Solar-Net:LeveragingTransformersfor EnhancedSolarPowerPrediction

MSc Research Project
MSc in Data Analytics

AKIF ROSHAN SHAIK
StudentID:23204231

SchoolofComputing
NationalCollegeofIreland

Supervisor: Vladimir Milosavljevic

National College of Ireland
Project Submission Sheet School of Computing

Student Name:	AKIF ROSHAN SHAIK
Student ID:	23204231
Programme:	MSc in Data Analytics
Year:	2024
Module:	MSc Research Project
Supervisor:	Vladimir Milosavljevic
Submission Due Date:	29/01/2025
Project Title:	Solar-Net: Leveraging Transformers for Enhanced Solar Power Prediction
Word Count:	7424
Page Count:	20

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature:	Akif Roshan Shaik
Date:	29 January 2025

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST:

Attach a completed copy of this sheet to each project (including multiple copies).	□
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	□
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	□

Assignments that are submitted to the Programme Coordinator office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	

Date:	
Penalty Applied (if applicable):	

Solar-Net: Leveraging Transformers for Enhanced Solar Power Prediction

Akif Roshan Shaik
23204231

Abstract

Accurate forecasting of electricity generation from solar systems is very critical for efficient grid balancing and integration of renewable energy sources. This study explores the application of machine learning (ML) and deep learning (DL techniques to predict solar power generation using various environmental data. Specifically, it evaluates the performance of various traditonal ML and DL algorithms and proposed Transformer-based model, Solar-Net. The dataset used for experimnets in this study has 21 independent variables including temperature, humidity, radiation, wind speed and the solar power output which is the dependent variable expressed in kilowatts.The results have shown that Solar-Net Transformer model had higher predictive accuracy than the traditional ML models with an R^2 score of 81.54%, the lowest MSE of 0.1906 and the lowest MAE of 0.2899. For Transformer-based models with attention, the study demonstrates the possibility of using this approach to increase the accuracy of forecasting of solar power generation based on time series data. Future work will include the consideration of more sources of information, improving the models' accuracy in real-time predictions, and extending the applicability of the models for different geographic areas and conditions. The findings of this research will be useful in the existing literature on solar energy forecasting as well as the use of deep learning in renewable energy systems.

1 Introduction

1.1 Background and Context

Global transition to renewable energy sources is one of the most effective measures in addressing climate change and achieving sustainable development. Solar power, because of its availability and advantages over conventional electricity sources, plays crucial role in this transition. However, the ability of solar power to enhance power generation is irregular and depends on weather factors, including solar irradiance, temperature as well as cloudiness. Variability of wind generation presents problems for grid operators who are required to balance supply and demand, typically with backup systems (Kim et al.; 2023; Azizi et al.; 2023). Precise solar power forecasting is critical for efficient operation of the solar farm, stability of the electrical grid, and overall reduction in the reliance on conventional generating sources. Statistical methods and classical machine learning algorithms have been used previously in solar power forecasting. However, these methods do not largely address the problem of non-linearity of the system under consideration, the relation between the output of solar power and the environment. The drawbacks of these models particularly in modeling long-term dependencies in time-series data call for better models (Gaboitaolelwe et al.; 2023).

Recent development in the deep learning models especially the Recurrent Neural Networks (RNN) and the Convolutional Neural Networks (CNN) has been seen to address these challenges. RNNs and CNNs show remarkable performance in dealing the local pattern and short-term temporal constraint learning in the available data. However, processing and modeling of long range dependencies are restricted due to architectural constraints and this aspect leads to sub-optimal forecast accuracy. Transformers, initially introduced for natural language processing, provide a solution for this issue, as they enable better capturing of long-range dependencies by means of self-attention mechanisms(Vaswani et al.; 2017). This makes transformers especially applicable for solar power prediction where accurate prediction depends on the temporal dependencies over long time periods.

1.2 Motivation

The rising utilization of solar energy means that efficiency of solar power forecasting is imperative in the energy market. Some of the current approaches fail to address the complexity and the stochastic nature of the solar power generation system because they cannot adequately model long term temporal dependencies or incorporate external environmental factors. As transformers are capable of handling sequential data in parallel and can capture the dependency over multiple time horizons, they might be a viable solution to the existing forecasting models.

This research presents Solar-Net, a custom transformer-based model for solar power prediction that has been developed in this work. Solar-Net able to capture temporal dependencies with the ability of transformer attention mechanism using external conditions like weather data. Solar-Net aims to eliminate the shortcomings of previous models and increase the effectiveness of forecasts. Use of the transformer architecture allows SolarNet to provide an innovative solution to the problem of variability of power generation from solar energy, which will help to develop more efficient and reliable systems.

1.3 Research Questions

The study is guided by the following research question:

"How can Solar-Net, a transformer-based model, revolutionise solar power prediction – capturing intricate temporal patterns of variability and in what ways does the use of deep learning enhance the accuracy of solar power forecasting?"

This question aims to find out how Solar-Net that uses the transformer architectures to advance the forecasting of the solar power generation and the ability of the deep learning techniques in improving the model. In answering this question, the study aims at filling this gap through investigating transformers' ability for solar power forecasting to enhance the design of more efficient renewable energy systems.

1.4 Research Objectives

The objectives of this research are:

- Develop Solar-Net, a customized transformer model for forecasting solar power utilizing the long range temporal dependencies in the data using attention mechanism of transformer based architectures.

- Comparison of the results of the proposed approach Solar-Net with other machine learning algorithms such as SVR, Random Forest Regression, GBM, LSTM, and deep learning models.

- Incorporate ensemble learning strategies in order to enhance Solar-Net's global stability and prediction abilities in various conditions and with various structures of test data sets.

- Assess the suitability of Solar-Net by performing simulations with realistic data of solar energy generation, thus proving the improvement of the reliability of the solar power systems.

Therefore, renewable energy especially solar power, is crucial to combating climate change and guarantee energy sustainability. Nevertheless, variability is an essential feature of solar power generation because the output is contingent on such factors as weather and seasons Hence, the forecasting of its output is characterized more by challenges than by certainty. Conventional methods of forecasting present a disadvantage in the sense that they are unable to model the relationship between solar energy generation and the weather factors in a manner that is non-linear. However, recent progress in deep learning especially the transformer models provide this solution because they are capable of modelling long-term dependencies and temporal features. This research presents Solar-Net, a transformer-based model to address the existing approaches' limitations and enhance the solar power forecast accuracy. By using the self attention mechanism of transformer architecture, Solar-Net will be a novel solution of improving the reliability and effectiveness of solar energy systems. The study aims at providing significant information on how the sophisticated deep learning methodologies can be applied to enhance the solar power prediction process and consequently advance renewable power systems and grid reliability.

2 Literature Review

2.1 Introduction

Solar energy is one of the most popular renewable energy sources because of its availability and the least effects on the environment. However, due to the variability of solar energy that is dependent on weather conditions, it is difficult to incorporate into power grid and, therefore, requires solar power forecasting. Many methods have been proposed for solar power forecasting, starting from simple statistical ones and ending up with machine learning and deep learning methods. This section provides a discussion of the state-ofart in solar power forecasting and the selection of deep learning models including RNNs, CNNs, and transformers. The review also points out the shortcomings of the existing strategies and suggests the possible areas of using the transformer models including the Solar-Net model developed in this work.

