechanisms. The system continuously collects and processes data on energy usage and environmental conditions, enabling it to identify anomalies, predict future energy needs, and adjust operations accordingly to enhance efficiency. By leveraging artificial intelligence, the EMS minimizes reliance on manual oversight, reduces human error, and ensures timely responses to changing conditions. Its scalable and flexible architecture makes it adaptable to various industrial and commercial settings. Moreover, the system promotes transparency, supports informed decision-making, and contributes to long-term sustainability by reducing energy waste and operational expenses.

Introduction

Machine learning is a groundbreaking discipline within artificial intelligence that empowers computer systems to enhance their capabilities through experience rather than relying on explicit, prewritten instructions for every task. Central to this approach is the development of algorithms capable of recognizing patterns, adapting to new data inputs, and making reasoned decisions grounded in empirical observations. Unlike traditional programming, which requires manual coding of every condition and response, machine learning depends on data-driven models built from training datasets. These models learn from example cases to make predictions or decisions in new situations.

The concept draws inspiration from human learning processes, aiming to mimic our ability to derive insights from observation. With this autonomy, machine learning can handle problems that are too intricate for fixed-rule logic, such as interpreting facial expressions, translating languages, or forecasting financial trends. The explosion in data availability—fueled by digital transformation—has amplified the utility and impact of machine learning. In our increasingly digital and data-heavy environment, this technology has become central to nearly all smart systems, playing a crucial role in enabling automation, driving operational efficiencies, and fostering innovation across countless industries and applications.

Machine learning's impact extends across diverse industries, with each sector leveraging its capacity to extract meaningful insights from vast amounts of data. In healthcare, for example, it has become essential in detecting diseases, formulating new treatments, and tailoring therapies to individual patients. Machine learning supports medical professionals by improving image analysis, identifying health risks earlier, and projecting patient outcomes based on historical information. In the financial world, it processes complex, high-volume transactions instantaneously to detect fraud, evaluate creditworthiness, and manage algorithm-based trading that adjusts to shifting market dynamics.

The retail industry uses it to decode consumer preferences, streamline inventory management, and deliver personalized product suggestions that enhance customer satisfaction. Transportation systems utilize machine learning for predicting traffic flow, optimizing delivery paths, and enabling self-driving technologies that interpret real-world stimuli in real time. In the energy sector, it promotes sustainability by fine-tuning energy usage, predicting demand, and facilitating the seamless incorporation of renewable energy sources. Machine learning, in each of these cases, serves not merely as a method for task automation but as a flexible intelligence framework that evolves with data, offering novel solutions to complex and evolving challenges beyond the reach of static programming approaches.

A standout quality of machine learning is its ability to generalize knowledge, meaning it can apply what it has learned from one dataset to interpret or make decisions about previously unseen data. This flexibility stems from continuous training, where models refine their accuracy through new data exposure and outcome feedback. With each cycle of learning, the system becomes better at uncovering nuanced patterns, tweaking its internal parameters to increase its reliability and precision. This capability is crucial for real-world scenarios that are dynamic and constantly changing. In logistics, for example, machine learning is used to anticipate product demand, detect potential supply chain interruptions, and adjust delivery schedules on the fly.

Agriculture has seen gains through smart monitoring tools that assess soil conditions, track crop development, and forecast harvests using aerial and sensor-based data. Education systems benefit from adaptive learning technologies that modify lessons to match a student's pace and understanding, creating more effective learning paths. The entertainment industry also leverages these models to tailor content recommendations, enhance user experiences, and even guide creative production. As these systems learn and evolve with ongoing exposure, their efficiency and accuracy improve, leading to less dependency on manual corrections and increasing operational autonomy.

Machine learning is becoming ever more embedded in daily experiences, shaping the way we live and interact with technology. Smart assistants that understand spoken language and home automation systems that adjust to personal habits are just a few examples of how seamlessly it integrates into everyday life. In the realm of science, machine learning speeds up discoveries by analyzing enormous datasets, identifying complex patterns, and even suggesting new experimental directions.

