1. INTRODUCTION

The modern energy landscape is marked by the intricate interplay of various factors, including dynamic environmental conditions, temporal dependencies, and complex interactions among diverse parameters. Within this context, accurate load forecasting emerges as a critical component of efficient energy management, guiding the prediction of future energy consumption through the amalgamation of historical data and advanced modeling techniques. The precision of load forecasting is key to optimizing resource allocation, ensuring resilient operational planning, and fostering the development of adaptive, sustainable energy infrastructures. This study delves into the realm of microgrid systems, aiming to enhance load prediction accuracy using machine learning algorithms such as K-Nearest Neighbors (KNN), Auto-Regressive Integrated Moving Average (ARIMA), Support Vector Machines (SVM), Artificial Neural Networks (ANN), and Long Short-Term Memory (LSTM). By meticulously navigating the complexities of data preprocessing, feature extraction, and hyperparameter optimization, the research provides valuable insights into the comparative strengths and weaknesses of each algorithm. As the energy landscape evolves, this study serves as a strategic guide, empowering microgrid operators and stakeholders in navigating dynamic environments and fostering resilience within the ever-changing energy sector.

1.1. Motivation

Load forecasting serves as a crucial element in the modern energy landscape, driven by the motivation to optimize resource allocation, enhance operational efficiency, and ensure grid stability. Accurate predictions enable utilities to allocate resources effectively, reducing waste and operational costs. Proactive planning, facilitated by load forecasting, allows for resilient management of supply-demand imbalances, preventing disruptions in energy supply. The integration of renewable energy sources and the promotion of sustainable practices further underline the significance of load forecasting. By encouraging demand-side management and guiding grid planning, load forecasting becomes a strategic tool for utilities, fostering a resilient, efficient, and sustainable energy ecosystem.

2. PROBLEM STATEMENT

Accurate load forecasting is a critical challenge in the evolving energy landscape. Conventional methods struggle to predict energy demand accurately, leading to inefficiencies and grid instability. Fluctuating factors like weather and the growing integration of renewable sources further complicate the forecasting process. This becomes particularly crucial with the rise of microgrids, where precise predictions are essential for managing distributed resources and ensuring reliable power supply. The project aims to address these challenges by leveraging advanced machine learning algorithms to improve the accuracy and adaptability of load forecasting, contributing to the efficiency and resilience of energy systems.

3. LITERATURE REVIEW

Table No 1: Literature review of load forecasting

AUTHOR	TITLE	REVIEW
[1] B. Farsi, M. Amayri, N. Bouguila and U. Eicker	On Short-Term Load Forecasting Using Machine Learning Techniques and a Novel Parallel Deep LSTM-CNN Approach.	This literature review explores the significance of accurate load forecasting in enhancing energy management and scheduling within industrial infrastructures. It presents a novel hybrid deep learning model, PLCNet, which combines LSTM and CNN architectures, demonstrating superior performance in short-term load forecasting compared to other traditional machine learning models, with notable improvements in accuracy, reaching up to 98.23% for Malaysian data and 91.18% for German data.
[2] M. M. Asiri, G. Aldehim, F. A. Alotaibi, M. M. Alnfiai, M. Assiri and A. Mahmud	Short-Term Load Forecasting in Smart Grids Using Hybrid Deep Learning.	This study proposes a Short-Load Forecasting scheme utilizing a Hybrid Deep Learning and Beluga Whale Optimization approach, aiming to enhance load prediction accuracy in Smart Grid environments. By employing convolutional bidirectional long short-term memory with autoencoder models and optimizing hyperparameters through the BWO algorithm, the proposed LFS-HDLBWO method demonstrates superior predictive performance compared to existing deep learning algorithms, as evidenced by notably reduced error rates in experimental evaluations.
[3] S. H. Rafi, Nahid-Al-Masood, S. R.	A Short-Term Load Forecasting Method Using	This article addresses the challenge of short-term load forecasting at the consumer level in microgrid energy

Deeba and E. Hossain	Integrated CNN and LSTM Network	distribution, proposing a robust model that combines random forest, support vector regressor, and long short-term memory techniques to handle the volatility and uncertainty in energy consumption. By dynamically assigning weights to each predictor based on forecasting efficacy, the proposed model achieves significant reductions in forecasting errors compared to existing models, demonstrating its suitability for microgrid energy management amidst highly inconsistent load patterns.
[4] G. Tziolis, A. Livera, J. Montes-Romero, S. Theocharides, G. Makrides, and G. E. Georghiou	Direct Short-Term Net Load Forecasting Based on Machine Learning Principles for Solar-Integrated Microgrids	This study addresses the vital need for accurate net load forecasting in solar-integrated microgrids to ensure efficient planning and integration of variable solar photovoltaic systems into modern power systems. Leveraging machine learning principles, the proposed methodology showcases promising results in achieving precise short-term net load forecasting, offering valuable insights for microgrid decision-making by utilities and operators.
[5] S. -V. Oprea and A.Bâra	Machine Learning Algorithms for Short-term Load Forecast in Residential Buildings Using Smart Meters, Sensors, and Big Data Solutions	This paper introduces a scalable Big Data framework coupled with machine learning algorithms for short-term load forecasting in residential buildings, addressing the growing importance of accurate electricity consumption prediction. By comparing the performance of various machine learning algorithms including feed-forward artificial neural networks, non-linear autoregressive models, and ensemble methods, it highlights the need for robust forecasting techniques to optimize energy management in smart buildings.

4. OVERVIEW OF TECHNOLOGY

4.1 INTRODUCTION

With the advancements of machine learning and deep learning, the pursuit of knowledge and expertise is an ongoing journey marked by curiosity, exploration, and discovery. Immersion in this domain unveils a dynamic landscape where state-of-the-art algorithms intersect with real-world applications, shaping the trajectory of technological advancement. The exploration of advanced machine learning and deep learning methodologies traverses a diverse terrain of techniques and models, from convolutional neural networks to recurrent neural networks, from reinforcement learning to generative adversarial networks. Each avenue of inquiry offers profound insights into the capabilities and potential of artificial intelligence.

This aims to capture the essence of this journey, providing a retrospective glimpse into the projects pursued, the methodologies embraced, and the outcomes realized. By dissecting the challenges encountered and the strategies devised, a comprehensive panorama of the practical applications of advanced machine learning and deep learning emerges across various domains. Moreover, beyond the technical intricacies lies a broader canvas upon which the impact of these technologies unfolds. From healthcare to finance, from autonomous systems to natural language processing, the transformative influence of advanced machine learning and deep learning reverberates across industries, reshaping paradigms and redefining possibilities. As we navigate this odyssey of exploration and discovery, it becomes apparent that advanced machine learning and deep learning are not mere tools but gateways to a future where intelligence is harnessed to unlock new frontiers of innovation and empowerment.

4.2 What Is Data Science

Data science is the field that deals with extracting knowledge and insights from structured and unstructured data. It involves various techniques such as data mining, data visualization, statistical analysis, and machine learning. Data scientists work with large volumes of data to uncover hidden patterns, trends, and correlations that can be used to make informed decisions and predictions. They often use programming languages like Python or R and tools like SQL and TensorFlow to manipulate and analyze data.

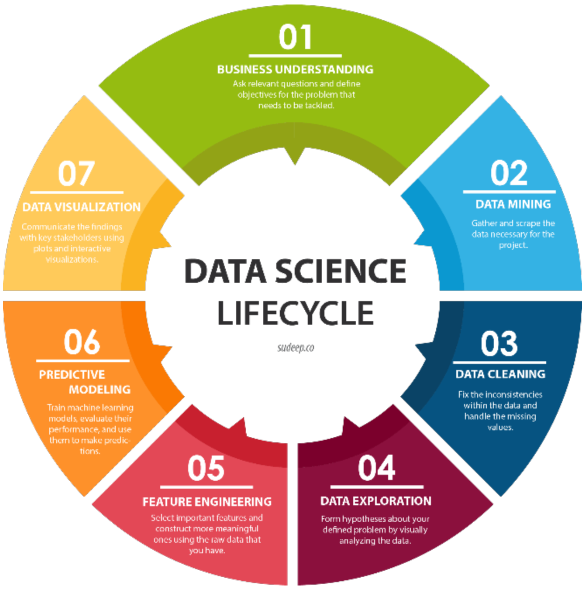

Figure 1: Data science lifecycle

4.3 What Is Artificial Intelligence

Artificial intelligence refers to the development of computer systems that can perform tasks that typically require human intelligence. These tasks include understanding natural language, recognizing objects in images, making decisions, and solving problems. AI systems can be classified into two types: narrow AI, which is designed for specific tasks like speech recognition or playing chess, and general AI, which aims to possess human-like intelligence across a wide range of tasks.

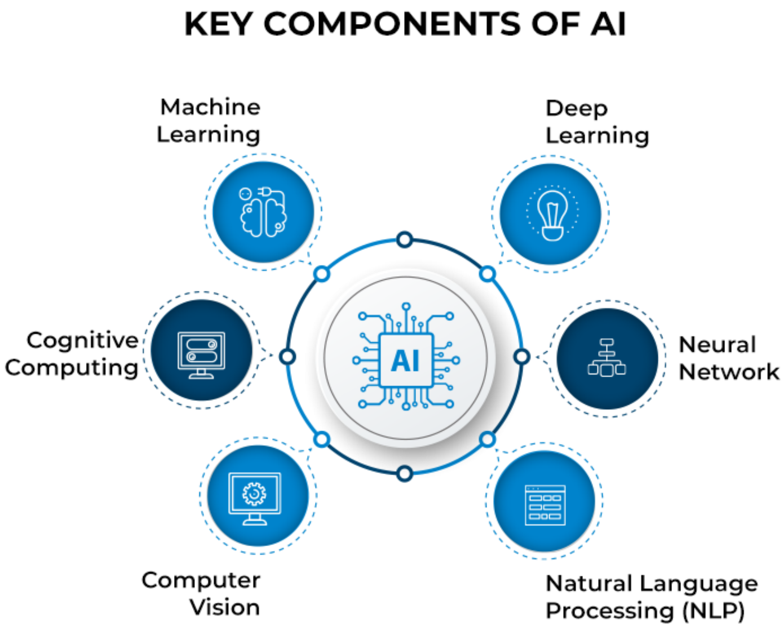

Figure 2: Components of Artificial Intellegence

4.4 What Is Machine Learning

Machine learning is a subset of artificial intelligence that focuses on algorithms and models that allow computers to learn from data and improve their performance over time without being explicitly programmed. In machine learning, algorithms are trained on a dataset to identify patterns and relationships, then used to make predictions or decisions on new data. Common machine-learning techniques include supervised learning, unsupervised learning, and reinforcement learning.

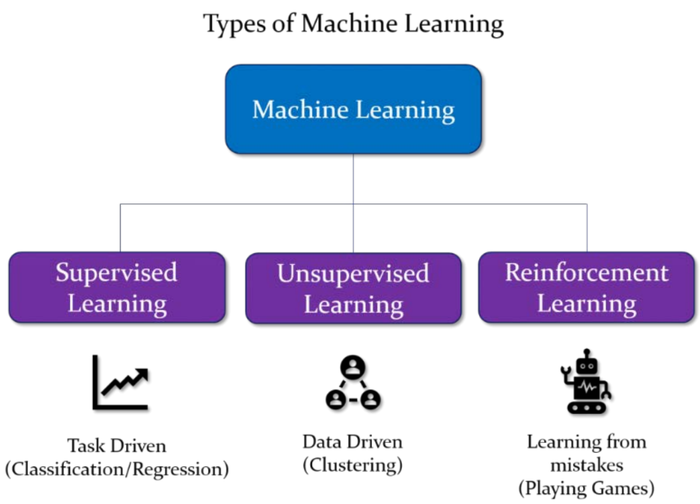

Figure 3: Types in Machine Learning

4.5 What Is Deep Learning

Deep learning is a subset of machine learning that utilizes artificial neural networks with many layers. These neural networks are inspired by the structure and function of the human brain and are capable of learning increasingly complex representations of data. Deep learning has achieved remarkable success in various tasks such as image and speech recognition, natural language processing, and autonomous driving. It requires large amounts of labeled data for training and significant computational resources for training deep neural networks. Popular deep learning frameworks include TensorFlow, PyTorch, and Keras.