2.2 Solar Power Prediction: Traditional Approaches

The initial approaches to the solar power forecasting were based on the statistical models including ARIMA and other time series models. These models are based on linear correlations and cannot reflect nonlinear dependencies as a result of various factors that affect solar power generation (Atique et al.; 2020). However, these traditional methods are easy to implement and require little computational resources, but they have low accuracy when it comes to providing forecasts for more distant time horizons (LaraBen´ıtez et al.; 2020).

Advanced ML techniques like SVR, Random forest regression, and GBM have revealed that the advance AI models perform better than the statistical benchmarks by capturing the non-linearity (Yadav et al., 2021). For instance, SVR has been proved to perform well when solving small data and time series data but fails to handle high-dimensional data and non-stationary series (Gaboitaolelwe et al.; 2023). Random Forest and GBM are two methods of ensemble learning which decrease the amount of overfitting and increase the predictability by combining a number of decision trees (Kim et al.; 2023). Nevertheless, these models have some drawback in modeling long term dependencies and temporal relationships in sequential data.

2.3 Deep Learning Models for Solar Power Forecasting

The deep learning models have been found to be quite useful for solar power forecasting because they do not need feature engineering. RNNs and its modifications LSTM and GRU has been used in many time series forecasting applications including solar power prediction (Salman et al.; 2024). Unlike other recurrent models, RNNs have hidden status where information from previous times is stored to cater for sequences. However, standard RNNs have a vanishing gradient problem, which makes it difficult for them to capture long term dependencies (Chung et al., 2014).

The LSTM and the GRU model do that by introducing memory cells and gating mechanisms to enable the model to utilize the input data for longer periods (Hochreiter & Schmidhuber, 1997). These models have been found to provide enhanced solar power forecasting by incorporating short term and long term temporal characteristics (Elizabeth Michael et al., 2022). However, LSTM and GRU models are computation intensive and have a problem with parallelism because of their sequential nature and therefore not very scalable (Vaswani et al., 2017).

Convolutional Neural Networks (CNNs) are normally applied in image processing, but have also been implemented in time series forecasting. CNNs are highly effective in capturing local patterns via convolution layer, thus they are good for such tasks which involve identification of short-term pattern (Azizi et al.; 2023). For solar power forecasting, CNNs have been demonstrated to learn localized temporal characteristics, but their effectiveness declines when it comes to learning long-term temporal patterns (Elsaraiti and Merabet; 2022).

2.4 Transformers and Their Application in Time-Series Forecasting

Transformers by (Vaswani et al.; 2017) formed a base for many new models expanding perspectives in natural language processing by considering long-term dependencies based on self-attention mechanisms. This is advantageous over RNNs and LSTMs for they have a capability of processing all the inputs in parallel and are thus efficient in capturing global dependency (Vaswani et al.; 2017). Since then, transformers have been used for different time series forecasting problems such as weather forecasting, traffic flow prediction, and energy load prediction (Liu and Fu; 2023).

The self-attention mechanism used in the transformer model exposes more significant onboard importance to various time steps in the input sequences, and it is therefore capable of performing what long-term dependencies are crucial in determining outcomes in different tasks (Zhou et al.; 2021). However, transformers are still relatively underutilized in the solar power forecasting domain despite their effectiveness in other time-series domains. The capability of transformers to encode both short-term and longterm dependencies makes them a potential solution to enhance the solar power predictions.

Recent research works have started to analyze the application of transformer-based models in renewable energy forecasting. (Liu and Fu; 2023) have proved that transformers are suitable for energy load forecasting and that the proposed models outperformed RNN and CNN models in terms of accuracy. In the latter work, (Zhou et al.; 2021) presented Informer, a transformer-based model for long-sequence time-series forecasting that outperformed other models in various energy prediction tasks. These studies indicate that transformers could capture temporal dynamics and that models based on transformers such as Solar-Net could have much better forecast accuracy for solar power.

2.5 Limitations of Existing Models

While using such deep learning models as RNNs, LSTMs and CNNs the accuracy of predictions of solar power has been enhanced, there are still the following limitations: The RNN-based models have problems with long-term dependencies because of the sequential structure of the architecture; it causes problems with time consumption and with finding global dependencies (Salman et al.; 2024). CNNs, even though can capture short-term temporal dynamics well, are not built conceptually for time-series learning and forecasting, and are less powerful in capturing long-range temporal relations (Azizi et al.; 2023).

2.6 Comparative Analysis of Literature on Solar Power Forecasting

To give a better picture of the current state of knowledge in the field of solar power forecasting and to distinguish between significant features of different methodologies, a brief comparison of research findings is presented in the following table. This table provides the authors, methodologies, models employed, key performance indicators, limitations, and the recommended future work of each research.

Table 1: Comparison of Existing Researches

Authos	**Methodology**	**Model Used**	**Limitations**	**Future Work**
Atique et al. (2020)	Time series forecasting using statistical methods (ARIMA)	ARIMA	Struggles with nonlinear relationships; limited ability to capture complex dependencies in solar data.	Explore machine learning and deep learning approaches for better accuracy.
Kim et al. (2023)	Ensemble learning with Random Forest and Gradient Boosting	Random Forest, GBM	Overfitting with smaller datasets; does not effectively handle long-term dependencies in timeseries data.	Integrate deep learning models to capture temporal patterns and improve generalization.
Gaboitaolelwe et al., (2023)	Machine learning comparison of SVR and Random Forest for solar prediction	SVR, Random Forest	SVR struggles with high-dimensional data; Random Forest faces scalability issues with larger datasets.	Investigate hybrid models combining SVR with deep learning models to address high-dimensional solar datasets.
Salman et al. (2024)	Hybrid deep learning with CNN and LSTM	CNN, LSTM	LSTM is computationally expensive; difficulty in scaling due to sequential nature of training.	Optimize LSTM models for reduced computational cost and faster training.
Azizi et al. (2023)	Adaptation of CNN for time-series forecasting of solar irradiance	CNN	CNN is effective at capturing local features but struggles with long-term dependency modeling.	Explore combining CNN with other models, such as transformers, to improve long-term dependency handling.
Elsaraiti et al. (2022)	Solar power prediction using deep learning methods	CNN, RNN	RNN struggles with vanishing gradient issues, and CNN cannot capture longterm dependencies.	Introduce attention mechanisms or transformer models to better handle temporal dependencies in solar data.
Lara-Ben´ıtez et al. (2020)	Review of time-series forecasting techniques for solar energy prediction	ARIMA, SVR, LSTM	Traditional methods fail to handle non-linear patterns; LSTM and SVR have scalability and long-term dependency challenges.	Explore transformer-based models to overcome longterm dependency issues in solar forecasting.

2.7 Discussion and Key Insights

The literature comparison table highlights several key insights and trends in solar power forecasting:

- Traditional Models such as ARIMA, SVR, do not work well with non-linear data and long-term dependencies, which is why they are not suitable for predicting complex solar weather patterns (Atique et al.; 2020; Gaboitaolelwe et al.; 2023). • Ensemble Learning Techniques including Random Forest and GBM work better than normal models by reducing the problem of overfitting; however, they also fail to address long-term temporal dependencies (Kim et al.; 2023).

