**Educational Administration: Theory and Practice**
2024, 30(5), 10018-10029
ISSN: 2148-2403
https://kuey.net/

**Research Article**

# Harnessing Multidimensional Insights and Advanced Machine Learning for Optimized Energy Efficiency: Revolutionizing Sustainable Systems through Predictive Optimization, Ensemble Learning and IoT Integration for Enhanced Heating and Cooling Load Management

Nishant Anand[1]*, Pritee Parwekar[2]*, Vikram Bali[3]

[1]Research Scholar, Department of CSE, SRM Institute of Science and Technology, Delhi-NCR Campus, Ghaziabad, India
na5835@srmist.edu.in
[2]Professor, Department of CSE, SRM Institute of Science and Technology, NCR Campus, Ghaziabad, India
priteep@srmist.edu.in,
[3]Professor, Department of CSE, IMS Engineering College, Ghaziabad, India
vikramgcet@gmail.com,

**Citation:** Nishant Anand[1](2024 Harnessing Multidimensional Insights And Advanced Machine Learning For Optimized Energy Efficiency: Revolutionizing Sustainable Systems Through Predictive Optimization, Ensemble Learning, And IoT Integration For Enhanced Heating And Cooling Load Management *Educational Administration: Theory and Practice, 30(5), 10018-10029*
*Doi: 10.53555/kuey.v30i5.4698*

**ARTICLE INFO**

**ABSTRACT**

Integrating Multidimensional Insights for Enhanced Feature Selection in Energy Transition Models presents a comprehensive approach to enhancing the energy efficiency of sustainable energy systems. The purpose of this research is to find the categorical features that can be boosted with ensemble learning for finding most relevant aspect in energy generation. The study leverages sophisticated machine learning techniques, including deep learning and ensemble methods, to improve the prediction and optimization of heating and cooling loads in systems using application of Advanced Machine Learning Algorithms. In this research article, we are trying to focus on critical energy consumption areas like heating and cooling. These are crucial aspects of building energy consumption, and the study's emphasis on these areas demonstrates an understanding of key factors in energy efficiency. This research represents a significant step forward in applying machine learning to sustainable design and energy savings. It underscores the potential of machine learning in transforming the way systems are designed and operated for better energy efficiency. Understanding the application of machine learning algorithms to cross-domain optimization, such as integrating building energy systems with electric vehicles and smart grid technologies, can create synergies that enhance overall energy efficiency. This holistic approach can lead to more significant energy savings by optimizing across multiple domains simultaneously. We also focus on improving the scalability and generalization capabilities of machine learning models to ensure they can be effectively applied across different types of buildings and geographic locations. It involves developing models that can adapt to diverse conditions without retraining. It enhances collaboration with IoT Devices and strengthening the collaboration between machine learning systems and IoT (Internet of Things) devices can enhance the granularity and precision of energy management. IoT devices can provide detailed, real-time data, which, when analyzed by advanced machine learning algorithms, can lead to more nuanced and effective energy-saving. The model is performing reasonably well, with the ability to predict values that correlate with the actual data. Feature Y1 is by far the most predictive of the model's output, which could mean that focusing on this feature could lead to improvements in the model's performance. The accuracy of our model is near 97% with further scope to improve with ensemble learning and XG boosting.

**INDEX TERMS:** Energy efficiency, Machine Learning, Sustainable Systems, Machine Learning algorithms, heating cooling loads, Sustainable design, Energy Management, Ensemble Learning, Prediction Systems, Deep Learning

# I. INTRODUCTION

In the dynamic world of energy research and transition, there is a growing emphasis on advancing technologies to address the challenges posed by escalating energy demands and the integration of renewable sources. One area is the development of high-energy rechargeable batteries, crucial for applications such as electric vehicles, with a particular focus on lithium-ion batteries. The recognition of the limitations in current lithium-ion batteries has spurred research into alternative materials, prominently lithium metal, as a promising anode material. These developments underscore the necessity for batteries with significantly higher specific energy, achievable through the use of advanced cathode materials. Machine Learning (ML) plays a pivotal role in enhancing the impact of such research by offering innovative solutions. In predictive modelling, ML algorithms contribute to the identification of high-capacity cathode materials and optimal anode compositions. Additionally, ML's ability to establish structure-property relationships aids in understanding and improving critical interactions within battery systems. Beyond material considerations, ML applications extend to optimizing battery components, predicting life cycles, and mitigating failure modes, ultimately enhancing the overall performance and longevity of high-specific-energy cells. Machine Learning for Renewable Energy Forecasting by Gaamouche et al. [25] and advances in Lithium-Ion Battery Technologies" by Grey et al. [26] presents a potential understanding of predictive modelling potential in Integrating multi-dimensional datasets, it also increases the efficiency of data transformation by processing the data through many neural networks and increases the accuracy in the process.

Parallel to advancements in energy storage, the broader energy transition towards renewable sources introduces complexities in the electricity sector. This transition, initially focused on establishing the viability of renewables, is now met with new challenges such as complex interactions between multiple technologies, disruptions to traditional business models, and economic and political considerations for key industry actors. Understanding and navigating these challenges require a holistic approach that combines technological innovation with effective policy and economic strategies.

ML applications contribute significantly to forecasting renewable output, optimizing grid operations, and conducting market trend analyses. These applications facilitate the seamless integration of renewables into existing grids, ensuring stability and reliability. Moreover, ML tools provide insights into the market trends influenced by renewable energy, assisting policymakers and industry stakeholders in adapting to changes brought about by these transitions. The simulation capabilities of ML models further empower policymakers to anticipate and understand the potential impacts of various policy decisions on renewable energy adoption and grid integration, thereby fostering more effective and informed decision-making.