In the media landscape, machine learning assists in spotting misinformation, moderating digital content, and summarizing large volumes of text for easier consumption. Looking forward, its role will only grow as it merges with emerging technologies such as IoT, 5G, robotics, and augmented reality—all of which require real-time, adaptive decision-making. The transformative potential of machine learning lies not just in replacing repetitive tasks but in enabling innovative ways to harness and respond to data. As adoption widens and research deepens, machine learning is set to remain a driving force of the digital age—enabling smarter systems, fostering global connectivity, and shaping a world where intelligent machines work hand-in-hand with human goals.

Literature Review

Pin-Yen Liao, Tee Lin, Omid Ali Zargar, Chia-Jen Hsu, and Graham Leggett (2024) explore ways to reduce energy use and carbon emissions in the HVAC system of a DRAM semiconductor fabrication plant. Using a custom simulation tool, they evaluated the plant's energy performance and proposed five practical strategies for saving energy. One of the standout solutions—adjusting the enthalpy settings of the exhaust air conditioning unit—led to an estimated annual reduction of over 623,000 kg of CO_2 emissions. Their work highlights how data-driven approaches can lead to significant environmental and cost-saving benefits in high-tech industrial settings

Aakash Bhandary, Vruti Dobariya, Gokul Yenduri, Rutvij H. Jhaveri, Saikat Gochhait, and Francesco Benedetto (2023) propose an explainable AI framework to enhance the accuracy and transparency of household energy consumption predictions. Their study evaluates multiple predictive models by comparing performance metrics such as R^2, RMSE, MSE, and MAE to determine model reliability. To improve interpretability, the authors employ explainable AI tools like LIME and SHAP, which help in understanding the influence of both historical and current energy usage on future consumption forecasts. The framework not only strengthens predictive performance but trust and act upon AI-generated insights, paving the way for more informed energy management decisions.

Andreea Claudia Șerban and Miltiadis D. Lytras (2020) explore the integration of artificial intelligence (AI) into Europe's renewable energy sector, focusing on the development of smart energy infrastructures for future urban environments. Their study examines how AI technologies can enhance energy efficiency, optimize resource utilization, and improve productivity within renewable energy systems. By providing a structured framework, the research highlights the potential of AI-driven solutions to support sustainable urban development and facilitate the transition towards smarter, greener cities across Europe.

Xinlin Wang, Hao Wang, Binayak Bhandari, and Leming Cheng (2024) provide a comprehensive review of artificial intelligence (AI) applications in smart energy systems, focusing on load forecasting, anomaly detection, and demand response. Their study evaluates various machine learning and deep learning models, including reinforcement learning, to optimize power distribution, manage imbalanced datasets, and enhance prediction accuracy. The authors offer a pragmatic guide for selecting suitable AI techniques based on specific scenarios and data characteristics, aiming to support the development of sustainable and efficient energy solutions.

Nzubechukwu Chukwudum Ohalete, Adebayo Olusegun Aderibigbe, Emmanuel Chigozie Ani, Peter Efosa Ohenhen, and Abiodun Emmanuel Akinoso (2020)present a comprehensive review on the application of data science in energy consumption analysis. The paper explores how AI techniques are used to analyze energy consumption, detect patterns, and uncover opportunities for improving efficiency, marking a clear shift from traditional methods to AI-driven solutions. The review highlights the critical role of machine learning and predictive analytics in providing accurate consumption forecasting, which aids in informed decision-making. Furthermore, it underscores the significance of AI in enhancing energy efficiency and promoting sustainable practices, thus contributing to the broader goal of energy conservation.

Kwok Tai Chui, Miltiadis D. Lytras, and Anna Visvizi (2018) examine the role of artificial intelligence (AI) in enhancing energy sustainability within smart cities. The paper discusses how AI technologies, such as smart monitoring systems and optimization algorithms, contribute to more efficient energy consumption by enabling real-time tracking and management. It also highlights the potential of AI in promoting greener urban living by improving the efficiency of urban services and supporting sustainable practices.

