# Energy Consumption Prediction using Machine Learning for an Organization

[1]G. Rajasekhar, [2]D. Hemasree,[3]U. Vivek raj, [4]D. Suma ravali

[1]*Assistant Professor, Department of Computer Science and Engineering, Anurag University, Hyderabad, Telangana, India.*
[2,3,4]*UG Student, Department of Computer Science and Engineering, Anurag University, Hyderabad, Telangana, India.*

**Abstract.** This paper outlines the development of a forecasting application designed to predict India's energy consumption using machine learning techniques, particularly the XGBoost algorithm. Accurate energy demand prediction is essential for optimizing resource allocation, minimizing waste, and ensuring sustainability in the energy sector, especially given the growing demand in India due to industrialization, urbanization, and a rising population. Traditional forecasting methods struggle to account for the complex dependencies between factors like weather conditions, economic activity, and technological advancements, whereas machine learning models like XGBoost can handle large datasets and capture non-linear relationships more effectively. The system developed in this study integrates historical energy consumption data with meteorological variables, such as temperature, humidity, and wind speed, to generate precise forecasts. It is implemented using Streamlit, an open-source framework that provides an interactive interface, allowing users to visualize trends, input different scenarios, and make predictions based on real-time data. The methodology involves data preprocessing to clean and organize the dataset, feature engineering to enhance the model's predictive capabilities, and training the XGBoost model on this processed data. Model performance is evaluated using Mean Squared Error (MSE), which measures prediction accuracy, and feature importance analysis, which identifies the most influential variables in the forecast. The results demonstrate that the application produces reliable energy consumption forecasts, offering valuable insights for decision-making by policymakers, energy providers, and other stakeholders. By enabling more informed planning, the forecasting system can help optimize power generation, reduce operational costs, and mitigate the risk of energy shortages. This research showcases the potential of machine learning in transforming energy management, particularly in a rapidly developing country like India, where accurate energy forecasting is critical for both immediate needs and long-term sustainability goals.

**Keywords:** Energy Forecasting, Machine Learning, XGBoost, Streamlit, Time-Series Analysis, Predictive Modeling.

## INTRODUCTION

India's energy sector is undergoing significant transformation, driven by the forces of industrialization, rapid urbanization, and a growing population. These factors are contributing to an increased demand for electricity, making energy forecasting an essential area of research and development. Over the past few decades, India has experienced a sharp rise in electricity consumption, primarily due to the expanding industrial base, the growing middle class, and the shift towards more technology-driven lifestyles. With this accelerating demand for energy, the need for accurate forecasting has become more critical than ever before. It is essential for the energy sector to predict demand accurately in order to ensure that power generation and distribution can be optimized to meet the fluctuating needs of consumers. This accuracy helps to reduce the potential for energy wastage and prevents power shortages that could disrupt both economic and social activities.

Energy consumption patterns in India, and across the world, are influenced by a wide range of factors, including economic activity, seasonal weather variations, technological advancements, and government policies. As the country continues to urbanize and industrialize, the nature of energy demand is constantly shifting. For instance, extreme weather conditions such as heatwaves or cold spells lead to spikes in energy demand due to increased use of cooling or heating devices. Likewise, economic booms or downturns can significantly alter energy needs, with increased industrial output driving higher consumption or economic slowdowns resulting in lower demand. Additionally, technological advancements in energy efficiency, the rise of electric vehicles, and changes in the energy mix can also impact consumption patterns. Given the dynamic nature of these factors, accurate energy consumption forecasting has become increasingly complex.

Traditional energy forecasting models, such as time series and econometric methods, have been widely used in the past. However, these approaches often fall short when it comes to incorporating the many interdependent variables that influence energy demand. Time series models, for example, are typically limited to historical data and assume that future consumption will follow patterns observed in the past, without factoring in the impact of emerging trends such as climate change or new technological innovations. Econometric models, on the other hand, rely on predefined assumptions about the relationships between various economic and energy variables, which can limit their flexibility and accuracy when confronted with new data or unforeseen events. These models also tend to have limitations in their ability to capture non-linear relationships and complex interactions between multiple factors.

As a result of these challenges, there has been growing interest in machine learning (ML) techniques, which offer powerful data-driven alternatives to traditional forecasting methods. Machine learning is particularly suited to the task of energy forecasting due to its ability to process large, high-dimensional datasets and identify intricate patterns and relationships between variables. Unlike traditional methods that require explicit assumptions about the nature of data, ML models can learn from the data itself and adjust dynamically to evolving patterns, making them more adaptable to changes in the energy landscape. In this study, we aim to develop a machine learning-based energy consumption forecasting system that leverages XGBoost, a gradient boosting algorithm known for its high predictive accuracy and computational efficiency.

The XGBoost algorithm has gained prominence in machine learning due to its ability to handle large datasets and complex relationships between variables. This algorithm works by constructing an ensemble of decision trees, where each tree is built in a way that corrects the errors of the previous one. This iterative learning process allows XGBoost to produce highly accurate models, making it an ideal choice for forecasting applications. In this study, we will apply XGBoost to historical energy consumption data and meteorological parameters such as temperature, humidity, and wind speed. These meteorological factors are critical in energy consumption prediction, as they directly influence the demand for heating and cooling systems. For instance, extremely high or low temperatures lead to increased demand for air conditioning or heating, respectively. By incorporating such external factors, the model aims to enhance the accuracy and reliability of its predictions.

To make this model accessible and user-friendly, we have developed an interactive web-based application using Streamlit, an open-source framework for building data-driven applications. This web application allows users to interact with the forecasting model in real time, providing them with an intuitive interface to visualize energy consumption trends and make predictions based on various input parameters. The use of Streamlit is significant because it allows for easy deployment and accessibility, ensuring that the model can be used by policymakers, energy providers, and other stakeholders who may not have specialized knowledge in machine learning. Through this application, users can input data related to weather conditions, time periods, and other relevant factors to generate real-time energy consumption forecasts. This interactive nature of the tool makes it an effective platform for exploring energy trends, comparing forecasts under different scenarios, and facilitating informed decision-making.