- Deep Learning Models, particularly LSTM and CNN, are among the most effective specifically in capturing short term and localized trends but they are not fully scalable and do not support long term dependency modeling well (Salman et al.; 2024; Azizi et al.; 2023). LSTM models are computationally expensive, while CNNs are not inherently designed for time-series tasks.
- Transformer Models (Liu and Fu; 2023; Zhou et al.; 2021) are a major advancement in addressing long-range dependencies and enabling the parallelization of the learning process, making them suitable for solar power forecasting. But in the domain of knowledge management, their use is still quite limited.
-

Motivation for Solar-Net: The current models have some drawbacks, this research study will investigate the effectiveness of transformer-based models in enhancing solar power forecasting. Solar-Net, the model introduced in this study, is designed to address these limitations by incorporating two key enhancements: Contextual Data Embedding and Multi-Scale Temporal Attention. These improvements enable Solar-Net to incorporate the temporal characteristics and external information better than other models. Due to the self-attention mechanism of transformers, Solar-Net will be more effective in shortterm and long-term solar power forecasting tasks compared to existing machine learning and deep learning models (Liu and Fu; 2023; Zhou et al.; 2021).

By summarizing the existing literature on solar power forecasting which includes discussing the weakness of traditional and machine learning models but the potential of deep learning models to enhance the prediction. As seen before, the RNNs, LSTMs, and CNNs have been able to capture temporal patterns, but the problems with longterm dependencies and external information incorporation make it essential to consider transformer-based models. The Solar-Net model suggested in this research is designed to solve these issues with the help of transformer architecture and new ideas, such as CDE and MSTL. The next section will describe the method by which Solar-Net has been developed and assessed with respect to solar power prediction.

3 Methodology

This research study employed a well-defined and step-by-step approach to predict solar power generation by employing a combination of conventional machine learning algorithms and deep learning models with an emphasis on the Customly Designed Transformer model. The methodology is a few main steps, which are data collection and exploration, preprocessing of Solar Power Generation Dataset, data transformation, feature engineering, model building, and model evaluation. In the following section, each of these steps will be explained in detail.

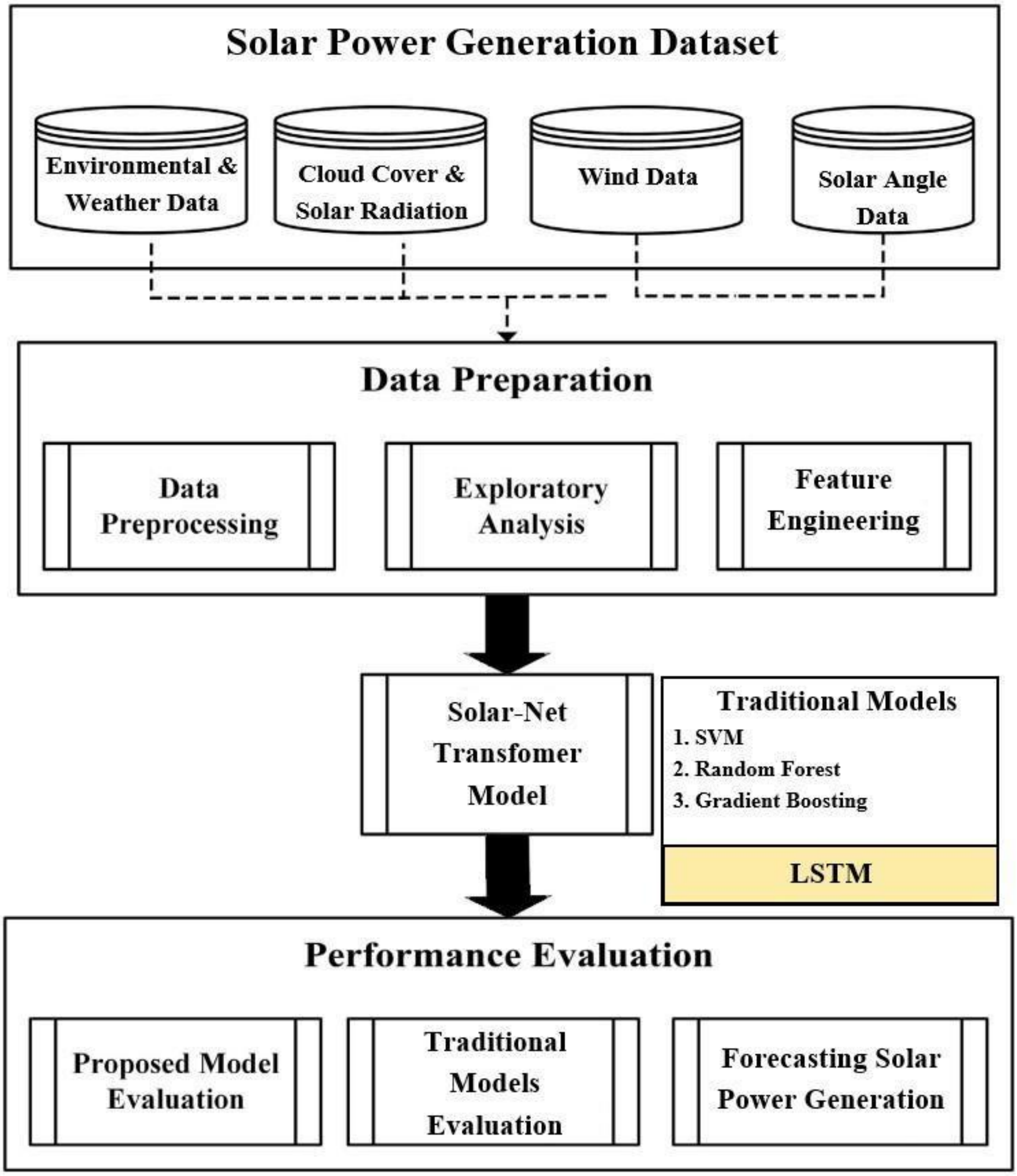

Figure 1: Research Methodology for Solar Power Prediction

3.1 Data Collection and Exploration

In this study, the data set used consists of a number of meterological and environmental variables that affect the generation of solar power. These are features like temperature, humidity, pressure, radiation, cloud cover and wind speed and direction and the target variable is the amount of electricity generated in kilowatts by the solar power plant. There are 21 features to the dataset and the dependent variable is ”generated power _kw”.

Firstly the solar power generation data is imported into a Pandas DataFrame which makes the analysis process easy to conduct. The structure and distribution of data are therefore first explored in a process known as Exploratory Data Analysis (EDA). This step involves the process of assessing the first few records often with a view to familiarizing ourselves with the data at hand as well as its summary characteristics such as types of features, presence of missing values, among others, as well as the assessment of the pattern of correlation between various features and the target variable. In this phase, there is a widespread use of visualizations.