The synergy between technological advancements and the strategic application of machine learning positions us on the brink of transformative changes in the energy landscape. The combination of cutting-edge research and sophisticated data-driven approaches holds the promise of not only addressing current challenges but also shaping a sustainable and efficient energy future.

## A. Next Phase of Energy Transitions

"The Next Phase of the Energy Transition and its Implications for Research and Policy" from Nature Energy by Markard, J. [27] discusses the evolving landscape of the electricity sector, focusing on the growing share of renewable energy technologies. Initially, the research and policy emphasis were on establishing renewables as viable options. Now, with renewables rapidly integrating into many electricity grids, new challenges and phenomena are emerging. Next, we will discuss the strategies and challenges in moving away from fossil fuels towards renewable energy sources, emphasizing the role of policy, technology, and market factors and the exnovation processes for renewable energy transition and sustainable energy paths, respectively by Maine et al. [6]. These include complex interactions of multiple technologies, decline of traditional business models, economic and political challenges for key actors like utility companies, and integration issues within the electricity sector. The paper compares the two phases of the energy transition and discusses implications for future research and policymaking.

The paper titled "Exergoeconomic and exergoenvironmental analysis and optimization of an integrated double-flash-binary geothermal system and dual-pressure ORC using zeotropic mixtures; multi-objective optimization" by Chet et al. [25] focuses on the optimization of an integrated geothermal system that combines double-flash and binary processes with a dual-pressure Organic Rankine Cycle (ORC). The study uses zeotropic mixtures for the working fluid in the ORC, aiming to improve the system's thermodynamic efficiency and environmental performance.

The research employs a multi-objective optimization approach to balance various performance metrics, including net present value (NPV), exergy efficiency, exergoeconomic factors (which combine thermodynamic and economic analyses), and exergoenvironmental impacts. The optimization process generates a Pareto frontier, illustrating the trade-offs between these objectives and identifying the set of optimal solutions where no single objective can be improved without worsening at least one other.

This study is significant for its comprehensive approach to optimizing geothermal energy systems, considering not just economic and thermodynamic efficiency, but also environmental impacts. It contributes to the field by providing insights into how zeotropic mixtures and advanced system integration can enhance the sustainability and viability of geothermal energy production by Jeong et al. [23].

### B. Machine Learning Promoting Sustainable Energy

Article "Machine Learning for a Sustainable Energy Future" published in Nature Reviews Materials by Ferraz-Caetano et al.[22] emphasizes the crucial role of machine learning (ML) in promoting sustainable energy. The necessity for transitioning to sustainable energy systems and policy in the next phase of the energy transition proposed by Geels [1]. It illustrates how ML methods can significantly improve the discovery and development of materials and systems that are energy efficient. The paper also demonstrates the efficacy of ML in optimizing energy systems and outlines its successful implementation across various energy-focused sectors. In discussing the various ML techniques used in sustainable energy research, the article includes a mention of the sklearn. Preprocessing module. Machine learning and artificial intelligence are becoming pivotal in optimizing energy systems, forecasting renewable energy production, and discovering new materials for energy applications. This module is vital for data normalization, involving the removal of mean values and scaling features to a uniform variance. Such normalization is essential in ML as it ensures that all features equally contribute to the predictive model. This is particularly critical for algorithms that are sensitive to the scale of input data. In this paper we improved on energy materials presenting additional modelling challenges. Machine learning (ML) could help in the representation of structurally complex structures, which can include disordering, dislocations and amorphous phases. Flexible models that scale efficiently with varied dataset sizes are in demand, and ML could help to develop robust predictive models with further scaling and deployment with IOT.

## II. BENCHMARK ANALYSIS OF EXISTING METHODOLOGY

**Enhanced Prediction and Optimization in Renewable Energy Systems:** ML algorithms that can more accurately predict renewable energy outputs (such as solar and wind energy) and optimize their integration into the power grid. This can include real-time adjustments to energy distribution based on predictive models. The aspects of improvement in machine learning have further new aspects with new methods and optimization for big datasets for example De Luna et al. [7] studies machine learning to find energy materials. Wetterstrand et al. [16] improves on DNA sequencing costs. NREL et al. [19] works studies best research-cell efficiency charts while Burger, B., et al. [20] paper further improves on a mobile robotic chemist.

**Advanced Energy Storage Techniques:** ML-driven methods to enhance the efficiency and longevity of energy storage systems, such as batteries. This can involve developing algorithms for better management and maintenance of these storage systems, optimizing charging/discharging cycles, and predicting battery life degradation. Improved Energy Efficiency in Industrial processes by Utilizing ML to analyze and optimize energy usage in industrial processes. Evaluating current methodologies in energy efficiency estimation and performance prediction as a benchmark for accurate energy performance estimation in residential buildings through statistical machine learning tools by Davidson et al. [3]. This can include predictive maintenance of machinery, optimization of process parameters for energy efficiency, and reduction of waste.

**Smart Grid Management:** Develop ML models to enhance the efficiency and reliability of smart grids, focusing on demand response management, fault detection, and automated restoration processes.

**Sustainable Urban Planning:** Apply ML to optimize energy usage in urban environments, including smart building management, efficient urban planning, and integration of renewable energy sources in cities.

**Enhanced Climate Modelling and Impact Assessment:** Utilize ML for more accurate and comprehensive climate modelling to better understand the impact of various energy policies on the environment.

