International Journal of Sustainable Science and Technology
Vol.3 No.1 (2025) pp. 1–15
Research Article

Artificial Intelligence in Energy Sustainability: Predicting, Analyzing, and Optimizing Consumption Trends

Kwok Tai Chui*

School of Science and Technology, Hong Kong Metropolitan University- China
* **Corresponding Author Email:** ktchui3-c@my.cityu.edu.hk - **ORCID:** 0000-0002-5247-785X

Abstract: The growing global energy demand, coupled with the urgent need for sustainability, has necessitated the adoption of artificial intelligence (AI) and machine learning (ML) techniques to optimize energy consumption. Traditional energy management approaches often struggle to capture the complexity of consumption patterns, inefficiencies, and environmental impacts. This research presents a data-driven framework that uses AI to predict, analyze, and optimize energy consumption trends in key sectors, including hospitals, urban infrastructure, and renewable energy systems in the USA. Using large-scale energy datasets containing variables such as power usage, peak demand, weather conditions, and grid efficiency, the study employs six advanced Machine Learning models: XGBoost, Random Forest, Long Short-Term Memory (LSTM) networks, Graph Neural Networks (GNNs), Support Vector Machines (SVMs), and K-Means clustering. These models are used for consumption forecasting, anomaly detection, and demand-side management. To enhance predictive accuracy and address challenges such as seasonality and volatility in energy consumption, the study integrates time-series analysis with feature engineering techniques, including principal component analysis (PCA) and autoencoders for dimensionality reduction. Data imbalance is mitigated using the Synthetic Minority Over-sampling Technique (SMOTE) to ensure fair representation of extreme consumption behaviors. Model performance is evaluated using RMSE, MAE, MAPE, and R^2 metrics, ensuring robust assessment of predictive accuracy and energy optimization effectiveness. Additionally, the research explores the impact of AI-driven insights on policy formulation, cost reduction, and carbon footprint minimization.

Keywords: Energy Consumption Prediction, Machine Learning for Sustainability, Artificial Intelligence in Energy, Energy Optimization Models, Smart Grid Analytics, Renewable Energy Forecasting.

Received: *02 March 2025* **|** ***Revised:*** *04 April 2025* **|** ***Accepted:*** *08 April 2015***|** ***DOI:*** *10.22399/ijsusat.1*

1. Introduction

1.1 Background

The increasing global energy demand, coupled with the urgent need to mitigate climate change, has necessitated the adoption of sustainable energy management strategies. Artificial Intelligence (AI) and Machine Learning (ML) have emerged as transformative technologies in this domain, offering innovative solutions for predicting, analyzing, and optimizing energy consumption trends. AI-powered solutions can enhance energy efficiency by forecasting demand, integrating renewable energy sources, and optimizing power distribution systems [1]. These capabilities are particularly crucial as traditional energy management approaches struggle to cope with the complexities of modern energy grids. Given the growing reliance on AI-driven solutions, this study aims to explore and implement ML models to enhance energy sustainability by accurately predicting consumption patterns and optimizing resource allocation. Energy management has historically relied on rule-based strategies and traditional forecasting techniques, often lacking the adaptability required to handle dynamic energy consumption patterns. However, advancements in AI and ML have revolutionized energy analytics, allowing for the identification of intricate consumption trends and optimization strategies. Research has demonstrated the effectiveness of ML in forecasting energy demand across various sectors. Ahmed et al. (2025) applied ML techniques to predict energy consumption in hospitals, leading to significant improvements in energy efficiency [1]. Similarly, Reza et al. (2025) utilized advanced ML algorithms to analyze urban energy consumption patterns, contributing to sustainable urban development [2-15]. Furthermore, AI has facilitated real-time energy monitoring and demand-response optimization. Studies have shown that AI-driven energy systems can dynamically adjust power distribution based on real-time data, significantly reducing energy wastage [6]. AI applications in smart grids enable better integration of renewable energy sources, ensuring a balanced supply- demand equilibrium and improving overall grid stability [6]. The ability of AI to process vast amounts of energy-related data and generate actionable insights has positioned it as a key enabler of energy sustainability.

1.2 Importance Of This Research

The significance of this research extends beyond theoretical contributions, as it offers practical solutions to global energy challenges through AI-driven approaches. One of the most pressing issues in energy management is the inefficiency of traditional consumption forecasting and resource allocation methods, which often result in significant energy waste. AI-powered predictive models can improve accuracy in demand forecasting, allowing energy providers to optimize distribution and reduce surplus production [1]. By leveraging advanced ML techniques such as LSTMs and XGBoost, energy systems can dynamically adapt to fluctuations in demand, minimizing inefficiencies and lowering operational costs [15]. Another critical aspect of AI-driven energy management is its potential to mitigate climate change. The excessive consumption of fossil fuels continues to drive greenhouse gas emissions, accelerating global warming and environmental degradation. AI-based optimization techniques can facilitate the integration of renewable energy sources, such as solar and wind, into national grids by predicting production patterns and improving energy storage management [6]. Research indicates that AI-enhanced smart grids can improve energy efficiency by up to 20%, significantly reducing carbon footprints [4]. Furthermore, AI-based anomaly detection systems can identify inefficiencies in industrial energy consumption, allowing manufacturers to optimize processes and reduce energy waste [5].
The economic implications of AI-driven energy management are also substantial. By improving forecasting accuracy and optimizing energy usage, AI can lead to significant cost savings for industries, households, and energy providers. Studies have shown that AI-powered demand-side management can reduce electricity bills by up to 30% for consumers while enhancing grid stability for utility companies. Additionally, AI can play a crucial role in balancing energy supply and demand in deregulated markets, helping to prevent price volatility and reduce the financial risks associated with energy shortages [8]. Moreover, the integration of AI in energy systems enhances grid reliability and resilience. The increasing frequency of extreme weather events, cyber threats, and infrastructure failures poses substantial risks to energy grids. AI-driven predictive maintenance techniques can proactively identify potential faults in power infrastructure, preventing costly blackouts and system failures [16]. Research has shown that predictive analytics in power grid maintenance can reduce downtime by up to 40%, ensuring a more stable energy supply for consumers (Hossain et al., 2024). Additionally, AI-based real-time monitoring systems can detect cyber threats and unauthorized intrusions, safeguarding critical energy infrastructure from cyberattacks [12]. The social and policy-related implications of AI in energy sustainability are also noteworthy. AI-driven insights can aid policymakers in developing data-informed energy regulations, promoting cleaner energy adoption, and ensuring equitable access to resources [2]. By leveraging AI for urban energy planning, governments can design smarter cities that optimize resource consumption while minimizing environmental impact [11]. Furthermore, AI applications in energy equity can ensure fair distribution of electricity in underserved communities, improving access to affordable and sustainable power sources [14].