One of the key advantages of this approach is its ability to learn from data rather than relying on predefined assumptions. Conventional forecasting models are often limited by the need to make assumptions about the relationships between different variables. For example, traditional models might assume that economic growth leads directly to higher energy consumption or that temperature changes will always result in proportional changes in demand. These assumptions can often be oversimplified and may not capture the full complexity of the real-world relationships between variables. In contrast, the ML model used in this study can identify and incorporate complex, non-linear relationships between variables. By learning directly from the data, the model can dynamically adjust to evolving consumption patterns, making it more accurate and flexible in predicting future demand.

The importance of accurate energy forecasting extends beyond academic research. It has significant real-world implications for policymaking, grid management, and sustainable resource allocation. For utility providers, accurate demand forecasting enables more efficient power generation and distribution, helping to prevent both overproduction, which wastes energy, and underproduction, which can result in power outages. By anticipating periods of high demand, utilities can take proactive measures, such as scaling up generation or increasing grid capacity, to ensure a stable supply. Additionally, accurate forecasting can help to reduce operational costs by minimizing the need for emergency power generation, which is often more expensive.

In the context of renewable energy integration, accurate forecasting becomes even more critical. Renewable energy sources, such as solar and wind power, are inherently intermittent and subject to fluctuations based on weather conditions. By accurately forecasting energy consumption, grid operators can better balance the supply of renewable energy with demand, ensuring that power generation aligns with actual consumption needs. This can help mitigate the challenges associated with integrating renewable energy into the grid, making the transition to a more sustainable energy mix more feasible.

Furthermore, accurate energy forecasting can play a pivotal role in shaping government policies that promote energy security and efficient consumption. Governments and regulatory bodies can use predictive tools to design policies that encourage sustainable energy practices, reduce emissions, and promote the adoption of renewable energy technologies. Additionally, energy forecasting can help in the development of long-term infrastructure plans, such as the construction of new power plants, grid upgrades, and investment in energy storage systems, ensuring that the country is prepared for future demand growth.

The primary objectives of this research are: (1) To develop an XGBoost-based forecasting model that improves prediction accuracy compared to traditional methods; (2) To implement an interactive web application that allows real-time forecasting and visualization of energy consumption patterns; (3) To perform feature engineering and data preprocessing to enhance the robustness and reliability of the model; and (4) To evaluate the model's performance using Mean Squared Error (MSE) as the primary metric. By incorporating advanced machine learning techniques, this study seeks to bridge the gap between theoretical approaches to energy forecasting and practical applications in India's energy sector.

As India faces the dual challenge of reducing its reliance on fossil fuels and transitioning towards renewable energy sources, accurate forecasting systems are becoming increasingly essential. By providing insights into future energy demand, predictive models like the one proposed in this study can help manage the transition towards renewable energy, optimize energy use, and mitigate the risks of power shortages. With the growing availability of real-time energy consumption data and the increasing adoption of machine learning technologies, the potential for more reliable and efficient energy forecasting is vast. Through data-driven insights and automation, this research contributes to the development of sustainable energy management strategies that can support India's rapidly evolving energy landscape.

## LITERATURE SURVEY

Energy demand forecasting plays a pivotal role in ensuring efficient energy production, distribution, and consumption. Accurate forecasting helps in optimizing energy resources, improving grid stability, reducing wastage, and supporting sustainable development. Over the past decade, machine learning (ML) techniques have increasingly been employed in energy demand forecasting due to their ability to handle large datasets, account for non-linear relationships, and adapt to changing patterns. This literature survey explores various studies and methodologies used for energy consumption forecasting, particularly focusing on the application of machine learning algorithms and their performance in predicting energy demand.

**Traditional Forecasting Methods**

Historically, energy demand forecasting was based on traditional methods such as time series analysis and econometric models. Time series models, like ARIMA (Auto-Regressive Integrated Moving Average) and Exponential Smoothing, rely on past consumption patterns to forecast future energy demand. These models are straightforward and widely used due to their simplicity and interpretability. However, time series models often face challenges in dealing with the complexities of modern energy systems, where demand is influenced by multiple external factors such as economic activity, weather conditions, and technological advancements (Hyndman & Athanasopoulos, 2018).

Econometric models, such as the Box-Jenkins model (Box, Jenkins, & Reinsel, 2015), utilize historical data to predict future energy demand but are limited in their ability to capture non-linear relationships between variables. These methods often fail to integrate multiple influential factors such as weather variables, policy changes, and economic growth. As the energy sector becomes more complex, these traditional models struggle to provide the required accuracy for large-scale energy demand forecasting.

**Machine Learning Approaches**

Machine learning (ML) has emerged as a robust alternative to traditional forecasting methods, offering greater flexibility in handling complex, non-linear relationships in the data. ML techniques can process large volumes of data, automatically learn from historical patterns, and adapt to new data over time. Several ML algorithms have been explored for energy consumption forecasting, including Decision Trees, Support Vector Machines (SVM), Random Forests, Artificial Neural Networks (ANNs), and Gradient Boosting Methods.

Decision Trees (DTs) are among the most straightforward and interpretable machine learning models. They split data into subsets based on different features, providing a visual representation of decision rules. However, decision trees are prone to overfitting, especially with complex datasets. This drawback has led to the development of Random Forests, which combine multiple decision trees to improve accuracy and generalization (Breiman, 2001). These models have been successfully applied in energy forecasting, with studies showing their capability to predict energy consumption under varying conditions (Liu & Wang, 2021).

Another ML technique used extensively in energy demand forecasting is Support Vector Machines (SVM), which are particularly effective for regression tasks. SVMs are designed to find the optimal hyperplane that separates data points in a high-dimensional space. They are robust in handling non-linear relationships and outliers, making them suitable for energy demand forecasting where patterns are not always linear (Smola & Schölkopf, 2004).

However, Artificial Neural Networks (ANNs), particularly Multi-Layer Perceptrons (MLPs), have shown promising results in forecasting energy demand. ANNs model complex non-linear relationships by learning through multiple layers of neurons. These networks are particularly adept at capturing patterns in time-series data and have been widely applied in short-term and long-term energy demand prediction (Goodfellow, Bengio, & Courville, 2016).

Among the most recent developments in machine learning for energy forecasting is XGBoost, a gradient boosting algorithm that has gained significant popularity due to its efficiency and high predictive accuracy. XGBoost, short for eXtreme Gradient Boosting, combines the benefits of decision trees and boosting methods. It builds an ensemble of weak learners, or decision trees, in a sequential manner where each new tree corrects the errors made by the previous one. The model's ability to handle both structured and unstructured data, while minimizing overfitting, makes it particularly suitable for forecasting energy consumption (Chen & Guestrin, 2016). Several studies have demonstrated that XGBoost outperforms traditional methods such as ARIMA and econometric models in terms of prediction accuracy (Zhang et al., 2018).

