成都理工大學 中英合作办学
CDUT Sino-British Collaborative Education

OXFORD BROOKES UNIVERSITY

UNDERGRADUATE PROJECT REPORT

Project Title:	**Solar Irradiance Prediction using Attention-SolarMeNet**
Surname:	**Duan**
First Name:	**Wenhui**
Student Number:	**202118010202**
Supervisor Name:	**Grace Ugochi Nneji**
Module Code:	**CHC 6096**
Module Name:	**Project**
Date Submitted:	**May 6, 2025**

Chengdu University of Technology Oxford Brookes College

Chengdu University of Technology

BSc (Single Honours) Degree Project

Programme Name: Projects

Module No.: CHC 6096

Surname: Duan

First Name: Cora

Project Title: Solar Irradiance Prediction using Attention-SolarMeNet

Student No.: 202118010202

Supervisor: Grace Ugochi Nneji

Date submitted: 6th May 2025

A report submitted as part of the requirements for the degree of BSc (Hons) in Computer Science

At

Chengdu University of Technology Oxford Brookes College

Declaration

Student Conduct Regulations:

Please ensure you are familiar with the regulations concerning Academic Integrity. The University takes this issue very seriously, and students have been expelled or had their degrees withheld for cheating in assessments. Students having difficulties with their work must seek help from their tutors rather than be tempted to use unfair means to gain marks. Students should not risk losing their degree and undermining all the work they have done towards it. You are expected to have familiarised yourself with these regulations.

https://www.brookes.ac.uk/regulations/current/appeals-complaints-and-conduct/c1-1/

Guidance on the correct use of references can be found on www.brookes.ac.uk/services/library, and also in a handout in the Library.

The full regulations may be accessed online at https://www.brookes.ac.uk/students/sirt/student-conduct/

If you do not understand what any of these terms mean, you should ask your Project Supervisor to clarify them for you.

I declare that I have read and understood Regulations C1.1.4 of the Regulations governing Academic Misconduct, and that the work I submit is fully in accordance with them.

Signature Cora WenhuiDuan Date ………6th May 2025……

REGULATIONS GOVERNING THE DEPOSIT AND USE OF OXFORD BROOKES UNIVERSITY MODULAR PROGRAMME PROJECTS AND DISSERTATIONS

Copies of projects/dissertations, submitted in fulfillment of Modular Programme requirements and achieving marks of 60% or above, shall normally be kept by the Oxford Brookes University Library.

I agree that this dissertation may be available for reading and photocopying in accordance with the Regulations governing the use of the Oxford Brookes University Library.

Signature Cora WenhuiDuan Date ………6th May 2025……

Acknowledgment

For my solar radiation forecasting project, I am extremely grateful to my supervisor, Dr. Grace Ugochi Nneji, and my module leader, Joojo Walker, for their valuable guidance and support. Dr. Grace provided profound feedback and strict supervision throughout the research process, which greatly improved the quality of my work. Although mainly responsible for the delivery of course materials, Joojo Walker also provided a lot of help, promoting my understanding and application of complex concepts. The joint support of both sides played an important role in the successful completion of my project. Not only that, I also want to thank my friend for providing me with inspiration for the experiment and offering constructive suggestions during the project. And I also thank my parents for giving me a very good life, allowing me to devote myself wholeheartedly to my studies. It offers solace to my soul when I am confused about life. Finally, I would like to thank Oxford Brookes College for providing me with good learning resources for my reference.

Table of Contents

Figure Content

Table Content

Abstract

With the increasing impact of fossil fuels on the environment, the world is gradually transitioning to clean energy and renewable energy, among which solar energy has received extensive attention. The advantages of solar power generation in the grid are to reduce greenhouse gas emissions and reduce electricity costs, but at the same time there are risks such as equipment aging, dependence on meteorological conditions, grid stability and land resource constraints. This project aims to improve the prediction accuracy by creating a hybrid model called Attention-SolarMeNet model to predict the output of solar Irradiance using time series data. The hybrid Attention-SolarMeNet model combines the feature extraction ability of Attention-SolarNet and the time-dependent capture ability of RadiMeNet, and further enhances the feature point extraction and prediction accuracy of the model by using the attention mechanism. The project uses the data set from GitHub, and through the steps of data preprocessing, model design, training and evaluation, it finally realizes the accurate prediction of solar power output. Experimental results show that the model achieves excellent performance, including dataset 1 includes an MAE of 0.065, an MSE of 0.009, an RMSE of 0.098, and an R2 of 0.852 and dataset 2 includes an MAE of 0.039, MSE of 0.006, RMSE of 0.081, and R2 of 0.829. indicators on different data sets, which verifies its effectiveness and feasibility in the field of solar energy prediction. The project not only improves the efficiency of energy management, promotes the integration of renewable energy sources, but also enhances the stability of the grid, reduces environmental pollution, and brings significant benefits to solar farm operators, grid regulators, energy exchange markets, and environmental organizations.

Keywords: solar energy forecasting; Attention-SolarMeNet model; Time series analysis; Attention mechanism; Renewable energy sources

Abbreviations

SolarNet: Solar Network

RadiMeNet: Radiance Memory Network

SolarMeNet: Solar Memory Network

CNN: Convolutional Neural Network

LSTM: Long Short-Term Memory

MAE: Mean Absolute Error

MSE: Mean Squared Error

RMSE: Root Mean Squared Error

R^2: Coefficient of Determination

SHAP: SHapley Additive exPlanations

IoT: Internet of Things

ARIMA: Auto-Regressive Integrated Moving Average

GPU: Graphics Processing Unit

CPU: Central Processing Unit

ANN: Artificial Neural Network

SVR: Support Vector Regression

NWP: Numerical Weather Prediction

PSO: Particle Swarm Optimization

MSE: Mean Squared Error

MAPE: Mean Absolute Percentage Error

PV: Photovoltaic

Glossary

Solar Power Generation: It refers to the process of converting sunlight into electricity through devices such as solar panels.

Convolutional Neural Network (CNN): A deep learning model that is particularly good at processing data with a grid-like structure, such as images.

Long Short-Term Memory (LSTM): A special type of recurrent neural network (RNN) that is capable of learning and remembering long-term dependency information.

Hybrid Model: Machine learning models that combine two or more different models to take advantage of the strengths of each.

Time Series Data: A chronologically ordered collection of data points used to analyze trends and patterns over time.

Feature Extraction: The process of identifying and extracting important features from raw data using algorithms.

Attention Mechanism: A deep learning technique that enables the model to focus on critical parts of the input data

MAE (Mean Absolute Error): The average of the absolute values of the prediction errors, which measures the accuracy of the model predictions.

Mean Squared Error (MSE): The average of the squared prediction errors, giving more weight to large errors.

Root Mean Squared Error (RMSE): The square root of the mean squared error, which measures the error in the same units as the original data

R-squared (Coefficient of Determination): It reflects how well the model explains the variation in the data, with a value closer to 1 indicating a better model fit.

SHAP Values: A method for interpreting the predictions of machine learning models, based on the Shapley value from cooperative game theory

IoT (Internet of Things): A network that interconnects various devices and services through the Internet.

ARIMA (Auto-Regressive Integrated Moving Average):A statistical model for time series forecasting.

Graphics Processing Unit (GPU): Computer hardware designed specifically to speed up graphics rendering and gaming

Central Processing Unit (CPU): The main processor of a computer, responsible for executing program instructions and managing computer operations.

ANN (Artificial Neural Network): A computational model that mimics the structure of the neural network of the human brain.

Support Vector Regression: A machine learning approach for regression analysis, using support vector machines

Numerical Weather Prediction (NWP): A method of forecasting the weather using mathematical models and computational techniques.

PSO Particle Swarm Optimization: An optimization algorithm based on group cooperation that simulates the social behavior of flocks of birds or schools of fish.