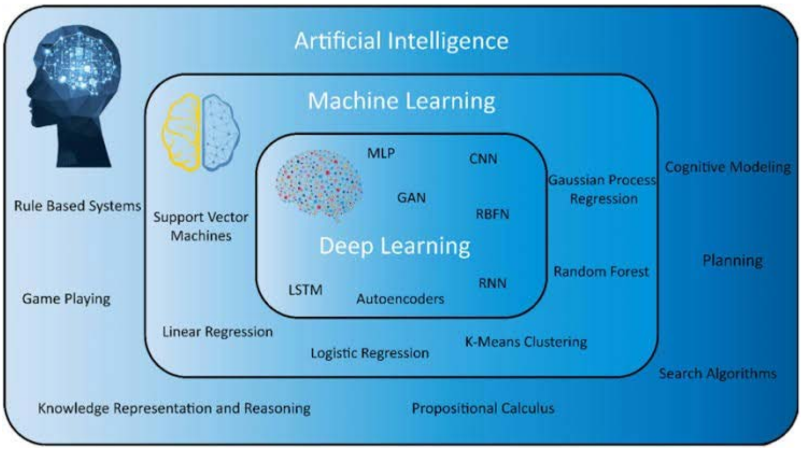

Figure 4: Types in Deep Learning

4.6 History Of Machine Learning And Deel Learning

The history of Machine Learning (ML) and Deep Learning (DL) spans several decades, witnessing transformative developments that have shaped their evolution into the powerful fields they are today. In the early stages, dating back to the 1950s and 1960s, ML laid its foundation with the introduction of early neural networks and decision tree algorithms, including the Perceptron by Frank Rosenblatt in 1957. However, the field faced limitations in computational power and algorithmic sophistication. Throughout the 1970s and 1980s, symbolic AI dominated ML research, focusing on expert systems and symbolic reasoning approaches. It wasn't until the 1990s that statistical learning methods gained prominence, with techniques such as Support Vector Machines (SVM) and Decision Trees becoming popular for classification and regression tasks. This resurgence marked the rise of deep learning, a subfield of ML focused on neural networks with many layers. Breakthroughs in training algorithms and architectures, such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs), propelled deep learning into prominence across various domains. Today, deep learning continues to advance rapidly, fueled by innovations in architectures, training techniques, and hardware, driving progress in research and applications across industries.

4.7 Deep Learning Trends

In recent years, deep learning has experienced a multitude of trends, reflecting both the rapid pace of innovation and the evolving landscape of artificial intelligence. One prominent trend is the rise of attention mechanisms and transformer architectures, which have revolutionized natural language processing tasks by allowing models to focus on relevant information within input sequences. Concurrently, self-supervised learning has emerged as a powerful approach for pretraining deep learning models, leveraging unlabeled data to learn meaningful representations. Continual learning and lifelong learning have also gained traction, addressing the challenge of adapting models to evolving datasets over time and enabling continuous learning without forgetting previously acquired knowledge. Adversarial robustness and security have become pressing concerns, driving research efforts to develop models resilient to adversarial attacks and enhance the trustworthiness of AI systems. Additionally, there is a growing emphasis on efficient model architectures and training techniques to enable deployment on resource-constrained devices. Multimodal learning, which integrates 4 information from multiple modalities, holds promise for improving model performance and enabling more natural interactions. Finally, ethical AI and responsible AI practices are increasingly prioritized, emphasizing fairness, transparency, accountability, and privacy in the development and deployment of AI technologies. These trends collectively shape the trajectory of deep learning, driving innovation and fostering responsible use of artificial intelligence for the benefit of society.

4.8 Future Prospects And Challenges In Deep Learning

On the positive side, deep learning could benefit a lot from new technologies like quantum and neuromorphic computing. These could make deep learning models even smarter and more efficient. Plus, deep learning is already making big waves in areas like healthcare, finance, and climate science. It's helping doctors diagnose diseases, make financial decisions smarter, and even predict extreme weather events. But there are hurdles to overcome. One big challenge is making deep learning models more reliable and easier to understand. Right now, they can be a bit like black boxes – we don't always know why they make the decisions they do. We also need to make sure these models are fair and unbiased, treating everyone equally. Then there are ethical concerns, like privacy and data security. We need rules and guidelines to make sure deep learning is used responsibly and ethically. And we can't forget about the environmental impact. Training these models takes a lot of energy, so finding ways to make them more efficient is crucial.

5. OPTIMIZERS

Optimizers are algorithms used to minimize the error or loss function during the training process of neural networks. The goal of optimization is to adjust the parameters (weights and biases) of the neural network in such a way that the model's predictions become closer to the actual targets. Various optimizers are available, each with its approach to updating the model parameters. Some common optimizers include:

Stochastic Gradient Descent (SGD): This is a basic optimizer that updates the parameters based on the gradient of the loss function concerning each parameter.

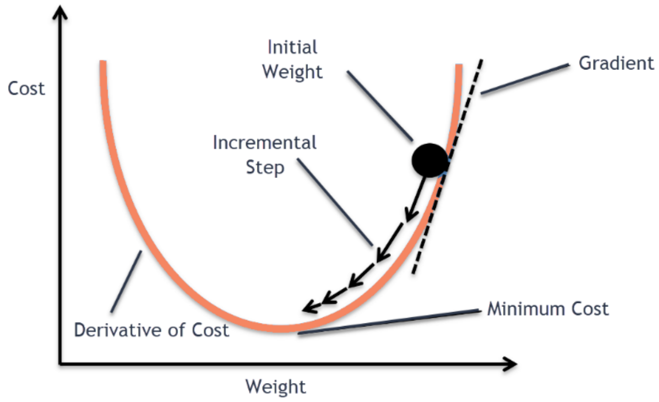

Figure 5: SGD working

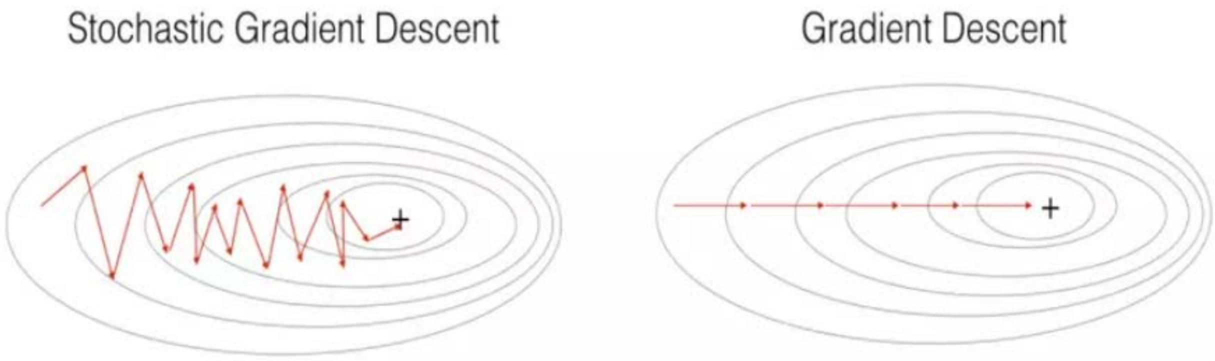

Figure 6: SGD vs GD working

Adam: Adam is an adaptive learning rate optimization algorithm that combines the advantages of both AdaGrad and RMSProp. It adapts the learning rates for each parameter based on past gradients and squared gradients.

RMSProp: RMSProp (Root Mean Square Propagation) is an adaptive learning rate method that scales the learning rate by dividing it by the exponentially decaying average of squared gradients.

Adagrad: Adagrad (Adaptive Gradient Algorithm) adapts the learning rate for each parameter based on the historical gradients.

Adadelta: Adadelta is a variant of Adagrad that improves its performance by replacing the learning rate with an exponentially decaying average of squared gradients.

Adamax: Adamax is a variant of Adam that is more robust to large gradients and sparse gradients

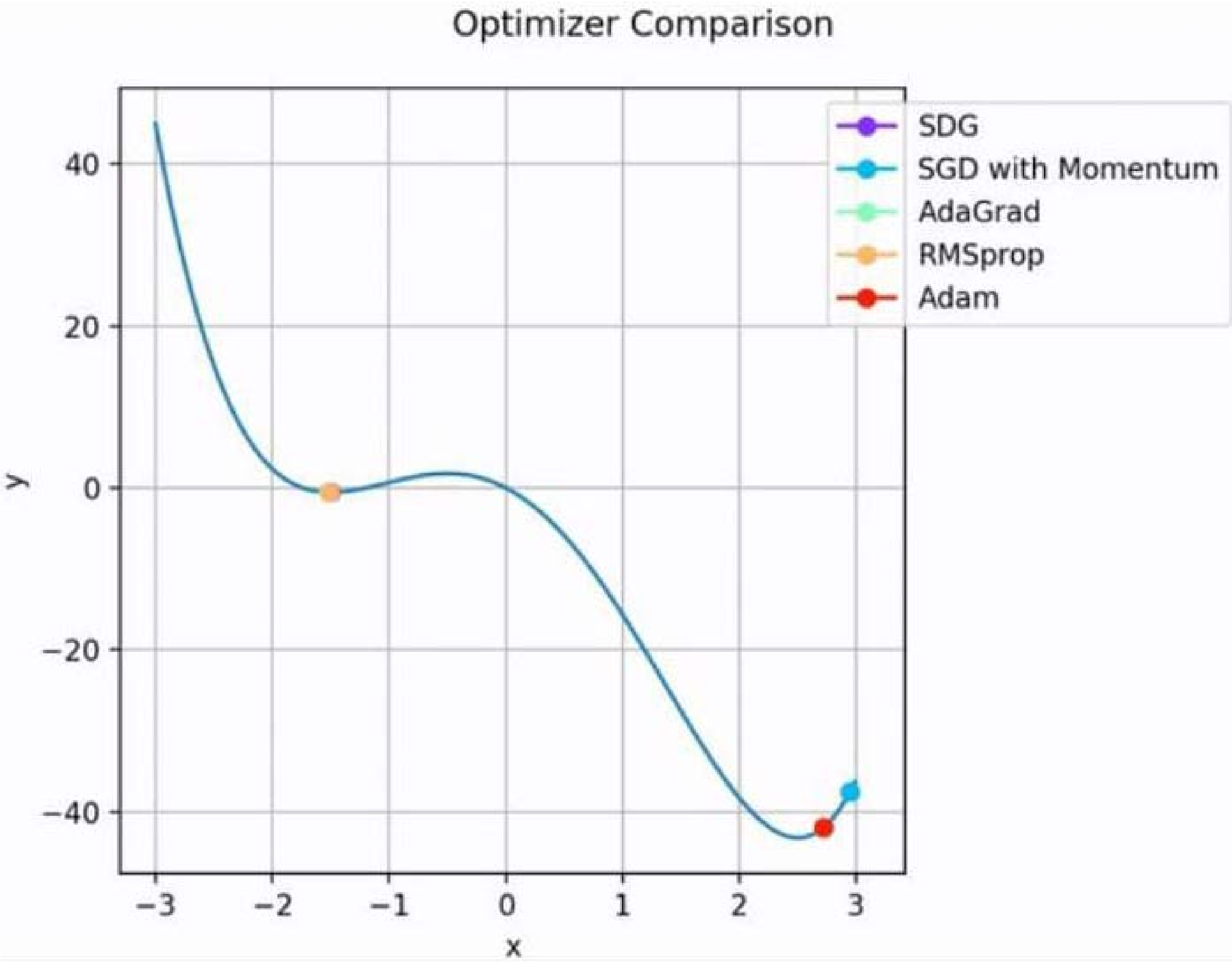

Figure 7: Comparison between optimizers

6. ACTIVATION FUNCTIONS

Activation functions introduce non-linearity into the neural network, allowing it to learn complex patterns in the data. They determine the output of a neuron given its input. Common activation functions include:

Sigmoid: The sigmoid function squashes the input values between 0 and 1, which is useful for binary classification problems.

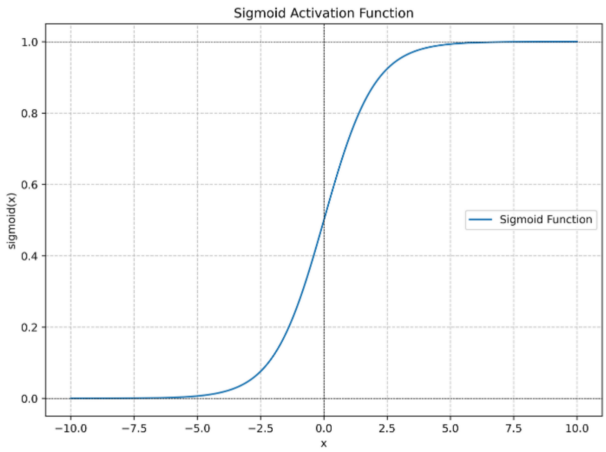

Figure 8: Sigmoid activation function

Tanh (Hyperbolic Tangent): Tanh squashes the input values between -1 and 1, making it suitablefor hidden layers of neural networks.

