DOI: 10.15849/IJASCA.250330.04

Int. J. Advance Soft Compu. Appl, Vol. 17, No. 1, March 2025
Print ISSN: 2710-1274, Online ISSN: 2074-8523

The Application of Machine Learning and Deep Learning Techniques for Global Energy Utilization Projection for Ecologically Responsible Energy Management

Pranavi Singh[1], Nilima Zade[2], Prashant Priyadarshi[3], Aditya Gupte[4]

[1]Computer Science and Engineering,
Symbiosis Institute of Technology Pune, Symbiosis International
(Deemed University) Pune, India
e-mail: pranavi.singh.btech2022@sitpune.edu.in

[2]Computer Science and Engineering,
Symbiosis Institute of Technology Pune, Symbiosis International
(Deemed University) Pune, India
e-mail: nilima.zade@sitpune.edu.in

[3]National Institute of Technology Patna-800005
Bihar, India
e-mail: prashantp.phd19.cs@nitp.ac.in
orcid id: 0000-0002-8205-1383

[4]Aviation Services Research Council,
Hong Kong Polytechnic University, Hong Kong

Abstract

Accurately estimating future energy consumption is critical as the world seeks alternatives to fossil fuels amidst rising energy demands. The research employs various prediction models for global energy prediction with GDP analysis in energy consumption context. These models include Regression models that are Linear, Polynomial, Bayesian, Tree, Extreme Gradient Boosting, K Nearest Neighbour, Stacked Model, Random Forest (RF), also Long Short-Term Memory (LSTM) and Convolution Neural Networks (CNN) methods. Models are employed to enhance global energy consumption modelling, analysing their adaptability to varying weather and social conditions. A comparative investigation shows that RF performs better than other Regression models. LSTM models perform better than RF in predicting the primary energy consumption per capita and GDP growth, with the lowest MSE value of 0.002 with comparatively higher time and processing complexity. However, RF outperforms in predicting renewable energy share, access to clean cooking fuel, CO2 emission and GDP per capita analysis. The study's novelty lies in its comprehensive evaluation of machine learning and deep learning methods across multiple geographic and temporal energy consumption patterns, emphasizing the superiority of advanced techniques in accurately modelling global energy usage.

Keywords: *Energy Consumption, Machine Learning Algorithm, Renewable Energy, Global Energy Management, Data Pre-Processing.*

1. Introduction

Global growth in the population, progress in the economy, and technology advancements are pushing energy demand to a greater peak. The fuel dependent industry is the one that has the biggest impact on the environment's rising temperatures. Environmental change, in turn, has the potential to overburden services, interfere with fuel distribution itself, and endanger the health of the public. For efficient energy management, progress in predicting energy consumption is essential. Forecasting the world's energy consumption requires the application of Machine Learning (ML) and Deep Learning (DL), both of which have demonstrated potential in many different fields. The primary goal of this investigation is to compare and thoroughly analyse performance of algorithms in order to determine how well they would be able to estimate energy usage, under variable socio-environmental scenarios. Furthermore, a thorough evaluation of the effectiveness of regression techniques is provided.

Predictions need to be refined and adjusted to local needs as sustainable development gains traction. As DL and ML continue to advance, we now have the means to transform energy management in more proactive ways than ever before. In order to finally make the world cleaner, this research aims to capitalize on these developments and, ideally, offer some guidance for the future that could affect new electricity providers and all governmental levels. Projecting demand serves as one of energy administration systems' main applications. The findings show how crucial it is for governments in developing countries to recognize the need to reduce the negative effects of energy use by implementing energy-related discourse strategies and controlling the flow of resources to recipient entities through the onset of globalization. Despite their significance, traditional approaches are unable to fully convey the complexity of energy usage. Moreover, the potential of contemporary technology, like ML, which has demonstrated impressive results in handling complicated data structures, might not ever be fully realized by conventional approaches [1]. In the times where demand for electrical power is rising and ecological issues are real, there is a growing need for precise estimates of energy usage. In order to create a more adaptable, feasible, and ecologically responsible electricity ecosystems; this research tackles the urgent need for a potent predictive energy consumption model.

This study's relevance extends beyond simple counting since it clarifies the complex interplay between the environment, economy, and society in relation to world energy consumption. Most of the research on energy prediction is explored on the localized data for a specific application. It has been observed that global energy prediction with socio economic context still has scope to explore. The main objectives of the research are to analyse and visualize the data and understand the relationship between different energy consumption factors; to conduct an in-depth examination of ML and DL techniques used to forecast global energy consumption patterns; to evaluate the performance of ML and DL models across diverse geographic regions and temporal dependencies, assessing their efficacy in predicting energy usage globally; to compare the effectiveness of ML and DL models in energy consumption forecasting, establishing the best approach for accurate predictions. Though, this all-encompassing strategy will not only address the current energy forecasting issue but also pave the way for a future that is even more robust and clean.

The article is organized in sections as follows. Introduction is explained in Section 1. Section 2 dedicated for literature review. Section 3 discusses about methodology implemented. Results are discussed in Section 4. Research is concluded in Section 5.

2. Literature Review

This in-depth review of the literature will address two aspects of the investigation. Begin by familiarizing with the idea of energy usage forecast and the various approaches available. Next, new methods for understanding patterns of utilization of energy are comprehended. To develop an efficient prediction model, ML and DL methods are examined for precise ways to calculate energy usage globally.

2.1 Forecast for Usage of Energy

In the realm of environmentally friendly energy management, prediction approach is essential since it forms the basis for well-informed decision-making processes, resource optimization, and sustainable development. Researchers in [2] claim that precise energy projections are essential for proactive management, which aids authorities and utilities in cutting down on wastage and better allocating resources. ML and DL techniques have become indispensable tools to tackle this difficult problem. The concept of prediction of utilization of energy, as presented by authors in [3], bases itself on the application of computer techniques and historical data to anticipate potential changes in energy use. Utilization trends were categorized to forecast energy production. Collaborative methodologies for learning were built to estimate hourly consumption of electricity. To evaluate the value of the proposed strategy, a number of training cases were taken into consideration. Researchers in [4] compare two modelling techniques, "All records" as well as "relevant records," to present a fresh methodology. The "all records" approach trains on all of the available data. To mitigate the challenge of creating nonlinear algorithms from past everyday weather trends, the "relevant records" approach focuses on smaller sample size day records. To handle complicated datasets and assess these two modelling strategies, Support Vector Machine (SVM) was used. The “Relevant records” method showed better accuracy in predicting energy consumption than the “All records” approach. However, the “relevant records” method may require additional pre-processing and parameter adjustments to work optimally. Research in [5], proposed a customized EnergyPlus model that serves a frozen goods store which was calibrated using operational data. System developed has an average inaccuracy of 2 kWh for predicting hourly energy use. However, it can be tested with the global applications.