SolarNet: SolarNet is a CNN model that is a part of hybrid model

RadiMeNet: RadiMeNet is an LSTM model that is a part of hybrid model

SolarMeNet: SolarMeNet is a hybrid model combining CNN and LSTM models. Used to learn solar energy predictions

Attention-SolarMeNet: Hybrid model combining Attention layer and SolarMeNet model. Further improve the accuracy and interpretability of the prediction

Chapter 1 Introduction

1.1 Background

With the adverse impact of fossil fuels on the environment becoming increasingly prominent which is Pollutants emitted from the burning of fossil fuels, such as sulfur dioxide (SO2), nitrogen oxides (NOx) and particulate matter (PM), have harmful effects on human health and the environment. These pollutants not only cause respiratory diseases but also increase the likelihood of severe weather events such as acid rain and smog, so that the world is gradually transitioning to clean energy and renewable energy [1]. Solar energy is getting a lot of attention. Solar energy is relatively in clean energy resources and sustainable development [2], so it has become the main means of power generation.

1.1.1 Risk and Factor

Moreover, the most significant advantage of solar power in the grid is to reduce greenhouse gas emissions and reduce the cost of electricity [3]. However, using solar energy also has some risks as following.

Equipment aging: Usually, after the solar station continues to work for a long time, there will be technical failure and the aging of the equipment as Figure 1, which will lead to the reduction of power generation efficiency, resulting in a certain negative impact [4].

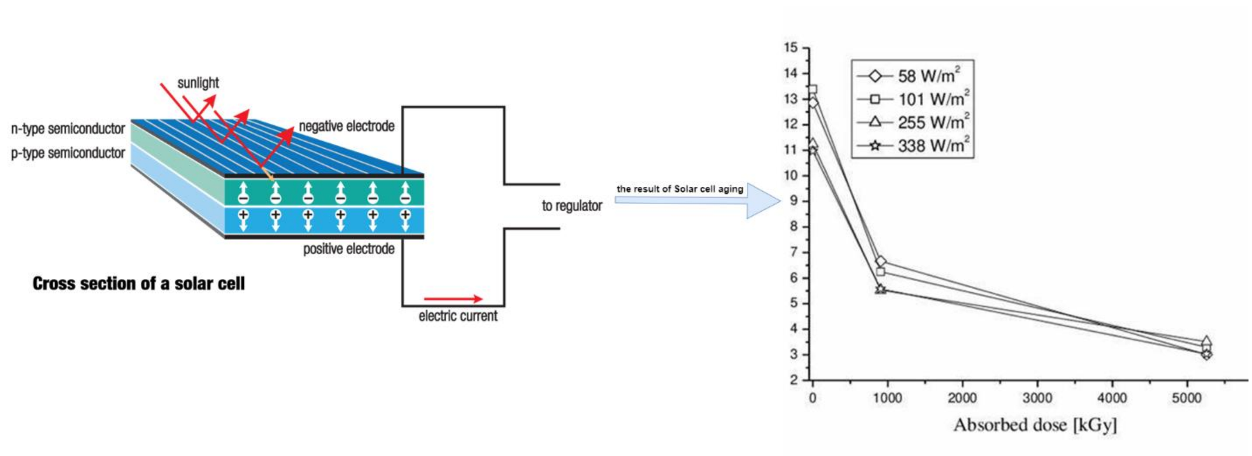


Figure 1: Equipment aging

Figure 2 shows how a solar panel works and how the efficiency of a solar cell decreases as the absorbed dose (amount of radiation) increases at different initial light intensifies. This indicates that during the long-term use of solar cells, their performance will gradually degrade due to factors such as radiation.

Meteorological conditions: Solar power generation is highly dependent on meteorological conditions, such as cloud cover, atmospheric transparency, temperature and humidity. The uncertainty of these factors can lead to fluctuations in solar radiation, which can affect power generation output [5]

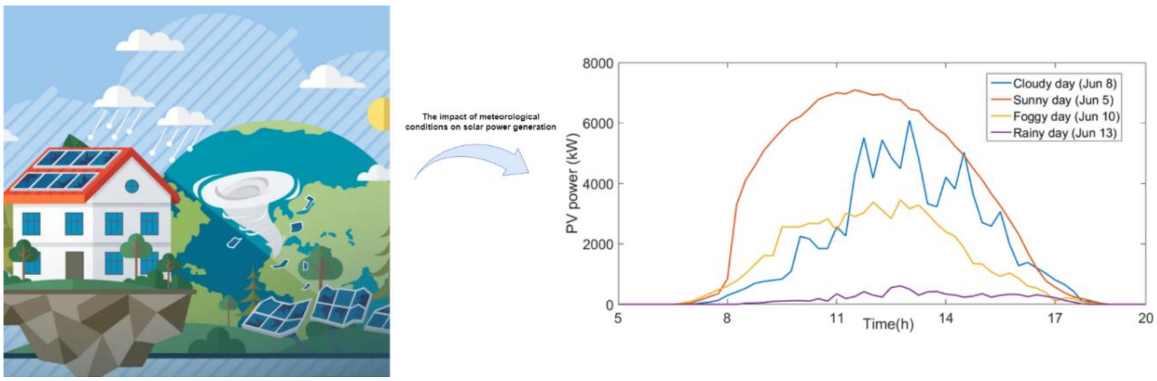


Figure 2: Impact of Meteorological Conditions

1.1.2 Challenge

While the transition to clean energy solar is beneficial, there are some hurdles that need to be overcome to ensure a stable and reliable power supply

Grid Stability: Despite these properties, the high penetration of solar energy in the grid can lead to serious instability [3]. As shown in Figure 3, solar PV varies with the saturation of distributed and utility-scale systems. Due to the mismatch between photovoltaic power generation and typical residential electricity demand, the power demand is duck-shaped. The volatility and uncertainties of solar power generation, such as the unpredictability of wind and light, pose challenges to the frequency stability and voltage stability of the power system [6]. Traditional synchronous generators are the biggest pillar to maintain stability, but new energy sources do not have the ability of traditional generators to maintain voltage and frequency. This limits the further expansion of solar power generation.