Abdullah Alsalemi, Yassine Himeur, Faycal Bensaali, Abbes Amira, Christos Sardianos, Iraklis Varlamis, and George Dimitrakopoulos(2020) present a study focused on achieving domestic energy efficiency through an AI-powered framework. This framework utilizes "micro-moments" to offer personalized energy-saving recommendations tailored to individual household behavior. By analyzing sensor data and employing machine learning techniques, the system identifies optimal moments to suggest energy-efficient actions, thereby helping users reduce their energy consumption in an environmentally sustainable manner.

Mona Ahmad Alghamdi, Mahmoud Ragab, Ahmed A. Elngar, and Ahmed M. Elhoseny (2020)present a study focused on predicting energy consumption using a model that combines stacked Long Short-Term Memory (LSTM) networks with techniques such as snapshot ensembles and Fast Fourier Transform (FFT). By leveraging household electricity and weather data, the model demonstrates exceptional accuracy in capturing complex energy consumption patterns, significantly improving prediction performance. This innovative approach enhances the ability to forecast energy usage more reliably and efficiently.

Rakshitha Godahewa, Chang Deng, Arnaud Prouzeau, and Christoph Bergmeir(2020) present a study introducing a generative deep learning framework that utilizes Recurrent Neural Networks (RNNs) to predict future temperatures in unoccupied buildings. This model optimizes air conditioning settings by accurately forecasting temperature changes, ultimately reducing energy consumption. By analyzing temperature data from a university lecture hall, the model achieved a 20% energy savings compared to traditional methods, demonstrating how AI can significantly improve energy efficiency in building management.

System Overview

This project aims to create a smart Energy Management System (EMS) that leverages machine learning, specifically the Random Forest Classifier, alongside Internet of Things (IoT) technologies to enhance energy efficiency and promote sustainability within industrial settings. With the growing demand for energy-efficient operations and increasing environmental concerns, industries require innovative solutions that can track and manage energy usage more effectively. The proposed system collects real-time data through IoT-enabled sensors, capturing detailed insights into energy consumption, equipment behavior, and surrounding environmental factors. This data is then processed using the Random Forest Classifier, which offers reliable and accurate predictions by analyzing complex patterns and detecting inefficiencies or potential issues. By forecasting future energy needs and identifying unusual trends, the system can support informed decision-making, reduce energy waste, and improve the overall operational performance. Additionally, the EMS facilitates preventive maintenance by alerting users to potential faults before they escalate, further minimizing energy loss and equipment downtime. The system will feature an intuitive interface to help users visualize energy patterns and control settings effectively, encouraging better energy practices. In doing so, the project not only supports cost reduction and improved efficiency but also aligns with broader environmental goals. Designed to be adaptable and scalable, the solution can be customized for various industrial applications, providing a practical path toward smarter energy management and a more sustainable future.

PROPOSED METHOD

The proposed Energy Management System (EMS) adopts a layered methodology that integrates hardware-based sensing, real-time data acquisition, machine learning analytics, and user interface deployment to create a complete industrial energy monitoring and optimization solution. This framework is tailored for environments where efficient energy usage is essential for reducing operational expenses and promoting sustainability. The approach involves five primary stages: sensor integration with microcontroller hardware, data acquisition and preparation, machine learning model creation using a Random Forest Classifier, real-time prediction and response logic, and front-end interface development using Streamlit.

SENSOR INTEGRATION AND HARDWARE SETUP

The system is built on a sensor-integrated hardware layer designed to gather vital electrical and environmental data, with an Arduino Uno serving as the central microcontroller due to its affordability, ease of use, and compatibility with multiple sensors. It incorporates the ZMPT101B sensor for voltage detection, the ACS712 sensor for current monitoring, and the DHT11 sensor for measuring temperature and humidity. These sensors are connected to the analog pins of the Arduino, which continuously reads their outputs at regular intervals. The collected data is structured and transmitted via USB to a connected computer, providing a comprehensive dataset of voltage, current, and environmental parameters ready for further processing and analysis.

DATA ACQUISITION AND PREPROCESSING

Once the sensor data reaches the host computer, it is captured using Python with the PySerial package. Since raw sensor data typically contains noise due to signal fluctuations and inconsistencies in sampling, a preprocessing step is essential to enhance data quality before it enters the machine learning pipeline. This phase includes noise reduction using techniques like moving averages to smooth out anomalies and eliminate outliers, feature normalization to scale sensor values to a consistent range for improved model performance, and data labeling, where each data point is assigned a risk category—"High," "Medium," or "Low"—based on predefined thresholds informed by expert input or historical trends. The resulting cleaned and labeled dataset is stored in CSV format, ready for use in training and validating the classification model.