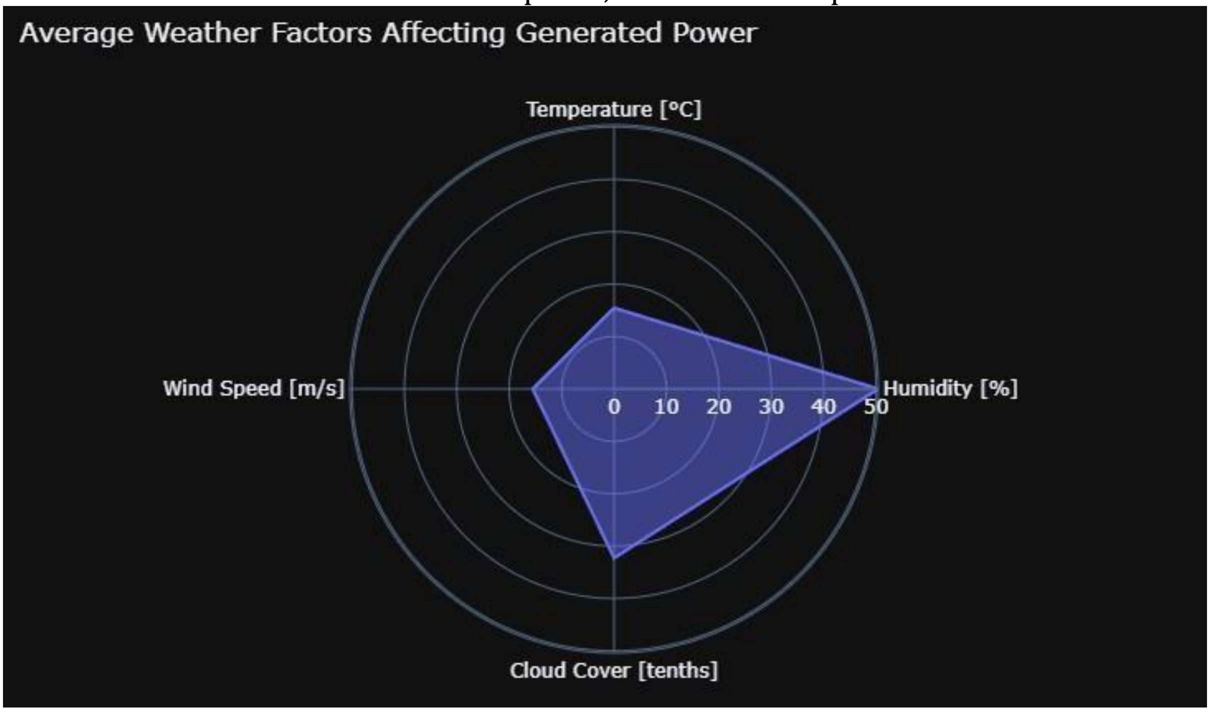

Figure 2: Average Weather Factors Affecting the Generation of Power

In the radar chart of Figure 2, the weather factors that influence the efficiency of solar power generation are identified with humidity at 50% and cloud cover at 32 tenths having the greatest impact as humidity affects the scattering of sunlight and cloud cover reducing the amount of solar radiation. It is also important as the temperature affects the efficiency of the solar panels, and as it is evident, the temperature is low (15°C) which is ideal for the solar panels. On the other hand, the wind speed (18 m/s) is higher than expected, which help cool the panels but require strong structures for installation due to the winds. These findings imply that temperature and wind speed are good supporting factors, however, the models that would be predictive or the operating strategies must consider the impacts of humidity and cloud cover to enhance the efficiency of solar power generation.

A correlation heatmap is produced to determine the degree of relationship between each feature and the target, which assists in choosing which predictors should be included in the models. Moreover, scatter and pair plots are employed in order to visualize the correlation between different meteorological variables and solar power generation, as well as to get a feeling of whether there is any pattern in the distribution.

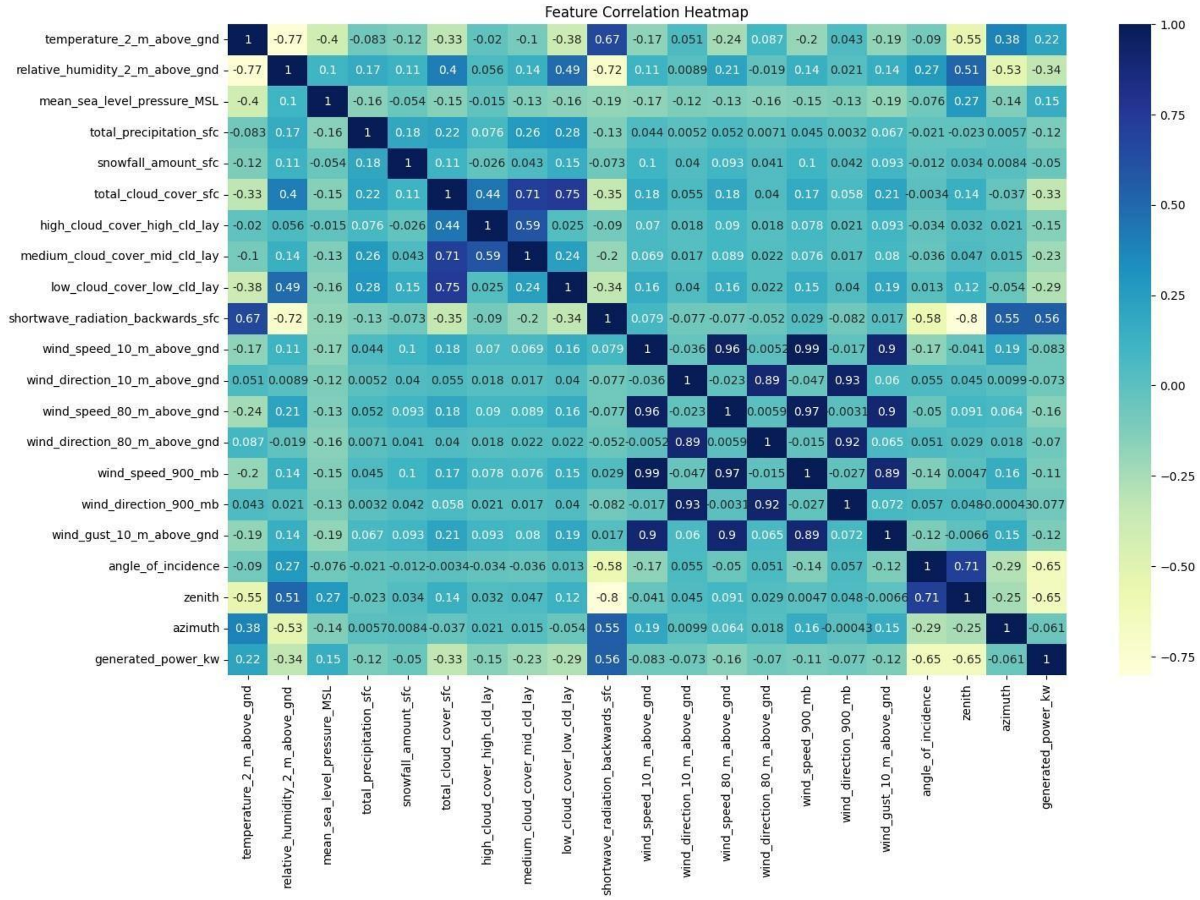

Figure 3: Correlation Heatmap for Solar Generation Power Dataset

3.2 Preprocessing of the Solar Power Generation Dataset

Data pre-processing is very important step as it prepares the data set for feeding into the model. It expedites data preparation which makes the data clean, complete and consistent for analysis. The first step under data preprocessing is how to handle missing values. If there are any missing values in the datasets they are replaced by the mean of the particular columns depending on the type of data it is. This helps to make certain that the database stays intact and no information got missing in the process. One of the other major task involves identifying and eradicating outliers which limit model efficiency and give wrong predictions. Outliers are located by Z-score, which is calculated as the number of standard deviations a specific value is away from the mean. An observation with a Zscore more than 3 is regarded as an outlier and is removed from the data set. This step is important so that the model is trained using data that only has typical cases.