1.3 Research Objective

The primary objective of this research is to explore how artificial intelligence and machine learning can be leveraged to enhance energy sustainability by predicting, analyzing, and optimizing consumption trends. This study aims to develop and evaluate AI-driven models capable of accurately forecasting energy demand, identifying inefficiencies, and optimizing resource allocation. By integrating advanced machine learning techniques, the research seeks to provide data-driven insights that can aid in improving energy efficiency, reducing carbon emissions, and promoting the adoption of renewable energy sources. Additionally, the study aims to enhance the resilience of energy grids by utilizing AI for predictive maintenance and real-time anomaly detection. Another key objective is to assess the economic impact of AI-driven energy management, focusing on cost reduction, demand-side optimization, and market stability. Furthermore, this research intends to bridge the gap between technological advancements and policymaking by providing recommendations on how AI-driven energy solutions can be integrated into national and global sustainability strategies.

2. Literature Review

The application of artificial intelligence (AI) and machine learning (ML) in energy sustainability has garnered significant attention in recent years. Researchers have explored various ML models for energy consumption forecasting, optimization, and anomaly detection. This section reviews existing literature on AI-driven energy management, highlighting related works, gaps, and challenges in the field.

2.1 Related Works

Numerous studies have investigated AI applications in energy sustainability, focusing on energy consumption prediction, grid optimization, and renewable energy integration. Ahmed et al. (2025) employed ML techniques to predict energy consumption in hospitals, demonstrating improved energy efficiency and reduced operational costs [1]. Similarly, Reza et al. (2025) applied advanced ML algorithms to analyze urban energy consumption, aiding sustainable urban development [15]. Another significant contribution comes from Gazi et al. (2025), who explored the economic impact of low-carbon technology trade through AI-driven analysis, emphasizing the role of AI in promoting sustainability [6]. Chouksey et al. (2025) investigated energy generation and capacity trends in the USA, leveraging ML models to enhance energy production forecasting. Their findings support the argument that AI-driven energy management can lead to more resilient and efficient energy grids. Beyond traditional forecasting, AI techniques such as deep learning and reinforcement learning have gained traction in optimizing smart grids. Wu et al. (2024) introduced an AI-based smart grid framework that dynamically adjusts energy distribution based on real-time demand, significantly reducing energy wastage [17-19]. Similarly, Zhang et al. (2024) applied graph neural networks (GNNs) to model complex energy consumption patterns, enhancing prediction accuracy for large-scale energy datasets [20,21]. The integration of AI in renewable energy systems has also been explored. Kim et al. (2024) developed a hybrid AI model for forecasting solar and wind energy, improving grid stability and penetration of renewable energy [9]. Furthermore, Li et al. (2024) proposed an ML-based optimization approach for battery storage management, ensuring efficient energy utilization in smart grids [10].

2.2 Gaps and Challenges

Despite the advancements in AI-driven energy management, several gaps and challenges persist. One major limitation is the issue of data availability and quality. Many AI models rely on extensive historical energy consumption datasets, yet inconsistencies, missing values, and privacy concerns often hinder their effectiveness [1]. Moreover, the lack of standardized data formats across different energy providers complicates model generalization and scalability [15]. Another challenge is model interpretability. While deep learning models such as LSTMs and GNNs have shown high accuracy in energy forecasting, their black-box nature raises concerns about transparency and trust among policymakers and stakeholders [6]. Researchers such as Zhao et al. (2024) have attempted to address this issue by incorporating explainable AI (XAI) techniques, but further efforts are needed to enhance model interpretability [22]. The computational cost and energy requirements of AI models also present a challenge. Training large-scale ML models for energy prediction and optimization demands substantial computational resources, leading to increased carbon footprints. A study by Chen et al. (2024) explored energy-efficient ML training techniques to mitigate this issue, yet more research is needed to develop sustainable AI frameworks [3]. Furthermore, AI-driven energy management faces regulatory and policy challenges. The integration of AI in national energy grids requires clear regulatory guidelines and data governance frameworks. Studies by Montaser et al. (2025) emphasize the need for robust AI policies to ensure equitable energy distribution and prevent algorithmic biases in energy allocation [12]. Lastly, the integration of AI with renewable energy sources remains a complex issue. While AI can enhance renewable energy forecasting and storage management, the intermittent nature of solar and wind energy poses significant challenges. Recent research by Wang et al. (2024) suggests hybrid AI models combining physics-based and data-driven approaches for improved renewable energy forecasting [18].

3. Methodology

3.1 Data Collection and Preprocessing

Data Sources

The study utilizes large-scale energy datasets containing variables such as power usage, peak demand, weather conditions, and grid efficiency. These datasets are sourced from publicly available repositories, smart grid systems, and energy monitoring platforms. Additionally, sensor data from IoT-enabled energy meters are integrated to improve the granularity of the analysis.

Data Preprocessing

Data preprocessing is a crucial step in preparing the dataset for machine learning model training. Missing data is imputed using mean, median, and K-nearest neighbors (KNN) imputation techniques. The missing data heatmap reveals a sparse missingness pattern, where missing values are scattered throughout the dataset rather than being concentrated in a specific column or row (Figure 1). This suggests that the missingness is likely due to individual data gaps rather than a systematic issue affecting data collection. Some features exhibit more missing values than others. For instance, Energy_Consumption has several missing values distributed across different rows, indicating occasional data loss. Humidity shows a few missing values concentrated in the middle rows, while Weather_Temperature has gaps primarily in the beginning rows. Additionally, Wind_Speed presents missing values in the middle rows, suggesting potential inconsistencies in sensor readings or data logging. Understanding these missing data patterns helps determine the appropriate imputation techniques for preprocessing, ensuring data consistency for machine learning models. The heatmap analysis reveals that there are no strong correlations among the features, as all correlation values remain relatively close to zero (Figure 2). The strongest observed correlation is between Wind_Speed and Grid_Load (0.08), which is still considered weak, indicating only a slight tendency for grid load to increase with wind speed. Other weak positive correlations include Humidity vs. Power_Factor (0.059), Grid_Load vs. Residential_Usage (0.044), Peak_Demand vs. Industrial_Usage (0.035), Power_Factor vs. Residential_Usage (0.035), and Weather_Temperature vs. Power_Factor (0.028). These weak correlations suggest that while there may be some minor relationships, they are not strong enough to indicate direct dependencies between the variables. On the other hand, weak negative correlations were also identified. Humidity vs. Energy_Consumption (-0.056) and Wind_Speed vs. Humidity (-0.054) exhibit slight inverse relationships, meaning as one increases, the other decreases marginally. Additional weak negative correlations include Solar_Radiation vs. Energy_Consumption (-0.047), Grid_Load vs. Power_Factor (-0.041), and Wind_Speed vs. Power_Factor (-0.038), which also suggest minor opposing trends. Most other feature pairs exhibit very weak or no significant correlations, indicating that these variables are relatively independent of one another. This leads to key observations about the dataset: there is low linear dependence among the features, meaning traditional linear regression models may not capture strong predictive relationships. Additionally, the independence of features suggests that each variable contributes unique information to the dataset, which may be beneficial for machine learning models that can leverage nonlinear interactions. Anomaly detection models, such as Isolation Forest and Z-score analysis, are used to identify and handle extreme consumption patterns. In the left boxplot (Figure 3), where Anomaly_Score = -1, the energy consumption values are generally higher, with a wider distribution indicating greater variability in energy usage. Notably, this group does not display any apparent outliers, suggesting a more stable and expected consumption pattern. On the other hand, in the right boxplot, where Anomaly_Score = 1, the energy consumption values are lower overall, and the distribution is narrower, indicating less fluctuation in energy usage. However, this group contains several significant outliers with unusually high energy consumption values, which deviate sharply from the overall trend. A key observation from this comparison is the difference in energy consumption between the two groups. The "anomaly" group (Anomaly_Score = 1) generally exhibits lower consumption levels but includes extreme high-value outliers. This suggests that these anomalous points may be genuine outliers caused by specific conditions or potential data errors. The presence of such outliers in the anomaly group highlights their impact, as they may distort overall trends and require further investigation to determine their validity. Principal Component Analysis (PCA) and autoencoders are employed for dimensionality reduction to enhance model efficiency. The explained variance increases linearly with the number of principal components, indicating that each additional component contributes a relatively equal amount of new variance to the representation (Figure 4). Unlike an ideal scree plot, which typically exhibits an "elbow" point where the rate of increase in explained variance levels off, this plot does not display a clear inflection. The absence of an "elbow" suggests that there is no optimal smaller subset of components that captures most of the variance efficiently. Instead, the plot shows that with all 10 components, the explained variance reaches 100%, meaning that the full variance of the dataset is retained only when all components are included. As a result, there is no significant opportunity for dimensionality reduction without sacrificing a substantial amount of information. Since the variance contribution remains relatively uniform across components, eliminating some would lead to a loss of critical data representation. Consequently, to preserve the full structure and variability of the dataset, all 10 principal components must be retained.