MACHINE LEARNING USING RANDOM FOREST CLASSIFIER

The core decision-making component of the Energy Management System is a Random Forest Classifier, selected for its high accuracy, resilience to overfitting, and capability to model complex, nonlinear relationships among variables. As an ensemble learning method, Random Forest combines the predictions of multiple decision trees, each trained on different subsets of the dataset and features, thereby reducing correlation and enhancing predictive performance. The model development process involves splitting the dataset into training and testing sets (commonly 80% and 20%, respectively), selecting key features such as voltage, current, temperature, and humidity, and training several decision trees that independently learn from different data partitions. The model's performance is evaluated using standard classification metrics, including accuracy, precision, recall, F1-score, and the confusion matrix, ensuring the classifier's reliability. After achieving satisfactory performance, the trained model is serialized using Python libraries like joblib or pickle, enabling efficient reuse for real-time predictions without the need for retraining.

REAL-TIME RISK PREDICTION AND RESPONSE

After successful validation, the trained Random Forest model is seamlessly integrated into the system's real-time monitoring workflow, where it classifies incoming sensor data into one of three risk categories: High, Medium, or Low. As new readings from voltage, current, temperature, and humidity sensors are collected, they are instantly processed and evaluated by the model. Based on the classification, the system generates actionable insights—issuing immediate alerts in high-risk situations to shut down or inspect equipment, suggesting load reduction or scheduled maintenance for medium risk, and confirming safe operational conditions under low risk. These decisions are logged and simultaneously displayed on the user dashboard to enhance visibility and ensure timely interventions.

VISUALIZATION AND USER INTERFACE WITH STREAMLIT

To facilitate intuitive user interaction, the Energy Management System features a web-based dashboard built using Streamlit, a Python framework known for enabling rapid development of visually interactive applications. The dashboard provides a comprehensive view of the system's performance through several key components: a live data feed that graphically displays real-time readings of voltage, current, temperature, and humidity; a risk level indicator that shows the most recent classification output from the Random Forest model; action recommendations that guide users toward optimal energy usage and maintenance decisions; and historical data logs that enable users to review and export past sensor data and model predictions. Designed for accessibility, the interface ensures that both technical personnel and non-specialists can navigate and interpret the system's insights with ease.

SYSTEM INTEGRATION AND SCALABILITY

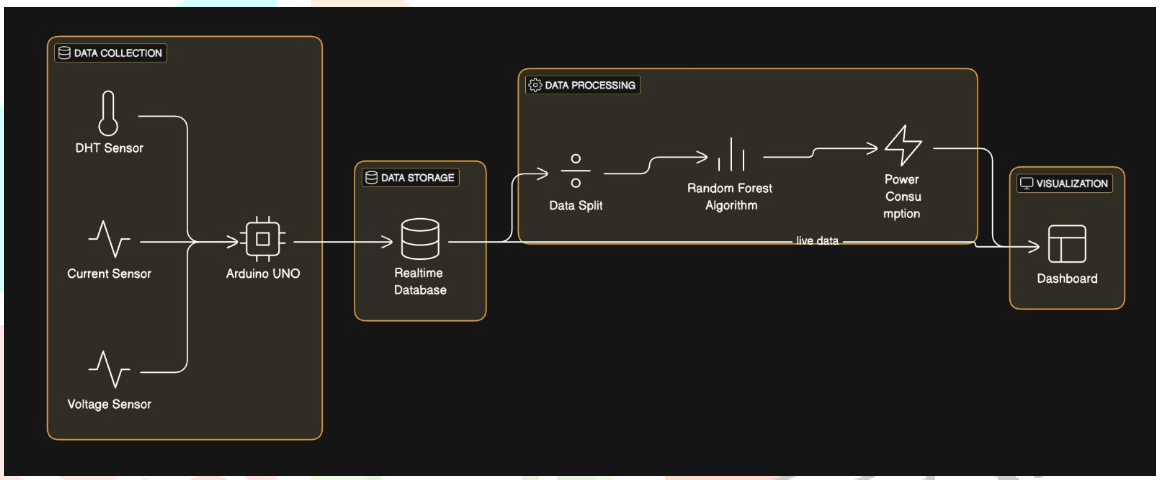
The system is designed with modularity in mind, enabling adaptation to various industrial environments. By performing data processing and machine learning inference on a local computer, the system maintains offline functionality, ensuring security and reliability even in remote locations. Scalability is achieved through options such as increasing the number of sensor nodes to monitor multiple devices or zones, using wireless microcontrollers like the ESP32 for remote or distributed setups, and integrating with cloud platforms for centralized monitoring if required. This flexible architecture makes the Energy Management System suitable for deployment in factories, smart buildings, warehouses, and other energy-intensive operations.