Feature scaling is another important process of preprocessing that needs to be discussed. Some algorithms are sensitive to the scale of input features and in order to make all features scale invariant, it is recommended to standardize continuous variables. This is done using StandardScaler which balances the features so that it has zero mean and unit standard deviation. Standardization is especially relevant for such models as SVR that use distance-based measures and enhances the model's quality.

3.3 Feature Engineering

Feature engineering means the steps taken to gather raw data and prepare it for transformation to a more easily understandable form by the model. The first stage of feature engineering involves the removal of the target variable, generating power KW from the features. This creates a feature matrix X consists of all the predictors, and a target vector y consists of the corresponding Solar Power Output values ("generated _ power kw"). The given dataset is then divided into training dataset and testing dataset. This step means that, the developed model can be tested on data that has not been used in the development of the model so as to determine its capability of performing well generically. Typically, 80% of the data is used for training, while the remaining 20% is set aside for testing. This split allows for both model training and validation, providing an unbiased evaluation of the model's performance.

3.4 Model Development and Training

In this stage, there are a number of machine learning and deep learning models trained and assessed. All the basic algorithms from traditional machine learning and all the state of the art deep learning like LSTM and Transformer models are used.

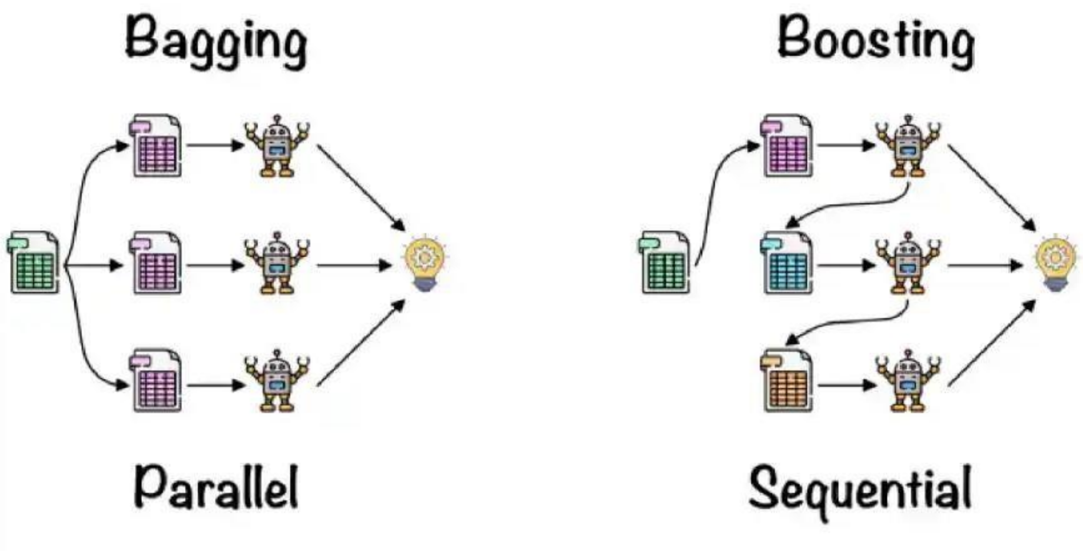

Figure 4: Working of Random Forest (Bagging) and Gradient Boosting Regression (Boosting) Models

3.4.1 Traditional Machine Learning Models • Support Vector Regression (SVR): SVR is used for regression problems, especially when more complex relationship between the features and the target needs to be modeled.The SVR model is refined using GridSearchCV with the type of kernel (linear, polynomial, or Use radial basis function), the value of C (regularization parameter) and epsilon (margin of tolerance) for its optimal performance (Yadav et al., 2021).

• Random Forest Regression (RF): Random Forest is another technique of the bagging model which combines several decision trees and makes the mean of their results to get more precise and less overfitting. The hyperparameters including the number of trees (estimators) and the minimum number of samples required to split a node are tuned at optimal values using GridSearchCV (Gaboitaolelwe et al., 2023). • Gradient Boosting Regression (GBM): Gradient Boosting is an ancestral technique of model ensemble in which models are trained sequentially in an iterative way in which each model attempts to minimize the errors of the previous model (Kim et al., 2023). Hyerparameters include learning rate, depth of tree, number of boosting iterations etc. have been tuned to give the best performance of the model.

3.5 Deep Learning Models

Then, deep learning models are discussed, with the emphasis on time-series prediction and the use of Transformer-based models for improved temporal relationships in the data.

- **Long Short-Term Memory (LSTM):** LSTM is a kind of RNN used for sequences data which include time series data. To address this in this study, the dataset is then reshaped to a 3D tensor to conform to the input format expected by the LSTM as stated by Salman et al. (2024). The model is trained using the Adam optimizer by minimizing the error and the dropout layer for the purpose of generalization.

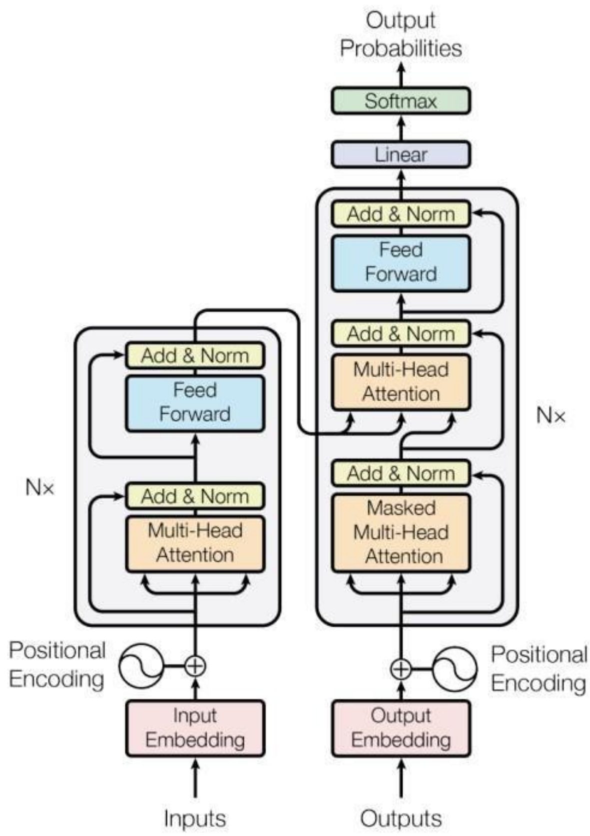

Figure 5: Architecture of Transformer Model

- **Transformer Model (Solar-Net):** The most emphasis of the research is given to the Transformer-based models, though. These models, especially widely used for natural language processing, are very efficient in the extraction of long-range dependencies in time series data. In the proposed Solar-Net Transformer model, multi-head attention is employed to enable the model to pay attention to the attention heads of the input sequence to grasp various relations among features. This design allows the model to achieve higher accuracy for solar power generation prediction than previous models. Hyperparameters for the Transformer model are tuned using KerasTuner's RandomSearch where several features such as the number of head attentions, feed-forward dimensions, and dropout rates are varied.