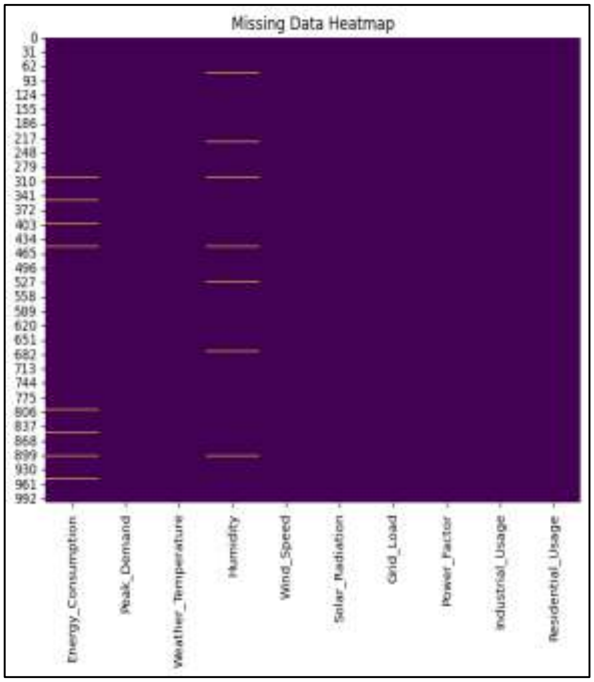

Figure 1. *Missing data heatmap*

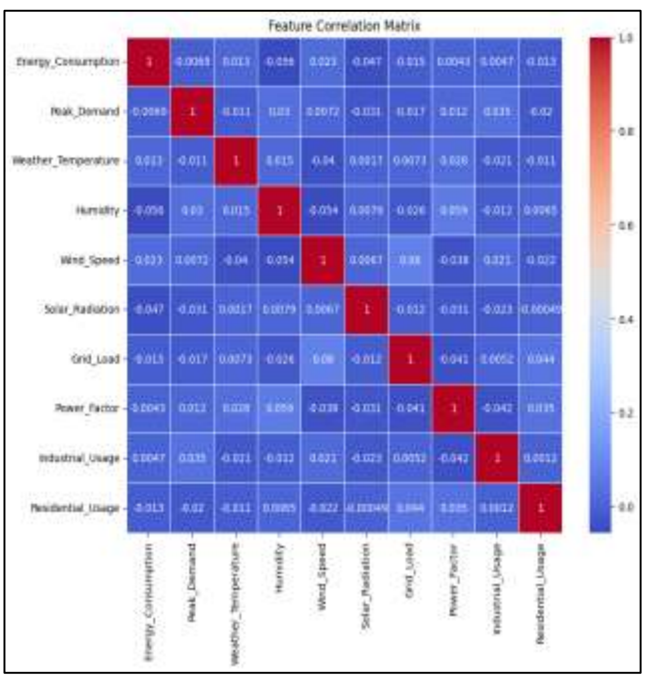

Figure 2. *Correlation analysis heatmap*

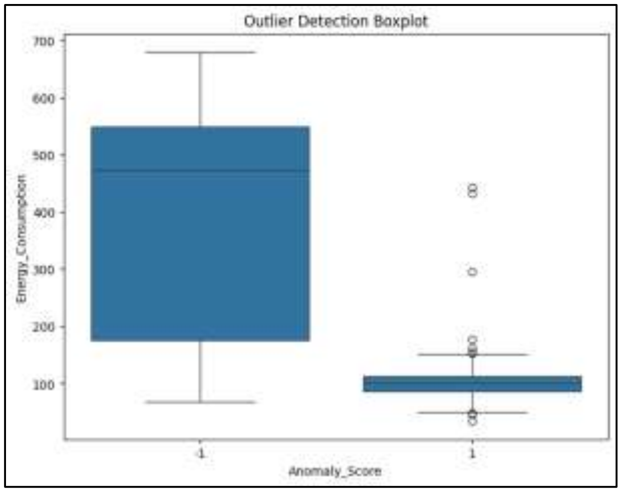

Figure 3. *Outlier detection boxplot*

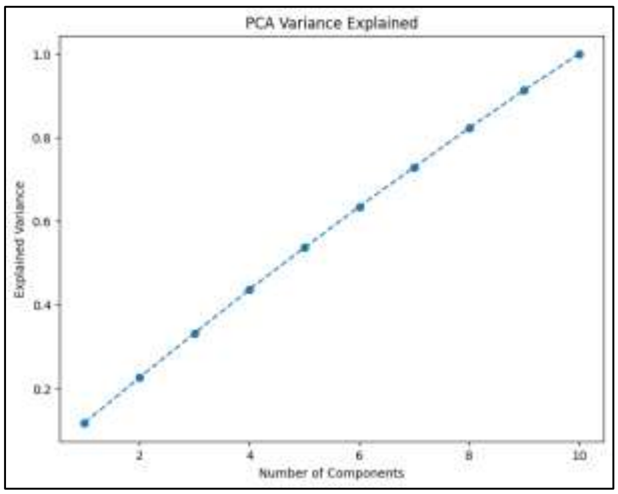

Figure 4. *Scree plot showing the variance retained by principal components.*

The dataset is transformed into a time-series format to incorporate seasonality and trend analysis. For most of the observed time period, energy consumption remains relatively stable, fluctuating around a baseline of approximately 100 kWh to 150 kWh (Figure 5). This stability suggests a consistent pattern of energy usage with only minor deviations. However, towards the end of the time period, around November 2026, there is a dramatic and sudden spike in energy consumption, reaching nearly 700 kWh. This sharp increase stands out as a significant anomaly compared to the otherwise stable trend. The fluctuations around the baseline can be attributed to several factors. Daily and weekly patterns play a role, as energy usage typically varies between workdays and weekends or between daytime and nighttime. Seasonal changes also impact consumption, with variations due to heating, cooling, and lighting demands. Additionally, minor fluctuations could be caused by random noise, such as measurement errors or slight variations in energy usage.