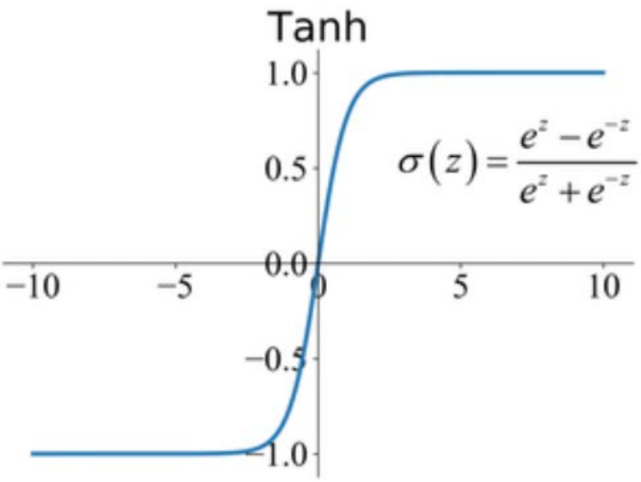

Figure 9: Tanh activation function

ReLU (Rectified Linear Unit): ReLU sets all negative input values to zero and leaves positive values unchanged. It is one of the most widely used activation functions due to its simplicity and effectiveness.

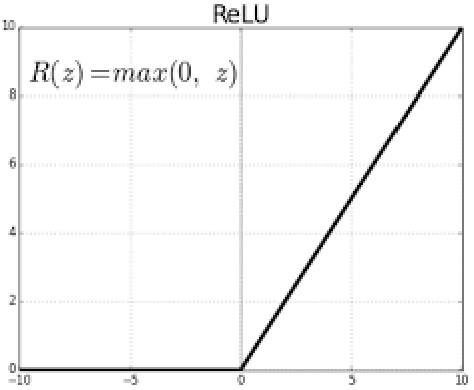

Figure 10: ReLU activation function

Leaky ReLU: Leaky ReLU is a variant of ReLU that allows a small gradient for negative input values to address the "dying ReLU" problem.

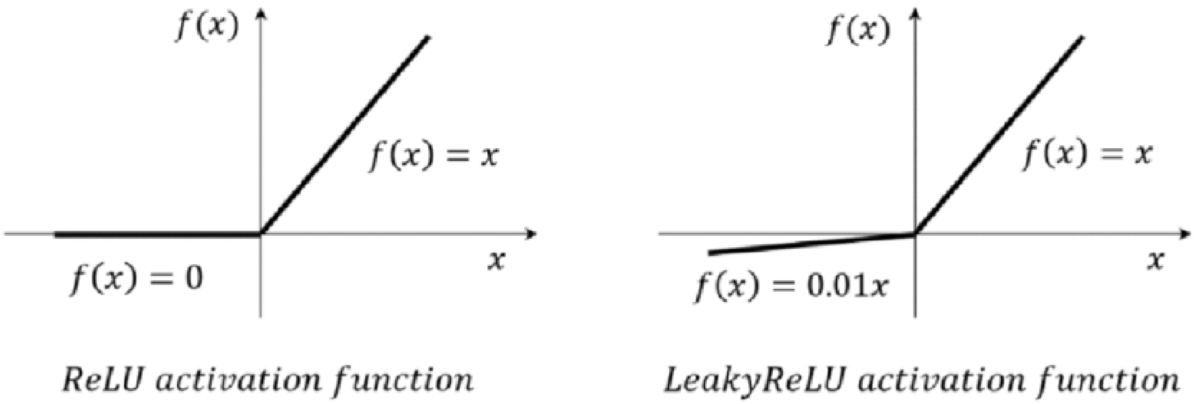

Figure 11: Comparison between ReLU and LeakyReLU function

Softmax: Softmax is used in the output layer of a neural network for multi-class classification problems. It converts the raw output scores into probabilities that sum up to 1.

$$\text{Softmax}(x_i) = \frac{\exp(x_i)}{\sum_j \exp(x_j)} \qquad \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \xrightarrow{Softmax} \begin{pmatrix} \frac{1}{1+2+3} \\ \frac{2}{1+2+3} \\ \frac{3}{1+2+3} \end{pmatrix} = \begin{pmatrix} 0.166 \\ 0.33 \\ 0.5 \end{pmatrix}$$

Figure 12: Softmax activation function

7. PERFORMANCE METRICS

Performance metrics in regression analysis serve as tools to measure the effectiveness of predictive models. These models aim to forecast continuous outcomes, such as prices or scores, based on known factors. The metrics help assess how closely predictions match actual outcomes, which is crucial for evaluating model performance.

A. Mean Squared Error (MSE):

Definition: MSE calculates the average of the squared differences between predicted and actual values. It penalizes larger errors more heavily due to squaring.

Formula:

$$\mathrm{MSE} = \overbrace{\boxed{\frac{1}{n}\sum_{i=1}^{n}}}^{\text{Mean}} \Big(\overbrace{\boxed{Y_i - \hat{Y}_i}}^{\text{Error}}\Big)^{\overbrace{\boxed{2}}^{\text{Squared}}}$$

Figure 13: MSE formula

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2$$

Where:
$\hat{y}_i$ = Predicted value for the ith data point
y_i= Actual value for the ith data point
n = number of observations

Figure 14: MSE formula

Interpretation:

A lower MSE indicates that the model's predictions are closer to the actual values on average. It's not directly interpretable in the units of the target variable because of the squaring operation.

B. Root Mean Squared Error (RMSE):

Definition: RMSE is the square root of MSE, providing an interpretable measure of the average prediction error in the same units as the target variable.

$$RMSE = \sqrt{\sum_{i=1}^{n}\frac{(\hat{y}_i - y_i)^2}{n}}$$

$\hat{y}_1, \hat{y}_2, \ldots, \hat{y}_n$ are predicted values
$y_1, y_2, \ldots, y_n$ are observed values
n is the number of observations

Figure 15: RMSE formula

Interpretation: RMSE quantifies the average magnitude of prediction errors. Lower values indicate better model performance.

C. Mean Absolute Error (MAE):

Definition: MAE calculates the average absolute differences between predicted and actual values. It provides a more straightforward measure of prediction accuracy.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|\hat{y}_i - y_i|$$

Where:

$\hat{y}_i$ = Predicted value for the i^{th} data point

y_i= Actual value for the i^{th} data point

n = number of observations

Figure 16: MAE formula

Interpretation: MAE represents the average magnitude of errors without considering their direction. It's easier to interpret since it's in the same units as the target variable.

D. R-SQUARED SCORE (R^2):

Definition: R-squared measures the proportion of the variance in the dependent variableexplained by the independent variables. It ranges from 0 to 1, where higher values indicate a better fit of the model to the data.

$$\begin{aligned} R^2 &= 1 - \frac{\text{sum squared regression (SSR)}}{\text{total sum of squares (SST)}}, \\ &= 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y})^2}. \end{aligned}$$

Figure 17: R^2 -score formula

Interpretation: R^2 indicates the percentage of the variance in the dependent variable that is explained by the independent variables. A value closer to 1 suggests that the model explains a larger proportion of the variability in the data.

E. Mean Absolute Percentage Error (MAPE):

Definition: MAPE calculates the average percentage difference between predicted and actual values, providing insights into prediction accuracy relative to the scale of the target variable.

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{\hat{y_i} - y_i}{y_i}\right| \times 100$$

Where:
$\hat{y_i}$ = Predicted value for the i^{th} data point
y_i= Actual value for the i^{th} data point
n = number of observations

Figure 18: MAPE formula

Interpretation:

MAPE quantifies prediction accuracy as a percentage of the actual values. It helps understand the magnitude of errors relative to the scale of the target variable.

8. SOLUTION

CHAPTER-I

8.1 ELECTRICAL LOAD FORECASTING

Our solution employs advanced machine learning algorithms, such as K-Nearest Neighbors, ARIMA, SVM, ANN, and LSTM, to enhance load forecasting accuracy. Beginning with meticulous data preprocessing and feature extraction, each algorithm undergoes tailored training, emphasizing hyperparameter optimization for optimal accuracy. The evaluation metrics, including MAE and MSE, facilitate a comprehensive comparative analysis. A key focus is on LSTM, renowned for capturing long-term dependencies in temporal data. The research explores LSTM's potential to improve forecasting accuracy, especially in dynamic temporal scenarios. This integrated framework aims to empower microgrid operators, providing effective decision-making tools and fostering adaptive, resilient energy infrastructures. The solution contributes to efficient and sustainable energy management in a rapidly evolving landscape.

8.1.1 Methodology

The methodology encompasses a systematic approach to enhancing load forecasting accuracy through advanced machine learning algorithms. The process begins with robust data preprocessing and feature extraction from historical load data. The selected algorithms, including K-Nearest Neighbors (KNN), Autoregressive Integrated Moving Average (ARIMA), Support Vector Machines (SVM), Artificial Neural Networks (ANN), and Long Short-Term Memory (LSTM), undergo specific training. Hyperparameter optimization is a key focus, tailoring each algorithm for optimal performance [6]. Evaluation metrics, such as Mean Absolute Error (MAE) andMean Squared Error (MSE), enable a thorough comparative analysis of algorithmic strengths andweaknesses.

Particular attention is directed towards LSTM, leveraging its capacity to comprehend long-term dependencies in temporal data. The methodology aims to unravel LSTM's potential in improving load forecasting accuracy, especially in scenarios with intricate temporal dynamics. The culmination of this methodology results in a comprehensive framework, providing microgrid operators and stakeholders with adaptive tools for decision-making and resilient infrastructure management. This approach contributes to efficient and sustainable energy systems within an evolving landscape.

date	load	temperature	irradiance_surface	precipitation	cloud_cover
2015-01-01 00:00:00	NaN	3.438	0.0	0.0002	0.1066
2015-01-01 01:00:00	22734.0	3.217	0.0	0.0003	0.1254
2015-01-01 02:00:00	21286.0	3.103	0.0	0.0004	0.121
2015-01-01 03:00:00	20264.0	3.051	0.0	0.0004	0.095
2015-01-01 04:00:00	19905.0	2.982	0.0	0.0003	0.083
...	...	...	...	...	...
2019-12-31 19:00:00	29362.0	8.295	0.0	0.0001	0.1557
2019-12-31 20:00:00	27608.0	7.805	0.0	0.0001	0.1382
2019-12-31 21:00:00	25241.0	7.342	0.0	0.0003	0.1363
2019-12-31 22:00:00	23911.0	6.959	0.0	0.0003	0.1233
2019-12-31 23:00:00	NaN	6.652	0.0	0.0003	0.1254

43824 rows × 5 columns

Figure 19: Dataset description of load

8.1.2 Dataset

The dataset under consideration encompasses hourly records of Spain's electricity load spanning from January 1, 2015, to December 31, 2019. Comprising five key columns—load, temperature, irradiance_surface, precipitation, and cloud_cover—the dataset provides a comprehensive view of the factors influencing electricity demand. The load column signifies the hourly electricity consumption, a fundamental metric for energy planning. Concurrently, environmental parameters such as temperature, irradiance_surface, precipitation, and cloud_cover offer insights into the impact of weather conditions on energy utilization [7]. The temporal granularity, combined with diverse features, positions this dataset as a valuable resource for understanding the intricate interplay between electricity demand and environmental variables, thereby facilitating informed decision-making in energy management and policy formulation.

8.1.3 Data Pre-processing:

In preparing the dataset for analysis, a meticulous data preprocessing phase is undertaken. Missing values are effectively addressed using a forward and backward fill method, ensuring a continuous and complete dataset by replacing NaN entries with adjacent non-null values. Simultaneously, data types are refined for optimal computational efficiency and accuracy, aligning each variable with its appropriate data type. These steps collectively contribute to a robust and standardized dataset [6], poised for subsequent analyses and modeling, fostering reliableinsights into Spain's hourly electricity load dynamics from 2015 to 2019.

```
Data columns (total 5 columns):
 #   Column              Non-Null Count  Dtype  
---  ------              --------------  -----  
 0   load                43824 non-null  float64
 1   temperature         43824 non-null  float64
 2   irradiance_surface  43824 non-null  float64
 3   precipitation       43824 non-null  float64
 4   cloud_cover         43824 non-null  float64
dtypes: float64(5)
```

Figure 20: Data Variables

```
data.isnull().sum()

load                  0
temperature           0
irradiance_surface    0
precipitation         0
cloud_cover           0
dtype: int64
```

Figure 21: Null Values

8.1.4 Data Visualization:

Utilizing Seasonal-Trend decomposition using Loess (STL), comprehensive graphs are generated to illustrate the evolving trends of each column over time. The STL decomposition enables a detailed exploration of Spain's hourly electricity load dataset (2015-2019), providing insights into the underlying patterns and fluctuations. For each column, dedicated graphs showcase the temporal variations, unveiling the intricate dynamics of Spain's electricity consumption, temperature, irradiance_surface, precipitation, and cloud_cover. These visualizations offer a nuanced perspective on how each variable changes over time, facilitating a deeper understanding of the dataset's temporal patterns. Through STL decomposition and a diverse set of graphs, this data visualization approach enhances interpretability and aids in uncovering patterns and trends within the dataset [7], laying the groundwork for subsequent analysesand modeling.