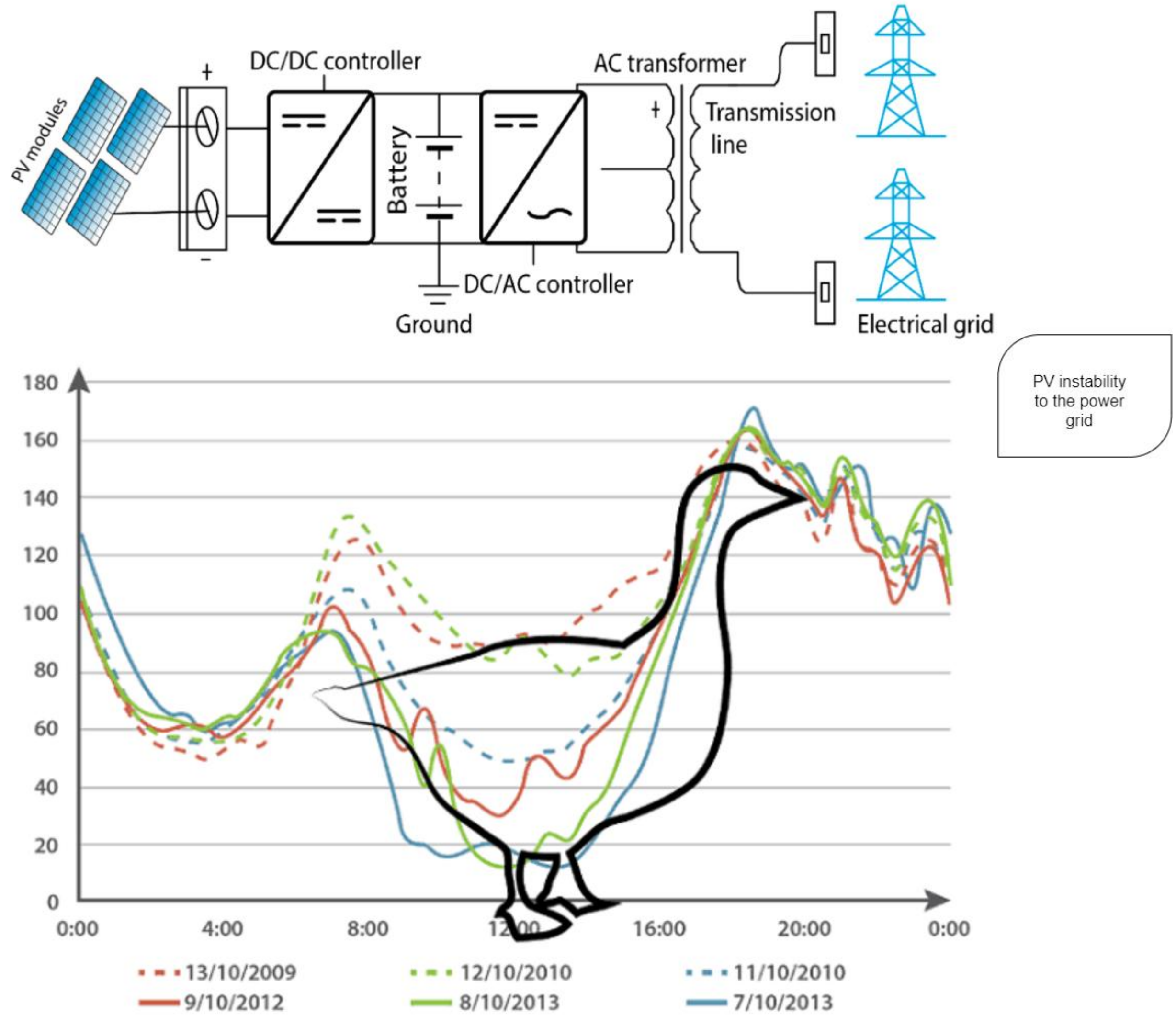


Figure 3: Duck Curve - Power Grid Instability

Land Resource Constraints: Land resources suitable for the construction of large-scale photovoltaic power plants are limited, especially in areas where land resources are tight [2]. This limitation is particularly pronounced in densely populated urban areas and areas of high agricultural or ecological value as Figure 4. The limited available land limits the further expansion of solar power generation, and alternative solutions need to be developed, such as optimizing land use, improving the efficiency of photovoltaic power generation, developing distributed photovoltaic power generation and other measures, which have become increasingly important methods to overcome land constraints.

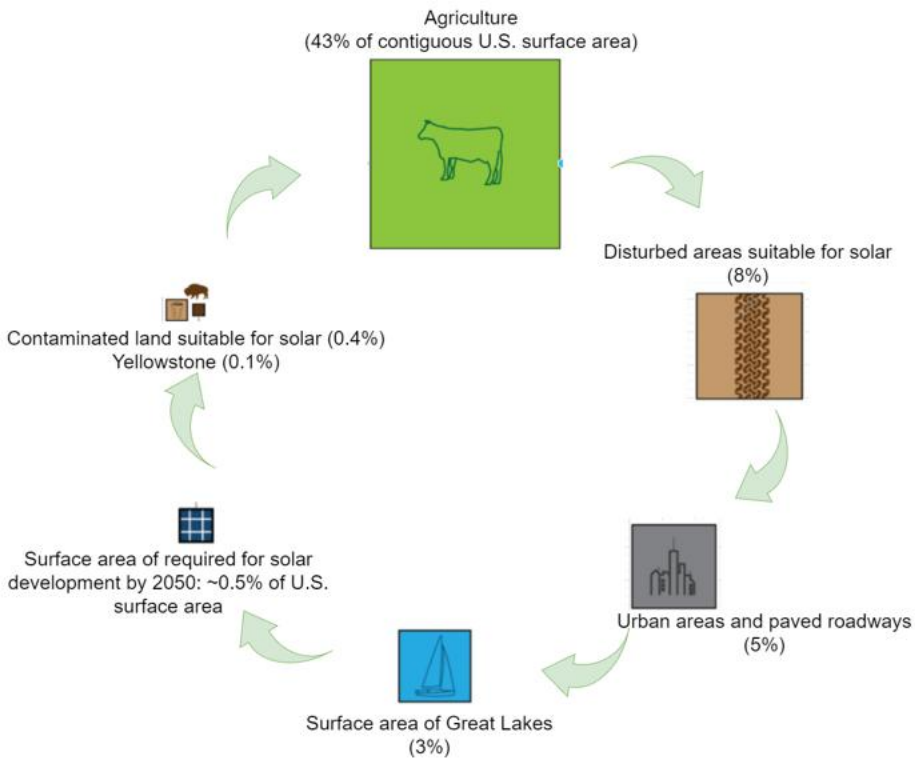


Figure 4: Land Resources Occupy Too Much

Predictive Challenges: The integration of solar energy into the grid presents significant forecasting challenges. As is shown in Figure 5, Unlike conventional energy sources, solar output is highly dependent on weather conditions such as cloud cover, temperature and sunlight intensity, which can make solar output volatile and unpredictable. This uncertainty complicates and changes the task of accurately predicting solar power generation, which is essential to maintaining grid stability and ensuring reliable energy supplies.

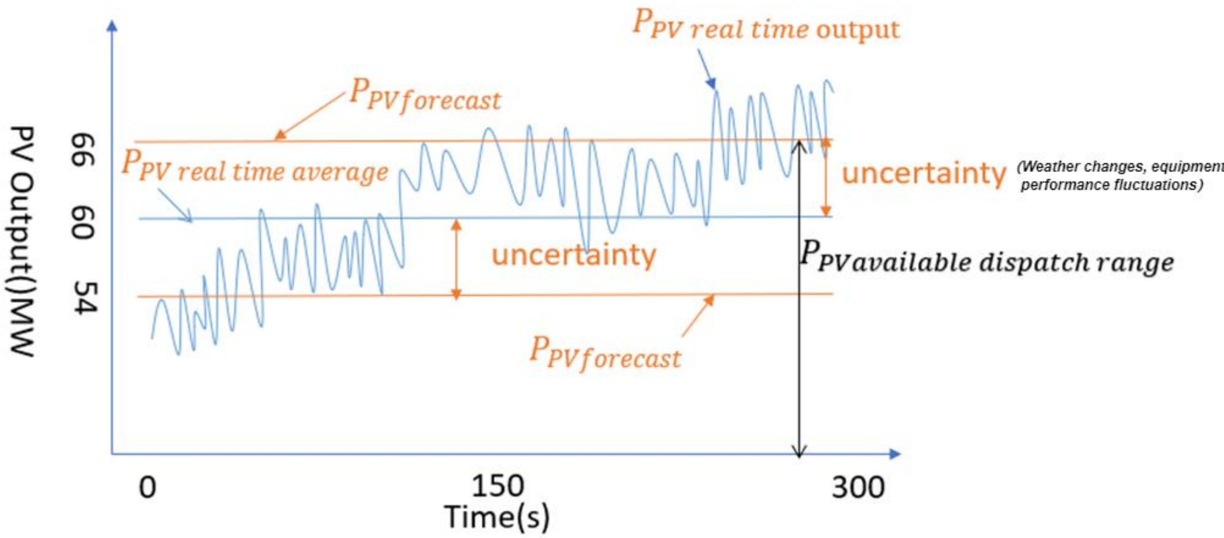


Figure 5: PV Forecast Uncertainty and Real-Time Output

1.2 Aim

In deep learning applications, a single model may have limitations that affect performance. The project aims to improve prediction accuracy by creating Attention-SolarMeNet model that uses time series data to predict solar farm output. A hybrid Attention-SolarMeNet model was developed to predict time series data, utilizing the capabilities of both convolutional and recurrent neural networks. The SolarNet layer efficiently extracts meaningful features from the sequence data, while the RadiMeNet layer captures the inherent time dependence of the data. This combination allows robust modeling of complex patterns and sequences. This model draws on the research results of Su Chang Lim et al., combines spatio-temporal data, greatly improves the accuracy and reliability of prediction, and overcomes the shortcomings of a single model.

1.3 Objectives

In order to achieve the above goals, it is divided into the following steps:

a) Meteorological and irradiance data are collected and utilized. Meteorological and irradiance data will be collected from 2014 to 2017. Merge them together to form a new dataset as dataset one. Another similar set of data is collected as dataset 2 to train the same model. It lays a foundation for the validity of the subsequent model.
b) Processing collected data sets includes processing missing values, normalizing data, and converting data types. At the same time, the unprocessed and processed data are visualized, and the data processing is more intuitive

c) The data set is divided into a training set and a validation set using a ratio of 80-20 to ensure that the model has enough data for training while also being able to verify its performance.

d) Then, the SolarNet model is designed to extract meaningful features, the RadiMeNet model is designed to capture the inherent time dependence of the data, the two adding attention are combined to form the Attention-SolarMeNet model

e) During model training, fine-tune hyperparameters: set batch size to 32, and adjust dropout to prevent overfitting. Use "ReLU" for the output layer, Adam optimizer, and "mean_squared_error" loss for evaluating the performance of regression model.

f) Finally develop a user interface that allows users to upload time series data of the solar farm. The trained Attention-SolarMeNet model is integrated into the user interface to realize the function of real-time prediction of solar output.