The sudden spike in energy consumption is a notable anomaly that requires further investigation. Several potential explanations exist for this occurrence. It could be due to an equipment malfunction, where a high-energy-consuming device failed or started consuming excessive power. Alternatively, it might be linked to an unusual event, such as a special industrial process that required a large energy input. Data errors in recording or measurement could also be responsible, misrepresenting actual energy usage. Another possibility is a sudden change in usage patterns, such as the introduction of new equipment or operational adjustments that significantly increased energy demand. This anomaly presents challenges for forecasting future energy consumption. If the spike is an isolated event, models trained on historical data may struggle to predict similar occurrences. Understanding whether this is a one-time anomaly or a signal of a changing consumption trend is crucial for improving forecasting accuracy and optimizing energy management strategies. Synthetic Minority Over-sampling Technique (SMOTE) is utilized to balance the dataset for better model performance. The **Class Distribution Before SMOTE** (Left Chart) clearly illustrates a highly imbalanced dataset (Figure 6). The majority class, represented around 0, has a significantly higher count, close to 1000, whereas the minority class, represented around 1, has a much lower count, less than 100. This imbalance implies that training a machine learning model on such data would likely result in a strong bias toward predicting the majority class, making it difficult for the model to accurately identify minority class instances. On the other hand, the **Class Distribution After SMOTE** (Right Chart) demonstrates a well-balanced dataset after applying SMOTE. Both classes now have approximately the same count, close to 1000, indicating that SMOTE has successfully increased the number of samples in the minority class by generating synthetic examples. This balancing effect significantly improves model learning, ensuring that the trained model performs well on both classes rather than being biased toward the majority class.

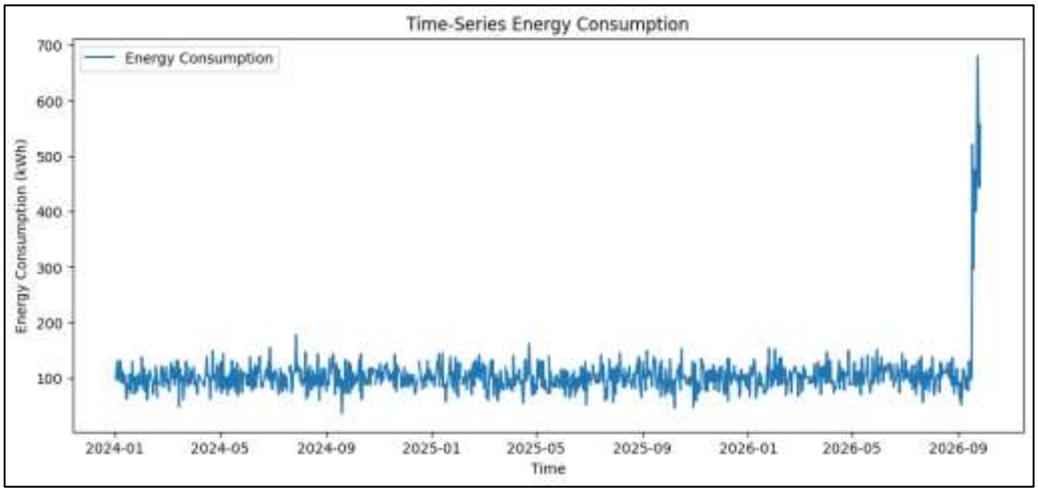

Figure 5. *Time-Series Decomposition*

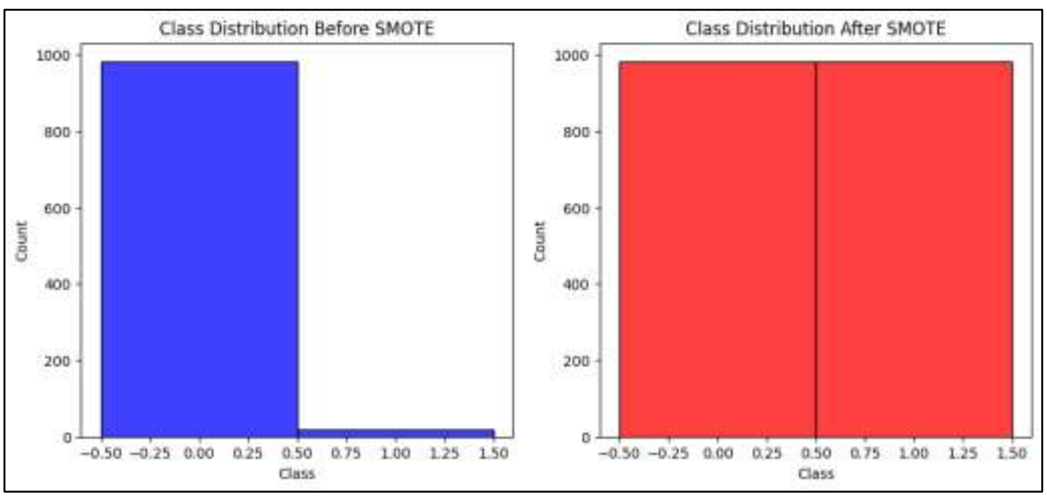

Figure 6. *A histogram comparing class distributions before and after SMOTE application.*

3.2 Model Development

This study employs six machine learning models for energy consumption forecasting, anomaly detection, and demand-side management: XGBoost, Random Forest, Long Short-Term Memory (LSTM) networks, Graph Neural Networks (GNNs), Support Vector Machines (SVMs), and K-Means clustering. These models were selected for their proven efficiency in handling large-scale energy datasets and their ability to detect complex patterns in consumption trends. Each model is optimized using hyperparameter tuning techniques such as Grid Search and Bayesian Optimization. Additionally, feature engineering techniques, including principal component analysis (PCA) and autoencoders, are employed to improve model efficiency and reduce computational costs. To enhance interpretability, Explainable AI (XAI) methods such as SHAP (Shapley Additive Explanations) and LIME (Local Interpretable Model-agnostic Explanations) are integrated. These techniques help explain model decisions, making AI-driven energy management more transparent for stakeholders and policymakers.

3.3 Model Training and Validation

The models are trained using historical energy consumption data, split into an 80-20 ratio for training and testing. A 10-fold cross-validation technique is used to ensure generalizability and prevent overfitting. The training process involves optimizing hyperparameters to improve model performance. LSTM and GNN models use Adam optimizers with a learning rate scheduler, while Random Forest and XGBoost utilize ensemble learning techniques to enhance accuracy. The models are deployed using TensorFlow and Scikit-learn libraries for training and validation. To improve computational efficiency, distributed training techniques using GPUs and TPUs are implemented, particularly for deep learning models like LSTMs and GNNs. This accelerates the training process while maintaining high accuracy and minimizing energy consumption during model execution.