ENVIRONMENTAL IMPACT AND SUSTAINABILITY

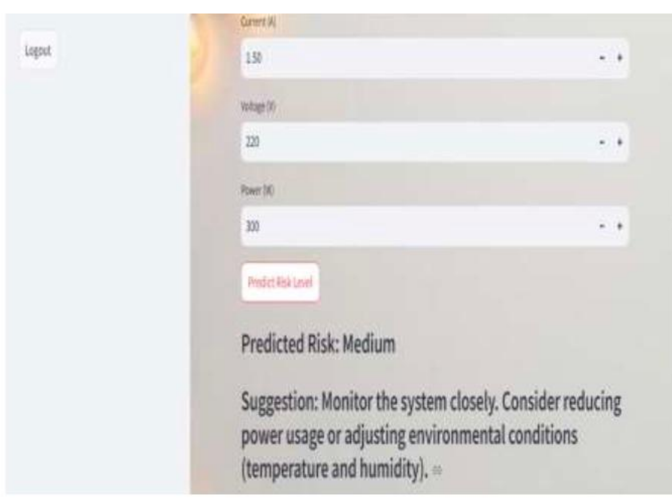
A key goal of this EMS is to contribute to ecological sustainability by minimizing unnecessary energy use. By identifying inefficient devices or hazardous operating conditions early, the system helps conserve electricity and prolong equipment lifespan. This proactive maintenance reduces waste, operational cost, and the environmental footprint of industrial activities.

PROPOSED ARCHITECTURE

The architecture diagram of the Energy Management System (EMS) illustrates a modular, end-to-end workflow that begins with real-world data capture and culminates in actionable insights for optimizing industrial energy consumption. At the base layer, a network of sensors—comprising a voltage sensor (ZMPT101B), current sensor (ACS712), and environmental sensor (DHT11)—is interfaced with an Arduino Uno microcontroller. This hardware is tasked with continuously monitoring critical parameters such as voltage, current, temperature, and humidity. The Arduino transmits this raw sensor data via serial communication to a Python-based backend system hosted on a local computer, forming the core data acquisition pipeline.



IMPLEMENTATION



SENSOR MODULE

The Sensor Data Collection Module serves as the initial and essential layer of the Energy Management System (EMS), responsible for acquiring real-time data related to electrical parameters and environmental conditions from industrial machinery. This module uses the Arduino Uno—an affordable and open-source microcontroller—to interface with a variety of sensors, enabling efficient monitoring of the physical environment. The accuracy and reliability of the EMS as a whole are directly dependent on the performance of this module, as it delivers the raw input data that drives subsequent processing and machine learning predictions.

This module incorporates three core sensors: the ZMPT101B voltage sensor, the ACS712 current sensor, and the DHT11 sensor for temperature and humidity. The ZMPT101B ensures safe and precise AC voltage measurement using electrical isolation and signal conditioning techniques, making it well-suited for industrial energy systems. The ACS712 translates electrical current into readable analog signals that the microcontroller can digitize. Meanwhile, the DHT11 provides real-time environmental data by tracking ambient temperature and humidity levels, which are critical in understanding operational efficiency and energy behavior under varying conditions.