1.4 Project Overview

Through this project, the use of clean energy can be increased, and the scope and beneficiaries of the project are also essential to predict the development of solar energy. Also according to Figure 6, it can provide a comprehensive overview of the project process. From initial data processing to model design and final evaluation testing, the core activities and outputs of each stage are presented in detail, providing important visual support and clear logic for the project.

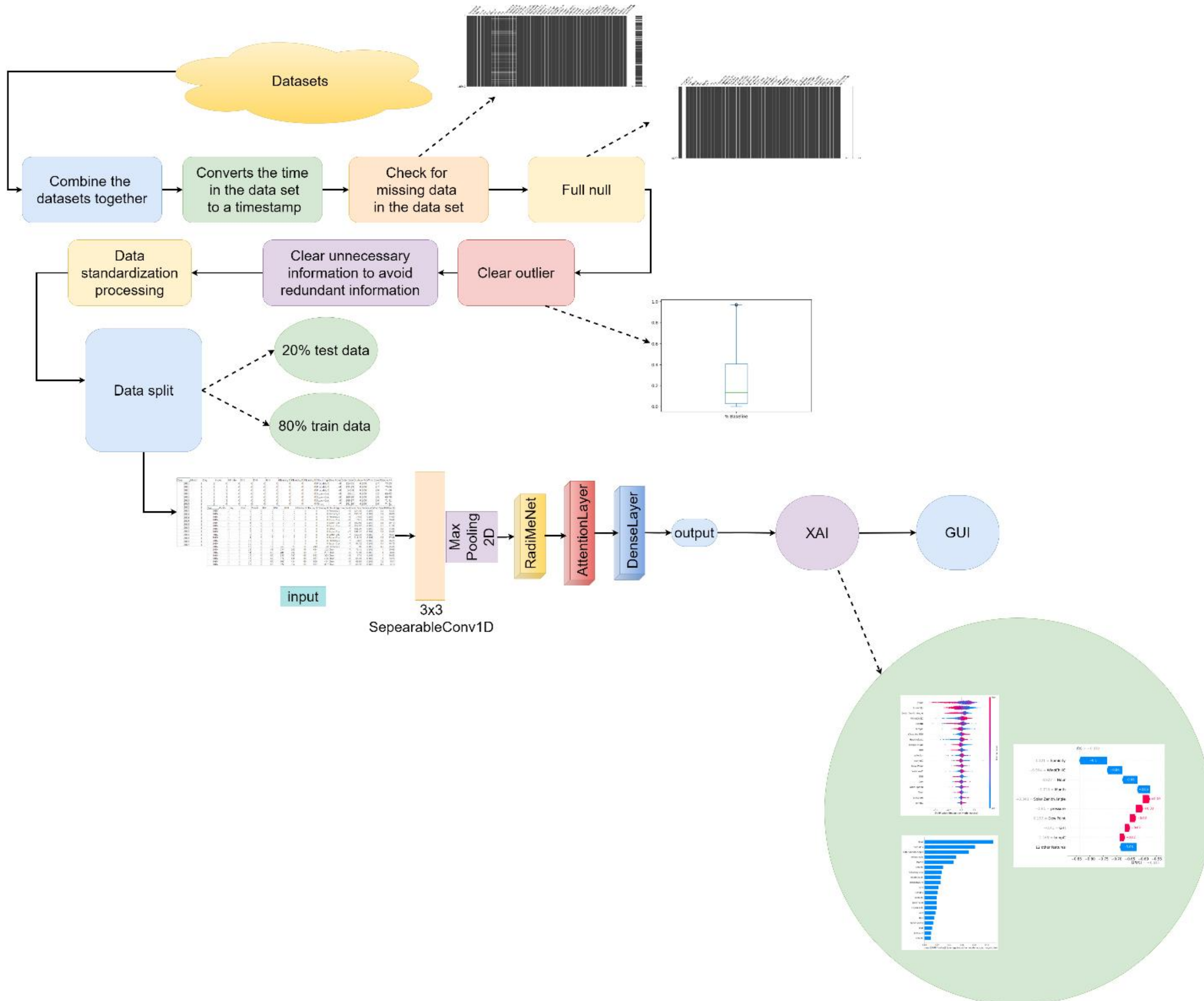


Figure 6: Project Overview

1.4.1 Scope

SolarNet model are effective for solar yield prediction, particularly in recognizing features that impact solar output [7]. However, they're not as strong as RadiMeNet in handling time series data. Attention-SolarMeNet can improve training speed by at least 40% [8] and enhance prediction accuracy and stability compared to using either model alone [9].

Here are the implications and potential contributions of the project:

- Improve energy management efficiency
- Promoting renewable energy integration
- Strengthen the stability of power grid
- Reduce environmental pollution

- Increase solar energy production
- Save equipment resource cost

1.4.2 Audience

The development of specialized solar production forecasting systems brings significant benefits to various stakeholders:

- Solar farm operators: Can optimize production and distribution, enhancing efficiency.
- Grid management agency: Can better stabilize electricity dispatch and grid operations
- Energy trading market: Can increase solar energy's market share, aiding in firm decision-making.
- Environmental protection organization: Predictions of solar power generation could improve its reliability and efficiency, which can promote renewable energy use and reduce fossil fuel dependence.

In conclusion, the proposed Attention-SolarMeNet is expected to improve the accuracy of forecast solar farm yields, bringing benefits to solar farm operators, grid regulators, energy trading markets, environmental organizations, and the broader research community.

Chapter 2 Background Review

Among the various methods to explore solar energy prediction, traditional methods provide us with an important starting point. These methods often rely on historical data to predict the output of solar energy. Although these methods may not be as precise as modern techniques, they played a key role in the early development of solar forecasting. In particular, observations of sunspots and auroras by ancient Chinese astronomers have not only provided valuable data for scientific research, but also provided a historical perspective for understanding the impact of solar activity on Earth's climate and energy supply.

2.1 Solar Energy Prediction Using Traditional Method

Hayakawa et al. [10] proposed that during the Song Dynasty, astronomers in China observed the sun through the naked eye at sunrise and sunset or through clouds, recording sunspots as dark spots or black gas in the sun and describing their number, shape and size in detail according to Figure 7. They also observe and record auroras at night, describing them as light, clouds, or gas, and noting features such as color, movement, and direction. These observations were carried out by trained experts at officially designated observatories for the purpose of not only scientific research, but also to provide the emperor with interpretations of celestial phenomena to guide political decisions. These observations were preserved in the official chronicle, The History of Song, providing valuable data on ancient solar activity for modern research

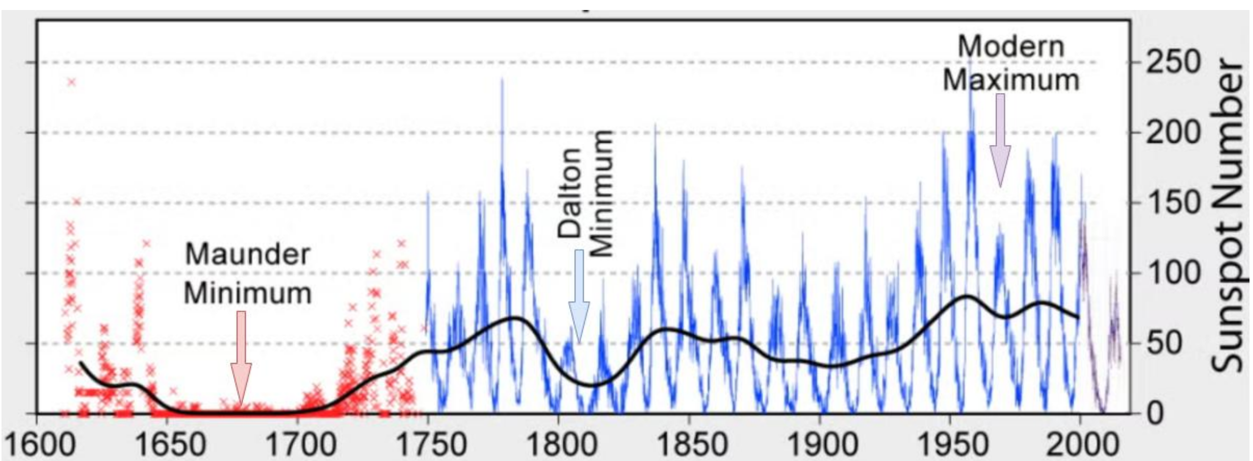


Figure 7: 400 Years of Sunspot Observations by Hayakawa et al. [10]