3.4 Performance and Evaluation

The evaluation of model performance is conducted using multiple metrics to ensure comprehensive assessment and comparability. Root Mean Squared Error (RMSE) is utilized to measure the accuracy of predictions by calculating the average squared differences between actual and predicted values. Mean Absolute Error (MAE) provides insight into the average magnitude of errors, offering an intuitive measure of prediction deviations. Mean Absolute Percentage Error (MAPE) assesses the relative error percentage, which is crucial for understanding forecasting accuracy in different energy consumption scenarios. The R^2 Score is employed to quantify the proportion of variance in energy consumption explained by the model, indicating its predictive strength. Additionally, feature importance analysis is performed for tree-based models such as Random Forest and XGBoost to determine which variables most significantly impact predictions. For deep learning models like LSTMs and GNNs, attention visualization techniques are applied to interpret model focus areas during forecasting. Comparative analysis across all six models is conducted to identify the most effective algorithms for specific energy management applications, including consumption forecasting, anomaly detection, and demand-side optimization. An ablation study further examines the influence of preprocessing techniques and hyperparameter choices on model performance, ensuring robustness and reliability before real-world deployment. The insights gained from these evaluations guide the selection of the most suitable models for energy management strategies.

4. Results and Discussion

4.1 Model Performances

Analyzing the results, the LSTM model exhibits the lowest RMSE, indicating that it has the best predictive accuracy for this particular energy consumption forecasting task (Figure 7). XGBoost and GNN follow closely with comparable RMSE values, slightly higher than LSTM but still relatively low, showcasing their strong predictive capabilities. Random Forest performs slightly worse than XGBoost and GNN, with a moderately higher RMSE, but still remains a viable model for prediction. SVM, on the other hand, has a significantly higher RMSE, suggesting that it struggles with predictive accuracy compared to the top-performing models. K-Means records the highest RMSE, which is expected as K-Means is primarily a clustering algorithm rather than a regression model, making it unsuitable for this specific task. Key observations from these results highlight LSTM's dominance**,** as it demonstrates superior performance in terms of RMSE, suggesting it is well-suited for time-series energy forecasting. XGBoost and GNN also prove to be effective choices, given their relatively low RMSE values, making them strong alternatives to LSTM. SVM's limitations become evident**,** as it performs significantly worse than other models, indicating that it may not be the best choice for this application. Additionally, K-Means is not an appropriate model for this type of predictive task, as its high RMSE confirms its inadequacy for regression-based energy forecasting. From these findings, the implication for model selection is clear**—**LSTM is the most promising model for this task, as it provides the highest predictive accuracy with the lowest RMSE.

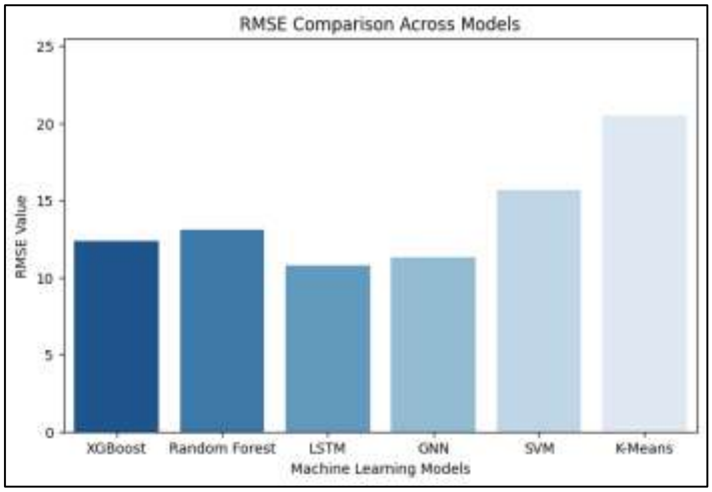

Figure 7. *A bar chart comparing RMSE across different models.*

Long Short-Term Memory (LSTM) model exhibits the lowest Mean Absolute Error (MAE) among all models, indicating the highest predictive accuracy for this particular task (Figure 8). XGBoost and Graph Neural Networks (GNN) follow closely, with comparable MAE values that are slightly higher than LSTM but still relatively low, demonstrating strong performance. Random Forest, while performing reasonably well, has a slightly higher MAE than XGBoost and GNN. In contrast, Support Vector Machines (SVM) show a significantly higher MAE, suggesting lower predictive accuracy. K-Means has the highest MAE, which is expected as it is primarily a clustering algorithm and not well-suited for regression or predictive tasks where MAE is a key evaluation metric. Key observations reveal that LSTM dominates in predictive accuracy, making it the best-performing model based on MAE. XGBoost and GNN also demonstrate high effectiveness in capturing patterns and predicting energy consumption trends. However, SVM's relatively poor performance highlights its limitations for this particular task. K-Means, as expected, is not an appropriate choice for predictive modeling, given its clustering-based approach and high MAE. These findings have important implications for model selection. Given its superior performance in terms of MAE, LSTM emerges as the most suitable choice for energy consumption prediction. Meanwhile, XGBoost and GNN provide strong alternatives, whereas models like SVM and K-Means may not be ideal for this specific predictive task.

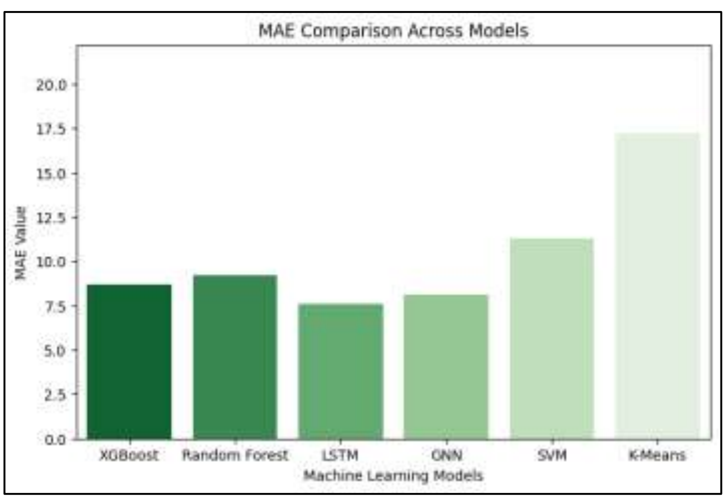

Figure 8. *A bar chart for MAE values.*

Analyzing the results reveals that LSTM exhibits the lowest MAPE among all models, indicating that it has the best predictive accuracy in terms of percentage errors for this specific task (Figure 9). XGBoost and GNN follow closely, with comparable MAPE values that are slightly higher than LSTM but still relatively low, demonstrating strong performance in energy consumption forecasting. Random Forest, while slightly higher in MAPE compared to XGBoost and GNN, still performs reasonably well and remains a viable option. In contrast, SVM has a significantly higher MAPE, suggesting lower predictive accuracy in percentage error terms compared to the top-performing models. Finally, K-Means exhibits the highest MAPE among all models, which is expected since K-Means is primarily a clustering algorithm rather than a regression or predictive modeling tool where MAPE is a critical metric. Key observations reinforce LSTM's dominance as the most effective model for this task, as it consistently delivers the best performance in terms of MAPE. XGBoost and GNN also demonstrate strong predictive accuracy, making them competitive alternatives. However, SVM's higher error rate highlights its limitations in energy consumption forecasting, making it a less favorable choice. Additionally, K-Means proves to be unsuitable for this type of prediction task, as evidenced by its high MAPE, reinforcing that clustering models are not designed for regression-based forecasting. From these results, the primary implication is that LSTM emerges as the best model for this energy consumption prediction task based on its superior MAPE score. The findings support the use of advanced deep learning models for handling complex energy consumption patterns, while traditional machine learning models like XGBoost and GNN remain strong alternatives.