3.6 Model Evaluation

Once the models are trained it is important to assess the performance of our model using the right evaluation metrics. In this research study, three key evaluation metrics are used to assess the predictive accuracy of the models:

After all models are trained and tested the desired performance evaluation metrics are compared (R2, MSE, and MAE). The transformer based Solar-Net, and other deep learning models are are expected to perform better than the traditional models because of their capability to capture temporal dependencies and non-linear relationships in the data. However, while choosing a model, one has to consider practical aspects such as computational complexity, model interpretability, and insensitivity to noise. The final model for predicting solar power generation is chosen as the model that minimizes prediction error while being at the same time feasible in terms of its complexity. Evaluating numerous machine learning and deep learning models for estimating the solar power generation. Hence, this research will seek to determine the appropriate model, whether the simple regression or the complex deep learning model, that yields the best results and can be used to optimize management of solar energy systems.

4 Model Evaluation Results, Findings, and Discussion

In this section, the performance of different machine learning and deep learning models that are applied to the problem of solar power prediction is presented. The evaluation utilized three primary regression metrics: From the results section, one can identify the R^2 score, Mean Squared Error (MSE), and Mean Absolute Error (MAE), which give a general insight into the model's predictive power and its capacity to explain the variation in the solar power output . These models were trained with different combinations of training and test set to determine the effect of different set sizes.

The models were evaluated across three different dataset splits: 70:30, 75:25, and 80:20. These splits enable one to compare the performance of the models when the proportion of data used for training is varied in order to determine the performance of the models on unseen data.

4.1 Model Evaluation Findings

The evaluation results is as follows: deep learning models especially the Transformerbased Solar-Net models surpasses the traditional ML models in all the splits of the dataset. These findings confirm the effectiveness of deep learning approaches – especially when it comes to addressing multiple-layered problems like those associated with solar power prediction.

Table 2: Evaluation Results of the Models (70:30 Ratio)

Model	R^2 Score (%)	MSE	MAE
SVM	73.45	0.268	0.411
RF	77.33	0.229	0.328
GBM	76.89	0.234	0.358
LSTM	79.48	0.207	0.320
Solar-Net Transformer	81.38	0.188	0.282

From Table 2, can observe that deep learning models outperformed ML models. Specifically, Solar-Net model able to outperform all the baselines which demonstrates the effectiveness of the transformer based architectures specifically due to able to capture range dependencies using self attention mechanism.

Table 3: Evaluation Results of the Models (75:25 Ratio)

Model	R^2 Score (%)	MSE	MAE
Support Vector Machine	76.39	0.244	0.376
Random Forest	77.36	0.234	0.329
Gradient Boosting	76.71	0.241	0.369
LSTM	78.98	0.217	0.326
Solar-Net Transformer	81.17	0.195	0.284

Similar trend have been observed in Table 3 adn Table 4 in results of 75:25 and 80:20 split ratio of data , can observe that deep learning models outperformed ML models. Specifically, Solar-Net model able to outperform all the baselines in all the splits demonstrates its ability.

Table 4: Evaluation Results of the Models (80:20 Ratio)

Model	R^2 Score (%)	MSE	MAE
SVM	74.77	0.260	0.406
RF	77.65	0.230	0.326
GBM	76.79	0.239	0.364
LSTM	80.24	0.204	0.312

Solar-Net Transformer	81.54	0.190	0.289

The lowest accuracy was achieved by the Support Vector Machine (SVM) model. For the 70:30 split, the R^2 of 73.45% meant that SVM only accounted for a minor proportion of the variance in the solar power generation. Again, the performance of SVM enhanced slightly with increased training data as it yielded an R^2 of 76.39% and 74.77% in 75:25 and 80:20 splits, respectively. However, these improvements still made SVM less favorable and less efficient as compared to the other models, especially the deep learning models. Furthermore, the MSE and MAE of SVM were higher, which suggested that the actual and the predicted values of this model were far from the real values. This implies that although SVM can be used as a starting reference for less complex tasks, it lacks the ability to model the temporal dependencies inherent in the solar power data.

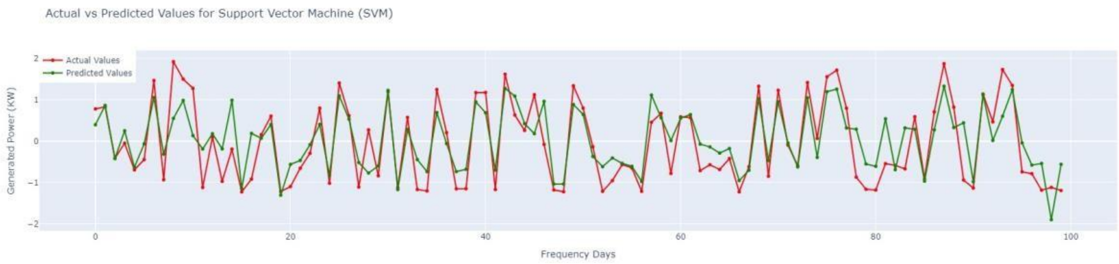

Figure 6: Actual vs Predicted Values for Support Vector Machine (SVM)

The Random Forest, an ensemble learning model which is a collection of multiple decision trees; the R^2 values this time varied from 77.33 to 77.65 across the splits of the dataset. Thus, Random Forest was found to have a better capacity of explaining more variance in the production of solar power. Also, as it applied to the MSE and MAE results, the proposed model made more accurate predictions than the SVM model. Random Forest outperforms other decision trees in terms of capturing the non-linear relationship between the features and is less sensitive to the overfitting problem. However, it is observed that it provides better accuracy than SVM but it was not as accurate as the deep learning models. This may be attributed to its poor ability to capture the temporal dependencies that are inherent in the solar power data.

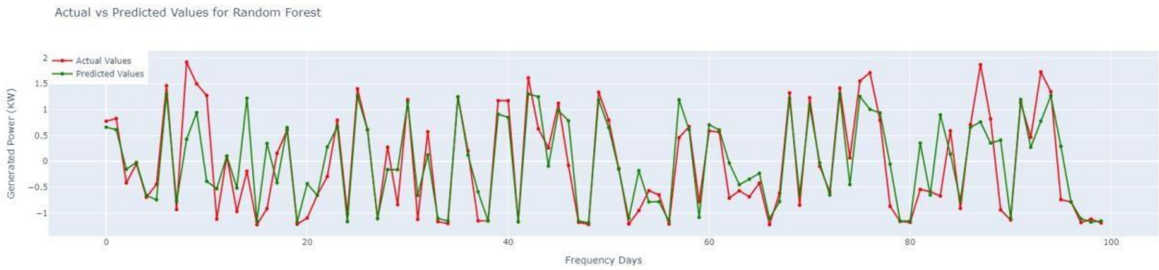

Figure 7: Actual vs Predicted Values for Random Forest

The another method of ensemble learning is Gradient Boosting, which yielded the result of 76.79% of R^2 that is just a little lower than Random Forest but still very good. Gradient Boosting also employs decision trees as the base model, but the model building process is very different from Random Forest – each tree added to the model tries to minimize the error made by the previous trees in the model. It can also enhance the model's calibration by providing more emphasis on the difficult-to-score observations. However, it is found that Gradient Boosting has MSE and MAE values which is higher than Random Forest, which signifies that the model was not able to predict the errors correctly. While Gradient Boosting surpassed the accuracy of SVM, it did not reach the level of accuracy provided by deep learning models, which again underlines the superiority of the deep learning method in dealing with time-series forecasting, like solar power prediction.