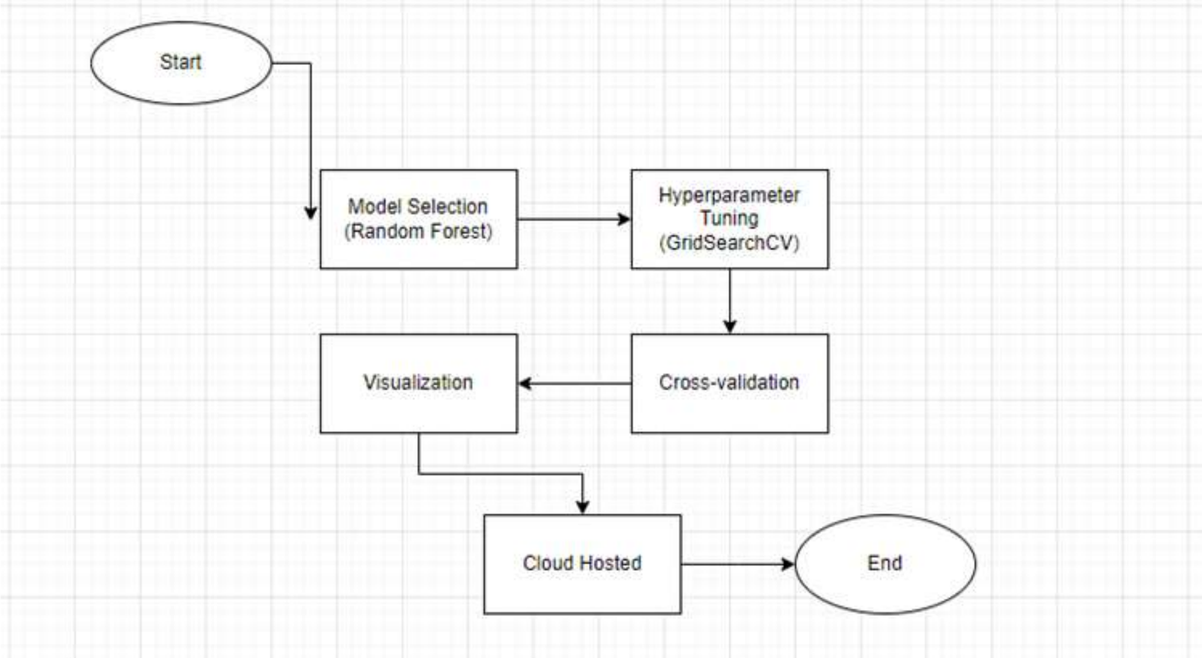



**FIGURE 1.** Proposed Methodology Flowchart

## II. PROPOSED FRAMEWORK

Introduce a novel framework that integrates advanced machine learning techniques with energy systems analysis to enhance the efficiency and sustainability of energy production and consumption. Leverage insights on using machine learning for material design and property prediction to support this framework by Rosen et al. and Jordan et al. [10, 11].

Linear regression is a linear approach to modelling the relationship between a dependent variable (target) and one or more independent variables (features) by fitting a linear equation to the observed data by Jordan [11]. The linear regression model is used to predict the target variable based on the selected feature variable.

Standard Scalar standardizes features by removing the mean and scaling to unit variance. This is often a requirement for the optimal performance of many machine learning algorithms, particularly those that assume data is normally distributed, like Support Vector Machines and k-nearest neighbours.

### A. Data Training:

```
X_train = torch.tensor(X_train, dtype=torch.float32)
X_test = torch.tensor(X_test, dtype=torch.float32)
y_train = torch.tensor(y_train, dtype=torch.float32)
y_test = torch.tensor(y_test, dtype=torch.float32)
```

**Purpose:** The torch.tensor function converts arrays (in this case, the standardized training and test sets for both features and target variables) into PyTorch tensors. Tensors are a specialized data structure similar to arrays and matrices. They are used in PyTorch to encode the inputs and outputs of a model, and the model's parameters.

dtype=torch.float32:

**Specification:** This argument specifies the data type of the tensor. torch.float32 is a common choice for input data in machine learning models, as it provides a good balance between precision and computation speed. It details the approach for training data within the proposed framework, emphasizing the importance of high-quality, diverse datasets for machine learning models. It highlights the relevance of benchmark datasets like the one from the UCI Machine Learning and the importance of key performance indicators for assessing the effectiveness of smart grid and charging infrastructure, respectively by Helmus et al. [13,] and Struck et al. [14].

### B. EDA and XGBoost

EDA is an important machine learning practice that allows us to better understand the context of our dataset, the relationships between variables. This is the basis for all decisions made in subsequent phases of the machine learning pipeline, from feature engineering to model selection and validation. It discusses the significance of initial data analysis for understanding dataset characteristics, with a nod to the importance of data quality and preparation as indicated by Wang et al. [8] in their energy efficiency dataset.

Model Training with XGBoost Harnessing the Power of Gradient Boosting Model training is the process by which machine learning algorithms learn from data. This is the heart of the machine learning pipeline and determines the model's ability to make accurate predictions.

XGBoost stands for extreme Gradient Boosting, a speed and performance implementation of gradient-boosted decision trees especially suited for structured data. XGBoost is a decision tree-based ensemble machine learning algorithm that uses a gradient boosting framework. This algorithm is known for its performance and speed, and often outperforms other algorithms for structured data tasks.
The reasons for its popularity are as follows.

**Efficiency at scale:** XGBoost is optimized for high efficiency and scalability across multiple CPUs and even GPUs.

**Regularization:** Includes L1 (lasso regression) and L2 (ridge regression) regularization to prevent overfitting and improve model performance.

- Handling missing values: XGBoost has built-in routines for handling missing values.
- Flexibility: Users can define custom optimization goals and criteria.
- Cross-validation: XGBoost implements internal cross-validation at each iteration of the boosting process, making it easy to obtain reliable estimates of model performance.