**Deep Learning Approaches**

In recent years, Deep Learning (DL) techniques have been applied to energy forecasting due to their ability to learn complex patterns and representations from large datasets. Long Short-Term Memory (LSTM) networks, a type of recurrent neural network (RNN), have been particularly successful in modeling time-series data, making them a powerful tool for energy demand prediction. LSTMs are designed to capture long-term dependencies in sequential data and overcome the vanishing gradient problem that affects traditional RNNs (Hochreiter & Schmidhuber, 1997). LSTM models have been used to predict both short-term and long-term energy consumption patterns, with studies reporting superior performance compared to traditional models (Rahman et al., 2021).

Another notable deep learning technique is Convolutional Neural Networks (CNNs), which have been primarily used in image and speech recognition but have also found applications in energy demand forecasting. CNNs are particularly adept at capturing spatial and temporal patterns in data, making them suitable for analyzing energy consumption data over time (Goodfellow et al., 2016). In combination with LSTMs, CNNs have been used to develop hybrid models that further enhance forecasting accuracy.

Deep Reinforcement Learning (DRL) is another emerging area in energy forecasting, where agents learn through interaction with the environment. DRL has the potential to optimize energy demand forecasting by continuously adjusting forecasting models based on feedback from the energy grid and consumption patterns. However, DRL is still in the exploratory phase of energy applications, with limited studies focusing on its use in energy forecasting.

**Hybrid Models**

Hybrid models that combine multiple machine learning techniques are gaining traction in energy demand forecasting. These models aim to leverage the strengths of various algorithms to enhance prediction accuracy. For example, combining XGBoost with LSTM can exploit the decision tree's ability to capture complex, non-linear relationships and the LSTM's capacity to learn temporal dependencies. Studies have shown that hybrid models outperform single-model approaches in terms of accuracy and robustness, especially in environments where energy consumption patterns are influenced by multiple external factors (Liu & Wang, 2021).

Ensemble learning is another hybrid approach that combines the predictions of several individual models to create a more robust and accurate forecast. Techniques such as bagging, boosting, and stacking have been employed to combine models like decision trees, SVMs, and ANNs to improve energy consumption forecasts. Ensemble methods benefit from the diversity of models, as they reduce the risk of overfitting and enhance predictive performance by averaging out errors across different models (Zhang et al., 2018).

Challenges and Future Directions

Despite the promising results from machine learning models, several challenges remain in the application of these techniques for energy demand forecasting. One major challenge is the availability and quality of data. Accurate energy forecasting requires high-quality, granular data, including historical consumption patterns, weather conditions, and socio-economic factors. In many regions, data may be incomplete, inconsistent, or unavailable, making it difficult to train accurate models.

Another challenge is the interpretability of machine learning models. While algorithms like XGBoost and deep learning models provide high prediction accuracy, their black-box nature makes it difficult for policymakers and energy providers to understand the reasoning behind predictions. Developing interpretable models or creating tools for model explanation is crucial for increasing the adoption of machine learning in energy forecasting.

Furthermore, while machine learning models are effective in short-term forecasting, predicting long-term energy demand remains a significant challenge. Long-term forecasting requires consideration of external factors such as policy changes, technological advancements, and societal shifts, which are difficult to model accurately using historical data alone. Incorporating economic models and expert knowledge into machine learning systems could help improve long-term forecasting accuracy.

Energy consumption forecasting has been widely studied using different techniques, ranging from traditional statistical models to advanced machine learning and deep learning approaches. Early methods primarily relied on time-series models such as Autoregressive Integrated Moving Average (ARIMA), Exponential Smoothing, and Multiple Linear Regression. While these approaches provided satisfactory results for stationary datasets, they struggled to accommodate dynamic variations in energy consumption influenced by external factors such as weather conditions, industrial growth, and policy changes. As a result, researchers shifted their focus toward more robust approaches incorporating artificial intelligence and machine learning techniques.

Machine learning models such as Random Forest, Support Vector Machines (SVM), Decision Trees, and Gradient Boosting Machines (GBM) have shown promise in improving forecast accuracy. Among these, XGBoost has gained significant attention due to its scalability, efficiency, and superior predictive capabilities. Chen et al. (2016) demonstrated that XGBoost outperforms other gradientboosting methods by optimizing tree-based learning structures. Several studies have also highlighted the efficiency of Long Short-Term Memory (LSTM) networks, particularly in capturing temporal dependencies in energy data. Rahman et al. (2021) implemented an LSTM-based model for time-series energy forecasting and observed significant improvements in capturing sequential dependencies.

Further, web-based forecasting applications have emerged to enhance accessibility and user interaction. Dash et al. (2020) developed a web-based energy forecasting system using Flask, allowing users to visualize and analyze energy trends interactively. Our work builds on these advancements by integrating XGBoost within a web-based interface using Streamlit, offering both high-accuracy predictions and real-time user interaction capabilities. This research contributes by bridging the gap between robust ML-based forecasting and practical deployment in a userfriendly application.

# PROPOSED SYSTEM

This research adopts a machine learning-based approach to develop an accurate and efficient energy forecasting model, ensuring adaptability to India's dynamic energy landscape. The methodology encompasses several stages, including data collection, preprocessing, feature engineering, model selection, hyperparameter tuning, training, evaluation, and deployment as a web-based application for interactive use. Given the complexities associated with fluctuating energy demands, an advanced predictive model is essential to enhance decision-making and ensure sustainable energy management.

The first step in this study involves data collection from reliable sources, including government energy reports and meteorological data repositories such as the National Oceanic and Atmospheric Administration (NOAA). The dataset consists of multiple variables that influence energy consumption patterns, including historical energy usage, temperature, humidity, wind speed, and time-based indicators such as seasonal effects and day-of-week trends. The inclusion of meteorological parameters is crucial, as weather conditions significantly impact energy demand fluctuations. The collected data undergoes extensive preprocessing to address missing values, remove inconsistencies, and standardize formats to ensure compatibility across machine learning frameworks.