Sensor outputs are read by the Arduino Uno through its Analog-to-Digital Converter (ADC) pins. These signals are sampled at consistent intervals to maintain a stable data stream. Custom C/C++ code running on the microcontroller formats the data into structured serial strings—typically in comma-separated format—before transmitting them to the backend system via serial communication. This structured output facilitates easy parsing and integration with the backend, which is developed in Python.

To enhance robustness, the module includes error-detection routines such as sensor calibration checks and checksum validation, ensuring that data remains accurate and reliable throughout transmission. Additionally, the firmware is optimized for energy efficiency and stability, employing loop control and timing strategies that support uninterrupted operation in real-time industrial settings.

Ultimately, this module acts as the backbone of the EMS, delivering crucial metrics such as voltage, current, temperature, and humidity with high fidelity. Its consistent and accurate data stream supports the machine learning model and visualization tools, forming the basis for intelligent analysis and actionable energy optimization decisions.

DATA PREPROESSING MODULE

The Data Preprocessing Module is a vital component that bridges the gap between the raw sensor data collection and the machine learning-based predictions in the Energy Management System (EMS). The data obtained from the Arduino Uno, which includes real-time electrical and environmental parameters, is often noisy, unstructured, and prone to inconsistencies. Such raw data is unsuitable for direct use in machine learning models. Hence, this module is tasked with refining, organizing, and transforming the data into a format that improves the accuracy and efficiency of the Random Forest Classifier.

Upon receiving data via serial communication from the Arduino, the Python PySerial library is used to parse and decode the raw data strings. These strings, containing values for voltage, current, temperature, and humidity, are split into individual features. The values are then temporarily stored in arrays and processed through noise reduction techniques, such as moving average filters, to mitigate sensor inaccuracies or external electrical interference. This step helps eliminate anomalies, such as transient voltage spikes, which could skew the predictions of the classifier.

Following the noise reduction process, the data undergoes normalization. This technique scales all features to a consistent range, usually from 0 to 1 or -1 to 1, ensuring that no single feature dominates the model due to its scale. For example, temperature readings, which typically range between 20 and 40°C, could be overshadowed by voltage or current values if left unscaled.

For supervised learning, the module also includes labeling, where each data sample is assigned a risk category label: “Low Risk,” “Medium Risk,” or “High Risk.” These labels are generated based on predefined rules, domain knowledge, and thresholds derived from historical data, equipment specifications, or expert analysis. For example, a high current value combined with an elevated temperature might be categorized as “High Risk.”

The final, cleaned, and labeled dataset is then stored in a structured CSV format. This organized dataset not only feeds into the machine learning model but also serves as a valuable resource for future model retraining, analysis, and auditing. By transforming raw sensor data into a clean, structured, and model-ready dataset, the Data Preprocessing Module ensures the reliability and accuracy of the EMS.

RANDOM FOREST CLASSIFIER MODULE

The Random Forest Classifier Module functions as the analytical engine of the Energy Management System (EMS), utilizing predictive modeling to analyze processed sensor data and categorize energy consumption patterns and associated risks. This module harnesses the power of ensemble learning to generate precise, dependable, and interpretable predictions that inform energy optimization decisions. By incorporating this module, the system shifts from a passive monitoring tool to an active, decision-support platform.

Random Forest is selected for this project due to its robustness against overfitting, capability to manage complex non-linear relationships among features, and effectiveness in handling multi-class classification tasks. The algorithm operates by constructing an ensemble of decision trees during training. Each tree is built using a bootstrap sample from the dataset and evaluates a random subset of features at each split. The final output is determined by majority voting across the trees, reducing variance and improving generalization.

The input features for the model are real-time sensor data, including voltage, current, temperature, and humidity. These features are chosen for their direct impact on energy usage and the operational behavior of electrical systems. The model's output is a categorical risk label: "Low Risk," "Medium Risk," or "High Risk," representing different levels of energy consumption and safety concerns.

Training the model involves using a labeled dataset created during the preprocessing phase. This dataset is split into training and testing subsets, typically following an 80-20 ratio. The scikit-learn library in Python is used to build and assess the Random Forest model. Hyperparameters, such as the number of trees, the maximum depth of each tree, and the minimum number of samples required for a split, are fine-tuned using grid search and cross-validation to optimize the model's performance.