2.2 Developing Energy Use Patterns

Researchers in [6] discuss that temperature has a greater influence on food stores energy use than humidity. When predicting energy use, utilizing multiple regression assessment is a versatile method to take into account. The findings indicate a sharp decline in gas consumption and a slight rise in electricity consumption. In particular, the study [7] considers a few building attributes along with the characteristics of a population, socioeconomics, and individual traits that define households in order to identify streamlined correlations during the evaluation of variables affecting modifications to energy use in homes. In order to display and compare data within two sets —the target set having smart meters when the controlled set is without smart meters—multiple regression modelling is utilized to distinguish across the many forms of typical heating that were found in an area of research. In order to describe energy use per sector, [8]

looked into Germany's energy supplies and needs. To investigate the connection amongst energy use and other important parameters, regression analysis is employed. The variables considered are GDP, population explosion, including growth rate of industry, using data from the energy usage statements. According to the approach, 2022 energy usage is expected to be 209018.1. In [9], three different approaches to Linear Regression (LR) are presented to estimate energy demand. Initially there is a straightforward multiple LR; secondly, there presents an economic modelling perception of the coefficients; and third, there is a double logarithm economical regression. Since it requires time to identify any abnormal consumption behaviour using traditional analysis methods, [10]'s effort will help the business identify its hotspots for wasted energy and reduction in consumption and enable quick action. Data processing will be done for Autoregressive Integrated Moving Average (AIMA) and multi-regression models in order to anticipate the year's quarterly usage. This need for high-quality data is addressed through careful data preparation. Researchers in [11] discuss the optimization of electricity consumption forecasting using LSTM-based models, demonstrating the model's efficacy in real-time prediction with high accuracy. Authors in [12], emphasize the limitations of Extreme Gradient Boosting (XGB) and Multi-Layer Perceptron (MLP) models in energy prediction, contrasting them with the better performance of memory-based architectures like LSTM in handling time series data. Authors in [13] explore energy demand forecasting in a neighbourhood context, finding that Deep Neural Network (DNN) and K Nearest Neighbour (KNN) algorithms offer more accurate predictions for energy balance, which is important for integrating solar panels and electric vehicles. In order to foresee electrical consumption, researchers in [14] suggest a hybrid method that blends adaptive weight updates with conventional neural networks. This allows the methodology to be flexible across various time scales and geographical locations. Additionally, researchers in [15] examine a variety of DL and ML methods for projecting renewable energy. They note that whereas sophisticated techniques comprising RF, SVM, along with XGB are better suited to examine the unpredictable patterns in the data, traditional approaches like LR are useful. The paper highlights how important it is to have strong methods that can deal with problems such as chaotic and inadequate data. The findings in [16] highlight the importance of building more resilient and flexible energy systems, emphasizing the need for a transition from non-renewable to renewable resources to better withstand future crises. The researchers in [17] delve into how data mining techniques, like K-Means clustering and Expectation Maximization (EM), can uncover patterns in global electricity consumption. Researchers in [18] explore how data is affects the performance of the advanced techniques. The study presented in [19] discusses application of RM, SVM, ANN and Regression for making predictions. It shows how ML and DL are revolutionizing the use of sensor data from smart devices in the Internet of Everything. Also, the study in [20] analyses tennis match data based on a supervised learning approach by using Decision Trees (DT) and LR. These serves to exemplify some of the methods that might be most suitable for predicting energy use cases. The work in [21] reveals that GDP per capita is a significant factor influencing life happiness scores. This insight illustrates how economic indicators can intersect with patterns of energy consumption, providing a broader context for understanding these complex relationships. All of the background studied is summarised with a comparison with the proposed work is tabulated in Table 1.

Table.1 Comparative Summary of the earlier studies

Ref No.	Dataset used	Objective	Methodology	Outcome	Scope of improvement

[2]	World Bank's World Developmen t data	Precision and proactive management for energy conservation is discussed.	It uses the Dynamic Seemingly Unrelated Regression (DSUR) method for panel estimation.	Next-11 countries, a 1% increase in financial development leads to a 0.156% rise in energy consumption	Broader evaluation using various models can be tested for improved accuracy in global predictions.
[3]	Building Data Genome	Effective application of historical data for pattern recognition is discussed.	SVR, ANN, and MLR, Ensemble models.	Improved RMSE by 11.9%–21.1% over other methods	Extensive temporal patterns and ML-DL comparisons can be used to test complexity with social and geographic variance.
[4]	Generated from Metronome software	Use of distinct data selection strategies for performance improvement.	SVM-based comparison of “all records” vs. “relevant records” approach.	The "relevant data" approach yielded better prediction accuracy for low-energy buildings, with R^2 of 0.98 and RMSE of 3.4	LSTM and RF can be tested for better handling sequential data.
[5]	From frozen food supermarket in the UK.	Importance of model customization for targeted energy prediction was discussed.	Customized EnergyPlus model for hourly energy predictions.	The EnergyPlus model predicted hourly energy use with a 2 kWh average error, and energy intensity was 1117.3 kWh/m^2 yearly due to high refrigeration loads.	Broader model evaluation; DL models for more generalizable global use rather than single-use applications can be implemented.
[8]	Based on Germany's energy balance sheet and national economic parameters.	Utility of economic indicators for accurate consumption forecasts.	Regression on Germany’s energy needs based on economic indicators like GDP and industry growth.	Predicted Germany’s energy consumption for 2022	The Study was limited to specific problems. It can be applied to global scale, including renewable energy focus.
[9]	Based on historical electricity consumption records of Bogotá's residential sector	Use of linear models for initial predictions and comparison.	Comparison of LR methods (simple, econometric, and double-logarithmic) to estimate energy demand.	The econometric regression model achieved the highest accuracy with a R^2 above 0.9	Incorporation of more advanced models like LSTM and CNN; broader dataset and varied predictive tasks.
[10]	Dubai Police Energy Conservation Department.	Exploration of abnormal consumption behaviour detection.	Business application of ARIMA and multi-regression models	An accuracy of 88% for predicting energy	Can be leveraged DL models for better real-time prediction, moving beyond traditional time series.

[11]	OpenEI data portal	Insight into LSTM's potential for time series forecasting.	LSTM optimization for real-time high-accuracy energy forecasts.	The research achieved 95% accuracy,	Can be tested with various ML-DL approaches
[12]	Private data	Benefits of memory-based architectures in energy data.	LSTM preferred for time series.	MAE of 0.21 for zone D1 and 0.20 for zone D5/2	Employed additional model optimization methods and DL models.
[13]	Houseplants site.	Adoption of DNN for accurate predictions in sector-specific contexts.	DNN and KNN for energy balance, focusing on renewable energy integration.	RMSE of 0.53,	Broader analysis with predictive tasks including renewable energy and CO2.
[14]	Electricity of Mayotte (EDM).	Adaptation of hybrid models to evaluate complex data.	Hybrid neural networks (CNN and LSTM) with adaptive weight updates for flexible predictions.	MAPE of 1.88% for 60-minute, 4.69% for 24-hour, and 5.31% for 7-day predictions.	Can be focused on global energy use beyond localized tasks.
[17]	Obtained from over 20,000 Sangli city consumers.	Clustering insights for data mining in consumption patterns.	Clustering techniques	achieves a 91.13% reduction in RMSE	Broader global application with energy utilization tasks beyond clustering.