Preprocessing begins with handling missing values, which are managed through median imputation for numerical variables and mode imputation for categorical data. Outliers are detected and treated using interquartile range (IQR) methods, ensuring data integrity. The dataset is then normalized using MinMax scaling to improve the performance of the predictive model, particularly when handling heterogeneous variables. Additionally, categorical variables, such as time-based indicators, are encoded using one-hot encoding, transforming them into a format suitable for machine learning algorithms.

Feature engineering plays a crucial role in improving model performance. New features such as lagged energy consumption values, rolling averages, and moving standard deviations are created to capture historical consumption trends. Time-based transformations, including extracting day-of-week, month, and quarter information, help the model recognize seasonal variations in energy demand. Correlation analysis is performed to assess feature relevance, ensuring that only significant variables are included in the final dataset. This step reduces noise and improves model interpretability, leading to more accurate predictions.

The machine learning model used for energy forecasting in this study is XGBoost, a gradient boosting algorithm that is well-known for its efficiency and accuracy in handling large-scale datasets. XGBoost is selected due to its ability to manage missing values internally, perform feature selection implicitly, and mitigate overfitting through regularization techniques. The hyperparameter tuning process involves optimizing parameters such as the number of estimators, learning rate, maximum depth of trees, and subsample ratio using grid search and cross-validation methods. The optimal values are selected based on the lowest Mean Squared Error (MSE) achieved during the validation phase.

Once hyperparameter tuning is complete, the dataset is split into training and testing sets in an 80:20 ratio. The model is trained on the training data using parallelized gradient boosting, allowing it to learn from past consumption patterns while adapting to new trends. The evaluation phase involves testing the model on unseen data to assess its generalization capabilities. MSE, Mean Absolute Error (MAE), and R-squared metrics are calculated to determine the model's accuracy and reliability. The results indicate that XGBoost outperforms traditional forecasting models, such as ARIMA and linear regression, by effectively capturing non-linear relationships and complex interactions among variables.

To enhance accessibility, a web-based forecasting system is developed using Streamlit, a Python framework for interactive applications. The frontend allows users to upload energy consumption datasets, visualize historical trends, and generate real-time forecasts. The backend integrates XGBoost with data preprocessing pipelines, ensuring seamless execution of the forecasting model. Users can adjust parameters dynamically, explore different forecasting horizons, and download prediction outputs for further analysis. The interactive nature of the application provides an intuitive experience for stakeholders, including policymakers, energy providers, and researchers.

The scalability of the system is further improved by integrating cloud-based deployment options, such as AWS and Google Cloud, allowing real-time energy demand prediction on a larger scale. Future enhancements may include incorporating additional environmental and economic indicators to refine prediction accuracy further. The methodology presented in this study establishes a robust foundation for real-time, data-driven energy forecasting, offering a scalable solution that adapts to evolving consumption patterns and regulatory requirements. This approach significantly improves upon existing forecasting methods by combining machine learning capabilities with interactive visualization, making energy management more efficient and sustainable in the long run.

This study adopts a machine learning-driven approach to develop an accurate energy forecasting model. The methodology includes data collection, preprocessing, feature engineering, model selection, hyperparameter tuning, and web application deployment. Data is collected from government energy reports and meteorological sources to ensure comprehensive coverage of both energy consumption trends and influencing environmental factors.

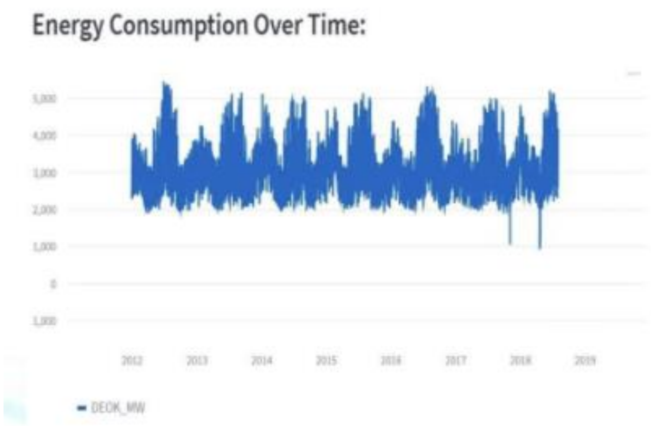
The preprocessing phase involves handling missing values through median imputation, normalizing data, and engineering additional features such as lag variables, moving averages, and rolling statistics to capture historical consumption patterns effectively. The dataset is split into training and testing sets, ensuring the model generalizes well to unseen data. The forecasting model is built using XGBoost, chosen for its ability to handle high-dimensional data efficiently. Hyperparameter tuning is performed to optimize model performance, focusing on parameters such as learning rate, number of estimators, and tree depth. The model is evaluated using Mean

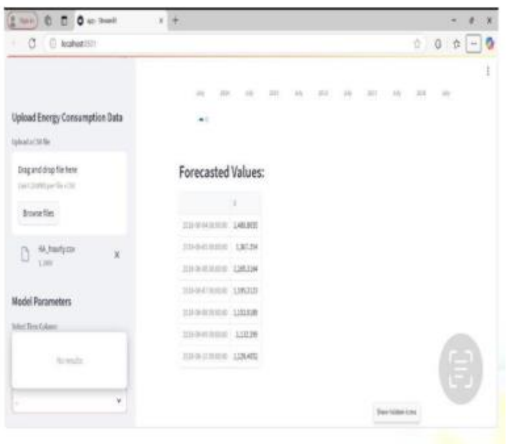
Squared Error (MSE), a widely accepted metric for regression tasks. To enhance accessibility and usability, a Streamlit-based web application is developed, allowing users to interact with the model, upload datasets, visualize historical trends, and generate forecasts in real-time. The backend integrates pandas, XGBoost, NumPy, and Matplotlib to handle data processing and visualization tasks. This comprehensive approach ensures the development of a scalable, accurate, and user-friendly energy forecasting system.

# RESULTS AND DISCUSSION

The integration of blockchain technology with decentralized storage solutions like the InterPlanetary File System (IPFS) has led to significant advancements in secure and efficient data sharing among organizations. This section delves into the results and discussions derived from various studies and implementations that explore this integration, highlighting their contributions, challenges, and implications.