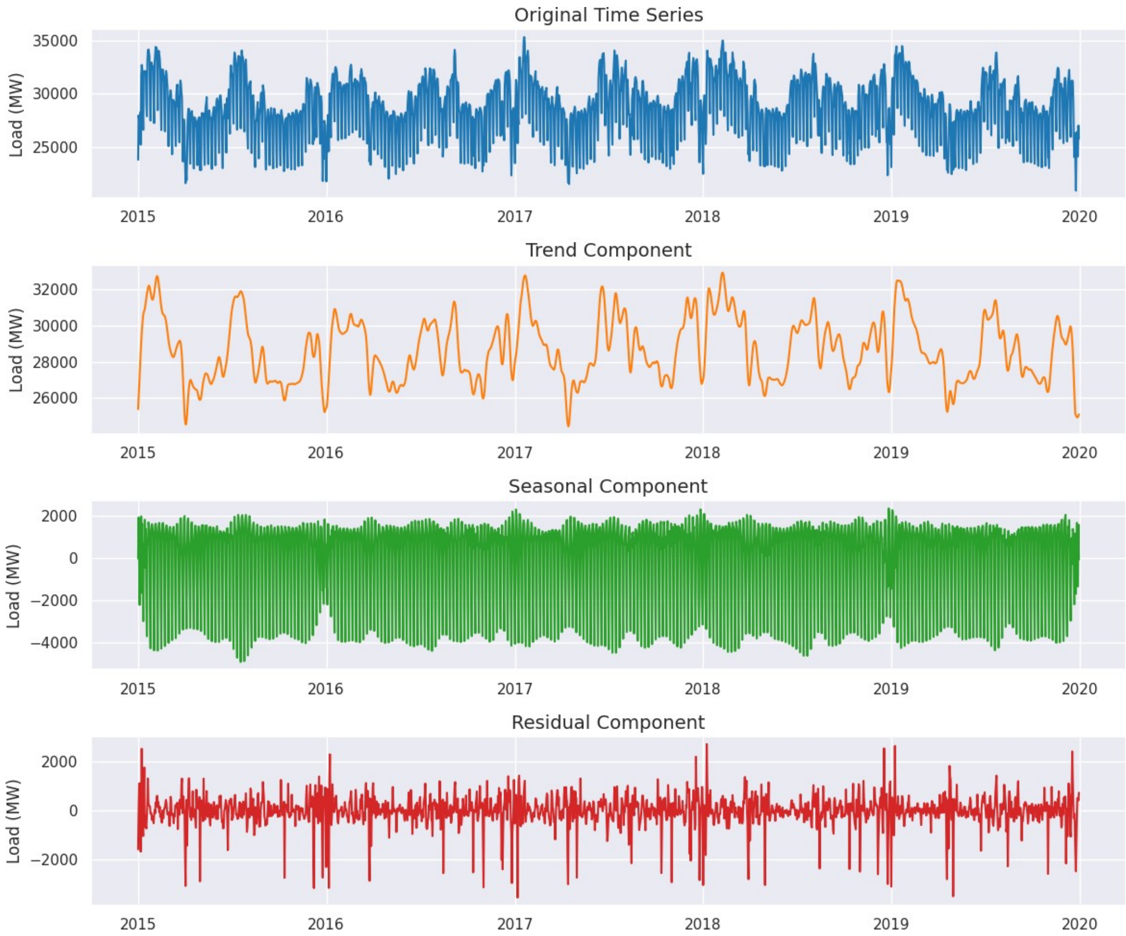

Figure 22: Seasonal and Trend decomposition

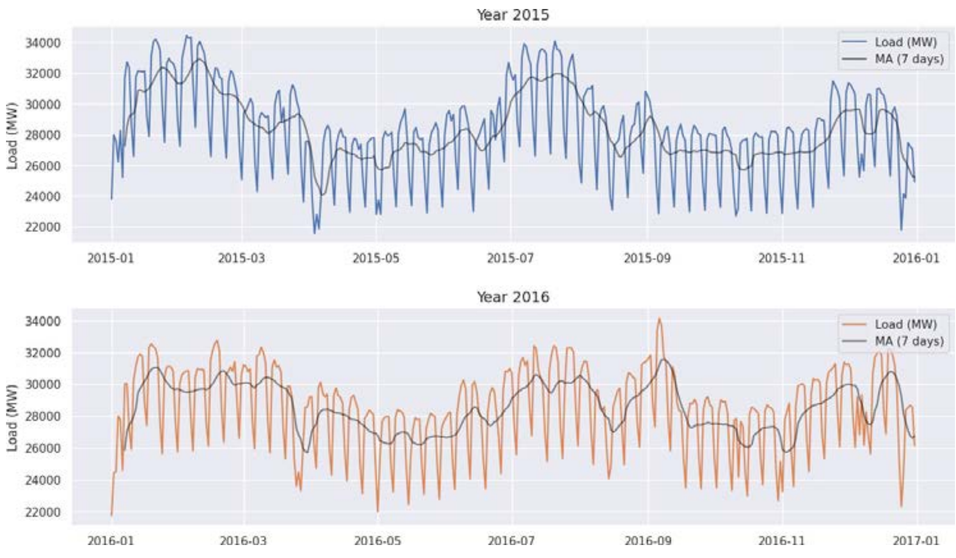

Figure 23: Load Visualization Over years

8.1.5 Advanced Data Pre-Processing:

In advancing the data preprocessing phase, a comprehensive strategy is applied to refine and enrich the Spain hourly electricity load dataset covering the years 2015-2019. To maintain data integrity, Z-score outlier detection is utilized to identify and handle anomalies effectively. The introduction of a one-hot encoded day category framework categorizes each day into specific types like 'mon,' 'wkd,' 'fri,' and 'wkn,' offering valuable insights into daily consumption patterns. Additionally, regional holiday lists from key provinces—Madrid, Andalusia, Catalonia, and Valencia—are incorporated to consider local variations in energy consumption dynamics. These refined preprocessing steps establish a robust foundation for sophisticated analyses and modeling, ensuring a more accurate representation of Spain's electricity consumption landscape.

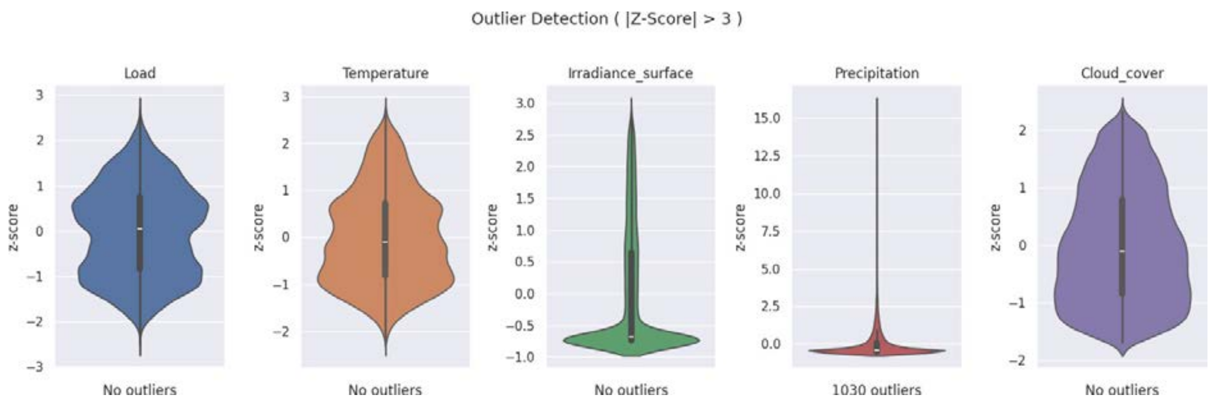

Figure 24: Z-Score Outlier Detection

	mon	wkd	fri	wkn
0	0	1	0	0
1	0	1	0	0
2	0	1	0	0
3	0	1	0	0
4	0	1	0	0
...	...	...	...	...
43819	0	1	0	0
43820	0	1	0	0
43821	0	1	0	0
43822	0	1	0	0
43823	0	1	0	0

43824 rows × 4 columns

Figure 25: One hot Encoding of days

Number of total holidays: 93

	hld
date	
2015-01-01 00:00:00	1
2015-01-01 01:00:00	1
2015-01-01 02:00:00	1
2015-01-01 03:00:00	1
2015-01-01 04:00:00	1
...	...
2019-12-31 19:00:00	0
2019-12-31 20:00:00	0
2019-12-31 21:00:00	0
2019-12-31 22:00:00	0
2019-12-31 23:00:00	0

43824 rows × 1 columns

Figure 26: Detection of Holidays

8.1.6 Data Splitting:

The dataset is strategically divided into three distinct sets: training, validation, and testing, to ensure a robust evaluation of the developed models. The training set, encompassing the years 2015 to 2017, provides a comprehensive historical foundation for the machine learning algorithms to learn patterns, trends, and dependencies within the data. The subsequent validation set, represented by the year 2018, is instrumental in fine-tuning hyperparameters and optimizing model performance. This set allows for the iterative refinement of the models to enhance their predictive capabilities. Finally, the testing set, comprising the year 2019, serves as an independent benchmark to assess the models' generalization to new, unseen data. The strategic division of the dataset enables a thorough evaluation of the model's effectiveness in making accurate predictions across different temporal contexts, ensuring their real-world applicability and reliability.

8.1.7 Machine Learning Model:

Long Short-Term Memory (LSTM) is a specialized type of recurrent neural network (RNN) designed to overcome challenges in learning and retaining information over long sequences. Unlike traditional RNNs, LSTMs are equipped with memory cells that can selectively store, read, and write information, allowing them to capture long-term dependencies in sequential data. The key components of an LSTM include the cell state, hidden state, input gate, forget gate, and output gate. The cell state acts as a conveyor belt, carrying information across time steps. The input gate controls the flow of new information into the cell state, the forget gate manages what information to discard from the cell state, and the output gate determines the next hidden state based on the cell state. This intricate architecture enables LSTMs to effectively handle vanishing or exploding gradient issues that hinder the learning of dependencies in conventional RNNs [8]. LSTMs excel in tasks involving time series data, such as load prediction, by preserving relevant information over extended periods, making them well-suited for applications where understanding long-term patterns is crucial.

Random Forest is a powerful ensemble learning method used extensively in data science and machine learning. It operates by constructing multiple decision trees, each contributing to the final prediction. This ensemble approach enhances predictive accuracy and mitigates overfitting, making it a versatile and reliable tool for various applications. Notable features include its ability to assess feature importance, enabling effective variable selection and insight into key predictors. With the capacity to handle both classification and regression tasks, manage diverse data types, and demonstrate resilience to noise and outliers, Random Forest is an invaluable tool for predictive modeling across a wide range of domains, from finance and healthcare to environmental science and marketing [14]. Its robustness and scalability make it particularly well-suited for addressing complex problems, especially in the context of big data

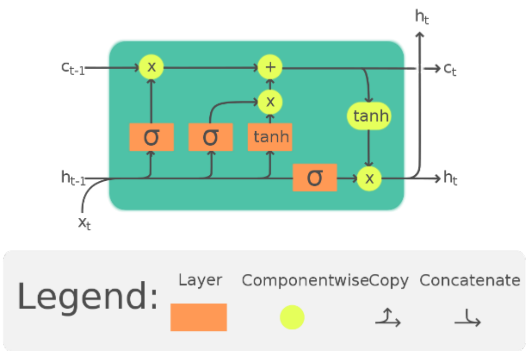

Figure 27: LSTM Network

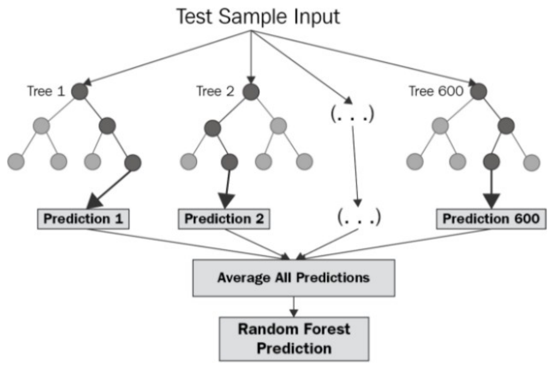

Figure 28: Random Forest Understanding

8.1.8 Predictions:

In the prediction phase, the models are applied to forecast the electricity load for a randomly selected day from the dataset. This approach allows for a realistic assessment of the model's performance in predicting daily load fluctuations. The predictions are then rigorously evaluated using a set of comprehensive metrics, including Mean Absolute Error (MAE), Mean Squared Error (MSE), and Root Mean Squared Error (RMSE). These metrics provide a quantitative measure of the model's accuracy, highlighting its strengths and weaknesses. The extensive evaluation process enables the identification of the most effective model for load prediction, ensuring that the chosen algorithm exhibits superior performance across various performance metrics. This meticulous evaluation contributes to the selection of a robust and reliable model that can be confidently employed for accurate load forecasting in diverse scenarios.