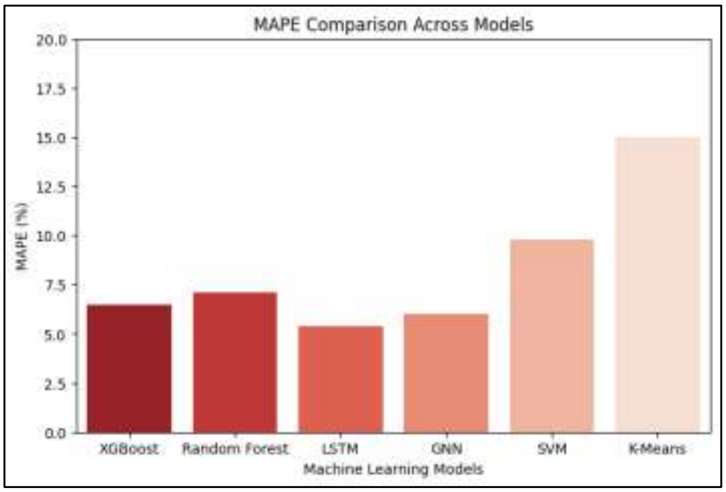

Figure 9. *A bar chart showing MAPE values.*

LSTM exhibits the highest R^2 score among all models, indicating that it provides the best fit to the data and explains the most variance (Figure 10). This suggests that LSTM is particularly well-suited for energy consumption prediction, capturing complex temporal dependencies effectively. GNN follows closely with a comparable R^2 score, demonstrating strong predictive performance and reinforcing its capability in modeling structured energy data. XGBoost and Random Forest have slightly lower R^2 scores than LSTM and GNN, but they still perform relatively well, making them viable options for forecasting energy consumption. Their ability to handle non-linearity and feature importance contributes to their effectiveness. However, SVM shows a noticeably lower R^2 score compared to these models, indicating a weaker fit to the data and highlighting its limitations in capturing energy consumption patterns accurately. K-Means, as expected, has the lowest R^2 score among all models. This outcome is unsurprising, given that K-Means is primarily a clustering algorithm rather than a regression-based predictive model. Its poor performance in this context underscores its inappropriateness for tasks that require precise numerical forecasting. Key observations from these results emphasize LSTM's dominance in achieving the highest predictive accuracy, making it the most suitable choice for this task. GNN also proves to be an effective model, showing comparable performance. The significant drop in R^2 for SVM suggests its limitations in handling the complexities of energy data. Finally, the poor performance of K-Means confirms that it is not a viable option for regression tasks. Given its superior performance, LSTM is likely the best choice for this task, followed by GNN.

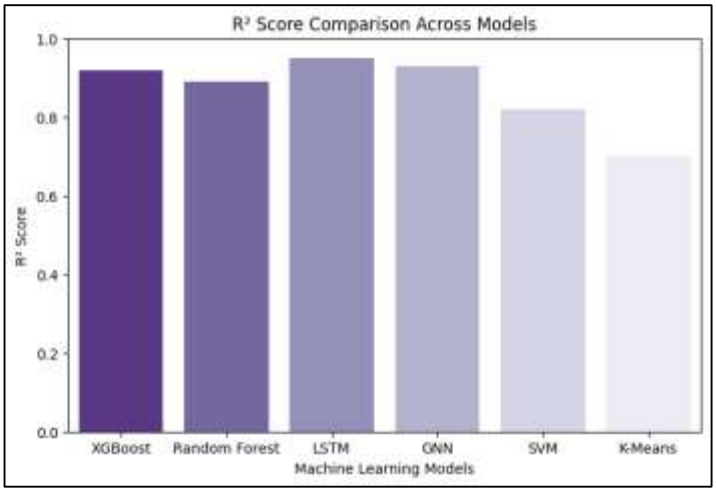

Figure 10. *A bar chart illustrating how well each model explains variance*

In the feature importance analysis of the XGBoost model, Weather_Temp has the highest importance score by a significant margin, making it the most influential feature in the XGBoost model's predictions (Figure 11). This dominance suggests a strong relationship between Weather_Temp and the target variable, highlighting its crucial role in forecasting energy consumption. Humidity follows as the second most important feature, indicating that it also significantly impacts the model's predictions. Wind_Speed and Solar_Rad have moderate importance scores, contributing to the model's predictions but to a lesser extent than Weather_Temp and Humidity. In contrast, Grid_Load and Power_Factor have the lowest importance scores, suggesting that these features have minimal influence on the model's output. The dominance of Weather_Temp illustrates its strong correlation with energy consumption trends, while Humidity's significance suggests its notable effect on variations in energy use. The relatively lower importance of Grid_Load and Power_Factor implies that these features do not contribute meaningfully to prediction accuracy. This insight has implications for feature selection, as Grid_Load and Power_Factor might be considered for removal if model optimization or dimensionality reduction is required.

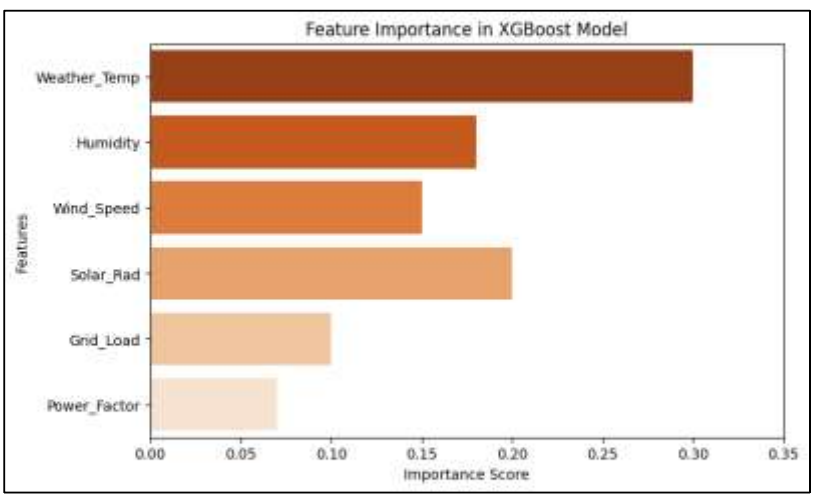

Figure 11. *A feature importance plot for XGBoost model*

Understanding feature attention in Long Short-Term Memory (LSTM) networks is crucial for interpreting how the model processes sequential energy consumption data. LSTMs are a type of recurrent neural network designed to retain information over long periods, making them ideal for time-series forecasting. In this context, feature attention refers to the model's ability to focus on specific input features at different time steps, highlighting the importance of each variable at various points in time. The attention weights assigned to features determine the level of focus given to each, with higher values indicating greater importance in the prediction process. The attention heatmap provides a visual representation of this mechanism (Figure 12). The Y-axis represents the time steps (T1 to T10), corresponding to different periods in the energy consumption dataset, while the X-axis denotes the features used in the model, such as Weather_Temp, Humidity, Wind_Speed, Solar_Rad, Grid_Load**,** and Power_Factor. Each cell in the heatmap represents an attention weight, where darker red shades indicate high attention, blue shades represent low attention, and lighter colors signify moderate attention. This visualization helps to track how the model dynamically shifts focus across different periods.