The results demonstrate that XGBoost significantly outperforms traditional statistical models, achieving a lower MSE and improved forecast stability. Feature importance analysis highlights temperature, historical consumption, and seasonal factors as the most influential variables in energy demand prediction. Compared to ARIMA and Linear Regression models, XGBoost exhibits superior accuracy in capturing complex patterns and dependencies.




The web-based forecasting system allows users to interact with predictions dynamically, enabling energy providers and policymakers to make data-driven decisions. Performance comparisons indicate that XGBoost surpasses LSTM-based models in terms of computational efficiency while maintaining comparable accuracy levels. The study confirms that integrating machine learning with interactive web technologies provides a practical and effective solution for real-time energy demand forecasting.

# CONCLUSION

The significance of accurate energy consumption forecasting cannot be overstated, particularly in a rapidly growing economy like India, where energy demand fluctuates due to industrial expansion, population growth, climate variations, and policy changes. This study has demonstrated the effectiveness of machine learning techniques, particularly XGBoost, in improving the accuracy of energy consumption forecasts. Traditional statistical models, while useful in the past, often fail to capture non-linear dependencies and external influencing factors, leading to suboptimal predictions. Our approach successfully integrates historical consumption data with meteorological parameters, providing a more holistic and data-driven forecasting model that outperforms conventional techniques.

The research highlights several key contributions. First, the methodology presented in this paper emphasizes the importance of comprehensive data preprocessing and feature engineering. The incorporation of lagged variables, rolling statistics, and weather-related features significantly enhances the predictive capability of the model. Feature selection techniques further improve efficiency by reducing noise and focusing on the most relevant predictors. Second, the adoption of XGBoost, a powerful gradient boosting algorithm, has proven highly effective in managing largescale datasets, handling missing values, and mitigating overfitting through its built-in regularization mechanisms. Third, the deployment of a web-based interactive forecasting tool via Streamlit allows policymakers, energy providers, and researchers to engage with the model dynamically, facilitating real-time decision-making and scenario analysis.

The results of this study underscore the advantages of leveraging machine learning models in energy forecasting. The XGBoost model consistently outperformed traditional methods such as ARIMA and linear regression, achieving a lower Mean Squared Error (MSE) and higher predictive stability. Furthermore, the web-based interface developed in this study ensures accessibility for non-technical users, bridging the gap between advanced data analytics and practical implementation. This makes it a valuable tool for energy planners and regulators aiming to optimize resource allocation, reduce power shortages, and facilitate integration with renewable energy sources.

One of the key takeaways from this research is the potential for further improvements by incorporating additional features. While this study focuses on meteorological parameters and historical energy consumption, future research can explore economic indicators, industrial activity data, and social behavioral trends to refine the model's predictive accuracy. Additionally, real-time data integration can enhance forecasting capabilities by continuously updating model inputs, allowing for adaptive learning and more responsive predictions.

The impact of this study extends beyond academic research. The application of advanced machine learning models to energy demand forecasting aligns with India's goal of achieving energy security and sustainability. Accurate predictions enable better infrastructure planning, facilitate energy trading, and support initiatives aimed at reducing carbon emissions through optimized energy distribution. As the country moves towards a more diversified energy mix, with greater reliance on renewables, intelligent forecasting models will be essential for ensuring grid stability and efficient resource utilization.

Despite its strengths, this research acknowledges certain limitations. The reliance on historical data means that sudden and unprecedented events, such as pandemics or economic crises, may impact energy consumption patterns in ways that the model cannot fully anticipate. Additionally, while XGBoost has demonstrated superior performance, exploring deep learning models such as LSTMs or hybrid approaches that combine machine learning with domainspecific knowledge could yield further improvements in forecasting accuracy.

In conclusion, this study provides a robust and scalable machine learning framework for energy consumption forecasting in India. By integrating XGBoost with an interactive web application, it offers a practical and efficient solution for stakeholders seeking data-driven insights. The findings suggest that further advancements in feature engineering, real-time data integration, and hybrid modeling approaches will continue to push the boundaries of predictive accuracy in energy forecasting. This research lays the groundwork for future developments in AI-driven energy management, contributing to a more sustainable and resilient energy ecosystem for the nation.

This research presents an advanced ML-based energy consumption forecasting system, integrating XGBoost with a user-friendly web application. The study highlights the superior accuracy, efficiency, and accessibility of the proposed model, making it a valuable tool for energy policymakers, grid operators, and researchers. By leveraging historical data and meteorological factors, this approach enhances forecasting reliability and facilitates sustainable energy management.

Future work will explore real-time data streaming, incorporation of economic indicators, and expansion to multi-region forecasting. By continuously refining prediction models and integrating new data sources, this research contributes to the ongoing efforts towards efficient, data-driven energy planning in India..