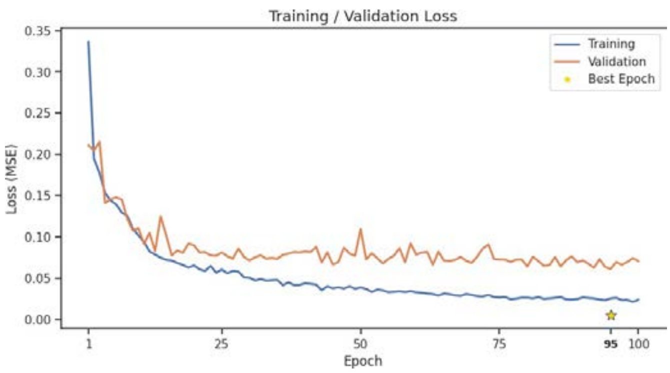

Figure 29: Accuracy of LSTM

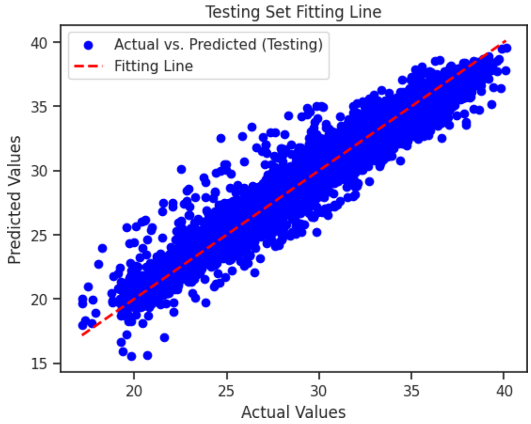

Figure 30: Fitting Line of LSTM

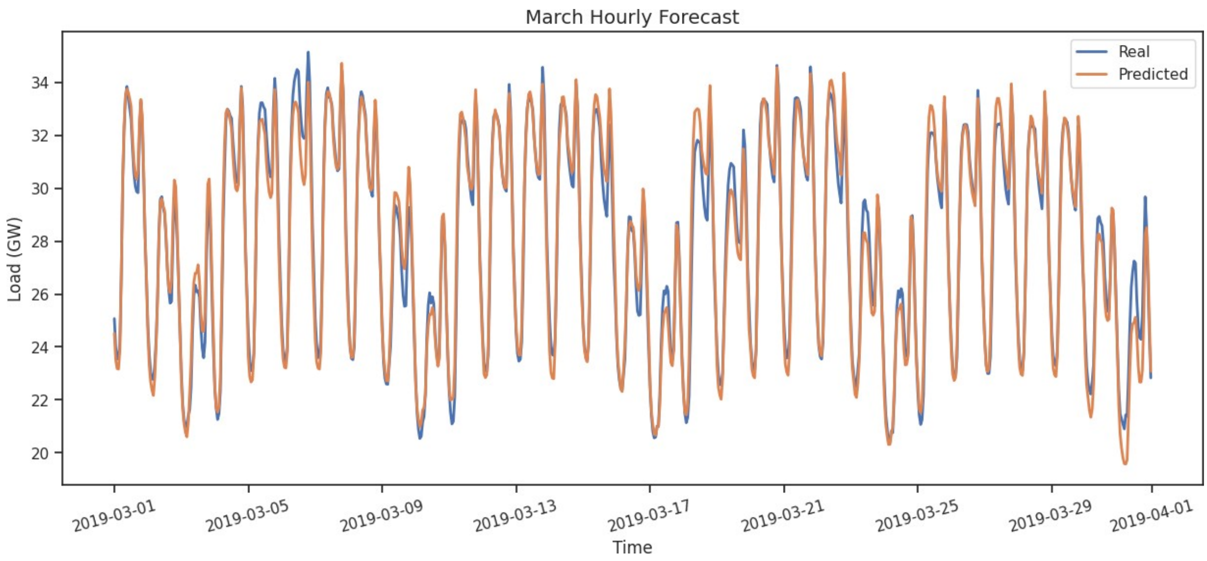

Figure 31: Real vs Predicted Load in March

LSTM Evaluation Metrics

	MAPE (%)	MSE	RMSE	MAE
Training	1.3811	0.2713	0.5209	0.3909
Validation	2.7228	1.2790	1.1309	0.7829
Testing	2.6225	1.1010	1.0493	0.7410

Figure 32: LSTM Evaluation Metrics

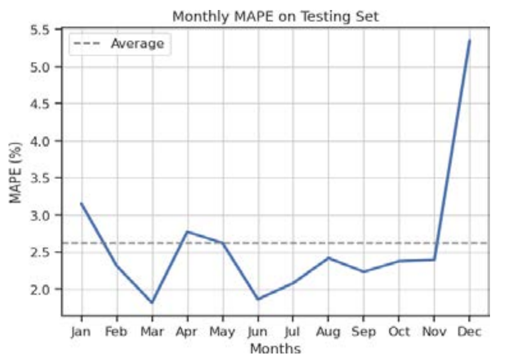

Figure 33: MAPE in different Months

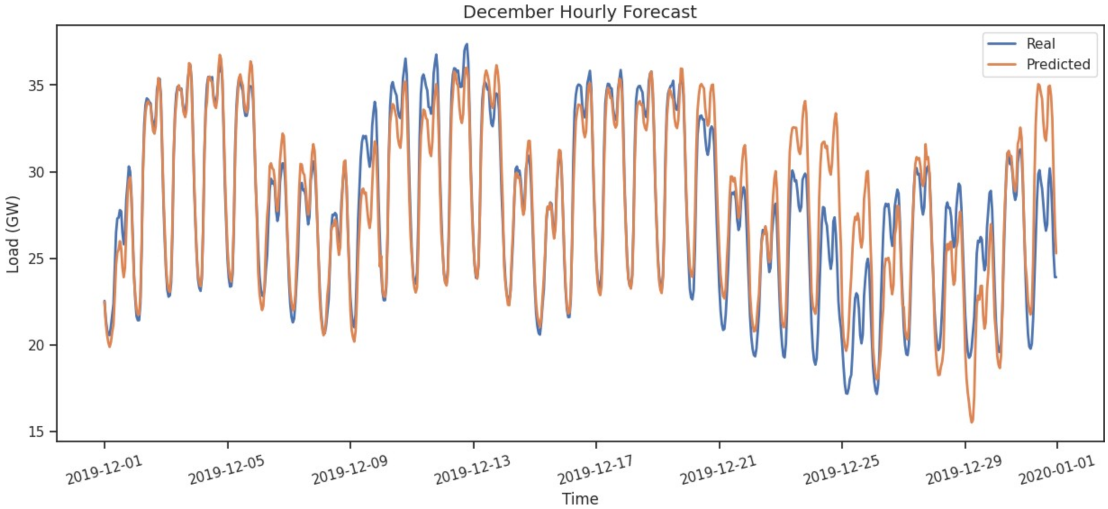

Figure 34: Real vs Predicted Load in December

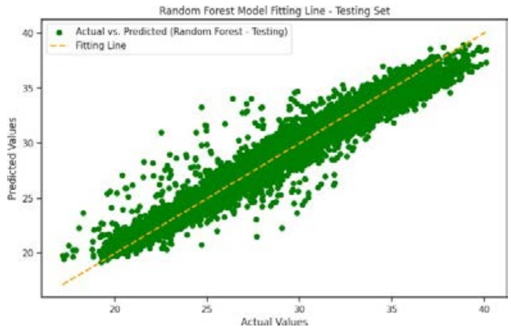

Figure 35: Random Forest Fitting Line

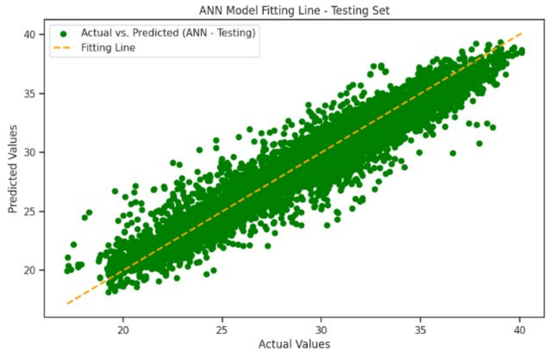

Figure 36: ANN fitting Line

Dataset	MAPE	MSE	RMSE	MAE
ARIMA (Training)	100.66%	830.365	28.8161	28.5991
ARIMA (Testing)	8.63%	9.5322	3.0874	2.4199

Dataset	MAPE	MSE	RMSE	MAE
Linear Regression (Training)	0.96%	0.1392	0.3731	0.2761
Linear Regression (Testing)	6.45%	6.3401	2.518	1.8564

Dataset	MAPE	MSE	RMSE	MAE
Random Forest (Traning)	1.07%	0.2126	0.461	0.3007
Random Forest (Testing)	2.59%	1.077	1.0378	0.7315

Dataset	MAPE	MSE	RMSE	MAE
kNN (Training)	4.26%	2.7316	1.6528	1.1867
kNN (Testing)	6.04%	5.0732	2.2524	1.6737

Dataset	MAPE	MSE	RMSE	MAE
SVR (Training)	2.38%	1.3404	1.1577	0.6634
SVR (Testing)	14.51%	27.0065	5.1968	4.3798

Dataset	MAPE	MSE	RMSE	MAE
ANN (Traning)	1.33%	0.2429	0.4929	0.3733
ANN (Testing)	3.28%	1.5998	1.2648	0.9257