### C. Model Training

Create an instance of the XGBoost classifier, which will be used for training on the dataset. Parameters: object='binary: logistics': Specifies the learning task and corresponding learning objective. Here it is defined for a binary classification task.seeds=42: Set random seeds to reproduce. Disable use of label encoder in XGBoost for label processing. This is relevant here because we've set the evaluation metric to logarithmic loss, suitable for binary classification problems. The XGBoost classifier (xgb_clf) is initialized with the specified parameters.
These parameters are chosen to ensure model compatibility with the data format and to specify how the model's performance will be evaluated.

### D. Confusion Matrix

This is an essential step in understanding the effectiveness of the model in classification tasks beyond simple accuracy. Confusion Matrix Visualization Objective is to provide a detailed view of the model's performance by showing the types of errors it makes, in addition to its correct predictions.

Implementation functions used Custom function plot_confusion_matrix uses confusion_matrix from sklearn.metrics and heatmap from seaborn
.
**Parameters:** y_test: Actual target value of the test data set.

The confusion_matrix function creates a matrix showing the number of true positive, true negative, false positive and false negative predictions. This matrix is then displayed as a heat map.

### E. Actual and Predicted Values

The model is performing reasonably well, with the ability to predict values that correlate with the actual data. Feature Y1 is by far the most predictive of the model's output, which could mean that focusing on this feature could lead to improvements in the model's performance. The chosen parameters suggest the model is complex and may be quite fitted to the current dataset. The use of confusion matrices in model evaluation, using references that discuss the evaluation of predictive models. It provides references that do not directly mention confusion matrices; the predictive modelling work in references like Yao et al. and Davidson et al. [9, 3] implies the use of such metrics.

### F. Feature Score

The importance is derived from a predictive model. Different models have different methods for calculating feature importance. While the specific concept of "feature score" might not be directly addressed, the machine learning and material discovery discussions in can be related to the importance of features in predictive modelling.
Importance Scores: The scores (reflected by the length of the bars) indicate how much each feature contributes to the predictive accuracy of the model. A longer bar means the feature is more important. In this chart, Y1 has the longest bar, suggesting it is the most important feature in the model.

### G. Training Parameters

The parameters { 'max_depth': 20, 'min_samples_split': 2, 'n_estimators': 100 } indicate that the model is a highly accurate form of ensemble learning algorithm, likely a Random Forest, given the parameters typical to this algorithm. The min_samples_split of 2 is the minimum number of samples required to split an internal node, which is quite low and again may lead to overfitting. The n_estimators of 100 indicates that the model uses 100 trees, which is a balance between computational efficiency and model performance.

**FIGURE 2. Actual vs Predicted Values**

This also explains the standard deviation in the actual and predicted values over range of fields in the dataset hereby comparing the overall accuracy.

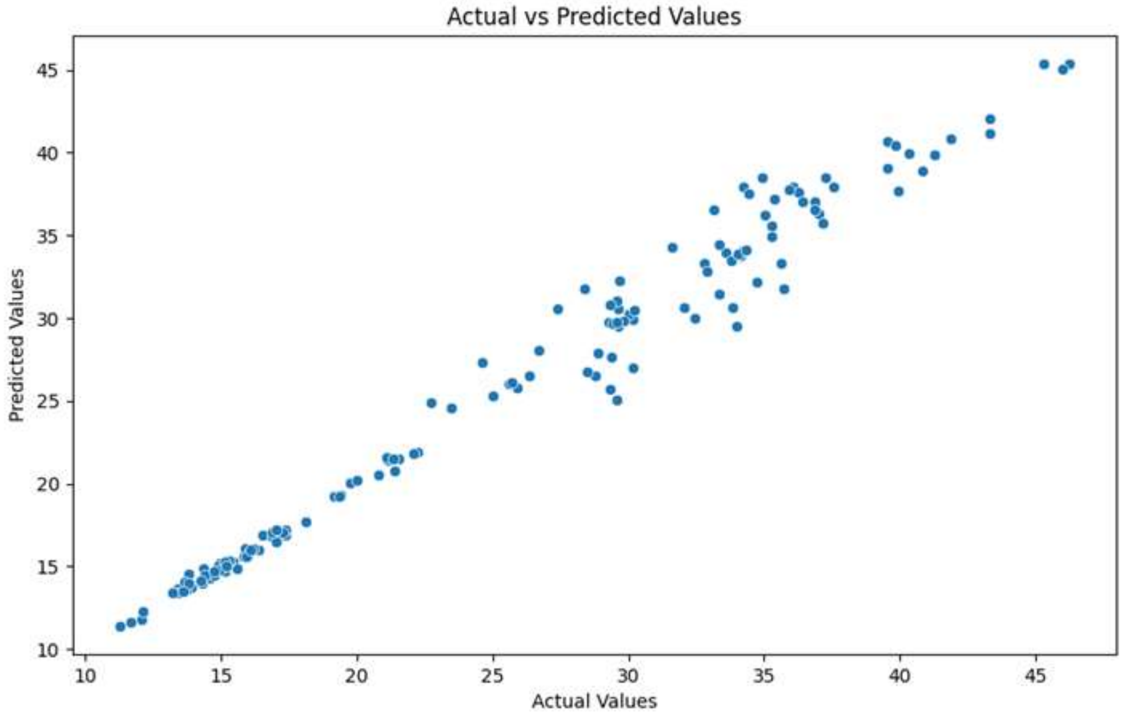



**FIGURE 3. Feature Importance**

**Here x1, x2 are the features described in the dataset that can be further studied for any energy categorical and predictive feature-based learning model.**