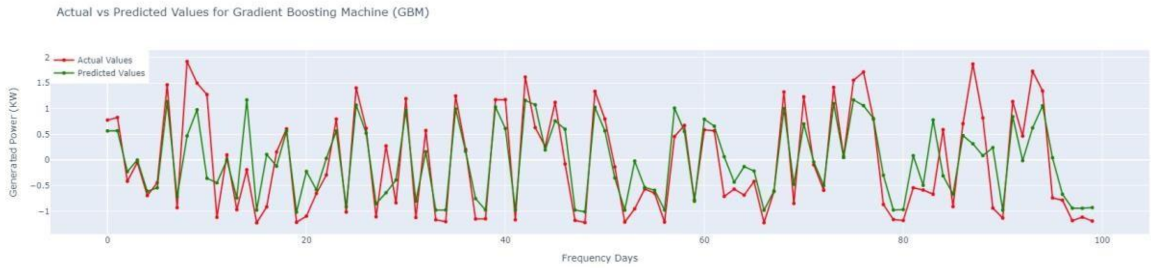

Figure 8: Actual vs Predicted Values for Gradient Boosting Machine (GBM)

The Long Short-Term Memory (LSTM) network intended for sequential data brought in more accuracy into the prediction. LSTM achieved R^2 scores of between 78.98% and 80.24% which meant that the network was able to predict a large portion of the variance in the solar power generation. This performance shows that LSTM is particularly suitable for the task of modeling time-series data and the temporal dependencies that are typical of the problem. The MSE and MAE for LSTM where lower than MSE and MAE of other traditional machine learning models, which confirmed that LSTM was more accurate in its prediction. Therefore, LSTM model was especially suitable for solar power generation forecasting as it could learn from the present time steps and forecast the future values based on the long-term dependencies.

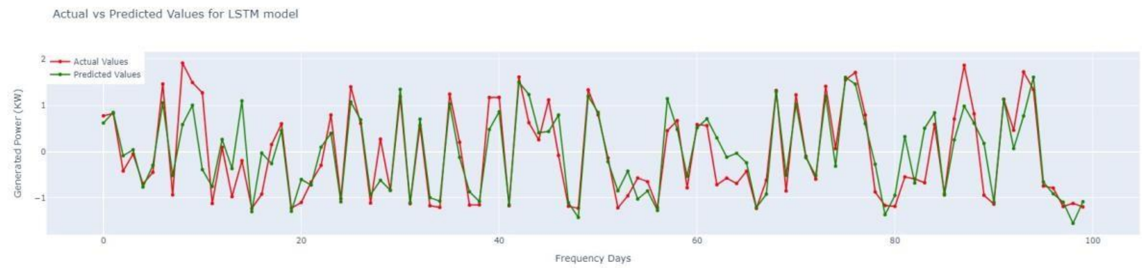

Figure 9: Actual vs Predicted Values for Long-Short Term Memory (LSTM)

Solar-Net Transformer came out as the best model overall, with an R^2 of 81.54% in 80:20 split. This result shows that the Transformer model achieved higher performance than LSTM model or any other traditional models tested herein to explain variance in solar power generation. The model's ability to attend to certain aspects of the data allowed it to capture intricate relations and dependencies between inputs and their responses, and therefore make more accurate predictions. Furthermore, the MSE and MAE scores for the Transformer model were even lower than those of the other models, which confirms the model's higher predictive accuracy. For the improvement of the performance, the Transformer model incorporated the attention mechanism where only the important features and time steps were considered. The hyperparameters including the number of attention heads, the dimension of feed-forward layer, and dropout ratio, were also tuned to achieve remarkable performance of the Transformer.

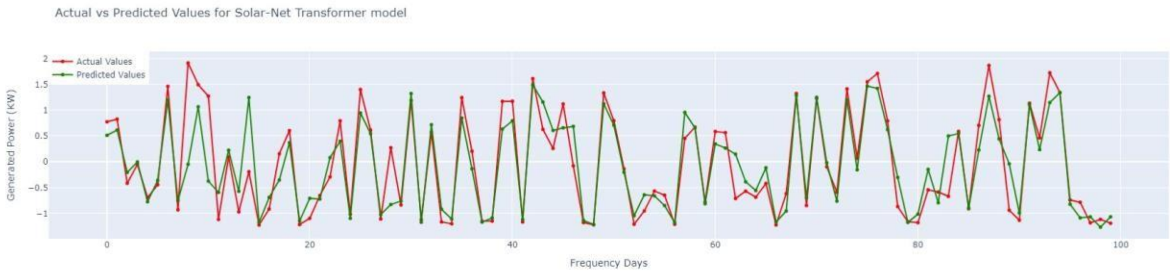

Figure 10: Actual vs Predicted Values for Solar-Net Transformer Model

4.2 Discussion & Interpretation of Findings

The evaluation results show that deep learning models are more effective than the traditional machine learning algorithms in the solar power prediction. Although models such as SVM, Random Forest, and Gradient Boosting gave fairly good results, the deep learning models including the LSTM and the Solar-Net models based on Transformer provided much better and more consistent forecasts. The above deep learning models performed well in modeling temporal and sequential characteristics of the data which are critical in applications like forecasting of solar power generation.

When comparing the models using three different splits of the dataset (70:30, 75:25 and 80:20), the trends were similar, with the deep learning models performing better than traditional models in all splits. In particular, the Solar-Net Transformer outperforms others for all three splits based on the higher values of R^2, the lower levels of MSE and MAE. As it was seen in the previous sections, the Transformer model had a slightly higher R^2 score when more data was given (from 70:30 split to 80:20 split, R^2 = 81.54%).

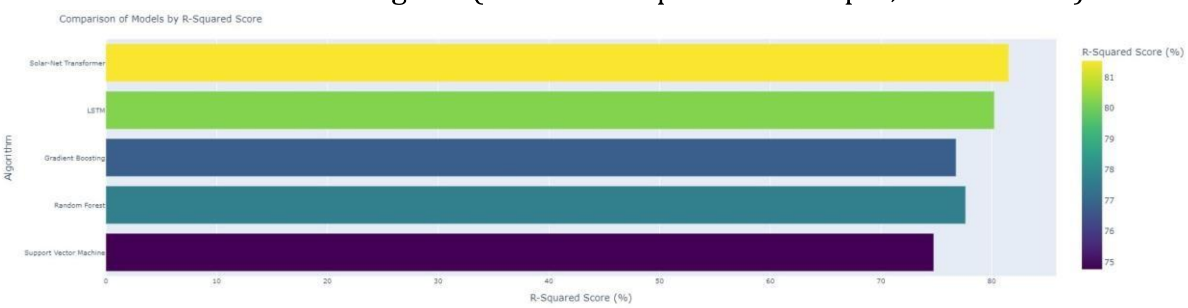

Figure 11: Comparison of R-Squared Score Among All Models

LSTM also appeared to do quite well, especially on data sequential in nature. The R^2 scores of its ranged from 79. 48% to 80. 24% across different splits of dataset revealing the capability of learning long term dependencies in the time series data. It was always consistently and slightly outperformed in every dataset split by the Solar-Net Transformer. This implies that the attention mechanism within the Transformer model is more appropriate to capture interactions and long-range dependencies that are essential for enhancing the accuracy of predicting solar power.