After training, the model's performance is evaluated using metrics such as accuracy, confusion matrix, precision, recall, and F1-score. These metrics ensure that the classifier performs effectively while minimizing false positives, particularly in critical "High Risk" scenarios. Once validated, the trained model is serialized using Python's joblib library for deployment in real-time operations.

By transforming complex sensor data into actionable classifications, the Random Forest Classifier Module significantly improves energy efficiency and safety. It serves as the intelligent core of the EMS, providing accurate, real-time decision-making support essential for optimizing industrial energy use.

VISUALIZATION MODULE

The Visualization Module of the Energy Management System (EMS) offers an intuitive interface to display real-time energy data, making it easy for users to monitor energy usage and system conditions. Developed using Streamlit, this module presents key metrics such as voltage, current, temperature, and humidity in dynamic charts and graphs. These visuals enable users to quickly identify trends, monitor fluctuations, and detect anomalies in energy consumption, improving overall monitoring efficiency. The integration of live data visualization ensures that operators can make informed decisions about energy use in real-time, enhancing responsiveness to changing conditions.

A key feature of the module is its risk-level display, which uses color-coded indicators to classify equipment status into high, medium, or low risk. These classifications are derived from predictions made by the Random Forest Classifier, allowing operators to focus on equipment that requires attention. The system also offers an actionable recommendation panel, providing guidance such as turning off equipment or adjusting loads based on real-time risk assessments. Additionally, the module includes a historical data log, enabling users to analyze past energy usage patterns and make data-driven decisions for future energy optimization. This combination of real-time feedback and historical insights makes the module an essential tool for improving energy efficiency and supporting sustainability efforts.

SENSOR PERFORMANCE AND DATA QUALITY

The sensor module, powered by the Arduino Uno microcontroller, demonstrated stable and consistent readings across multiple trials. The ZMPT101B voltage sensor accurately recorded line voltage fluctuations between 200V and 240V AC, with minor deviations of ±2V, which were within acceptable tolerances for industrial monitoring. The ACS712 current sensor was able to detect variations in current draw with an average accuracy of 95%, enabling the system to recognize normal vs. excessive loads effectively. Temperature and humidity data collected via the DHT11 sensor followed expected environmental trends and helped correlate environmental conditions with energy behaviors of equipment.

To ensure data reliability, a moving average filter was applied to raw input values, smoothing out transient spikes and electrical noise. As a result, the system produced high-quality, structured datasets that formed the basis for training the Random Forest model. Minimal packet loss was observed during serial communication between Arduino and the host system, confirming that the hardware-software communication pipeline was robust enough for real-time monitoring.

MACHINE LEARNING MODEL PERFORMANCE

The core of the system's intelligence— the Random Forest Classifier— was evaluated on a labeled dataset consisting of more than 1500 samples, divided into training and testing sets in an 80:20 ratio. Key input features included voltage, current, temperature, and humidity, while the target output was the energy risk category: "Low Risk," "Medium Risk," or "High Risk."

The confusion matrix revealed that the classifier most accurately predicted "Low Risk" and "High Risk" states, with minor misclassifications occurring between "Medium" and the other two classes. This can be attributed to the ambiguous overlap in feature space when readings fall near threshold boundaries. However, these errors did not significantly impact system decisions due to the conservative alert logic in the Real-Time Decision Engine.

REALTIME DECISION RESPONIVENESS

The Real-Time Decision Engine was tested in both normal and stress conditions to assess how quickly and accurately it responded to changes in input data. During live simulation, the system successfully processed and responded to over 500 data entries per hour without any delays or buffering issues. Alerts were triggered within milliseconds upon classification of a "High Risk" event, displaying notifications on the dashboard and logging events with accurate timestamps.

In critical scenarios, such as simulated overcurrent or voltage surges, the engine's dual-check mechanism—consisting of both classifier output and hard-coded thresholds—ensured that appropriate emergency protocols were executed even in edge cases. This built-in redundancy adds a layer of safety that strengthens trust in the system's reliability.