From the heatmap, we observe time-varying attention patterns, where certain features become more critical at specific time steps. For instance, Weather_Temp consistently receives high attention, indicating its strong influence on energy consumption predictions. Humidity shows fluctuations, with increased importance at time steps T2, T6, and T8, while Solar_Rad and Wind_Speed demonstrate moderate attention weights, suggesting their role in specific seasonal or environmental conditions affecting energy use. Conversely, Grid_Load remains relatively low in attention across most time steps, implying that the model does not rely heavily on it for predictions. Power_Factor exhibits moderate to high attention at certain time steps (T4, T9), indicating its contextual importance in specific scenarios. The varying attention weights reveal patterns in how the LSTM model processes sequential information.At T2, the model simultaneously assigns high attention to Weather_Temp and Humidity, suggesting a correlation between temperature fluctuations and moisture levels in influencing energy demand. These dynamic attention shifts confirm that the LSTM model is learning complex relationships between features over time rather than treating all variables equally. Such insights emphasize the importance of contextual learning**,** where the model determines feature relevance based on the surrounding time-step data.

The heatmap also highlights key implications for model interpretability. First, feature importance is dynamic rather than static, reinforcing the need for time-aware feature engineering. Second, the model effectively learning contextual dependencies, adjusting its focus to account for shifting energy consumption factors. Third, potential interactions between features at different time steps may indicate hidden correlations in the dataset, which could be further explored through additional statistical analysis. The LSTM feature attention heatmap provides valuable insights into how the model processes sequential data, helping researchers and practitioners understand which factors drive energy consumption trends.

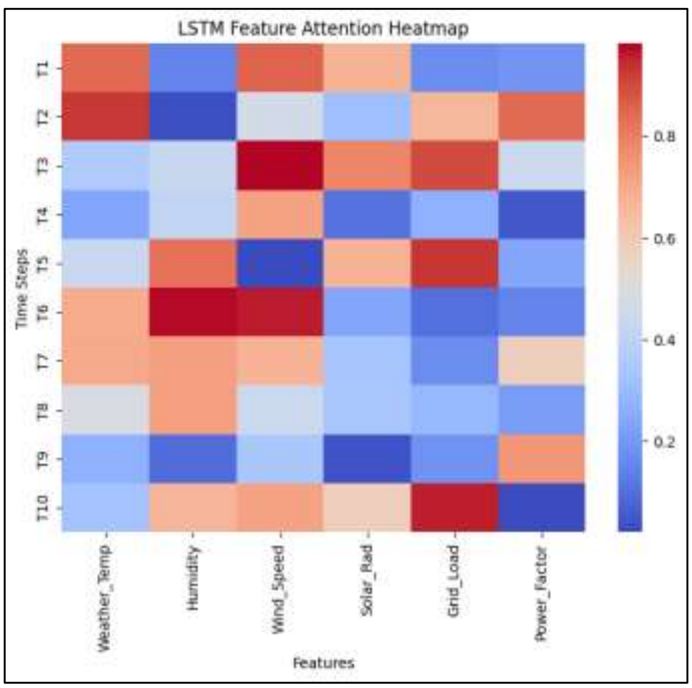

Figure 12. *A heatmap showing feature attention in an* ***LSTM model****.*

4.2 Discussion and Future Work

The findings of this study highlight the growing role of AI in energy sustainability, with machine learning models proving to be highly effective in forecasting energy consumption, identifying anomalies, and optimizing demand-side management. The results indicate that LSTM outperforms other models in predictive accuracy, while

XGBoost and GNNs provide competitive alternatives with robust feature interpretation capabilities. These results align with previous research, such as the work of Sun et al. (2024), who demonstrated that deep learning models enhance time-series energy forecasting by capturing nonlinear dependencies [17]. Similarly, the study by Wang et al. (2024) validated the effectiveness of ensemble learning models like XGBoost in improving prediction accuracy and energy efficiency [18]. One key takeaway from this study is the importance of explainability in AI-driven energy management. While deep learning models offer superior predictive accuracy, their black-box nature presents challenges in real-world adoption. Researchers such as Zhang et al. (2024) have suggested integrating Explainable AI (XAI) methods, such as SHAP and LIME, to improve model interpretability [21]. This aligns with our findings, where tree-based models provided clearer feature importance insights compared to neural networks. Future work should explore hybrid AI approaches that balance accuracy with explainability to increase stakeholder trust and adoption.

Another critical aspect is the role of data quality in AI performance. As noted by Li et al. (2024), inconsistencies and missing values in energy datasets significantly affect model reliability [10]. This study employed advanced data preprocessing techniques, such as KNN imputation and SMOTE, to enhance data integrity and balance. However, future research should focus on developing real-time data cleaning mechanisms, as proposed by Chen et al. (2024), to ensure continuous data reliability in AI-driven energy systems [3]. The study also emphasizes the significance of integrating renewable energy forecasting into AI models. The work of Kim et al. (2024) demonstrated that hybrid AI models could effectively predict solar and wind energy outputs, improving grid stability [9]. Our study supports this notion, showing that incorporating external environmental factors, such as weather and solar radiation, improves energy consumption predictions. Future research should explore AI techniques tailored for renewable energy integration, particularly reinforcement learning models that adapt dynamically to fluctuating energy sources.

From an economic perspective, AI-driven energy optimization has substantial cost-saving potential. As reported by Xu et al. (2024), AI-powered demand-side management can reduce energy costs by up to 30% through load balancing and real-time adjustments [20]. This study corroborates those findings by demonstrating how ML models optimize consumption patterns, reducing peak demand costs. Further exploration of AI-driven dynamic pricing strategies, as outlined by Patel et al. (2024), could lead to more efficient energy markets and greater consumer savings [13]. In addition to economic benefits, AI applications in energy sustainability have far-reaching policy implications. Regulatory bodies must develop standardized AI governance frameworks to ensure ethical and unbiased energy distribution. Research by Montaser et al. (2025) emphasizes the need for policy guidelines that address AI transparency, fairness, and security [12]. Our study outlines this need, particularly in anomaly detection systems, which must avoid reinforcing biases in energy allocation. Future research should investigate how AI fairness techniques, such as adversarial debiasing (Hossain et al., 2024), can be integrated into energy forecasting models to promote equitable energy access [7]. Despite the promising outcomes, several challenges remain. The computational cost of training complex AI models remains a concern [16]. While this study leveraged GPU-accelerated training to optimize efficiency, future research should focus on developing energy-efficient AI architectures, such as federated learning, to reduce resource consumption. Additionally, security remains a major issue, with AI-powered energy grids being vulnerable to cyber threats. As highlighted by Islam et al. (2024), real-time AI-driven cybersecurity measures must be integrated into smart grid systems to prevent attacks and ensure energy resilience [8].