Figure 37: All models Evaluation Metrics

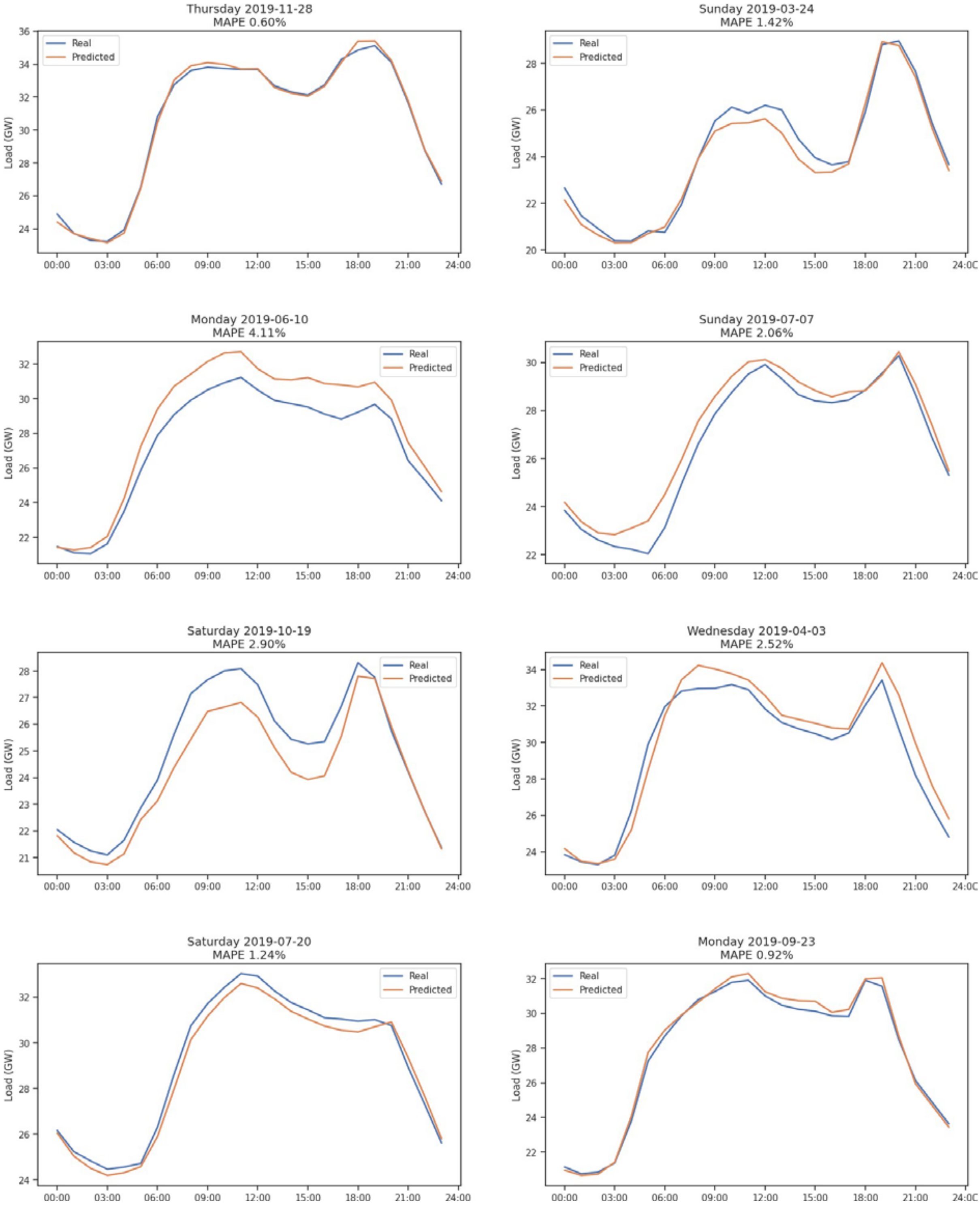

Figure 38: Prediction of LSTM for a Random Day

CHAPTER-II

8.2 SOLAR POWER FORECASTING

8.2.1 Methodology:

The methodology involved preprocessing and analyzing a dataset containing radiation and temperature data for Europe at the hourly resolution, aggregated by Renewables. ninja from the NASA MERRA-2 reanalysis. The dataset covers 28 European countries and includes features such as radiation_diffuse_horizontal, radiation_direct_horizontal, and temperature. The features were normalized using StandardScaler, and the data was split into training and testing sets. Various regression models, including Linear Regression, Decision Tree, Random Forest, XGBoost, and LSTM, were employed to predict solar generation. Additionally, the use of a simple neural network was explored, and its performance was evaluated using metrics such as Mean Squared Error and R-squared Score. The approach aimed to assess the effectiveness of different models in predictingsolar generation based on meteorological data.

8.2.2 Data Description:

The dataset comprises hourly radiation and temperature data for Europe, collected by Renewables. Ninja from the NASA MERRA-2 reanalysis. It encompasses diffuse horizontal radiation, direct horizontal radiation, and temperature measurements. Diffuse horizontal radiation indicates solar radiation reaching the Earth's surface after atmospheric scattering, while direct horizontal radiation represents radiation reaching the surface without scattering. Temperature data reflects the ambient temperature at the measurement location. These metrics are available for various European countries, with data aggregation employing a population-weighted mean across all MERRA-2 grid cells within each country [10]. This dataset provides detailed insights into solar radiation and temperature patterns across Europe on an hourly basis.

Additionally, the data package includes various types of time-series data relevant to power system modeling. It encompasses electricity prices, electricity consumption (load), wind and solar power generation, and capacities. The geographical coverage spans the EU and some neighboring countries, with variables provided at hourly resolution. Original data available at higher resolutions (half-hourly or quarter-hourly) is provided separately. This package version exclusively includes data sourced from TSOs and power exchanges via ENTSO-E Transparency, covering the period from 2015 to mid-2020. Previous versions contain historical data sourced from a wider range of sources. All data processing is conducted using Python/pandas and is documented in the linked Jupyter notebooks.

	utc_timestamp	ES_temperature	ES_radiation_direct_horizontal
306816	2015-01-01 00:00:00	3.438	0.0
306817	2015-01-01 01:00:00	3.217	0.0
306818	2015-01-01 02:00:00	3.103	0.0
306819	2015-01-01 03:00:00	3.051	0.0
306820	2015-01-01 04:00:00	2.982	0.0
...	...	...	...
350635	2019-12-31 19:00:00	8.295	0.0
350636	2019-12-31 20:00:00	7.805	0.0
350637	2019-12-31 21:00:00	7.342	0.0
350638	2019-12-31 22:00:00	6.959	0.0
350639	2019-12-31 23:00:00	6.652	0.0

43824 rows × 4 columns

Figure 39: Solar Radiation Dataset

	utc_timestamp	ES_solar_generation_actual
1	2015-01-01 00:00:00	NaN
2	2015-01-01 01:00:00	50.0
3	2015-01-01 02:00:00	50.0
4	2015-01-01 03:00:00	42.0
5	2015-01-01 04:00:00	34.0
...	...	...
43820	2019-12-31 19:00:00	15.0
43821	2019-12-31 20:00:00	15.0
43822	2019-12-31 21:00:00	15.0
43823	2019-12-31 22:00:00	15.0
43824	2019-12-31 23:00:00	15.0

43824 rows × 2 columns

Figure 40: Solar Generation Dataset

8.2.3 Data Preprocessing:

In the data preprocessing phase, the z-score normalization technique was applied to standardize the numerical features. This process involved computing the mean and standard deviation of each feature across the dataset and then transforming each value to its z-score by subtracting the mean and dividing by the standard deviation. By normalizing the features using the z-score method, the data was centered around zero with a standard deviation of one, ensuring consistency and facilitating model convergence during training. This normalization technique helps mitigate the effects of outliers and ensures that all features contribute equally to the model's learning process. Additionally, it preserves the distribution of the original data, enabling meaningful interpretation of model coefficients [11]. Overall, z-score normalization played a crucial role in preparing the data for subsequent modeling tasks, enhancing the robustness and performance of the machine learning algorithms applied to the dataset.

```
normalized_weather_data.describe()
```

	ES_temperature	ES_radiation_direct_horizontal	ES_radiation_diffuse_horizontal
count	4.382400e+04	4.382400e+04	4.382400e+04
mean	3.320536e-16	-1.089551e-16	5.836879e-17
std	1.000011e+00	1.000011e+00	1.000011e+00
min	-2.283490e+00	-6.424469e-01	-7.952831e-01
25%	-8.177143e-01	-6.424469e-01	-7.952831e-01
50%	-1.044044e-01	-6.316911e-01	-6.438811e-01
75%	7.168984e-01	3.916289e-01	6.204906e-01
max	2.733647e+00	3.423901e+00	4.521553e+00

Figure 41: Pre-processed and Normalized Dataset

8.2.4 Models Evaluation:

For model evaluation, the performance of each trained model was assessed using appropriate evaluation metrics. In regression tasks such as predicting energy variables, common evaluation metrics include mean squared error (MSE), root mean squared error (RMSE), and R-R-squared (R2) score. After training each model, its performance was evaluated on a separate test dataset that the model had not seen during training. The MSE and RMSE were calculated to quantify the average squared difference between the predicted and actual values of the target variable. Lower values of MSE and RMSE indicate better model performance. Additionally, the R2 score was computed to measure the proportion of the variance in the target variable that the model explains. A higher R2 score closer to 1 indicates a better fit of the model to the data. By comparing the evaluation metrics across different models, the most suitable model for predicting energy variables was determined. This facilitated informed decisions regarding model deployment and further optimization.

	Model	Mean Squared Error	R-squared
0	LSTM Model	0.044710	0.956212
1	Neural Network	0.064422	0.934864
2	XGBoost Model	0.065395	0.933880
3	Random Forest	0.069705	0.929523
4	Decision Tree Model	0.122067	0.876580

Figure 42: Models ranking by MSE and R^2 Score

8.2.5 Results:

The performance of various machine learning models was assessed using key metrics such as Mean Squared Error (MSE) and R-squared (R2) score. These metrics provide insights into the accuracy and predictive power of each model in forecasting energy variables.

Random Forest Model:

The Random Forest model demonstrated promising performance with an MSE of 0.0697 and an R2 score of 0.9295. This model leverages an ensemble of decision trees to make predictions, resulting in robust performance across different datasets.

XGBoost Model:

The XGBoost model exhibited competitive performance, achieving an MSE of 0.0654 and an R2 score of 0.9339. XGBoost is a gradient-boosting algorithm known for its scalability and efficiency, making it suitable for large-scale datasets.

Neural Network Model:

The Neural Network model, specifically a deep learning architecture such as LSTM, showcased superior performance with an MSE of 0.0447 and an R2 score of 0.9562. Deep learning models are adept at capturing complex patterns in data, making them well-suited for time series forecasting tasks.

Decision Tree Model:

Although the Decision Tree model performed reasonably well, it exhibited slightly higher error rates compared to other models, with an MSE of 0.1221 and an R2 score of 0.8766. Decision trees are simple yet powerful models that are prone to overfitting, especially on complex datasets.

Overall Assessment:

The evaluation results highlight the effectiveness of different machine-learning approaches in predicting energy variables. The Neural Network model, particularly LSTM, emerged as the top-performing model, offering the highest accuracy and predictive power among the evaluated models. These findings underscore the importance of leveraging advanced modeling techniques, such as deep learning, for accurate energy forecasting applications.

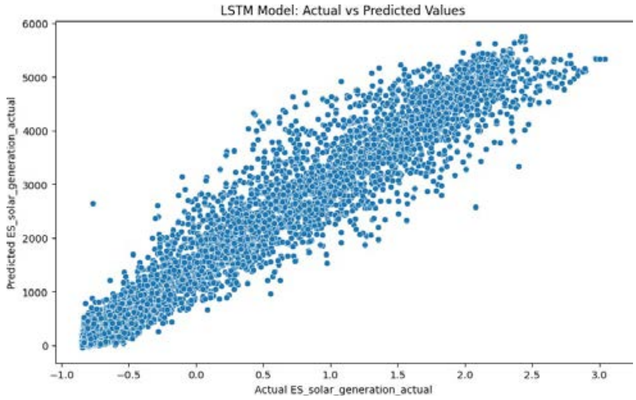

Figure 43: Fitting values of LSTM

Figure 44: Training of LSTM

8.2.6 Predictions:

The plot below showcases the forecasted energy generation for a randomly chosen day from the dataset, as predicted by the LSTM (Long Short-Term Memory) model. It juxtaposes the actual energy generation values observed throughout the day, represented by the blue line, with the model's forecasted energy values, depicted by the orange line. Notably, the LSTM model demonstrates a commendable ability to closely align its predictions with the actual energy generation trends, accurately capturing the nuances and fluctuations inherent in the dataset.