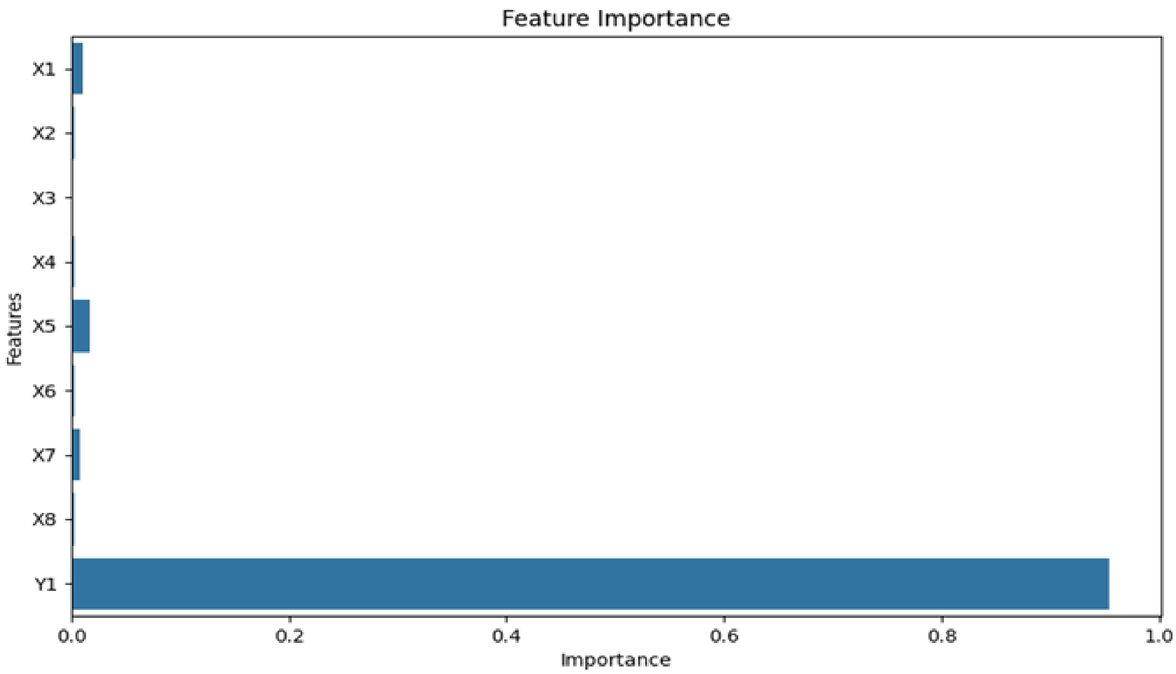


**FIGURE 4. Distribution of Actual and Predicted Values**

The following graph takes into the understanding of datasets to simulate the frequency with different variables.