The classical machine learning algorithms including SVM, Random Forest, and Gradient Boosting showed good performance, but could not capture the complex dependencies which existed in the solar power generation data well. SVM, for instance, presented the lowest performance in every split, with R^2 = 73.45% - 74.77%, and MSE and MAE, indicating potential large deviations from real solar power generation. The two algorithms that had the best performance were Random Forest and Gradient Boosting with R^2 values of between 77.33% - 77.65% for Random Forest and 76.71%-76.89% for Gradient Boosting. However, even in their best state, the two models still could not match the performance of the deep learning models, which showed the limitations of the two models in capturing the temporal dependencies and long-range interactions that are important in accurate prediction of solar power.

The results obtained on all three dataset splits reaffirm the need to use more complex models like deep learning based models for sequence and time series data like solar power forecasting. Random Forest and Gradient Boosting can work with non-linear correlation properly, but, they cannot model the sequential features present in the data as LSTM and the Transformer-based Solar-Net model does.

In conclusion, the Solar-Net Transformer model had the highest R^2 score, the lowest MSE and MAE for all three dataset splits, which makes it the best model to use for predicting solar power generation in this study. Some of the advantages of this model include its capability to model long-range dependencies, and non-linear temporal interdependencies which makes it the best model for solar power forecasting. The further studies could investigate the deeper improvements of the Transformer model, real-time predictions, and incorporation of more data streams as these improvements could illustrate the utility of the model in the solar energy systems. Moreover, the ability of the proposed model to be implemented in real life practice like grid management and energy optimization can also be tested.

5 Conclusion and Future Work

5.1 Conclusion

The purpose of this study was to build and assess machine learning and deep learning models for forecasting the solar power generation given meteorological and environmental attributes. A wide set of models was explored, including more classical machine learning methods like SVM, Random Forest, and Gradient Boosting, as well as deep learning architectures like LSTM and Transformer-based model Solar-Net. The evaluation of these models was performed using key regression metrics: The second is the R^2 score, the third is Mean Squared Error (MSE), the fourth is Mean Absolute Error (MAE).

The results showed that deep learning models are superior to conventional machine learning methods in all criteria, mainly the Transformer-based Solar-Net model. The Solar-Net Transformer model returned the highest R^2 of 81.54%, meaning that the model was able to account for most of the variability in the amounts of solar power generated. It also had the lowest MSE of 0.1906 and the MAE of 0.2899, which proves that the present model was the most accurate and efficient in its prediction. The Long Short-Term Memory (LSTM) model was also tested and achieved an R^2 score of 80.24% which, however, was slightly lower than the Transformer model. Nonetheless, conventional machine learning such as Support Vector Machine (SVM), Random Forest, and Gradient Boosting were revealed to have lower R^2 values and higher prediction errors. Although these models were useful in identifying some of the patterns in the data, they were not as successful in doing this as the deep learning models, especially the Transformer model which performed well in the analysis of the sequential data provided by the time series.

In summary, the results show that Transformer-based models with attention modules are quite suitable for forecasting solar power generation since they perform better and are more accurate than conventional models. This research adds to the existing literature on the use of deep learning methods in renewable energy forecasting, especially solar energy, and underlines the significance of such models as Solar-Net Transformer for realtime solar power prediction and energy management.

5.2 Future Work

While this study demonstrates the effectiveness of Transformer-based models for solar power prediction, there are several areas for future research and improvement. • RealTime Solar Power Forecasting: One of the primary challenges in solar power forecasting is the need for real-time predictions that can be used for grid management and energy distribution. Future work could focus on deploying the Transformer model in a real-time forecasting system, where the model continuously ingests new weather data and generates short-term solar power predictions. This would provide immediate and actionable insights for optimizing the usage of solar energy.

- Scalability and Generalization: This study used a single dataset to train and evaluate the models. Future work could focus on testing the generalizability of the models by applying them to other solar power datasets from different geographic regions. This would assess whether the models can be scaled effectively and generalized across various climates and locations.

- Hybrid Models: As an extension of the current study, hybrid models that combine multiple machine learning techniques, such as the integration of Random Forest or Gradient Boosting with Transformer models, could be explored. This approach might leverage the strengths of both models—such as the interpretability of treebased models with the deep learning capabilities of Transformers—leading to even more robust and accurate forecasting.

- Transfer Learning and Pretrained Models: Another promising direction for future research is exploring the use of transfer learning. Pretrained models trained on large solar datasets or other time-series forecasting tasks could be fine-tuned to specific solar power prediction tasks. This would reduce training time and potentially improve model performance by leveraging existing knowledge in the field.

By addressing these areas in future studies, the accuracy, scalability, and real-world applicability of solar power prediction models can be significantly enhanced, contributing to more efficient solar energy management and optimization systems.

References

Atique, S., Noureen, S., Roy, V., Bayne, S. and Macfie, J. (2020). Time series forecasting of total daily solar energy generation: A comparative analysis between arima and machine learning techniques, *2020 IEEE Green Technologies Conference (GreenTech)*, IEEE, pp. 175–180.

Azizi, N., Yaghoubirad, M., Farajollahi, M. and Ahmadi, A. (2023). Deep learning based longterm global solar irradiance and temperature forecasting using time series with multistep multivariate output, *Renewable Energy* **206**: 135–147.

Elizabeth Michael, N., Mishra, M., Hasan, S. and Al-Durra, A. (2022). Short-term solar power predicting model based on multi-step cnn stacked lstm technique, *Energies* **15**(6): 2150.

Elsaraiti, M. and Merabet, A. (2022). Solar power forecasting using deep learning techniques, *IEEE Access* **10**: 31692–31698.

Gaboitaolelwe, J., Zungeru, A. M., Yahya, A., Lebekwe, C. K., Vinod, D. N. and Salau, A. O. (2023). Machine learning based solar photovoltaic power forecasting: a review and comparison, *IEEE Access* **11**: 40820–40845.

Kim, Y. J., Kim, N. H., Park, S. Y., Kim, C. K., Oh, M., Kim, H. G. and Lee, Y. S. (2023). Developing prediction models for solar photovoltaic energy generation using statistical and machine learning methods, *Energy* .

Lara-Ben´ıtez, P., Carranza-Garc´ıa, M., Luna-Romera, J. M. and Riquelme, J. C. (2020). Temporal convolutional networks applied to energy-related time series forecasting, *Applied Sciences* **10**(7): 2322.

Liu, J. and Fu, Y. (2023). Renewable energy forecasting: A self-supervised learning-based transformer variant, *Energy* **284**: 128730.

Salman, D., Direkoglu, C., Kusaf, M. and Fahrioglu, M. (2024). Hybrid deep learning models for time series forecasting of solar power, *Neural Computing and Applications* pp. 1–18.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N. and Polosukhin, I. (2017). Attention is all you need, *Proceedings of the 31st International Conference on Neural Information Processing Systems (NeurIPS)*, pp. 5998–6008.

Wu, H., Liu, Q. and Wang, Y. (2021). Transformer-based models for renewable energy forecasting, *Energy AI* **3**: 100042.

Zhou, H., Zhang, S., Peng, J., Zhang, S., Li, J., Xiong, H. and Zhang, W. (2021). Informer: Beyond efficient transformer for long sequence time-series forecasting, *Proceedings of the AAAI conference on artificial intelligence*, Vol. 35, pp. 11106–11115.