After reviewing earlier studies in methods of data mining enabling the analysis of electrical usage data and the improvement of energy efficiency. It has been observed that energy prediction work is explored on the specific local data but global energy prediction applied with GDP analysis in energy consumption context, emphasizing economic variables is unexplored. A thorough investigation was conducted to find trends, correlations, and regulations in the global usage of energy. The research findings demonstrate how sophisticated ML and DL methods are being used more often to handle the challenges associated with energy governance. Finding models that increase precision in forecasting and more effectively incorporate energy from natural sources into daily use are the main areas of focus in the research.

3. Methodology

3.1 Data Processing and Visualizing

The initial data resources and the procedures followed to get the data ready for analysis are covered in this section. The dataset covers important variables and a thorough set of sustainable energy measures covering the years 2000–2020. It was acquired by the academics from Kaggle. The dataset has 3649 number of rows and 21 number of columns. Important variables like carbon dioxide emissions, utilization of energy, monetary investments, expansion of the economy, and availability to electricity—especially from renewable sources—are the focus of the research. Utilized it to compare countries by tracking progress toward sustainable growth and gain insightful knowledge about historical patterns in the world's energy consumption. In the preprocessing phase common to all models, column names are adjusted for better clarity and ease of use, while missing

values are removed from the dataset. The 'Year' column, which is deemed irrelevant for analysis, is dropped. Non-numeric columns, such as 'Country,' are excluded to avoid complications in model development. Pearson correlation coefficients are computed between numerical columns and rounded for simplicity. To ensure that all numerical features are on the same scale, Z-score normalization is applied.

For regression models, after completing the common preprocessing steps, polynomial feature transformation is applied to generate interaction terms between features, enhancing the model's predictive capabilities. The dataset is then split into training and testing subsets for further analysis and evaluation. RF is chosen for its ability to handle large datasets with diverse features effectively. The model combines multiple DTs and averages their outputs, which helps prevent overfitting and results in more accurate predictions during regression tasks. In the case of CNN models, while these networks are typically used for unstructured data, a transition to a fully connected neural network (Dense layers) is recommended for structured datasets. For LSTM models, additional steps are required to prepare the data for time series analysis. After the general preprocessing tasks, the numerical data is scaled and arranged into sequences suitable for LSTM training.

A correlation matrix was then calculated to examine relationships between variables, ensuring the dataset was refined and ready for further analysis Fig. 1a, 1b, 2a, and 2b illustrate the temporal dependencies from 2000 to 2020. Fig. 1a shows that in 2000, many countries in Sub-Saharan Africa and parts of South Asia had very low electricity access, often below 20%. Regions in Latin America, Southeast Asia, and North Africa had moderate access levels, ranging from 40% to 60%.

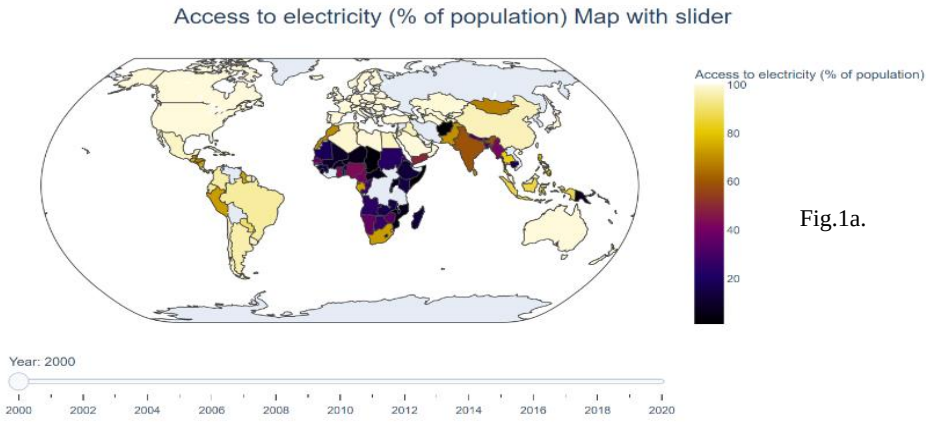

Fig.1a.

Fig. 1a. Access to electricity to the percent population across the globe in the year 2000

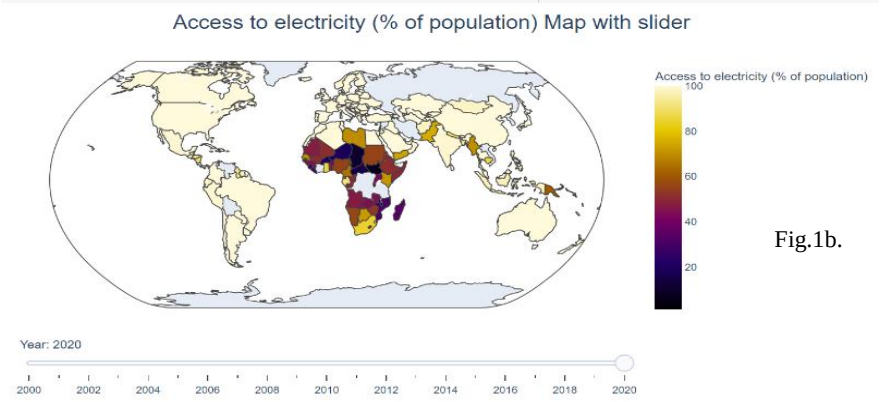

Fig.1b.

Fig. 1b. Access to electricity to the percent population across the globe in the year 2020

In contrast, developed regions such as North America, Europe, and East Asia enjoyed high electricity access, exceeding 80%. By 2020, as illustrated in Fig. 1b, electricity access improved significantly worldwide. Sub-Saharan Africa and South Asia saw increases, with access levels rising to the 40% to 60% range. Many countries in Latin America, North Africa, and Southeast Asia surpassed 80% access, while North America, Europe, and East Asia continued to maintain nearly universal access, approaching 100%.

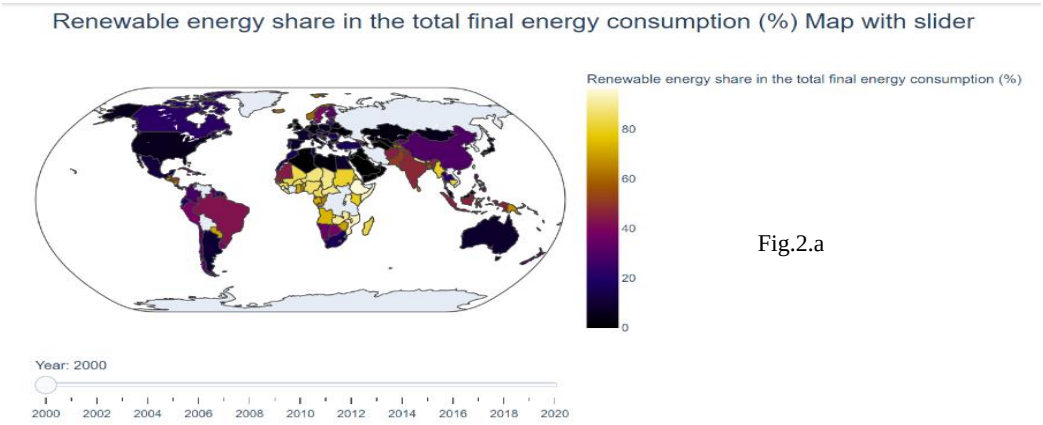

Fig.2.a

Fig. 2a. The percentage of renewable energy within the total amount of energy used in 2000