2.2 Solar Energy Prediction Using Machine Learning Techniques

Ledmaoui et al. [11] describes how, by comparing six machine learning algorithms, an artificial neural network (ANN) performed best in predicting solar output in the city of Benguerir, Morocco, where the square of R was up to 99%, but the study was limited by the limited access to data and the model may not be directly generalizable to other regions. Gottwald et al. [12] analyzed the use of the XGBoost model to predict solar power generation output, and compared various models, The XGBoost model has a high accuracy. In the next 31 days, the highest possible value is 1310.03 kW and the lowest possible value is 628.50 kW. Theocharides et al. [13] put forward the prediction of solar power output by ANN, SVR, etc. Figure 8 illustrates the flow chart of a typical method for generating short-term, day-ahead PV production forecasts from NWP and machine learning models, and the ANN model has the smallest error and the highest accuracy result: MAPE of 0.61%, nRMSE of 0.76%, and SS of 92.22%. Boutahir et al. [14] compared various models and finally concluded that ANN model has high accuracy in predicting solar energy output. And the square value of R is as high as 99.6%. Karimah et al. [15] proposed an integrated solar energy IoT system, which used an ARIMA (11,2,4) model to forecast solar energy. The final accuracy was 68%. In one year, the highest average possible value is 214.48 W. The lowest average possible value is 70W.

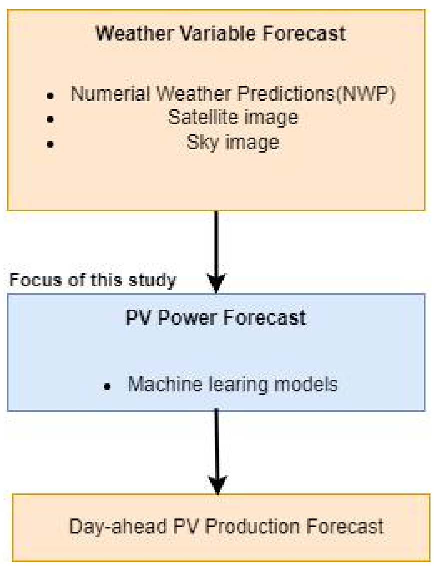


Figure 8: Prediction Flow Chart by Theocharides et al. [13]

2.3 Solar Energy Prediction Using Deep learning Techniques

2.3.1 Using CNN Method

Abdoos et al. [16] used a convolutional neural network (CNN) model as Figure 9 to predict solar power generation in the Mediterranean region up to 2050. By the summer of 2050, Spain's solar capacity is forecast to reach 42,547,680 watt-hours, while Turkey is expected to reach 20,528,640 watt-hours. However, the availability of the data may be limited because it is mentioned in the article that the data will be available upon request, which may limit further analysis and verification of the data by other researchers

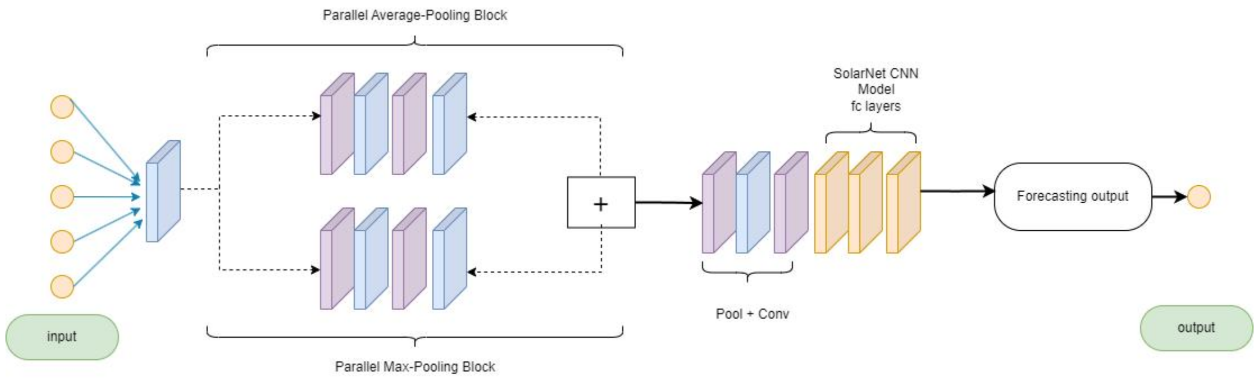


Figure 9: Solar CNN Model by Abdoos et al. [16]

2.3.2 Using LSTM Method

Zhenga et al. [17] proposed the LSTM model and PSO method to predict the output of solar. The average solar output per week is 200MW, which is very efficient, but the time is short. Campos et al. [18] used a long short-term memory network (LSTM) to predict solar output. The model structure is shown in Figure 10. The results show that MAE, RMSE and R^2 are 16.47, 31.18 and 0.84 respectively in winter. In spring, there are MAE of 9.44, RMSE of 19.76 and R^2 of 0.92. In summer, there are MAE of 8.49, RMSE of 18.03 and R^2 of 0.92. In autumn, there are also MAE of 12.99, RMSE of 30.78, and R^2 of 0.76. However, due to computer limitations, the entire data set was not used for training, which may affect the effect of model training

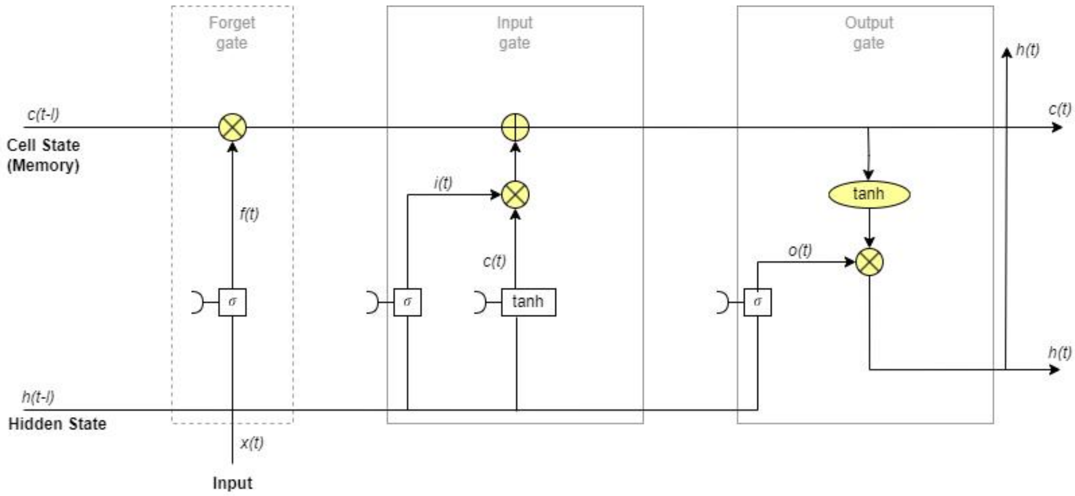


Figure 10: Architecture of an LSTM cell by Campos et al. [18]

2.3.3 Using Hybrid CNN-LSTM Methods

Jalali et al. [3] proposed a deep learning framework called MSCA-CLSTM, which combines CNN-LSTM and uses an improved sine and cosine algorithm to optimize the architecture, achieving higher results in RMSE of 0.0524, Pearson of 0.9637, and MAE of 0.0312. Alharkan et al. [19] introduced a deep learning-based two-flow convolutional neural network (CNN) and long short-term memory (LSTM) networks, followed by a network of self-attention mechanisms (DSCLANet). The model structure is shown in Figure 11. The prediction of solar power generation performed well, with error rates reduced by MSE of 0.0136 (mean square error), 0.0304 MAE of 0.0304 (mean absolute error), and RMSE of 0.0458 (root mean square error) compared to the latest advanced methods. However, there are limitations including reliance on specific data sets. AI-Ali et al. [20] proposed hybrid CNN-LSTM-Transformer model that has higher accuracy in solar production forecast, with have conclusion for the MAPE of 0.041, MAE of 0.393, and RMSE of 0.344. However, there are still some limitations. The model training process is complicated. A summary of the different researchers and their findings and possible results can be found in Table 1.