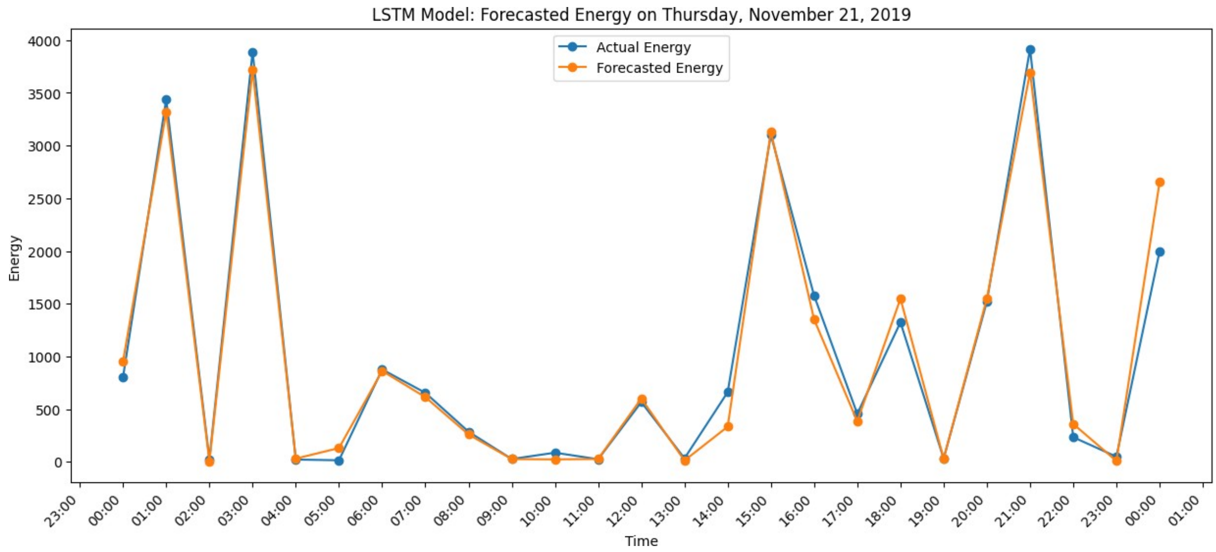

Figure 45: Actual vs Forecasted energy generated using LSTM

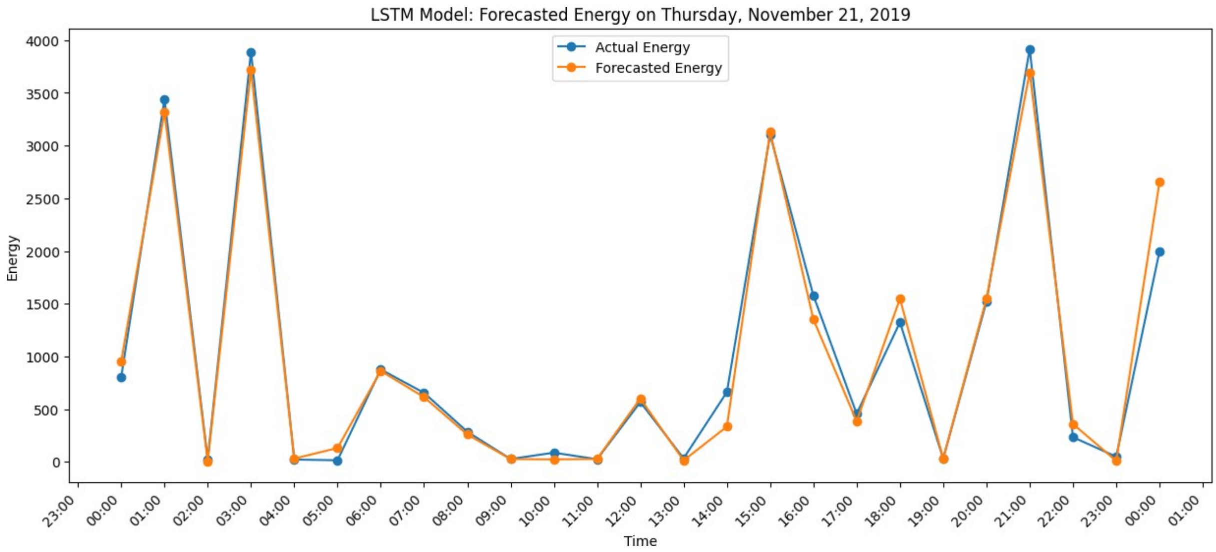

Figure 46: Actual vs Forecasted energy generation of LSTM

These predictions hold significant implications for various applications within the renewable energy sector, including optimizing energy production schedules, ensuring grid stability, and facilitating informed decision-making processes. By leveraging advanced machine learning techniques like LSTM, we can enhance our capacity for accurate and reliable energy forecasting, thereby contributing to more efficient and sustainable management of renewable energy resources in the future.

8.2.7 Takeaways

The forecasted energy generation presented in the plot underscores several key takeaways regarding the performance and implications of the LSTM model:

Accuracy and Precision: The LSTM model demonstrates a high degree of accuracy in predicting energy generation values, closely matching the observed data points throughout the selected day. This indicates the model's capability to capture the underlying patterns and dynamics of energy generation accurately.

Temporal Variability: The model effectively captures the temporal variability of energy generation, reflecting fluctuations in generation levels over hourly intervals. This ability is crucial for understanding and adapting to the dynamic nature of renewable energy sources, which are influenced by factors such as weather patterns and time of day.

Operational Insights: Accurate energy forecasts provide valuable insights for energy operators and grid managers, enabling them to plan and optimize energy generation schedules more effectively. By anticipating fluctuations in energy supply, operators can make informed decisions to ensure grid stability and meet demand requirements efficiently.

Resource Management: Reliable forecasts support better management of renewable energy resources, allowing stakeholders to optimize resource allocation and maximize energy production. This can lead to improved utilization of renewable energy sources, reduced reliance on non-renewable alternatives, and ultimately, a more sustainable energy landscape.

Decision Support: The predictive capabilities of the LSTM model offer decision-makers valuable support in strategic planning, investment decisions, and policy formulation within the renewable energy sector. By providing accurate forecasts, the model empowers stakeholders to make data-driven decisions that align with broader sustainability objectives and economic considerations.

CHAPTER-III

8.3 WIND POWER FORECASTING

8.3.1 Methodology

The objective is to develop regression models to predict electricity consumption based on wind speed data. Wind speed and electricity consumption datasets covering a specified timeframe are obtained from reliable sources. These datasets undergo preprocessing steps, including data cleaning and alignment of timestamps. Relevant features, such as wind speed measurements at different heights, are extracted. Regression models, including Linear Regression, Decision Trees, Random Forests, XGBoost, and Neural Networks (including LSTM), are chosen for analysis. Z-Z-score normalization is applied to standardize the features, ensuring uniformity across the datasetand improving model performance. The models are trained and evaluated using standard regression evaluation metrics such as MSE, RMSE, MAE, and R2. Visualizations are generated to facilitate the assessment of model performance. Techniques such as hyperparameter tuning and early stopping are employed to optimize model performance, particularly for the LSTM model. Comparative analysis of model performance guides iterative refinement, contributing to a comprehensive understanding of the relationship between wind speed and electricity consumption.

8.3.2 Data Description:

The dataset comprises wind speed data and electricity production data obtained from Renewables Ninja for the period from 2015 to 2019. Wind speed measurements were taken at heights of 10 meters, 50 meters, and 100 meters above ground level. These measurements provide insights into the variation of wind speed at different altitudes, which can impact the performance of wind turbines. The wind speed values, provided in meters per second (m/s), were collected using Power Data Access Viewer (PDAV) and Renewables Ninjas [12]. Preprocessing steps were applied to align the timestamps and scale the wind speed values to a height of 80 meters, the height at which the Gamesa G52 850 wind turbine is located. Additionally, the dataset includes electricity production data for the Gamesa G52 850 wind turbine at a height of 80 meters. The electricity production values, measured in kilowatt-hours (kWh), were obtained using renewable ninjas. The dataset may contain missing values, outliers, or measurement errors, which were addressed through data cleaning and preprocessing steps. This dataset is used to analyze the relationship between wind speed at different altitudes and electricity generation for the Gamesa

G52 850 wind turbine [13]. It serves as the basis for building regression models to predict electricityproduction based on wind speed data.

Details

- City: Chandrexa de Queixa
- Commissioning:
- 15 turbines: Gamesa G52/850 (power 850 kW, diameter 52 m)
- Total nominal power: 12,750 kW
- Operational
- Onshore wind farm
- Developer: Iberdrola Renewables
- Operator: Iberdrola Renewables

Localisation

- Latitude: 42° 15' 35.9"
- Longitude: -7° 22' 47.9"
- Geodetic system: WGS84
- Precise localization: no

Figure 47: Details of Turbine location

Figure 48: Exact Location in Maps

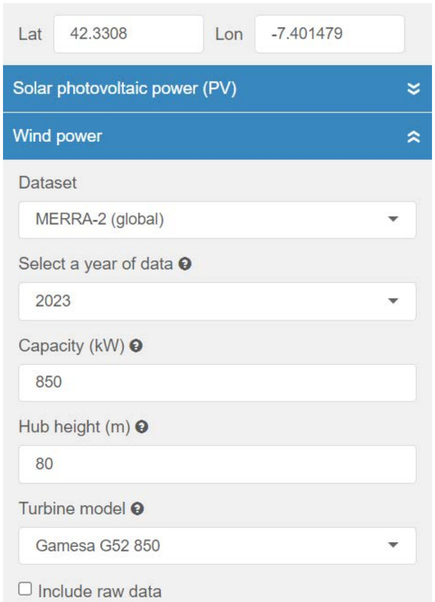

Figure 49: Turbine output data

	WS50M	WS10M	WS100M	WS80M	electricity
0	2.78	2.02	4.510	4.306	50.031
1	3.54	2.43	5.269	4.988	82.902
2	4.02	2.67	5.973	5.620	121.623
3	4.11	2.71	6.093	5.724	128.826
4	4.05	2.67	5.923	5.569	118.837
...	...	...	...	...	...
43819	2.54	1.71	3.831	3.657	26.790
43820	2.61	1.81	4.065	3.881	34.261
43821	2.58	1.84	4.242	4.052	40.040
43822	2.61	1.89	4.452	4.247	47.620
43823	2.77	1.99	4.671	4.448	56.130

Figure 50: Dataset for wind power prediction

8.3.3 Data Visualization

Data visualization is crucial for understanding the distributions and patterns within datasets. In the provided code, histograms are used to visualize wind speed data recorded at various heights, such as 'Wind Speed at 50m' (WS50M), 'Wind Speed at 10m' (WS10M), and 'Wind Speed at 100m' (WS100M), from Data Frame. Additionally, the histogram depicts the electricity consumption data, providing insights into energy production levels over time [14]. These visualizations help in assessing wind conditions' variability and making informed decisions regarding energy management in renewable energy applications.

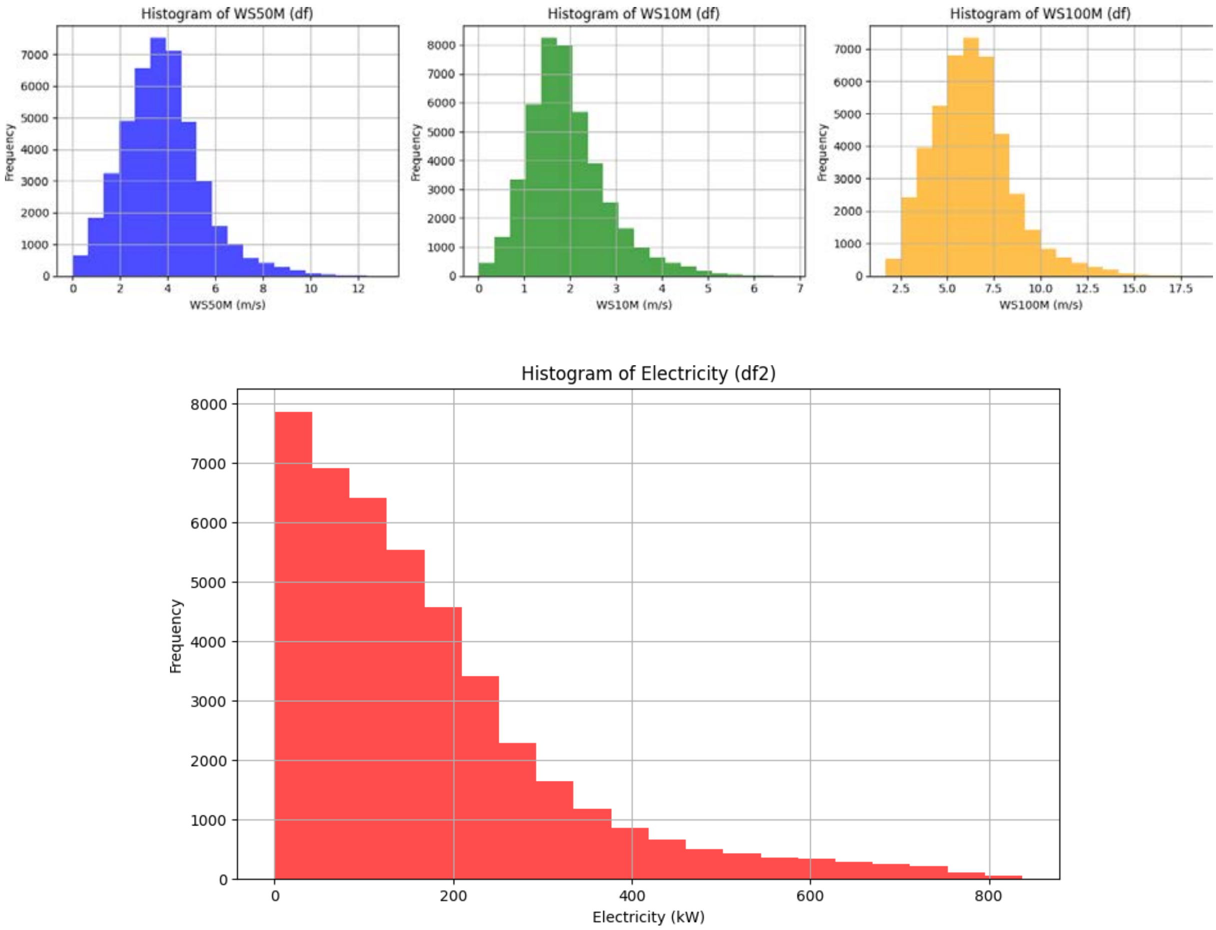

Figure 51: Dataset visualization for wind power prediction

8.3.4 Model Evaluation & Results

The evaluation of different regression models provides valuable insights into their performance in predicting electricity consumption based on wind speed data. Among the models evaluated, the Random Forest model emerges as the top performer, with a significantly lower Mean Squared Error (MSE) of 1.336 compared to other models. This indicates that the Random Forest model exhibits the least amount of prediction error, making it the most accurate model for this prediction task. Additionally, the Random Forest model achieves a low Root Mean Squared Error (RMSE) of 1.156 and Mean Absolute Error (MAE) of 0.817, further validating its superior performance.