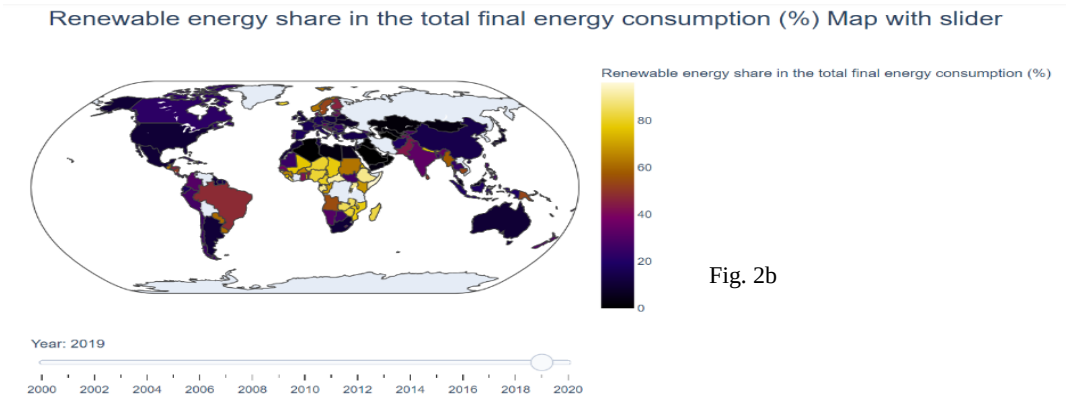

Fig. 2b

Fig. 2b. The percentage of renewable energy within the overall amount of energy consumed in 2019

Fig. 2a indicates that in 2000, the proportion of renewable energy within entire utilization of energy was generally low, particularly in North America, Europe, and parts of Asia, where many countries were below 20% share. Some regions in South America and Africa had a moderate share (20% to 40%), while only a few isolated areas exceeded 60%. Fig. 2b shows that by 2020, the share of renewable energy climbed dramatically, particularly in Europe, where a number of nations had reached the range of 40% to 60%. South America and parts of Africa maintained a moderate to high share, with countries like Brazil leading due to bioenergy and hydroelectric power. Nonetheless, a few areas—most notably those in Asia and North America—kept having a small percentage, frequently around 20%. From these plots it can be observed that there is a positive correlation between increased access to electricity and the rise in the renewable energy share, suggesting that improvements in electricity infrastructure often coincide with greater incorporation of renewables. Regions like Sub-Saharan Africa and South Asia, which had

low electricity access in 2000, showed significant progress by 2020, alongside a greater adoption of renewable energy sources, indicating a shift towards sustainable energy. The data visually confirms that global efforts to improve electricity access and increase renewable energy share are yielding positive results. Significant progress over the past two decades, particularly in developing regions, highlights the impact of sustainable energy investments and policies. Continued focus on these areas is essential for achieving universal electricity access and a higher renewable energy share, supporting climate change mitigation and sustainable development goals. Fig. 3 shows a weak negative correlation, indicating that the share of renewable energy is often lower in nations with higher levels of energy usage intensity. This may be due to reliance on fossil fuels, which are more traditional and established, and the higher upfront costs associated with renewable energy, posing a challenge for countries with lower GDPs. However, there are exceptions. Some high-energy usage-intensity countries are investing in renewables to reduce fossil fuel dependency, while some lower-energy usage-intensity countries continue to rely heavily on fossil fuels.

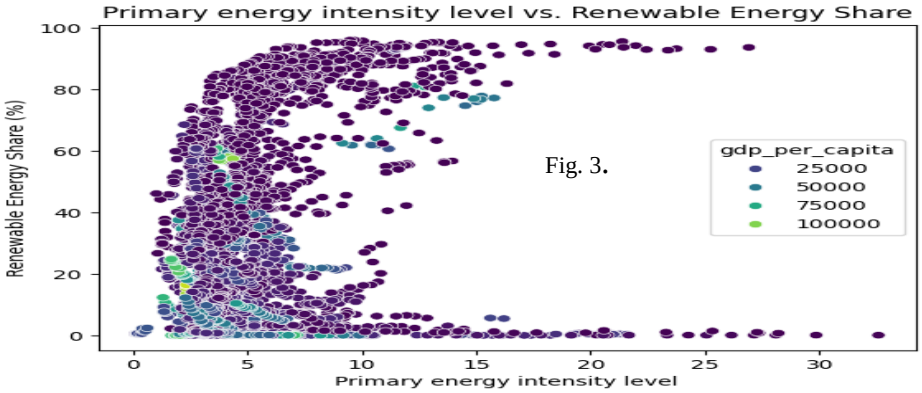

Fig. 3.

Fig. 3. Energy Consumption intensity vs. Fraction of Renewable Energy according to a Scatter Plot

Fig. 4 indicates that there is a positive correlation between the amount of green energy produced and primary energy consumption per capita, assuming that primary energy consumption includes all three types that is fossil, renewable, and nuclear energy consumption. Developed countries, which typically have higher energy consumption, are also more likely to invest in renewables. Additionally, countries with high renewable energy consumption share may use more overall other types of energy to ensure reliable supply. A positive correlation exists between the percentage of low-carbon (renewable, and nuclear) electricity and primary energy consumption per capita. This reflects that countries with greater low-carbon electricity use, tend to have higher energy consumption. A weak positive correlation is observed between electricity access and primary energy consumption per capita, indicating that more developed economies with greater electricity access also have higher energy consumption. A marginally positive correlation exists between GDP growth and primary energy consumption per capita, suggesting that countries with higher GDP growth tend to have greater energy consumption, linking economic development with energy use.

A positive correlation is seen between CO2 emissions per capita and primary energy consumption. This relationship likely arises because fossil fuels, which are major sources of CO2 emissions, dominate global energy consumption. Heat map indicated that the GDP per capita, together with GDP growth, population density and accessibility of electricity are highly positively correlated. This implies that

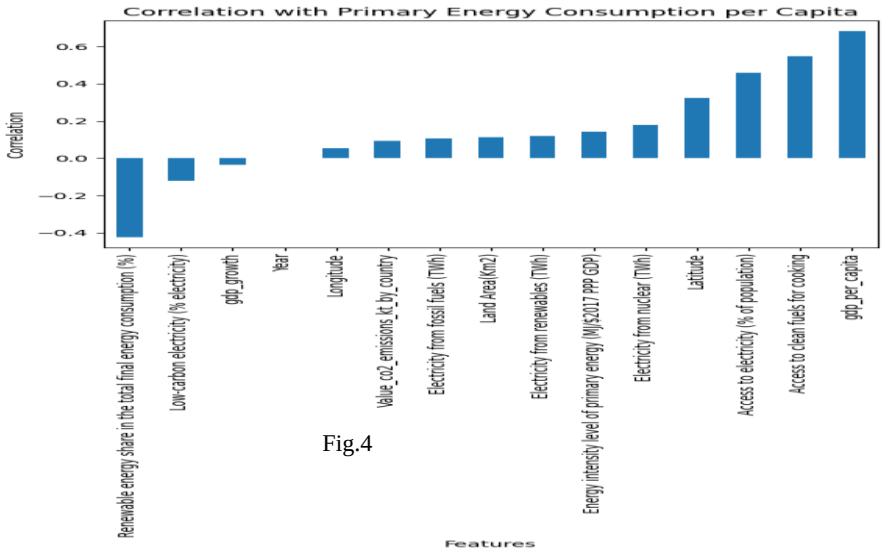

Fig.4

Fig. 4. Correlation with Primary Energy Consumption per Capita (Bar Graph)

countries with higher economic development and population density generally have better electricity access. Utilization of green electricity and CO2 emissions are strongly negatively correlated, meaning that nations that use more renewable energy usually produce lesser CO2 emissions. Furthermore, it appears that more advanced economies utilize energy more efficiently because there is a negative correlation among GDP per capita with energy usage intensity. Weak correlations are observed between some factors, such as renewable energy consumption and access to clean fuels for cooking, indicating a less pronounced relationship between these variables.