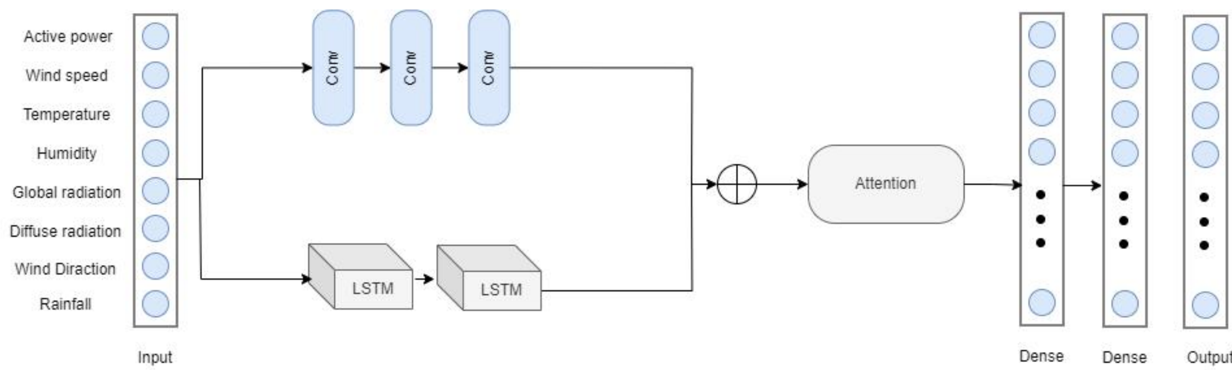


Figure 11: Proposed DSCLANet-LSTM Model by Jalali et al. [3]

Table 1: Summary of Related Works

Author	Dataset	Moules & Methods	Result	Limitation
Hayakawa et al. [10]	null	Gross observation	null	The data set has some defects
Ledmaoui et al. [11]	Benguerir, Morocco	ANN	R^2 = 0.99	Model performance and accuracy depend on the specific dataset
Gottwald et al. [12]	kaggle	XGBoost	Around 628 kW-1310kW in month	The dataset may have missing values
Theocharides et al. [13]	University of Cyprus, UCY	ANN & SVR & RT & PM	Around average 1150W in 5 days	The prediction accuracy of the model may vary under different weather conditions

Boutahir et al. [14]	A solar power plant in Benguerir, Morocco	ANN	Around average 100,000kW in one month	Models are unable to fully capture solar energy variations under extreme weather conditions
Karimah et al. [15]	Big Data centre	ARIMA(11,2,4)	Average 70W-214.48W in one year	Model performance depends on the choice of parameters p, d, and q in the ARIMA model
Abdoos et al. [16]	solar power generation in the Mediterranean region	CNN	Summer: 42,547,680 watt-hours	The initial model overfits due to its high complexity
Zhenga et al. [17]	Asian region dataset & Weather bureau	LSTM	Average 200MW in per week	The parameter Settings of PSO algorithm affect the optimization results of the model
Campos et al. [18]	UK	LSTM	Spring: MAE: 9.44 RMSE: 19.76 R^2: 0.92 Summer:	The adaptability of the model to different seasons is

			MAE: 8.49 RMSE: 18.03 R^2: 0.92	limited
Jalali et al. [3]	National Renewable Energy Laboratory	CNN-LSTM & MSCA	RMSE: 0.0524 Pearson:0.9637 MAE: 0.0312	Despite its good performance on the three datasets, its generalization ability to other regions or performance under different climatic conditions has not been verified
Alharkan et al. [19]	US	CNN-LSTM	MSE: 0.0136 MAE: 0.03 RMSE: 0.0458	The data set is limited and only the DKASC data set from the Alice Springs region of Australia was used
Al-Ali et al. [20]	National dataset	CNN - LSTM - Transformer	MAPE: 0.041 MAE: 0.393 RMSE: 0.344	Transformer models still face the problem of computational complexity and memory usage

when dealing
with long
sequences

Chapter 3 Methodology

In the project, it is necessary to effectively process the dataset and put it into the created model for training, which will be introduced below.

3.1 Approach

The proposed model is a hybrid model combining a convolutional neural network and a long short-term memory network. By analyzing time series, predict the output of solar farms. The following aspects will be followed in this project:

The GitHub dataset is used. The first dataset contains from 2014 year to 2017 year solar irradiance data and weather data which are combined them and divided 70% training data and 30% testing data. And the second dataset contains 6217 lines of solar data. Data pre-processing for the first data includes changes to the data type such as changing the time to timestamp form, filling in null values in the data, removing outliers to make the data cleaner, and finally modifying the data dimensions so that the model can be used correctly. Processing the second data involves normalizing it. The model will use the Attention-SolarMeNet module, which can extract feature points in data sets with time series and make effective predictions

3.2 Proposed Model Structure

3.2.1 Solar Network (SolarNet)

The SolarNet model described in Figure 12 consists of a convolutional layer, a pooling layer, and a Dropout layer. Firstly, the Conv1D layer uses 64 convolution kernels of size 3 and extracts local features from single-channel time series data by the ReLU activation function. Subsequently, the MaxPooling1D layer down-samples the convolutional features with a pooling window of size 2 to reduce the feature dimension while retaining key information. Finally, the Dropout layer randomly discards the outputs of some neurons with a 20% drop rate to prevent the model from overfitting. The role of the whole SolarNet part is to extract local features of the input data and perform dimensionality reduction processing. This model is especially suitable for capturing feature and has high efficiency [7].

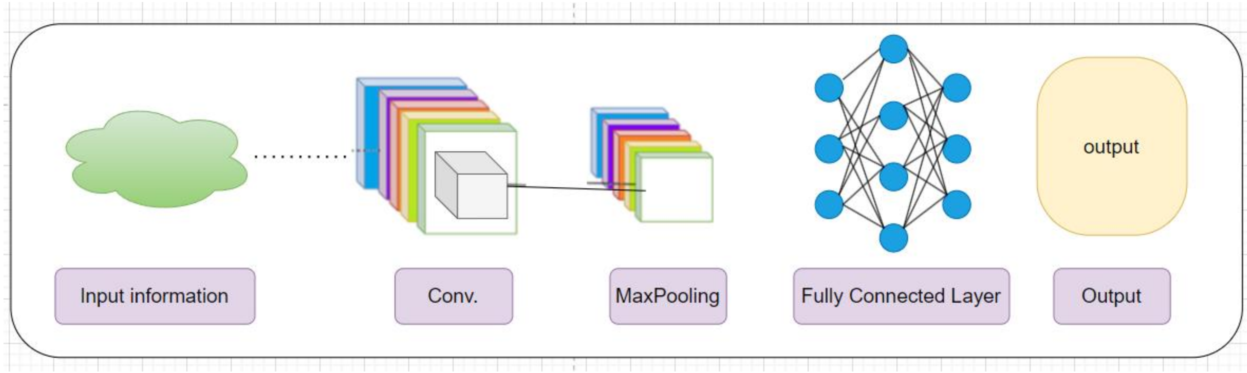


Figure 12: SolarNet Architecture

3.2.2 Radiance Memory Network (RadiMeNet)

In Figure 13, the data is passed to a RadiMeNet layer, which is configured to output the sequence, meaning that it will output information about the entire sequence instead of only the last time step of the sequence. In RadiMeNet, three main operations are performed at each time step in Figure 14, the forget operation, which decides what information to discard; Input operations that determine what new information needs to be added to the state Output operations that decide what information to output. These operations are implemented through a gating mechanism that allows the model to efficiently capture and exploit long-term dependencies [21].