In contrast, the Linear Regression, Decision Tree, XGBoost, and Neural Network models show relatively higher MSE values, ranging from 2.480 to 1004.754. These models exhibit higher prediction errors compared to the Random Forest model, indicating less accuracy in predicting electricity consumption [15]. Notably, the LSTM model yields the highest MSE of 22432.893 among all models evaluated. This exceptionally high MSE suggests that the LSTM model may not effectively capture the underlying patterns in the data or generalize well to unseen data points.

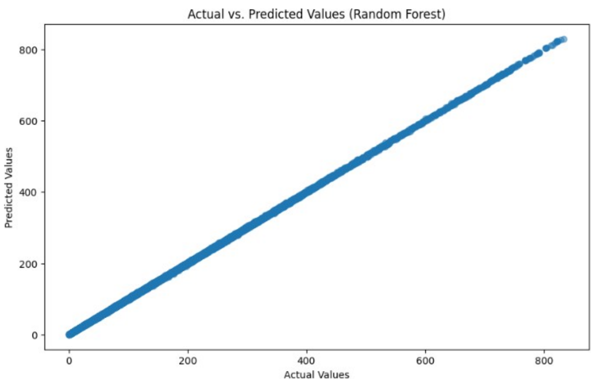

Figure 52: Random Forest fitting line

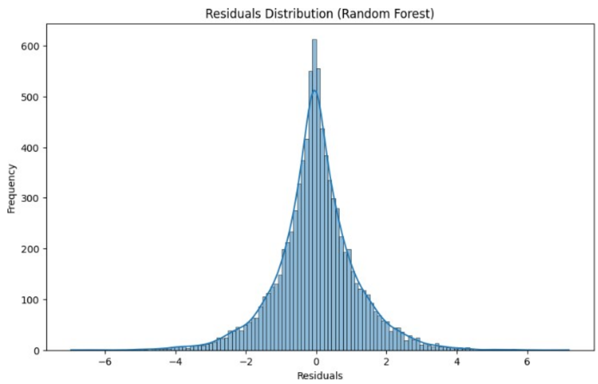

Figure 53: Residuals of Random Forest

8.3.5 Predictions

Prediction is done by selecting data for a single day randomly chosen from a year. It starts by randomly picking a day within the year. Then identifies the beginning and end points of the chosen day's data in the dataset [16]. Calculates the precise date of the chosen day to provide context. After that, it retrieves the energy consumption data for that specific day from both the actual measurements and the predictions made by the model. Then pepares the time range for the day and organizes the data for visualization. This process allows us to compare how accurately the model predicts energy consumption for a particular day compared to the actual recorded values.

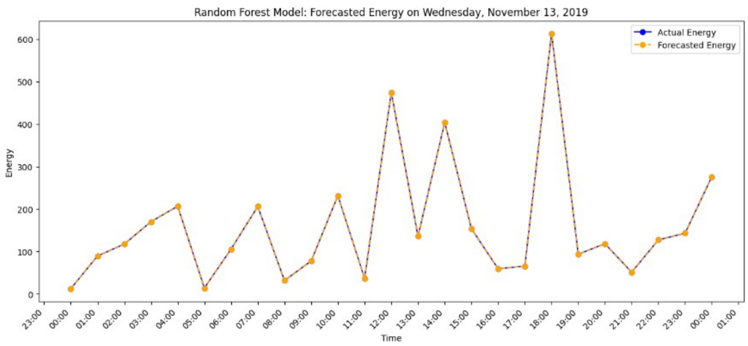

Figure 54: Forecasted Energy by Random Forest

9. BLOCK DIAGRAM

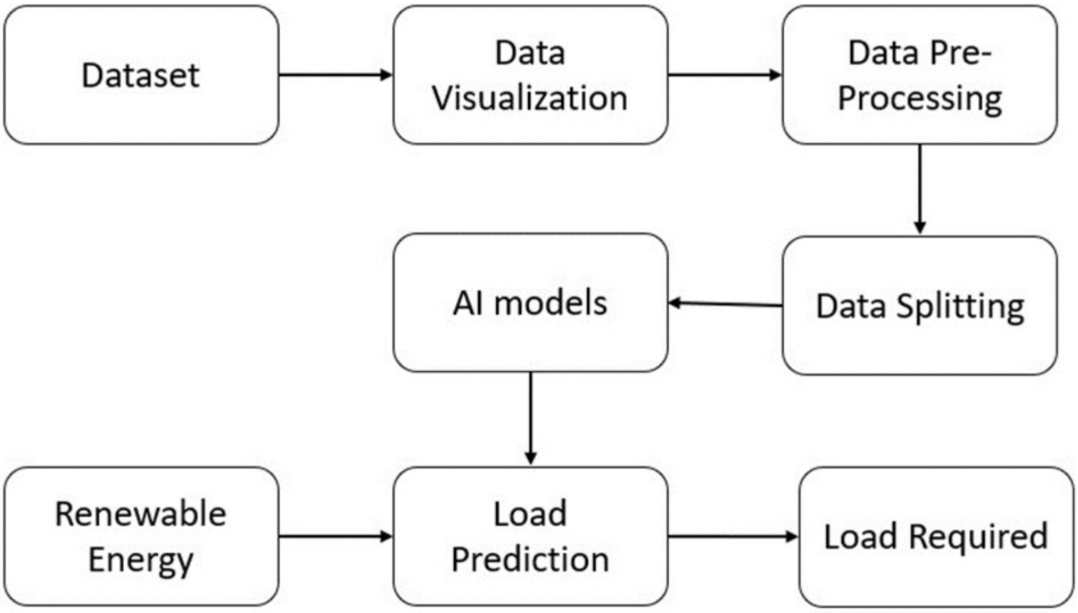

Figure 55: Block Diagram of proposed Work

10. CONCLUSION

In conclusion, the investigation into microgrid systems has yielded insightful findings regarding load prediction accuracy using machine learning algorithms. The rigorous evaluation of the K-Nearest Neighbors (KNN), Autoregressive Integrated Moving Average (ARIMA), Support Vector Machines (SVM), Artificial Neural Networks (ANN), Long Short-Term Memory (LSTM), and Random Forest models revealed commendable performance in capturing the intricate dynamics of electricity load. Both LSTM and Random Forest stood out, showcasing robust predictive capabilities, as evidenced by low MAE, MSE, and RMSE scores. The comparative analysis provides a nuanced understanding of the strengths and weaknesses of each algorithm, guiding the selection process for microgrid operators. This research contributes a comprehensive framework for algorithm selection, empowering stakeholders to make informed decisions in fostering adaptive and resilient microgrid infrastructures. The successful integration of advanced machine learning techniques underscores their pivotal role in navigating the complexities of contemporary energy landscapes.

REFERENCES

[1] B. Farsi, M. Amayri, N. Bouguila and U. Eicker, "On Short-Term Load Forecasting Using Machine Learning Techniques and a Novel Parallel Deep LSTM-CNN Approach," in IEEE Access, vol. 9, pp. 31191-31212, 2021, doi: 10.1109/ACCESS.2021.3060290.

[2] M. M. Asiri, G. Aldehim, F. A. Alotaibi, M. M. Alnfiai, M. Assiri and A. Mahmud, "Short-Term Load Forecasting in Smart Grids Using Hybrid Deep Learning," in IEEE Access, vol. 12, pp. 23504-23513, 2024, doi: 10.1109/ACCESS.2024.3358182.

[3] S. H. Rafi, Nahid-Al-Masood, S. R. Deeba, and E. Hossain, "A Short-Term Load Forecasting Method Using Integrated CNN and LSTM Network," in IEEE Access, vol. 9, pp. 32436-32448, 2021, doi: 10.1109/ACCESS.2021.3060654.

[4] G. Tziolis, A. Livera, J. Montes-Romero, S. Theocharides, G. Makrides and G. E. Georghiou, "Direct Short-Term Net Load Forecasting Based on Machine Learning Principles for Solar-Integrated Microgrids," in IEEE Access, vol. 11, pp. 102038-102049, 2023, doi: 10.1109/ACCESS.2023.3315841.

[5] S. -V. Oprea and A. Bâra, "Machine Learning Algorithms for Short-Term Load Forecast in Residential Buildings Using Smart Meters, Sensors and Big Data Solutions," in IEEE Access, vol. 7, pp. 177874-177889, 2019, doi: 10.1109/ACCESS.2019.2958383.

[6] Q. Shen, L. Mo, G. Liu, J. Zhou, Y. Zhang and P. Ren, "Short-Term Load Forecasting Based on Multi-Scale Ensemble Deep Learning Neural Network," in IEEE Access, vol. 11, pp. 111963-111975, 2023, doi: 10.1109/ACCESS.2023.3322167.

[7] Open Power System Data. 2020. *Data Package Weather Data*. Version 2020-09-16. https://doi.org/10.25832/weather_data/2020-09-16.

[8] T. A. Farrag and E. E. Elattar, "Optimized Deep Stacked Long Short-Term Memory Network for Long-Term Load Forecasting," in IEEE Access, vol. 9, pp. 68511-68522, 2021, doi: 10.1109/ACCESS.2021.3077275.

[9] Q. Chen, M. Xia, T. Lu, X. Jiang, W. Liu and Q. Sun, "Short-Term Load Forecasting Based on Deep Learning for End-User Transformer Subject to Volatile Electric Heating Loads," in IEEE Access, vol. 7, pp. 162697-162707, 2019, doi: 10.1109/ACCESS.2019.2949726.

[10] "Data access viewer, for weather data at a single point, https://power.larc.nasa.gov/data-access-viewer/

[11] I. M. Mehedi et al., "Intelligent Machine Learning with Evolutionary Algorithm Based Short Term Load Forecasting in Power Systems," in IEEE Access, vol. 9, pp. 100113-100124, 2021, doi: 10.1109/ACCESS.2021.3096918.

[12] “Renewable ninja for solar, wind and weather data”, https://www.renewables.ninja/

[13] P. Popovski, G. Veljanovski, M. Kostov and M. Atanasovski, "Optimizing Short Term Load Forecast: A study on Machine Learning Model Accuracy and Predictor Selection," 2022 57th International Scientific Conference on Information, Communication and Energy Systems and Technologies (ICEST), Ohrid, North Macedonia, 2022, pp. 1-4, doi: 10.1109/ICEST55168.2022.9828783.

[14] M. Bhatnagar, V. Dwivedi, D. Singh and G. Rozinaj, "Comprehensive Electric load forecasting using ensemble machine learning methods," 2022 29th International Conference on Systems, Signals and Image Processing (IWSSIP), Sofia, Bulgaria, 2022, pp. 1-4, doi: 10.1109/IWSSIP55020.2022.9854390.

[15] [3] E. A. El Aziz Ahmed, R. A. Ibrahim and A. K. Abdelsalam, "A Comparative Analysis for Machine Learning-based Short-Term Load Forecasting Techniques," 2023 IEEE 6th International Electrical and Energy Conference (CIEEC), Hefei, China, 2023, pp. 1166-1171, doi: 10.1109/CIEEC58067.2023.10165934.

[16] D. Smith, K. Jaskie, J. Cadigan, J. Marvin and A. Spanias, "Machine Learning for Fast Short-Term Energy Load Forecasting," 2020 IEEE Conference on Industrial Cyberphysical Systems (ICPS), Tampere, Finland, 2020, pp. 433-436, doi: 10.1109/ICPS48405.2020.9274781.