3.2 Predictive Model implementation

The dataset was prepared by defining the features (X) as all columns except the target column, 'Primary_Energy_Consumption_per_Capita_kWh', which was set as the target (y). The data was then split, with 80% data for training and remaining 20% for testing. LR, Polynomial Regression, and Bayesian LR, Extra Trees Regressor, DT Regressor, KNN Regressor, and XGB Regressor models were trained on the training data set, and their performance was evaluated using Mean Squared Error (MSE), R-squared (R^2) Score, and Root Mean Squared Error (RMSE) metrics.

Each technique is given a comprehensive description outlining the precise steps that must be taken to put it into practice. The particulars of the dataset plus the objectives of the study will choose what regression model to utilize. These are the justifications for using every single one of the three regression algorithms. LR is a good option if the variables which are independent and the variables that are dependent on them have a linear relation. It offers coefficients that show how each variable that is independent and the variable that depends on it are related to one another. LR is a suitable method for forecasting continuous numerical quantities, such as "Primary energy consumption per capita (kWh/person)" or "Access to electricity (% of population)." Polynomial regression is able to capture more intricate, nonlinear patterns in the interactions between the independent and dependent variables, which are not strictly linear. Variables such as "Renewable-electricity-generating-capacity-per-capita" and "Energy intensity level of primary energy (MJ/$2017 PPP GDP)" may benefit from this. When working with data that shows curves or bends in relationships, polynomial regression is a useful technique

because it allows for more accurate modelling of these higher-order trends. Bayesian LR, on the other hand, incorporates uncertainty into the model, which is particularly helpful when dealing with data that may have measurement errors or when you need to quantify the uncertainty of your predictions. One advantage of Bayesian LR is that it includes regularization by default, which helps prevent overfitting and improves the model's generalization, especially in high-dimensional datasets. This method is also beneficial when dealing with multi-collinearity because it provides posterior distributions for coefficients, making it easier to interpret the effects of correlated variables. The XGB Regression model, which demonstrated the best performance with the lowest Cross-Validation (CV) MSE, was optimized using the hyperparameters such as feature sampling to control the fraction of features considered for each tree split, learning rate to balance step size and overfitting, tree depth to define the maximum depth of each tree, number of boosting rounds to determine the number of boosting rounds, and subsampling to specify the fraction of training data used for each tree. Hyperparameter tuning through Randomized Search CV and Bayesian optimization using Hyperparameter yielded results comparable to a stacking model of all regressors, demonstrating effective optimization and model performance.

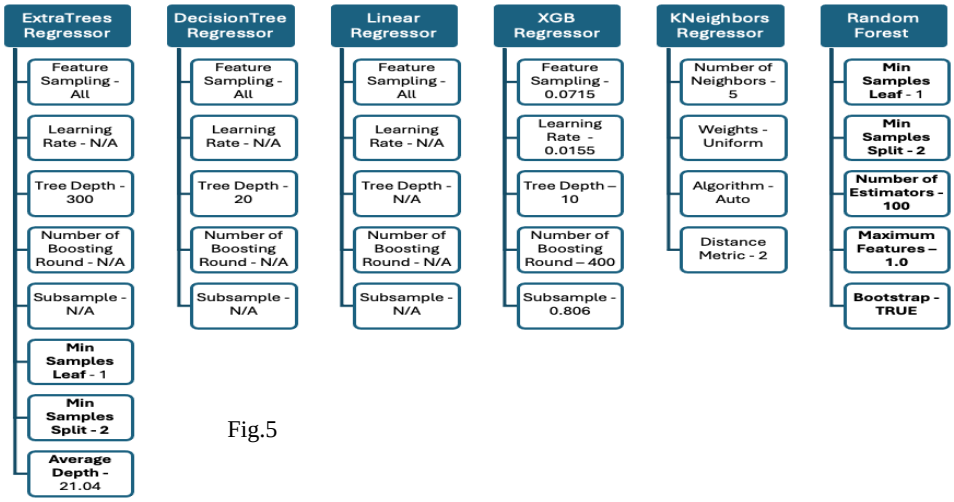

Fig.5

Fig. 5. Implementation details of machine learning models

LR is well-suited for analysing straightforward linear relationships between independent variables, such as energy consumption, and dependent variables like GDP per capita, offering a basic starting point for comparison. Polynomial regression builds on this by modelling non-linear relationships, making it effective for capturing trends, such as those between GDP growth and energy consumption. Logistic regression, also known as binomial regression, is ideal for binary classification tasks, like distinguishing between low and high energy consumption. Finally, the DT Regressor is a non-parametric model that divides data into regions and makes predictions based on the mean value within each region, effectively capturing non-linear relationships without the need for feature scaling. RF Regressor builds multiple DTs and merges their predictions which provides accurate results and robustness against overfitting. Extra Trees Regressor, similar to RF, uses different splitting criteria, offering faster training and potentially better accuracy by reducing both bias and variance. KNN Regressor predicts target values based on the average of the nearest k neighbours, capturing local relationships and performing well with non-linear feature interactions. Lastly, XGB Regressor builds trees sequentially, with each new tree correcting errors from the previous ones, providing high performance and

accuracy through effective optimization of loss functions and handling of large datasets. Fig. 5 shows the Implementation details of machine learning models.

Additionally RF and deep learning models that are CNNs and LSTM have been implemented. The study utilizes CNNs, LSTM networks, and RF models due to their complementary strengths in analysing the dataset. CNNs are chosen for their ability to capture spatial hierarchies and local patterns, as it is particularly beneficial for image data or structured data with inherent spatial relationships. LSTMs are employed to model sequential data and capture long-term dependencies, making them essential for analysis where the order of data points is critical. RF effectively handles high-dimensional data. It provides valuable insights by highlighting which features are most important, making it effective for both classification and regression tasks. Implementation details of CNN and LSTM model are given in Table.2. The performance of both deep learning models was evaluated using MSE and RMSE metrics Average Training Loss and Average Validation loss.