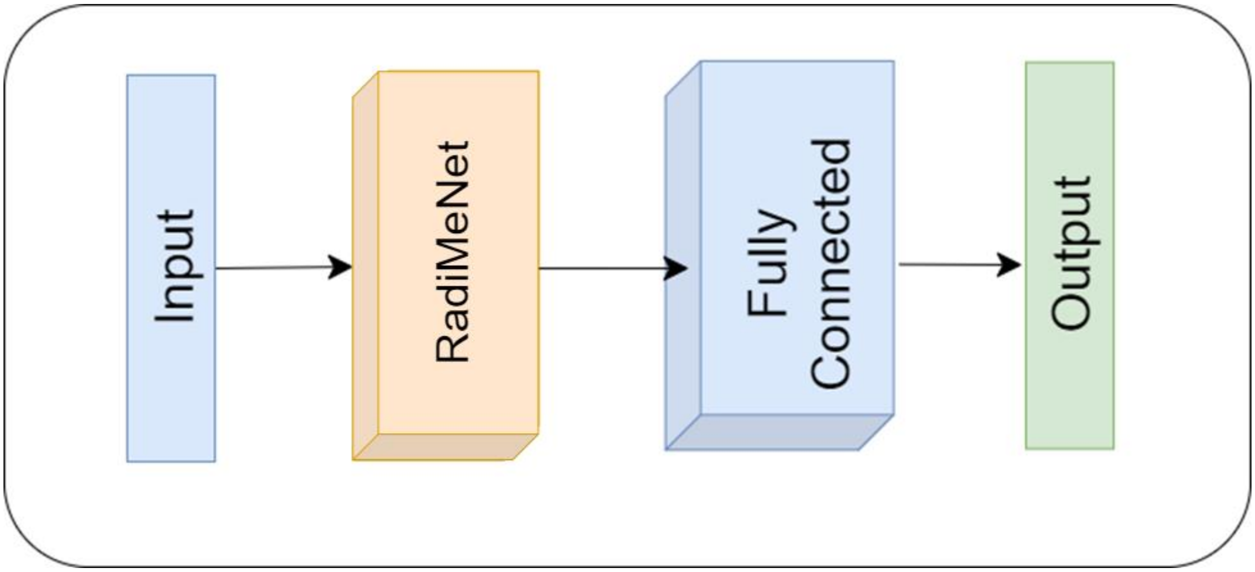


Figure 13: Architecture of the proposed RadiMeNet

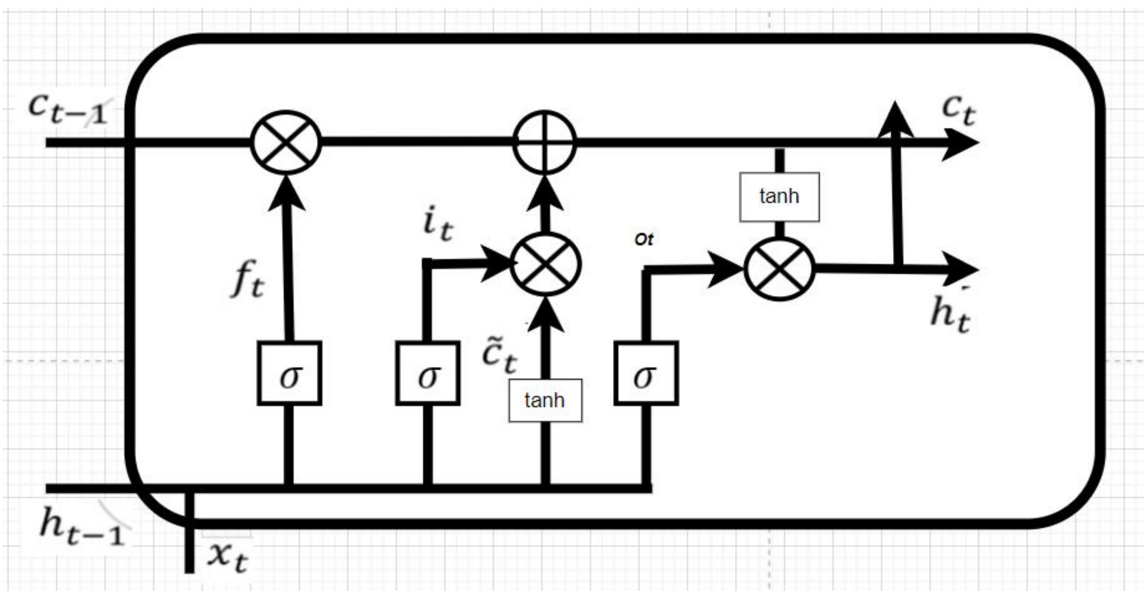


Figure 14: Structure of a basic RadiMeNet cell

Solar Memory Network (SolarMeNet)

This hybrid model is used to process sequential data. The model first extracts the local features of the input data through a one-dimensional convolutional layer (Conv1D), which uses 64 convolution kernels of size 3 and a ReLU activation function. Then, the model uses the MaxPooling1D layer to reduce the dimension of the convolutional features, so as to reduce the number of parameters and computational complexity while retaining important information. Then, a subset of neurons is randomly discarded through a Dropout layer to prevent overfitting. After that, the data is fed into the RadiMeNet layer that is configured to return the entire sequence. Means that the RadiMeNet layer outputs information about the entire time series, not just the output of the last time step. This allows the model to capture long-term dependencies in the time series. Finally, the model further processes the features through two fully connected Dense layers, the first Dense layer uses the ReLU activation function, and the second Dense layer uses the linear activation function to output the final prediction result.

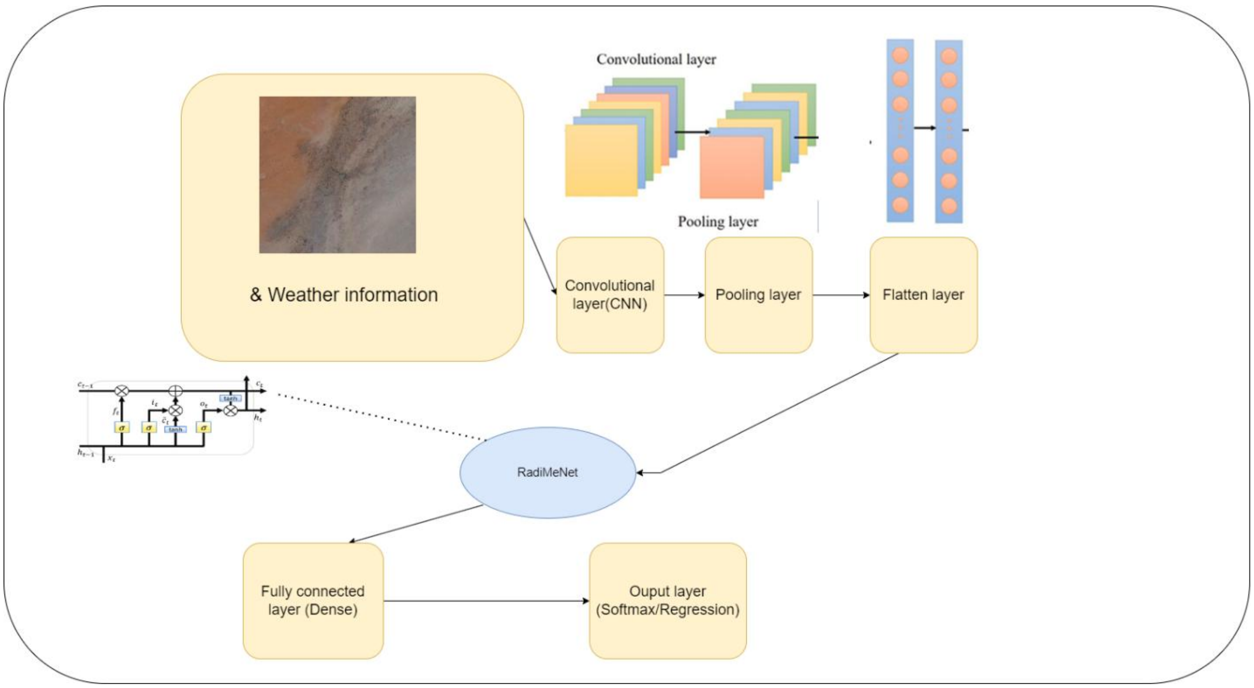


Figure 15: The structure of SolarMeNet Model

3.2.3 Attention – Solar Memory Network (Attention - SolarMeNet)

Attention Mechanism is a technique in deep learning that allows a model to dynamically focus on key parts of the input sequence while processing sequence data, thereby improving the ability to capture important information. This mechanism is particularly important in areas such as natural language processing and image recognition, as it enables models to make more efficient use of contextual information, especially in the processing of long sequences of data, such as tasks such as machine translation, text summarization and speech recognition. By calculating the attention weight, the model can learn the contribution of different input elements to the output, thereby optimizing the decision-making process and improving the overall performance [22]. The attention used in the model is shown in Figure 16.