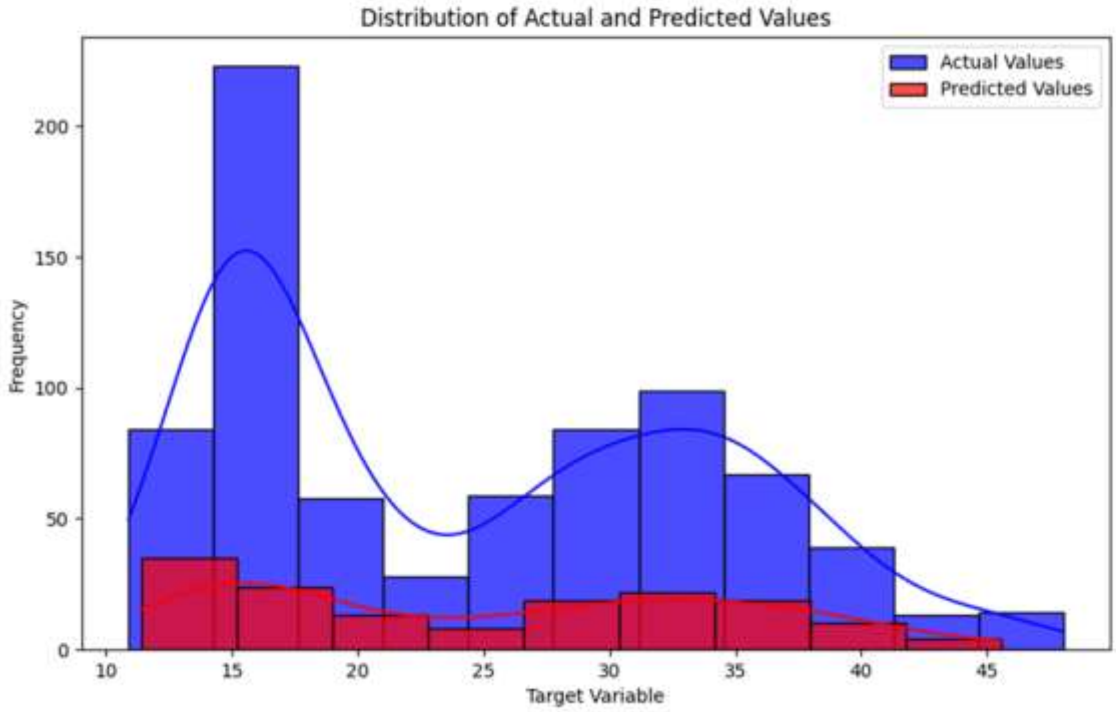



**FIGURE 5. Cumulative Feature Importance**

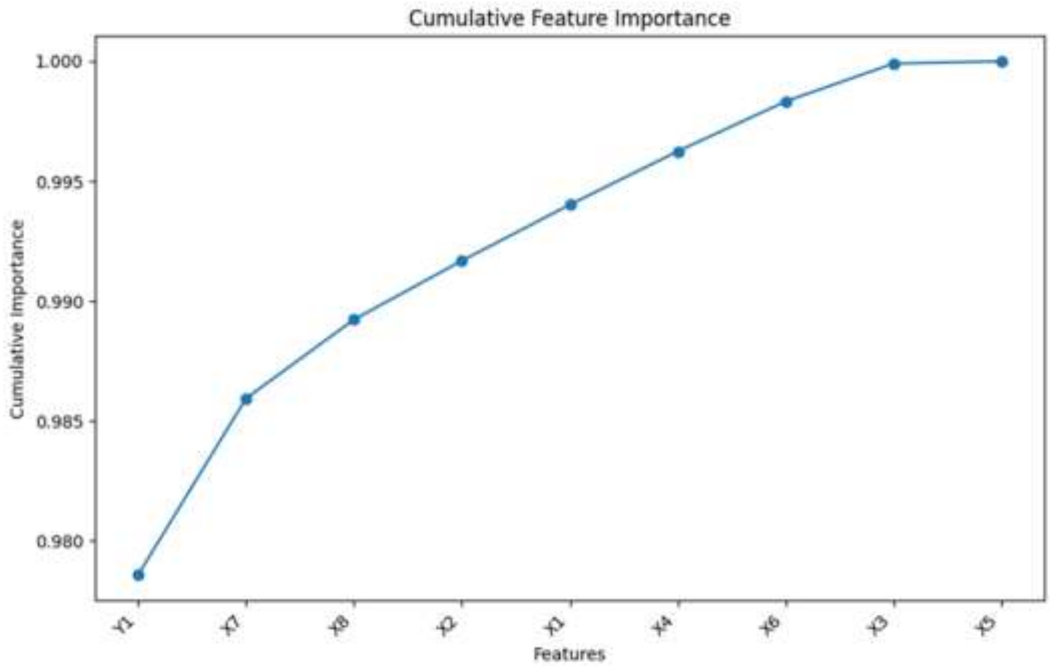
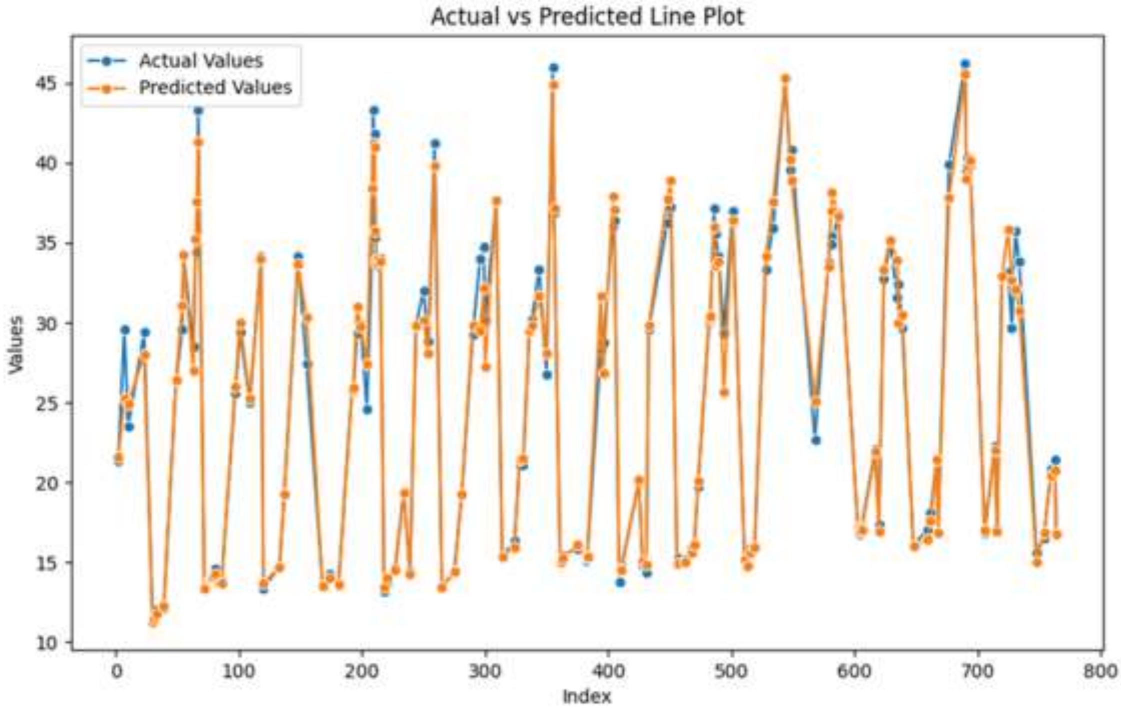


**FIGURE 6: Actual vs Predicted Line Plot**




## III. RESULTS AND DISCUSSIONS

### A. Experimental setup

**nn.Module**: This is PyTorch's base class for all neural network modules.

**init Function:** This is the constructor for your class. Here, we initialize the layers of the model. In this case, there is only one layer - nn.Linear.

**nn.Linear:** This represents a linear transformation. For a simple linear regression, you need a single linear layer with one input feature and one output feature (hence nn.Linear(1, 1)).

**Forward Method:** This method defines the forward pass of the model.

**TABLE I Actual vs Predicted Plot.**

| **Actual vs Predicted Plot** | **Scatter plot of actual vs predicted values.** | **Visualizing the accuracy of model predictions.** |
|---|---|---|
| **Correlation Heatmap** | Heatmap of correlation matrix. | Understanding relationships and dependencies between features. |
| **Box Plots** | Box plots for different qualification levels. | Analyzing the distribution and outliers in the data. |
| **Time Series Plot** | Trends over time for qualifications. | Observing changes or trends in data over time. |
| **Histogram** | Distribution of qualifications. | Understanding the frequency distribution of data points. |
| **Scatter Plot Matrix** | Pairwise relationships in data. | Exploring bivariate relationships in the dataset. |

**Model Initialization:** model = Linear Regression Model () creates an instance of our linear regression model. The importance of tuning training parameters for optimal model performance, using references that discuss machine learning methodologies and their applications in energy systems. The general discussions on machine learning in can provide a backdrop for understanding the significance of parameter tuning in model training.