Table.2 Implementation details of the deep learning models

Implementation	**CNN**	**LSTM**
Convolution Layers	1 layer with 64 filters, kernel size of 3	None
Max Pooling Layers	1 layer with a pool size of 2	None
Dropout Layers	2 layers, with dropout rates of 0.25 and 0.5.	None
Flatten Layers	1 layer	None
Dense Layers	3 layers.	1 Dense layer with 1 unit and 1 LSTM layer with 50 units and ReLU activation
Loss Function	Mean Squared Error	Mean Squared Error
Batch Size	10	32
Optimizer	Adam with a learning rate of 0.001	Adam with default parameters
Validation On split %	20% of the training data	20% of the training data
Steps per epoch	Calculated as the length of training data divided by batch size	Calculated as the length of training data divided by batch size

4. Results and Discussion

The analysis involved comparing several regression models to predict primary energy consumption per capita based on a set of features. The results indicated in Table 2 show that both polynomial regression and Bayesian LR generally outperformed simple LR (degree 1). Specifically, polynomial regression models of degrees 2, 3, and 4 showed improvements in R^2 values and reductions in MSE and RMSE compared to the baseline model. Similarly, Bayesian LR models of degrees 1 and 2 exhibited significant enhancements in R^2 values and decreases in MSE and RMSE, suggesting a better fit to the data. However, Bayesian LR with degree 4 displayed a negative R^2 value, indicating an inadequate fit to the dataset. Based on these findings, the Bayesian LR model with degree 2 emerged as the optimal choice. It demonstrated the highest R^2 value (0.966) and the lowest MSE, RMSE with reasonable training time and minimum prediction time among the three models tested, indicating superior predictive performance for the given dataset. Table.3 indicates that simpler approaches like LR, Extra Trees Regressor, DT Regressor,

KNN Regressor or XGB Regressor which rely on a single tree are not found to be effective to the most. In all these regression models, the XGB Regressor performed best, showing the lowest CV MSE. XGB Regressor is better than other tree regressors because it uses gradient boosting, regularization, and efficient computation to achieve higher accuracy and reliability.

Table.3 Comparison of the evaluation parameters of the ML and DL models to predict primary energy consumption per capita based on a set of features

Model	**Degree**	**CV MSE**	**MSE**	**R^2**	**RMSE**	**Training Time (sec.)**	**Prediction Time (sec.)**
Linear Regressor	-	-	0.167	0.809	0.409	0.0138	0.0021
Polynomial Regressor	1	-	0.167	0.809	0.409	0.045113	0.003720
	2	-	0.160	0.817	0.400		
	3	-	0.042	0.952	0.206		
	4	-	0.041	0.953	0.203		
	5	-	0.097	0.889	0.312		
Bayesian Linear Regressor	1	-	0.164	0.813	0.405	0.4420	0.0009
	2	-	0.030	0.966	0.172		
	3	-	0.124	0.859	0.352		
	4	-	6.385	6.291	2.527		
Extra Trees Regressor	-	38.05	17.74	2.07	0.25	17.0265	0.1282
DT Regressor	-	85.31	28.32	2.83	0.2	0.2544	0.0074
KNN Regressor	-	45.5	19.65	2.55	0.16	0.0559	0.2682
XGB Regressor	-	23.98	20.92	2.47	0.11	1.6301	0.1502
Stacking Model	-	24.37	19.2	2.42	0.18	5.7462	0.1264
Bayesian optimization	-	24.09	19.43	2.46	0.13	5.06	0.475
RF	-	0.662	0.0165	0.0704	0.981	1.02	0.0591
CNN	-	0.026	0.0062	-	0.079	13.83	0.2361
LSTM	-	0.004	0.0021	-	0.0539	14.30	0.0999

Hyperparameter tuning with Randomized Search CV and Bayesian optimization using hyper parameter were applied with XGB Regressor, yielding results comparable to a stacking model of all repressors. Best result is seen with XGB Regressor with CV MSE equal to 23.98 and reasonable training and prediction time. Table 3 indicates that RF outperformed in all the ML models. RF models rely on ensemble learning with multiple DTs rather than a single mathematical equation, making their predictions robust and resistant to overfitting. In the RF model, the predictions are computed on how often each class appears across the trees. In the Regression model, predictions are made by taking

the average of the outputs from all the trees. When all the models are tested for prediction tasks, RF model demonstrated improved accuracy by leveraging different DTs, making RF models more effective. Table 3 also indicates that LSTM and CNN have lower CV MSE compared to the RF. However, when compared with higher time and processing complexity, RF can be best suited for predictive tasks 1 to 4 as explained in the next section.

The different predictive tasks to which the RF model and LSTM were tested and the results obtained are given in Table. 4. Predictive Task 1 that is forecasting renewable energy share which forecasts the percentage share of renewable energy in total energy consumption. CV MSE is almost same for both models. However, the results indicates that the RF model explains 75.1% of the variance in renewable energy share, indicating a reasonable fit. If the R^2 score is 0.751, it indicates that 75.1% of the variance in the dependent variable (target variable) is explained by the independent variables in the model, the other part that is 1-75.1=24.9 percentage of the variance is not explained by the model. This reveals a strong relationship between energy access, economic growth, and CO2 emissions. The model for predicting GDP per capita that is Task 1 shows a positive correlation with actual values, though predictions are not perfect, with some overestimations and underestimations. Data points generally align with the diagonal line as shown in Fig. 6, reflecting the RF model's predictive accuracy.

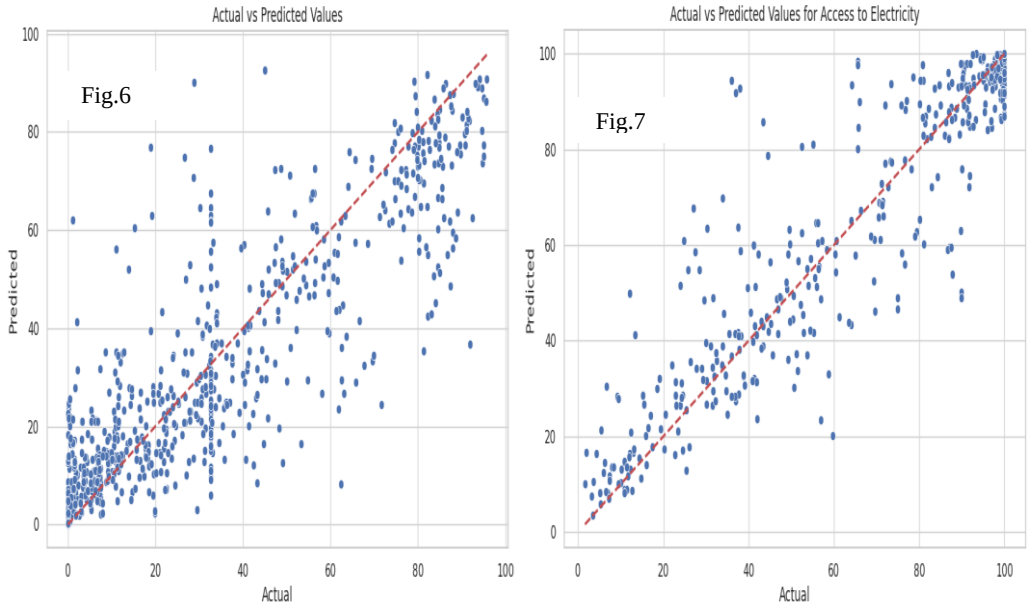

Fig. 6. Actual values Vs predictive values for Task 1
Fig. 7. Actual values Vs predictive values for access to electricity Task 2