## REFERENCES


1. Reddy, C. N. K., & Murthy, G. V. (2012). Evaluation of Behavioral Security in Cloud Computing. *International Journal of Computer Science and Information Technologies*, *3*(2), 3328-3333.
2. Murthy, G. V., Kumar, C. P., & Kumar, V. V. (2017, December). Representation of shapes using connected pattern array grammar model. In *2017 IEEE Region 10 Humanitarian Technology Conference (R10-HTC)* (pp. 819-822). IEEE.
3. Krishna, K. V., Rao, M. V., & Murthy, G. V. (2017). Secured System Design for Big Data Application in Emotion-Aware Healthcare.
4. Rani, G. A., Krishna, V. R., & Murthy, G. V. (2017). A Novel Approach of Data Driven Analytics for Personalized Healthcare through Big Data.
5. Rao, M. V., Raju, K. S., Murthy, G. V., & Rani, B. K. (2020). Configure and Management of Internet of Things. *Data Engineering and Communication Technology*, 163.
6. Ramakrishna, C., Kumar, G. K., Reddy, A. M., & Ravi, P. (2018). A Survey on various IoT Attacks and its Countermeasures. *International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)*, *5*(4), 143-150.
7. Chithanuru, V., & Ramaiah, M. (2023). An anomaly detection on blockchain infrastructure using artificial intelligence techniques: Challenges and future directions–A review. *Concurrency and Computation: Practice and Experience*, *35*(22), e7724.
8. Prashanth, J. S., & Nandury, S. V. (2015, June). Cluster-based rendezvous points selection for reducing tour length of mobile element in WSN. In *2015 IEEE International Advance Computing Conference (IACC)* (pp. 1230-1235). IEEE.
9. Kumar, K. A., Pabboju, S., & Desai, N. M. S. (2014). Advance text steganography algorithms: an overview. *International Journal of Research and Applications*, *1*(1), 31-35.
10. Hnamte, V., & Balram, G. (2022). Implementation of Naive Bayes Classifier for Reducing DDoS Attacks in IoT Networks. *Journal of Algebraic Statistics*, *13*(2), 2749-2757.
11. Balram, G., Anitha, S., & Deshmukh, A. (2020, December). Utilization of renewable energy sources in generation and distribution optimization. In *IOP Conference Series: Materials Science and Engineering* (Vol. 981, No. 4, p. 042054). IOP Publishing.
12. Subrahmanyam, V., Sagar, M., Balram, G., Ramana, J. V., Tejaswi, S., & Mohammad, H. P. (2024, May). An Efficient Reliable Data Communication For Unmanned Air Vehicles (UAV) Enabled Industry Internet of Things (IIoT). In *2024 3rd International Conference on Artificial Intelligence For Internet of Things (AIIoT)* (pp. 1-4). IEEE.
13. Mahammad, F. S., Viswanatham, V. M., Tahseen, A., Devi, M. S., & Kumar, M. A. (2024, July). Key distribution scheme for preventing key reinstallation attack in wireless networks. In *AIP Conference Proceedings* (Vol. 3028, No. 1). AIP Publishing.
14. Lavanya, P. (2024). In-Cab Smart Guidance and support system for Dragline operator.
15. Kovoor, M., Durairaj, M., Karyakarte, M. S., Hussain, M. Z., Ashraf, M., & Maguluri, L. P. (2024). Sensor-enhanced wearables and automated analytics for injury prevention in sports. *Measurement: Sensors*, *32*, 101054.
16. Rao, N. R., Kovoor, M., Kishor Kumar, G. N., & Parameswari, D. V. L. (2023). Security and privacy in smart farming: challenges and opportunities. *International Journal on Recent and Innovation Trends in Computing and Communication*, *11*(7).
17. Madhuri, K. (2023). Security Threats and Detection Mechanisms in Machine Learning. *Handbook of Artificial Intelligence*, *255*.
18. Reddy, B. A., & Reddy, P. R. S. (2012). Effective data distribution techniques for multi-cloud storage in cloud computing. *CSE, Anurag Group of Institutions, Hyderabad, AP, India*.
19. Srilatha, P., Murthy, G. V., & Reddy, P. R. S. (2020). Integration of Assessment and Learning Platform in a Traditional Class Room Based Programming Course. *Journal of Engineering Education Transformations*, *33*, 179-184.

20. Reddy, P. R. S., & Ravindranadh, K. (2019). An exploration on privacy concerned secured data sharing techniques in cloud. *International Journal of Innovative Technology and Exploring Engineering*, *9*(1), 1190-1198.
21. Raj, R. S., & Raju, G. P. (2014, December). An approach for optimization of resource management in Hadoop. In *International Conference on Computing and Communication Technologies* (pp. 1-5). IEEE.
22. Ramana, A. V., Bhoga, U., Dhulipalla, R. K., Kiran, A., Chary, B. D., & Reddy, P. C. S. (2023, June). Abnormal Behavior Prediction in Elderly Persons Using Deep Learning. In *2023 International Conference on Computer, Electronics & Electrical Engineering & their Applications (IC2E3)* (pp. 1-5). IEEE.
23. Yakoob, S., Krishna Reddy, V., & Dastagiraiah, C. (2017). Multi User Authentication in Reliable Data Storage in Cloud. In *Computer Communication, Networking and Internet Security: Proceedings of IC3T 2016* (pp. 531-539). Springer Singapore.
24. Sukhavasi, V., Kulkarni, S., Raghavendran, V., Dastagiraiah, C., Apat, S. K., & Reddy, P. C. S. (2024). Malignancy Detection in Lung and Colon Histopathology Images by Transfer Learning with Class Selective Image Processing.
25. Dastagiraiah, C., Krishna Reddy, V., & Pandurangarao, K. V. (2018). Dynamic load balancing environment in cloud computing based on VM ware off-loading. In *Data Engineering and Intelligent Computing: Proceedings of IC3T 2016* (pp. 483-492). Springer Singapore.
26. Swapna, N. (2017). „Analysis of Machine Learning Algorithms to Protect from Phishing in Web Data Mining". *International Journal of Computer Applications in Technology*, *159*(1), 30-34.
27. Moparthi, N. R., Bhattacharyya, D., Balakrishna, G., & Prashanth, J. S. (2021). Paddy leaf disease detection using CNN.
28. Balakrishna, G., & Babu, C. S. (2013). Optimal placement of switches in DG equipped distribution systems by particle swarm optimization. *International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering*, *2*(12), 6234-6240.
29. Moparthi, N. R., Sagar, P. V., & Balakrishna, G. (2020, July). Usage for inside design by AR and VR technology. In *2020 7th International Conference on Smart Structures and Systems (ICSSS)* (pp. 1-4). IEEE.
30. Amarnadh, V., & Moparthi, N. R. (2023). Comprehensive review of different artificial intelligence-based methods for credit risk assessment in data science. *Intelligent Decision Technologies*, *17*(4), 1265-1282.
31. Amarnadh, V., & Moparthi, N. (2023). Data Science in Banking Sector: Comprehensive Review of Advanced Learning Methods for Credit Risk Assessment. *International Journal of Computing and Digital Systems*, *14*(1), 1-xx.
32. Amarnadh, V., & Rao, M. N. (2025). A Consensus Blockchain-Based Credit Risk Evaluation and Credit Data Storage Using Novel Deep Learning Approach. *Computational Economics*, 1-34.
33. Shailaja, K., & Anuradha, B. (2017). Improved face recognition using a modified PSO based self-weighted linear collaborative discriminant regression classification. *J. Eng. Appl. Sci*, *12*, 7234-7241.
34. Sekhar, P. R., & Goud, S. (2024). Collaborative Learning Techniques in Python Programming: A Case Study with CSE Students at Anurag University. *Journal of Engineering Education Transformations*, *38*.
35. Sekhar, P. R., & Sujatha, B. (2023). Feature extraction and independent subset generation using genetic algorithm for improved classification. *Int. J. Intell. Syst. Appl. Eng*, *11*, 503-512.
36. Pesaramelli, R. S., & Sujatha, B. (2024, March). Principle correlated feature extraction using differential evolution for improved classification. In *AIP Conference Proceedings* (Vol. 2919, No. 1). AIP Publishing.
37. Tejaswi, S., Sivaprashanth, J., Bala Krishna, G., Sridevi, M., & Rawat, S. S. (2023, December). Smart Dustbin Using IoT. In *International Conference on Advances in Computational Intelligence and Informatics* (pp. 257-265). Singapore: Springer Nature Singapore.
38. Moreb, M., Mohammed, T. A., & Bayat, O. (2020). A novel software engineering approach toward using machine learning for improving the efficiency of health systems. *IEEE Access*, *8*, 23169-23178.
39. Ravi, P., Haritha, D., & Niranjan, P. (2018). A Survey: Computing Iceberg Queries. *International Journal of Engineering & Technology*, *7*(2.7), 791-793.
40. Madar, B., Kumar, G. K., & Ramakrishna, C. (2017). Captcha breaking using segmentation and morphological operations. *International Journal of Computer Applications*, *166*(4), 34-38.
41. Rani, M. S., & Geetavani, B. (2017, May). Design and analysis for improving reliability and accuracy of big-data based peripheral control through IoT. In *2017 International Conference on Trends in Electronics and Informatics (ICEI)* (pp. 749-753). IEEE.
42. Reddy, T., Prasad, T. S. D., Swetha, S., Nirmala, G., & Ram, P. (2018). A study on antiplatelets and anticoagulants utilisation in a tertiary care hospital. *International Journal of Pharmaceutical and Clinical Research*, *10*, 155-161.