## B. Performance estimation in the first iteration

- Setting Epochs: epochs = 100 defines the number of times the entire training dataset will pass through the model.
- Training Loop: The loop iterates over the dataset multiple times (defined by epochs).
- model.train(): Puts the model in training mode, which can affect certain layers like dropout or batch normalization if they were used in the model.
- optimizer.zero_grad(): Clears old gradients; otherwise, gradients would be accumulated with each backpropagation.
- Forward Pass: outputs = model(X_train) computes the predicted outputs by passing the training data through the model.
- Compute Loss: loss = criterion(outputs, y_train) calculates the loss between the predictions and the actual targets.
- Backward Pass: loss.backward() computes the gradient of the loss with respect to the model parameters.
- Update Parameters: optimizer.step() updates the model parameters based on the computed gradients.
- Logging: Every 10 epochs, the script prints the current epoch number and the loss at that epoch.

**TABLE II Number of training iterations (epochs) and preparing the model for training.**

| Step | Description | Purpose in Training Process |
|---|---|---|
| **Epochs Setting** | epochs = 100 | Defines the number of complete passes through the dataset. |
| **Training Mode** | model.train() | Prepares the model for training. |
| **Clearing Gradients** | optimizer.zero_grad() | Resets gradients from previous iterations. |
| **Forward Pass** | outputs = model(X_train) | Computes predictions using the current model state. |
| **Loss Computation** | loss = criterion(outputs, y_train) | Calculates the difference between predictions and actuals. |
| **Backward Pass** | loss.backward() | Computes gradients of the loss function w.r.t parameters. |
| **Parameters Update** | optimizer.step() | Adjusts model parameters based on calculated gradients. |
| **Logging** | print(f'Epoch[{epoch+1}/{epochs}] | Outputs training progress |

**TABLE III Feature Extraction from Input**

This table explains the purpose and description of each aspect in the feature extraction of the input.

| Aspect | Code in Script | Purpose & Description |
|---|---|---|
| Feature Extraction | X = np_data[:, 7].astype(float) | Selecting the 8th column as the feature variable and converting it to a float. |
| Target Extraction | y = np_data[:, 1].astype(float) | Selecting the 2nd column as the target variable and converting it to a float. |
| Reshaping X | X = X.reshape(-1, 1) | Reshaping X into a 2D array. The shape becomes (n_samples, 1). |
| Reshaping y | y = y.reshape(-1, 1) | Reshaping y into a 2D array. The shape becomes (n_samples, 1). |

**TABLE IV Model Training**

This table explains the purpose and description of each aspect in the Model Training and visualization.

| Step | Description | Purpose & Importance |
|---|---|---|
| Model Evaluation | Setting model to eval mode and predicting. | Assessing the model's performance on unseen data. |

| | | |
|---|---|---|
| Actual vs Predicted Plot | Scatter plot of actual vs predicted values. | Visualizing the accuracy of model predictions. |
| Correlation Heatmap | Heatmap of correlation matrix. | Understanding relationships and dependencies between features. |
| Box Plots | Box plots for different qualification levels. | Analysing the distribution and outliers in the data. |
| Time Series Plot | Trends over time for qualifications. | Observing changes or trends in data over time. |
| Histogram | Distribution of qualifications. | Understanding the frequency distribution of data points. |
| Scatter Plot Matrix | Pairwise relationships in data. | Exploring bivariate relationships in the dataset. |

```
Best Parameters: {'max_depth': 20, 'min_samples_split': 2, 'n_estimators': 100}
```

**TABLE V Comparison Results**

| **Feature** | **Paper A: "Advances in Lithium-Ion Battery Technologies" by Grey et al. [24]** | **Paper B: "Machine Learning for Renewable Energy Forecasting by Gaamouche et al."[25]-** | **Our Research: "Optimizing Energy Efficiency Using ML for Sustainable Systems"** |
|---|---|---|---|
| **Main Focus** | Development of high-energy rechargeable batteries | Predictive modelling for renewable energy outputs | Optimizing heating and cooling loads in sustainable energy systems |
| **Algorithms Used** | N/A (focus on materials science)<br><br>Grey et al. [24] | Time series forecasting algorithms, e.g., ARIMA, LSTM<br>Gaamouche et al. [25] | Deep learning, ensemble methods (e.g., XGBoost) |
| **Results** | Identification of lithium metal as a promising anode material for higher specific energy. | Improved accuracy in predicting solar and wind energy outputs. | Enhanced prediction and optimization of energy consumption in heating and cooling. |
| **Improvement Area** | Energy storage capacity and efficiency | Accuracy of renewable energy forecasts | Energy efficiency in heating and cooling systems |
| **Future Work** | Alternative materials and advanced cathode technologies. | Integrating more diverse data sources and advanced ML models. | Applying ML techniques to other areas of energy consumption and integrating IoT for real-time monitoring. |

**Comparison Results with Existing Algorithm**

We were able to improve on previous existing models like ARIMA, LSTM and other time series materials by ensemble methods such as XGBoost which enhanced prediction and optimization of energy consumption in heating and cooling.

## IV. CONCLUSION

We had progressive research in energy feature categorization and efficient energy system that may lay path for future research on bigger datasets for understanding the variation and improvement with diverse datasets. In the course various algorithms like ARIMA, LSTM and XGBoost were studied and eventually we selected XGBoost. The model is performing reasonably well, with the ability to predict values that correlate with the actual data. Feature Y1 is by far the most predictive of the model's output, which could mean that focusing on this feature could lead to improvements in the model's performance. Figure1 shows the flowchart of the proposed methodology, outlining the steps involved in applying machine learning techniques to enhance energy efficiency. Figure2 compares actual vs. predicted values, demonstrating the accuracy of the machine learning model in forecasting energy consumption or efficiency. Figure 3 displays the importance of different features used in the model, highlighting which variables significantly impact energy efficiency predictions. Figure4 illustrates the distribution of actual vs. predicted values, providing insights into the model's performance and prediction errors. Figure5 presents the cumulative importance of features, indicating how each contributes to the model's predictive capability cumulatively. Figure6 shows the correlation between actual and predicted values, offering a visual representation of the model's accuracy over a range of data points.