Table.4 Performance evaluation of the RF and LSTM model for predictive tasks

RF Model	**Model**	**CV MSE**	**MSE**	**MAE**	**R^2**	**Explained Variance (%)**
Predictive Task 1 (forecasting renewable energy share)	RF	0.0439	**0.**0212	0.1005	0.751	75.1
	LSTM	0.0402	0.0493	0.1742	0.323	32.3
Predictive Task 2 (access to electricity or clean cooking fuels)	RF	0.0331	0.0154	0.0733	0.900	90.0
	LSTM	0.0213	0.0230	0.1147	0.482	48.2

Predictive Task 3 (Estimating CO2 Emissions)	RF	0.0017	0.0003	0.007	0.871	87.1
	LSTM	0.0026	0.0107	0.0511	0.961	96.1
Predictive Task 4(a) GDP Per Capita Analysis	RF	0.0128	0.0023	0.0245	0.892	89.2
	LSTM	0.0419	0.5568	0.5779	0.519	51.9
Predictive Task 4(b) GDP Growth	RF	35.875	4.865	3.6808	-0.007	-0.7
	LSTM	0.0039	0.4687	0.5093	0.112	11.2

From Table 4 it can be observed that for the predictive Task 2 which predicts access to electricity or clean cooking fuels. CV MSE by LSTM is marginally better than RF. However the RF model demonstrate better accuracy, explaining 90% of the variance. There is a positive correlation between actual and predicted percentages of the population with access to electricity. However, data points are scattered around the diagonal line as shown in Fig. 7, indicating that while the model's predictions are generally accurate, there are instances of both underestimation and overestimation.

The predictive Task 3 RF and LSTM showed strong predictive power with an R^2 score of 87.1% and 96.1% and CV MSE 0.0017 and 0.002 respectively. This indicate that overall RF demonstrate better results. But it can be observed that there is no clear linear correlation between actual and predicted CO2 emissions per capita, indicating that the model's predictions are inconsistent. Data points are widely scattered as shown in Fig. 8, suggesting that factors influencing CO2 emissions may not be fully captured by the model. For GDP Per Capita Analysis, done by predictive Task 4(a), the RF performed well compared to LSTM, explaining 89.2% of the variance in GDP per capita with CV MSE 0.012. There is a positive correlation between actual and predicted GDP per capita, indicating that higher actual GDP per capita generally corresponds to higher predicted values. However, as per Fig. 9 the correlation is weak, and data points are scattered around the diagonal line, showing that the model's predictions are not always accurate, with some countries' GDP per capita being either underestimated or overestimated. GDP Growth Analysis i.e. predictive Task 4(b) is GDP and Energy Consumption Relationships where LSTM performs better. The RF performed poorly in predicting GDP growth, suggesting issues with feature selection or model choice. That is a weak positive correlation between actual and predicted GDP growth rates, indicating that while higher actual growth rates generally correspond to higher predicted rates, the relationship is not strong. This explains -0.7% of the variance in GDP Growth. Data points are scattered around the diagonal line as shown in Fig. 10, showing that predictions are often inaccurate, with some countries' GDP growth rates being either underestimated or overestimated.

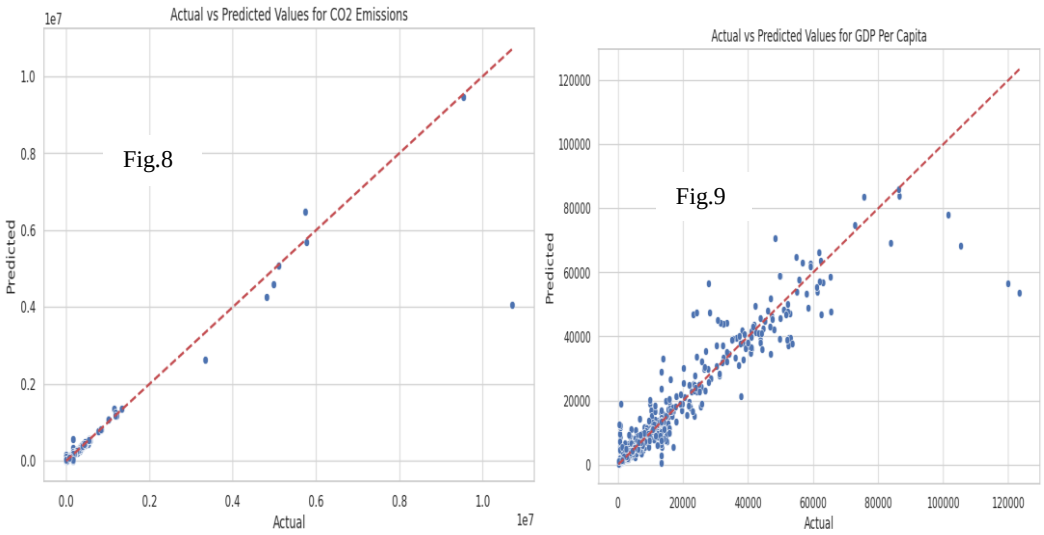

Fig. 8. Actual values Vs predictive values for CO2 emissions per capita Task 3
Fig. 9. Actual values Vs predictive values for GDP per capita Task 4(a)

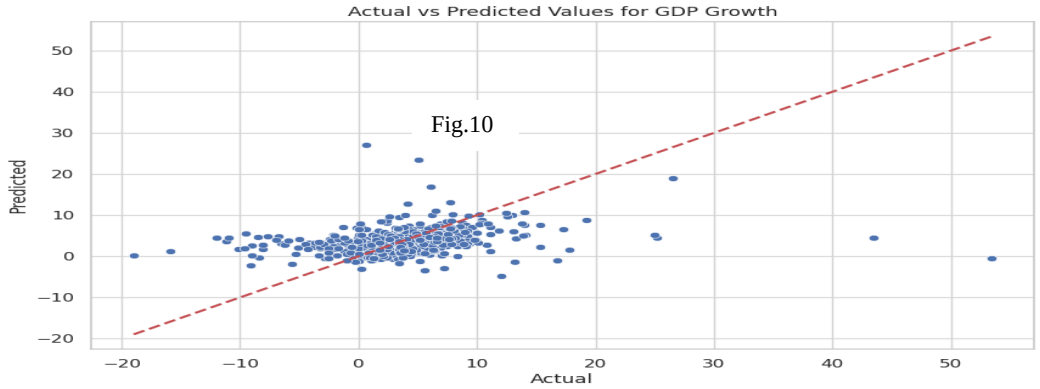

Fig. 10. Actual values Vs predictive values for GDP growth Task 4(b)

5. Conclusion

Several predictive models for forecasting primary energy consumption per capita based on given features have been evaluated. All the regression models have been tested, to enhance the performance various tree regressors were implemented. Among all regressors, XGB Regressor demonstrated CV MSE of 23.89. To optimize the performance, the RF model was implemented resulting in a CV MSE of 0.662 which improved the performance by 97.23% compared to the XGB Regressor. The LSTM model, when compared with ML models, achieved the lowest MSE value at 0.0021, surpassing the RF model improving the performance by 87.27% but with increased time complexity. When comparing the performance among DL models, the LSTM model's MSE is lower than that of the CNN model. The performance of DL models especially the LSTM model shows the least prediction error in predicting the primary energy consumption per capita and GDP growth at the cost of a little higher training time and reasonable prediction time. However RF outperforms in predicting renewable energy share, access to clean cooking fuel, CO2 emission and GDP per capita analysis. Based on the performance metrics, the RF and LSTM models emerges as the optimal choice among the models tested. These models demonstrates its strengths in handling temporal

dependencies and sequential data patterns effectively. Applications of this research include helping governments shape energy policies, aiding energy companies in resource management, and contributing to climate change efforts. Future research directions could involve integrating real-time data for up-to-date forecasts, collaborating with experts from various fields, and improving the interpretability of ML and DL models. Special attention could be given to renewable energy and its impact on economic growth and sustainability.