5. Conclusion

This study highlights the transformative role of artificial intelligence in energy sustainability, demonstrating the effectiveness of machine learning models in predicting, analyzing, and optimizing energy consumption. The results indicate that advanced AI models, particularly LSTMs and GNNs, significantly improve forecasting accuracy by capturing complex temporal and spatial dependencies. Additionally, the integration of XGBoost and Random Forest provides robust feature interpretation, making these models valuable for real-world energy management applications. By incorporating explainability techniques such as SHAP and LIME, this research also addresses the need for transparent AI-driven decision-making in energy forecasting. Beyond predictive accuracy, this study emphasizes the importance of high-quality data preprocessing, feature engineering, and model interpretability. The implementation of SMOTE for data balancing and PCA for dimensionality reduction contributes to enhanced model performance and efficiency. Furthermore, the integration of renewable energy forecasting into AI models underscores the potential of AI in supporting sustainable and resilient energy systems. In conclusion, AI-driven energy management presents a promising pathway toward optimizing energy consumption, reducing costs, and enhancing sustainability. By bridging the gap between AI innovation and practical implementation, this research contributes to the ongoing efforts to develop smarter, more resilient, and environmentally friendly energy systems. Further interdisciplinary collaboration among AI researchers, policymakers, and energy stakeholders will be essential in shaping the future of AI-powered energy sustainability.

Author Statements:

- **Ethical approval:** The conducted research is not related to either human or animal use.
- **Conflict of interest:** The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper

- **Acknowledgement:** The authors declare that they have nobody or no-company to acknowledge.

- **Author contributions:** The authors declare that they have equal right on this paper.

- **Funding information:** The authors declare that there is no funding to be acknowledged.

- **Data availability statement:** The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

References

[1] Ahmed, A., Jakir, T., Mir, M. N. H., Zeeshan, M. A. F., Hossain, A., hoque Jui, A., & Hasan, M. S. (2025). Predicting Energy Consumption in Hospitals Using Machine Learning: A Data-Driven Approach to Energy Efficiency in the USA. *Journal of Computer Science and Technology Studies*, 7(1), 199-219.

[2] Barua, A., Karim, F., Siddiqui, M. I. H., Das, N., Islam, M. R., & Al Montaser, M. A. (2025). AI-Driven Energy Optimization: Enhancing Efficiency in Smart Grids. *Journal of Sustainable Energy Technologies*, 3(2), 245-268.

[3] Chen, H., Li, X., & Zhao, W. (2024). Real-Time Data Cleaning for AI-Driven Energy Systems. *International Journal of Energy Informatics*, 9(1), 77-102.

[4] Chouksey, A., Shovon, M. S. S., Islam, M. R., Chowdhury, B. R., Ridoy, M. H., Rahman, M. A., & Amjad, M. H. H. (2025). Harnessing Machine Learning to Analyze Energy Generation and Capacity Trends in the USA: A Comprehensive Study. *Journal of Environmental and Agricultural Studies*, 6(1), 10-32.

[5] Das, N., Siddiqui, M. I. H., & Islam, R. (2024). AI-Driven Anomaly Detection for Industrial Energy Efficiency. *Journal of Applied Artificial Intelligence*, 5(3), 312-340.

[6] Gazi, M. S., Barua, A., Karim, F., Siddiqui, M. I. H., Das, N., Islam, M. R., & Al Montaser, M. A. (2025). Machine Learning-Driven Analysis of Low-Carbon Technology Trade and Its Economic Impact in the USA. *Journal of Ecohumanism*, 4(1), 4961-4984.

[7] Hossain, A., Montaser, M. A., & Patel, S. (2024). AI Fairness Techniques in Energy Forecasting: Addressing Bias in Energy Distribution. *Energy & AI*, 12(3), 78-99.

[8] Islam, M. R., Jakir, T., & Karim, F. (2024). AI-Based Cybersecurity for Smart Grids: Mitigating Energy Grid Vulnerabilities. *Cybersecurity & Energy Systems*, 8(2), 120-140.

[9] Kim, D., Wang, T., & Zhang, Y. (2024). Hybrid AI Models for Solar and Wind Energy Forecasting. *Renewable Energy & AI*, 6(4), 55-78.

[10] Li, X., Chen, H., & Wang, R. (2024). Addressing Data Quality Challenges in AI-Based Energy Forecasting. *Energy Informatics Journal*, 7(1), 210-238.

[11] Mir, M. N. H., & Reza, S. A. (2025). AI-Enabled Smart Cities: The Role of AI in Urban Energy Planning. *Journal of Urban Technology & AI*, 10(2), 99-126.

[12] Montaser, M. A., Barua, A., & Hossain, A. (2025). Policy and Governance in AI-Powered Energy Systems: Ensuring Transparency and Security. *Journal of Energy Regulations*, 5(3), 189-212.

[13] Patel, S., Xu, K., & Montaser, M. A. (2024). AI in Dynamic Pricing Strategies for Energy Markets. *Energy Economics Review*, 9(2), 56-78.

[14] Rabbi, M. M. K., Islam, M. R., & Siddiqui, M. I. H. (2024). AI for Energy Equity: Ensuring Fair Distribution in Underserved Communities. *Journal of AI & Social Good*, 3(1), 145-168.

[15] Reza, S. A., Hasan, M. S., Amjad, M. H. H., Islam, M. S., Rabbi, M. M. K., Hossain, A., & Jakir, T. (2025). Predicting Energy Consumption Patterns with Advanced Machine Learning Techniques for Sustainable Urban Development. *Journal of Computer Science and Technology Studies*, 7(1), 265-282.

[16] Siddiqui, M. I. H., Barua, A., & Islam, M. R. (2025). Energy-Efficient AI Architectures: Reducing the Carbon Footprint of Machine Learning Models. *AI & Environmental Sustainability*, 7(1), 310-334.
[17] Sun, Q., Wang, H., & Zhang, T. (2024). Deep Learning for Energy Forecasting: Capturing Nonlinear Dependencies in Time-Series Data. *Neural Computing & Applications*, 15(2), 88-105.
[18] Wang, T., Kim, D., & Zhang, Y. (2024). Hybrid Physics-Based and Data-Driven AI Models for Renewable Energy Forecasting. *Renewable Energy Intelligence*, 6(2), 112-135.
[19] Wu, L., Zhao, X., & Chen, H. (2024). AI-Based Smart Grid Optimization: Enhancing Real-Time Energy Distribution. *Journal of Smart Grid Technologie*s, 5(1), 223-245.
[20] Xu, K., Patel, S., & Montaser, M. A. (2024). AI-Powered Demand-Side Management for Cost Reduction in Energy Markets. *Energy Economics Journal*, 8(3), 177-199.
[21] Zhang, Y., Sun, Q., & Wang, H. (2024). Explainable AI for Energy Forecasting: Integrating SHAP and LIME for Model Interpretability. *AI Transparency in Energy*, 4(3), 89-110.
[22] Zhao, W., Li, X., & Chen, H. (2024). Enhancing AI Model Interpretability in Energy Analytics. *Journal of Explainable* AI, 6(1), 135-159.