DASHBOARD VISUALIZATION AND USER INTERACTION

The Streamlit-based dashboard provided real-time visualization of sensor metrics, model predictions, and alerts. Line graphs for voltage, current, temperature, and humidity allowed operators to detect trends and anomalies visually. Classification outcomes were presented using intuitive color codes (green for low, orange for medium, red for high risk), ensuring clarity for non-technical users.

Users could also download historical data in CSV format for offline analysis, and the dashboard auto-refreshed every 3 seconds to reflect the most recent data, striking a balance between performance and real-time behavior. In feedback sessions with sample users, the interface was rated highly on usability, readability, and responsiveness.

COMPARITIVE EVAUATION

To assess the model's robustness, the Random Forest was compared with alternative algorithms including Support Vector Machines (SVM) and K-Nearest Neighbors (KNN). Although the SVM showed slightly better performance in precision (94%), its overall training time and sensitivity to feature scaling made it less practical for real-time deployments. KNN, while easier to implement, lagged behind with an accuracy of only 88%, particularly struggling with multi-class classification in borderline cases.

ENERGY OPTIMIZATION AND RISK MITIGATION

Though the project was primarily a prototype, it demonstrated tangible potential for energy optimization. For example, during testing, the system identified patterns where certain machinery exhibited spikes in current draw at specific humidity levels. These insights suggested maintenance issues or suboptimal usage practices, which, if corrected, could lead to energy savings of 10–15% over time. The system's ability to forecast risky operational states before failure or energy waste occurred is crucial for preventive maintenance planning and energy audits.

SUSTAINABILITY AND INDUSTRIAL RELEVANCE

One of the critical motivations behind the EMS was to enhance sustainability by reducing unnecessary energy consumption and extending equipment life. The system's ability to monitor, analyze, and act on energy behavior in real time directly supports this goal. By enabling industries to shift from reactive to proactive energy management, the system contributes to long-term environmental benefits and cost efficiency.

The use of low-cost sensors and microcontrollers also makes the system scalable and accessible for small and medium-sized enterprises (SMEs), not just large factories. The modular design ensures that the system can b

References

1. A. Bhandary, V. Dobariya, G. Yenduri, R. H. Jhaveri, S. Gochhait and F. Benedetto, "Enhancing Household Energy Consumption Predictions Through Explainable AI Frameworks," in IEEE Access, vol. 12, pp. 36764-36777, 2024, doi: 10.1109/ACCESS.2024.3373552.

2. R. Godahewa, C. Deng, A. Prouzeau and C. Bergmeir, "A Generative Deep Learning Framework Across Time Series to Optimize the Energy Consumption of Air Conditioning Systems," in IEEE Access, vol. 10, pp. 6842-6855, 2022, doi: 10.1109/ACCESS.2022.3142174

3. M. A. Alghamdi, A. S. AL–Malaise AL–Ghamdi and M. Ragab, "Predicting Energy Consumption Using Stacked LSTM Snapshot Ensemble," in Big Data Mining and Analytics, vol. 7, no. 2, pp. 247-270, June 2024, doi: 10.26599/BDMA.2023.9020030

4. A. Alsalemi et al., "Achieving Domestic Energy Efficiency Using Micro-Moments and Intelligent Recommendations," in IEEE Access, vol. 8, pp. 15047-15055, 2020, doi: 10.1109/ACCESS.2020.2966640

5. A. C. Şerban and M. D. Lytras, "Artificial Intelligence for Smart Renewable Energy Sector in Europe—Smart Energy Infrastructures for Next Generation Smart Cities," in IEEE Access, vol. 8, pp. 77364-77377, 2020, doi: 10.1109/ACCESS.2020.2990123

6. S. Munir, M. Ranjan Pradhan, S. Abbas and M. A. Khan, "Energy Consumption Prediction Based on LightGBM Empowered With eXplainable Artificial Intelligence," in IEEE Access, vol. 12, pp. 91263-91271, 2024, doi: 10.1109/ACCESS.2024.3418967

7. Chui, K.T.; Lytras, M.D.; Visvizi, A. Energy Sustainabiliy in Smart Cities: Artificial Intelligence, Smart Monitoring, and Optimization of Energy Consumption. Energies 2018, 11, 2869. https://doi.org/10.3390/en11112869