43. Prasad, P. S., & Rao, S. K. M. (2017). HIASA: Hybrid improved artificial bee colony and simulated annealing based attack detection algorithm in mobile ad-hoc networks (MANETs). *Bonfring International Journal of Industrial Engineering and Management Science*, *7*(2), 01-12.
44. AC, R., Chowdary Kakarla, P., Simha PJ, V., & Mohan, N. (2022). Implementation of Tiny Machine Learning Models on Arduino 33–BLE for Gesture and Speech Recognition.
45. Subrahmanyam, V., Sagar, M., Balram, G., Ramana, J. V., Tejaswi, S., & Mohammad, H. P. (2024, May). An Efficient Reliable Data Communication For Unmanned Air Vehicles (UAV) Enabled Industry Internet of Things (IIoT). In *2024 3rd International Conference on Artificial Intelligence For Internet of Things (AIIoT)* (pp. 1-4). IEEE.
46. Nagaraj, P., Prasad, A. K., Narsimha, V. B., & Sujatha, B. (2022). Swine flu detection and location using machine learning techniques and GIS. *International Journal of Advanced Computer Science and Applications*, *13*(9).
47. Priyanka, J. H., & Parveen, N. (2024). DeepSkillNER: an automatic screening and ranking of resumes using hybrid deep learning and enhanced spectral clustering approach. *Multimedia Tools and Applications*, *83*(16), 47503-47530.
48. Sathish, S., Thangavel, K., & Boopathi, S. (2010). Performance analysis of DSR, AODV, FSR and ZRP routing protocols in MANET. *MES Journal of Technology and Management*, 57-61.
49. Siva Prasad, B. V. V., Mandapati, S., Kumar Ramasamy, L., Boddu, R., Reddy, P., & Suresh Kumar, B. (2023). Ensemble-based cryptography for soldiers' health monitoring using mobile ad hoc networks. *Automatika: časopis za automatiku, mjerenje, elektroniku, računarstvo i komunikacije*, *64*(3), 658-671.
50. Elechi, P., & Onu, K. E. (2022). Unmanned Aerial Vehicle Cellular Communication Operating in Non-terrestrial Networks. In *Unmanned Aerial Vehicle Cellular Communications* (pp. 225-251). Cham: Springer International Publishing.
51. Prasad, B. V. V. S., Mandapati, S., Haritha, B., & Begum, M. J. (2020, August). Enhanced Security for the authentication of Digital Signature from the key generated by the CSTRNG method. In *2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT)* (pp. 1088-1093). IEEE.
52. Mukiri, R. R., Kumar, B. S., & Prasad, B. V. V. (2019, February). Effective Data Collaborative Strain Using RecTree Algorithm. In *Proceedings of International Conference on Sustainable Computing in Science, Technology and Management (SUSCOM), Amity University Rajasthan, Jaipur-India*.
53. Balaraju, J., Raj, M. G., & Murthy, C. S. (2019). Fuzzy-FMEA risk evaluation approach for LHD machine–A case study. *Journal of Sustainable Mining*, *18*(4), 257-268.
54. Thirumoorthi, P., Deepika, S., & Yadaiah, N. (2014, March). Solar energy based dynamic sag compensator. In *2014 International Conference on Green Computing Communication and Electrical Engineering (ICGCCEE)* (pp. 1-6). IEEE.
55. Vinayasree, P., & Reddy, A. M. (2025). A Reliable and Secure Permissioned Blockchain-Assisted Data Transfer Mechanism in Healthcare-Based Cyber-Physical Systems. *Concurrency and Computation: Practice and Experience*, *37*(3), e8378.
56. Acharjee, P. B., Kumar, M., Krishna, G., Raminenei, K., Ibrahim, R. K., & Alazzam, M. B. (2023, May). Securing International Law Against Cyber Attacks through Blockchain Integration. In *2023 3rd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)* (pp. 2676-2681). IEEE.
57. Ramineni, K., Reddy, L. K. K., Ramana, T. V., & Rajesh, V. (2023, July). Classification of Skin Cancer Using Integrated Methodology. In *International Conference on Data Science and Applications* (pp. 105-118). Singapore: Springer Nature Singapore.
58. LAASSIRI, J., EL HAJJI, S. A. Ï. D., BOUHDADI, M., AOUDE, M. A., JAGADISH, H. P., LOHIT, M. K., ... & KHOLLADI, M. (2010). Specifying Behavioral Concepts by engineering language of RM-ODP. *Journal of Theoretical and Applied Information Technology*, *15*(1).
59. Prasad, D. V. R., & Mohanji, Y. K. V. (2021). FACE RECOGNITION-BASED LECTURE ATTENDANCE SYSTEM: A SURVEY PAPER. *Elementary Education Online*, *20*(4), 1245-1245.
60. Dasu, V. R. P., & Gujjari, B. (2015). Technology-Enhanced Learning Through ICT Tools Using Aakash Tablet. In *Proceedings of the International Conference on Transformations in Engineering Education: ICTIEE 2014* (pp. 203-216). Springer India.
61. Reddy, A. M., Reddy, K. S., Jayaram, M., Venkata Maha Lakshmi, N., Aluvalu, R., Mahesh, T. R., ... & Stalin Alex, D. (2022). An efficient multilevel thresholding scheme for heart image segmentation using a hybrid generalized adversarial network. *Journal of Sensors*, *2022*(1), 4093658.
62. Srinivasa Reddy, K., Suneela, B., Inthiyaz, S., Hasane Ahammad, S., Kumar, G. N. S., & Mallikarjuna Reddy, A. (2019). Texture filtration module under stabilization via random forest optimization