References

[1] N. Zade, K. Gupta, S. Mutha, O. Mengshetti, G. Joshi and R. K. Iyer, "Technical Analysis of Stock Market Trends using LSTM for Price Prognosis," 2023 7th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT), Ankara, Turkiye, 2023, pp. 1-5, doi: 10.1109/ISMSIT58785.2023.10304934.

[2] Danish, Saud, S., Baloch, M. A., & Lodhi, R. N. "The nexus between energy consumption and financial development: estimating the role of globalization in Next-11 countries". Environmental Science and Pollution Research, 25, 18651-18661. 2018. https://doi.org/10.1007/s11356-018-2069-0

[3] Dong, Z., Liu, J., Liu, B., Li, K., & Li, X. "Hourly energy consumption prediction of an office building based on ensemble learning and energy consumption pattern classification". Energy and Buildings, 241, 110929. 2021. https://doi.org/10.1016/j.enbuild.2021.110929

[4] Paudel, S., Elmitri, M., Couturier, S., Nguyen, P. H., Kamphuis, R., Lacarrière, B., & Le Corre, O. "A relevant data selection method for energy consumption prediction of low energy building based on support vector machine". Energy and Buildings, 138, 240-256. 2017. https://doi.org/10.1016/j.enbuild.2016.11.009

[5] Mylona, Zoi, Maria Kolokotroni, and Savvas A. Tassou. "Frozen food retail: Measuring and modelling energy use and space environmental systems in an operational supermarket." Energy and Buildings 144 (2017): 129-143. 2017. https://doi.org/10.1016/j.enbuild.2017.03.049

[6] Braun, M. R., H. Altan, and S. B. M. Beck. "Using regression analysis to predict the future energy consumption of a supermarket in the UK." Applied Energy 130 (2014): 305-313. 2014. https://doi.org/10.1016/j.apenergy.2014.05.062

[7] Laicane, Ilze, Dagnija Blumberga, Andra Blumberga, and Marika Rosa. "Comparative multiple regression analysis of household electricity use in Latvia: using smart meter data to examine the effect of different household characteristics." Energy Procedia 72 (2015): 49-56. 2015. https://doi.org/10.1016/j.egypro.2015.06.008

[8] Tabasi, Sanaz, Alireza Aslani, and Habib Forotan. "Prediction of energy consumption by using regression model." Computational Research Progress in Applied Science & Engineering 2, no. 3 (2016): 110-115.

[9] Peña-Guzmán, Carlos, and Juliana Rey. "Forecasting residential electric power consumption for Bogotá Colombia using regression models." Energy Reports 6 (2020): 561-566. https://doi.org/10.1016/j.egyr.2019.09.026

[10] Alqasim, Amal Rashid. "Using Regression Analysis for Predicting Energy Consumption in Dubai Police." (2022). https://scholarworks.rit.edu/theses

[11] Qureshi, Momina, Masood Ahmad Arbab, and Sadaqat ur Rehman. "Deep learning-based forecasting of electricity consumption." Scientific Reports 14, no. 1 (2024): 6489. https://doi.org/10.1038/s41598-024-56602-4

[12] Morcillo-Jimenez, Roberto, Jesús Mesa, Juan Gómez-Romero, M. Amparo Vila, and Maria J. Martin-Bautista. "Deep learning for prediction of energy consumption: an applied use case in an office building." Applied Intelligence 54, no. 7 (2024): 5813-5825. https://doi.org/10.1007/s10489-024-05451-9

[13] Mirjalili, Mohammad Amin, Alireza Aslani, Rahim Zahedi, and Mohammad Soleimani. "A comparative study of machine learning and deep learning methods for energy balance prediction in a hybrid building-renewable energy system." Sustainable Energy Research 10, no. 1 (2023): 8. https://doi.org/10.1186/s40807-023-00078-9

[14] Taleb, Ihab, Guillaume Guerard, Frédéric Fauberteau, and Nga Nguyen. "A flexible deep learning method for energy forecasting." Energies 15, no. 11 (2022): 3926. https://doi.org/10.3390/en15113926

[15] Benti, Natei Ermias, Mesfin Diro Chaka, and Addisu Gezahegn Semie. "Forecasting renewable energy generation with machine learning and deep learning: Current advances and future prospects." Sustainability 15, no. 9 (2023): 7087. https://doi.org/10.3390/su15097087

[16] Momina Shaheen, Muhammad Junaid Anjum, Faizan Ahmad, Aimen Anum, "Computational Data Analysis on Global Energy and COVID-19 Pandemic", International Journal of Information Engineering and Electronic Business (IJIEEB), Vol.15, No.6, pp. 1-17, 2023. DOI:10.5815/ijieeb.2023.06.01

[17] Rahman, A., Abahussin, H.K., Alghamdi, M.H., Alkhwaja, A.A., Alfawaz, F.A., Alkhwaja, I.A., Albugami, M.N., Youldash, M., Iqbal, T., Bakry, A., Al-Musallam, H.A. "Predicting global energy consumption through data mining techniques". International Journal of Design & Nature and Ecodynamics, Vol. 19, No. 2, pp. 397-406. (2024). https://doi.org/10.18280/ijdne.190205

[18] Masoud M. Z., Manasrah A., Jaradat Y. and N. A. A. Shaban, "How Data can Mislead Machine Learning Prediction Process: Case Study of Building Cooling and Heating Loads," 2023 International Conference on Information Technology (ICIT), Amman, Jordan, 2023, pp. 709-714, doi: 10.1109/ICIT58056.2023.10225944.

[19] Masoud, M., Jaradat, Y., Manasrah, A., & Jannoud, I. (2018). "Sensors of Smart Devices in the Internet of Everything (IoE) Era: Big Opportunities and Massive Doubts". Journal of Sensors, 2019(1), 6514520. https://doi.org/10.1155/2019/6514520

[20] Moaiad Ahmad Khder, Samah Wael Fujo, "Applying Machine Learning-Supervised Learning Techniques for Tennis Players Dataset Analysis", International Journal of Advances in Soft Computing and its Application, 14, 3(2022), 189-214. doi: 10.15849/IJASCA.221128.13

[21] Moaiad Ahmad Khder, Mohammad Adnan Sayfi, and Samah Wael Fujo, "Analysis of World Happiness Report Dataset Using Machine Learning aproaches", International Journal of Advances in Soft Computing and its Application, 14, 1(2022), 14-34.