The trained model is uploaded to a cloud server. Services like AWS Sage Maker, Google AI Platform, or Azure Machine Learning can be used for this purpose. They provide tools for deploying, monitoring, and managing machine learning models and their implications for sustainable energy, referencing the overarching themes of energy transition, machine learning in energy applications, and the potential for future research as discussed by Geels et al. [1], Tsanas et al. [2], Maine et al. [6], Wang et al. [8], and Yao et al. [9] These citations offer a broad perspective on the intersection of technology, policy, and sustainability in energy systems.

## REFERENCES

1. Geels, F. W., Sovacool, B. K., Schwanen, T., & Sorrell, S. (2018). The next phase of the energy transition and its implications for research and policy. Nature Energy, 3(8), 628–633. https://doi.org/10.1038/s41560-018-0171-7.
2. Tsanas, A., & Xifara, A. (2012). Accurate quantitative estimation of energy performance of residential buildings using statistical machine learning tools. UCI Machine Learning Repository. https://archive.ics.uci.edu/ml/datasets/energy+efficiency
3. Davidson, D. J. (2019). Exnovating for a renewable energy transition. Nature Energy, 4(3), 254–256.
4. Horowitz, C. A. (2016). Paris agreement. International Legal Materials, 55, 740–755.
5. International Energy Agency. (2018). 2018 World Energy Outlook: Executive Summary. Retrieved from https://www.iea.org/reports/world-energy-outlook-2018
6. Chu, S., Cui, Y., & Liu, N. (2017). The path towards sustainable energy. Nature Materials, 16, 16–22.
7. Maine, E., & Garnsey, E. (2006). Commercializing generic technology: the case of advanced materials ventures. Research Policy, 35, 375–393.
8. De Luna, P., Wei, J., Bengio, Y., Aspuru-Guzik, A., & Sargent, E. (2017). Use machine learning to find energy materials. Nature, 552, 23–27.
9. Wang, H., Lei, Z., Zhang, X., Zhou, B., & Peng, J. (2019). A review of deep learning for renewable energy forecasting. Energy Conversion and Management, 198, 111799-111814.
10. Yao, Z., et al. (2021). Inverse design of nanoporous crystalline reticular materials with deep generative models. Nature Machine Intelligence, 3, 76–86.
11. Rosen, A. S., et al. (2021). Machine learning the quantum-chemical properties of metal–organic frameworks for accelerated materials discovery. Matter, 4, 1578–1597.
12. Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: trends, perspectives, and prospects. Science, 349, 255–260.
13. Personal, E., Guerrero, J. I., Garcia, A., Peña, M., & Leon, C. (2014). Key performance indicators: a useful tool to assess Smart Grid goals. Energy, 76, 976–988.
14. Helmus, J., & den Hoed, R. (2016). Key performance indicators of charging infrastructure. World Electric Vehicle Journal, 8, 733–741.
15. Struck, M.-M. (1996). Vaccine R&D success rates and development times. Nature Biotechnology, 14, 591–593.
16. Moore, G. E. (1965). Cramming more components onto integrated circuits. Electronics, 38, 114–116.
17. Wetterstrand, K. A. (2020). DNA sequencing costs: data. NHGRI Genome Sequencing Program (GSP). Retrieved from http://www.genome.gov/sequencingcostsdata
18. Rajkomar, A., Dean, J., & Kohane, I. (2019). Machine learning in medicine. New England Journal of Medicine, 380, 1347–1358.
19. Jeong, J., et al. (2021). Pseudo-halide anion engineering for α-$FAPbI_3$ perovskite solar cells. Nature, 592, 381–385.
20. NREL. (2021). Best research-cell efficiency chart. Retrieved from https://www.nrel.gov/pv/cell-efficiency.html
21. Burger, B., et al. (2020). A mobile robotic chemist. Nature, 583, 237–241.
22. Clark, M. A., et al. (2009). Design, synthesis and selection of DNA-encoded small-molecule libraries. Nature Chemical Biology, 5, 647–654.
23. Ferraz-Caetano, J., Teixeira, F., & Cordeiro, M. N. D. S. (2023). Explainable Supervised Machine Learning Model To Predict Solvation Gibbs Energy. Journal of Chemical Information and Modeling, https://doi.org/10.1021/acs.jcim.3c00544
24. Grey, C. P., & Hall, D. S. (2020). Prospects for lithium-ion batteries and beyond—a 2030 vision. Nature Communications, 11(1), Article 6279. https://doi.org/10.1038/s41467-020-19991-4
25. Gaamouche, R., Chinnici, M., Lahby, M., Abakarim, Y., & El Hasnaoui, A. (2022). Machine Learning Techniques for Renewable Energy Forecasting: A Comprehensive Review. In Computational Intelligence

Techniques for Green Smart Cities (Green Energy and Technology). Springer. First Online: 22 April 2022.
26. Chen, Y., Liu, Y., Nam, E.-Y., Zhang, Y., & Dahlak, A. (2023). Exergoeconomic and exergoenvironmental analysis and optimization of an integrated double-flash-binary geothermal system and dual-pressure ORC using zeotropic mixtures; multi-objective optimization. *Energy*, 283. https://doi.org/10.1016/j.energy.2023
27. Markard, J. (2018). The next phase of the energy transition and its implications for research and policy. Nature Energy, 3, 628–633.