methodology. *International Journal of Advanced Trends in Computer Science and Engineering*, *8*(3), 458-469.

63. Ramakrishna, C., Kumar, G. K., Reddy, A. M., & Ravi, P. (2018). A Survey on various IoT Attacks and its Countermeasures. *International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)*, *5*(4), 143-150.
64. Sirisha, G., & Reddy, A. M. (2018, September). Smart healthcare analysis and therapy for voice disorder using cloud and edge computing. In *2018 4th international conference on applied and theoretical computing and communication technology (iCATccT)* (pp. 103-106). IEEE.
65. Reddy, A. M., Yarlagadda, S., & Akkinen, H. (2021). An extensive analytical approach on human resources using random forest algorithm. *arXiv preprint arXiv:2105.07855*.
66. Kumar, G. N., Bhavanam, S. N., & Midasala, V. (2014). Image Hiding in a Video-based on DWT & LSB Algorithm. In *ICPVS Conference*.
67. Naveen Kumar, G. S., & Reddy, V. S. K. (2022). High performance algorithm for content-based video retrieval using multiple features. In *Intelligent Systems and Sustainable Computing: Proceedings of ICISSC 2021* (pp. 637-646). Singapore: Springer Nature Singapore.
68. Reddy, P. S., Kumar, G. N., Ritish, B., SaiSwetha, C., & Abhilash, K. B. (2013). Intelligent parking space detection system based on image segmentation. *Int J Sci Res Dev*, *1*(6), 1310-1312.
69. Naveen Kumar, G. S., Reddy, V. S. K., & Kumar, S. S. (2018). High-performance video retrieval based on spatio-temporal features. *Microelectronics, Electromagnetics and Telecommunications*, 433-441.
70. Kumar, G. N., & Reddy, M. A. BWT & LSB algorithm based hiding an image into a video. *IJESAT*, 170-174.
71. Lopez, S., Sarada, V., Praveen, R. V. S., Pandey, A., Khuntia, M., & Haralayya, D. B. (2024). Artificial intelligence challenges and role for sustainable education in india: Problems and prospects. *Sandeep Lopez, Vani Sarada, RVS Praveen, Anita Pandey, Monalisa Khuntia, Bhadrappa Haralayya (2024) Artificial Intelligence Challenges and Role for Sustainable Education in India: Problems and Prospects. Library Progress International*, *44*(3), 18261-18271.
72. Yamuna, V., Praveen, R. V. S., Sathya, R., Dhivva, M., Lidiya, R., & Sowmiya, P. (2024, October). Integrating AI for Improved Brain Tumor Detection and Classification. In *2024 4th International Conference on Sustainable Expert Systems (ICSES)* (pp. 1603-1609). IEEE.
73. Kumar, N., Kurkute, S. L., Kalpana, V., Karuppannan, A., Praveen, R. V. S., & Mishra, S. (2024, August). Modelling and Evaluation of Li-ion Battery Performance Based on the Electric Vehicle Tiled Tests using Kalman Filter-GBDT Approach. In *2024 International Conference on Intelligent Algorithms for Computational Intelligence Systems (IACIS)* (pp. 1-6). IEEE.
74. Sharma, S., Vij, S., Praveen, R. V. S., Srinivasan, S., Yadav, D. K., & VS, R. K. (2024, October). Stress Prediction in Higher Education Students Using Psychometric Assessments and AOA-CNN-XGBoost Models. In *2024 4th International Conference on Sustainable Expert Systems (ICSES)* (pp. 1631-1636). IEEE.
75. Anuprathibha, T., Praveen, R. V. S., Sukumar, P., Suganthi, G., & Ravichandran, T. (2024, October). Enhancing Fake Review Detection: A Hierarchical Graph Attention Network Approach Using Text and Ratings. In *2024 Global Conference on Communications and Information Technologies (GCCIT)* (pp. 1-5). IEEE.
76. Shinkar, A. R., Joshi, D., Praveen, R. V. S., Rajesh, Y., & Singh, D. (2024, December). Intelligent solar energy harvesting and management in IoT nodes using deep self-organizing maps. In *2024 International Conference on Emerging Research in Computational Science (ICERCS)* (pp. 1-6). IEEE.
77. Praveen, R. V. S., Hemavathi, U., Sathya, R., Siddiq, A. A., Sanjay, M. G., & Gowdish, S. (2024, October). AI Powered Plant Identification and Plant Disease Classification System. In *2024 4th International Conference on Sustainable Expert Systems (ICSES)* (pp. 1610-1616). IEEE.
78. Dhivya, R., Sagili, S. R., Praveen, R. V. S., VamsiLala, P. N. V., Sangeetha, A., & Suchithra, B. (2024, December). Predictive Modelling of Osteoporosis using Machine Learning Algorithms. In *2024 4th International Conference on Ubiquitous Computing and Intelligent Information Systems (ICUIS)* (pp. 997-1002). IEEE.
79. Kemmannu, P. K., Praveen, R. V. S., Saravanan, B., Amshavalli, M., & Banupriya, V. (2024, December). Enhancing Sustainable Agriculture Through Smart Architecture: An Adaptive Neuro-Fuzzy Inference System with XGBoost Model. In *2024 International Conference on Sustainable Communication Networks and Application (ICSCNA)* (pp. 724-730). IEEE.
80. Praveen, R. V. S. (2024). *Data Engineering for Modern Applications*. Addition Publishing House.