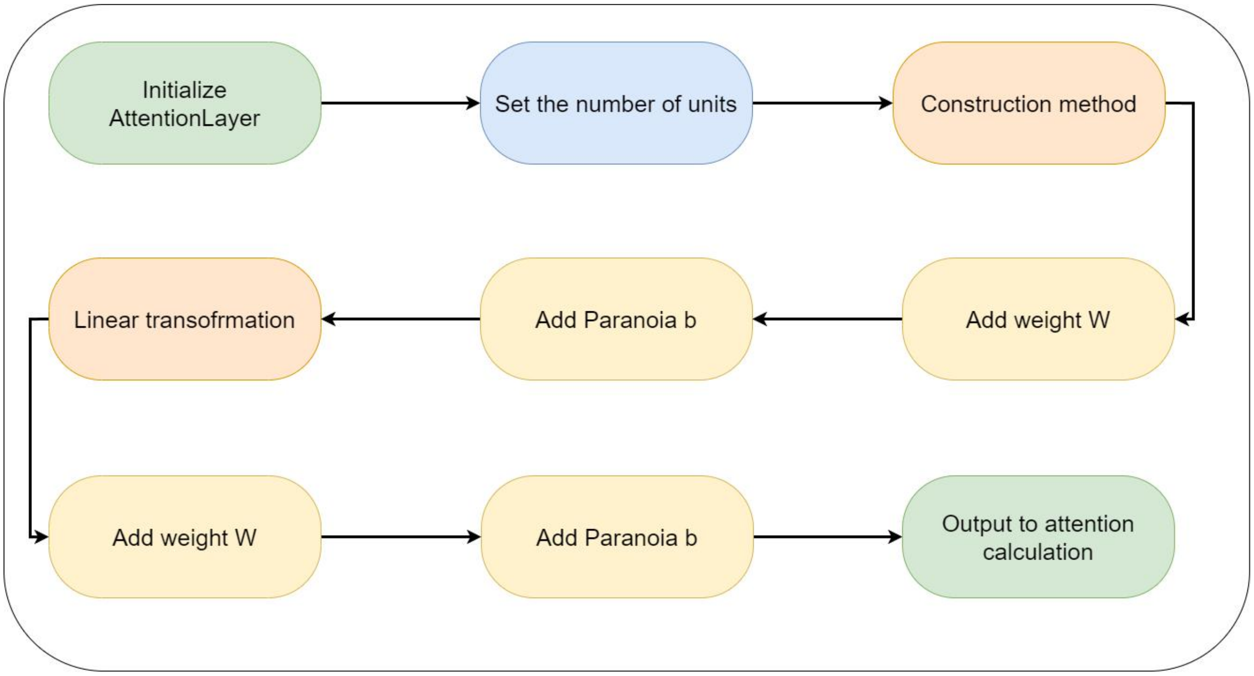


Figure 16: Attention Layer

Attention-SolarMeNet model is a combination of the above two models, as shown in Figure 17. With this model, different kinds of datasets can be trained together, using time series datasets from GitHub. In addition, the model greatly shortens the training time and improves the accuracy. During the building of the model, use two kinds of model formulas (1-9).

Convolution operation:

$$(f * g)[n] = \Sigma_{m=-\infty}^{\infty} f[m] \cdot g[n-m] \tag{1}$$

Pooling operation:

$$max\ pooling(X) = max(Xi, j) \tag{2}$$

Activation function:

$$ReLU(x) = max(0, x) \tag{3}$$

Forgotten door:

$$ft = \sigma(Wf \cdot [ht-1, xt] + bf) \tag{4}$$

Input gates and candidate memory units:

$$it = \sigma(Wi \cdot [ht-1, xt] + bi) \tag{5}$$

$$Ct = tanh(WC \cdot [ht - 1, xt] + bC \quad (6)$$

Memory unit update:

$$Ct = ft \cdot Ct - 1 + it \cdot \widehat{C}t \quad (7)$$

Output gate:

$$ot = \sigma(Wo \cdot [ht - 1, xt] + bo) \quad (8)$$

$$ht = ot \cdot tanh(Ct)ht = ot \cdot tanh(Ct) \quad (9)$$

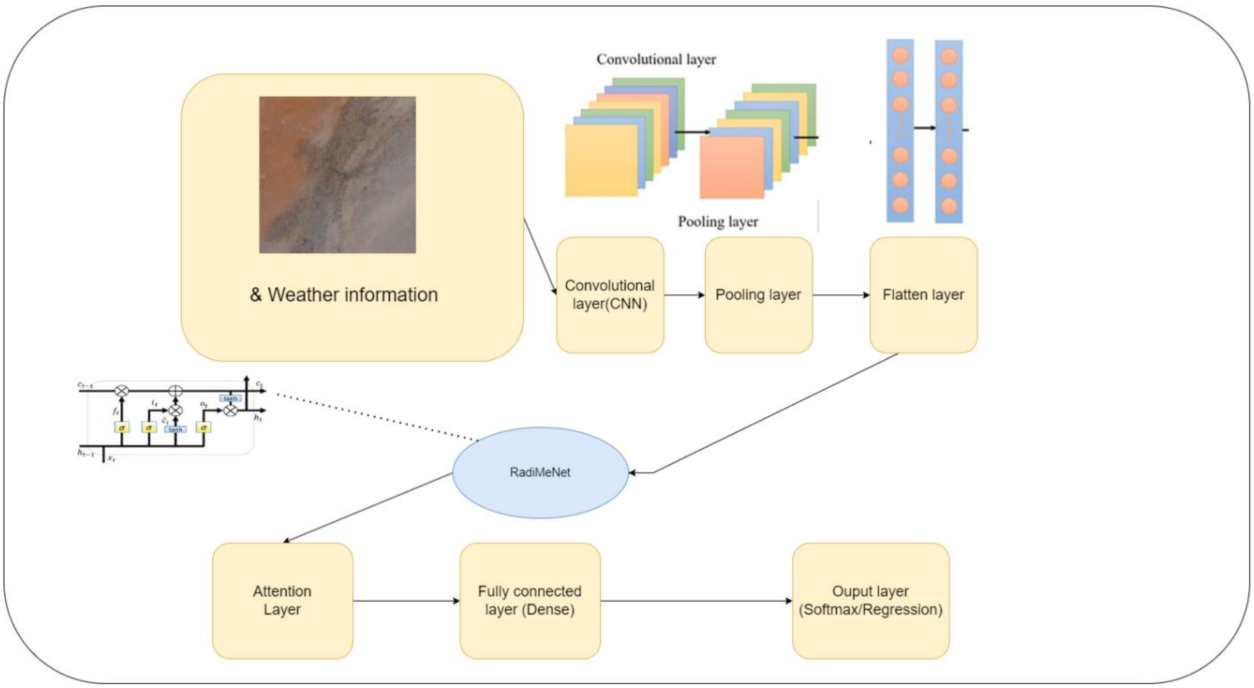


Figure 17: The Structure of Attention-SolarMeNet Model

3.3 Dataset

In this project, in order to ensure that the developed model not only performs well on a specific dataset, but also generalizes to other similar datasets, we adopted a strategy of using two different datasets for model training and validation. The purpose of this approach is to evaluate and improve the generalization ability of the model, ensuring that it remains accurate and robust in the face of new, unseen data.

3.3.1 Dataset 1

The collection of datasets is from GitHub https://github.com/VinTanz/Solar-Energy-Output-Prediction-using-CNN-LSTM, which uses time-stamp columns to merge solar irradiance data and weather data. The combined data is integrated with training and test datasets to ensure a comprehensive dataset for analysis.

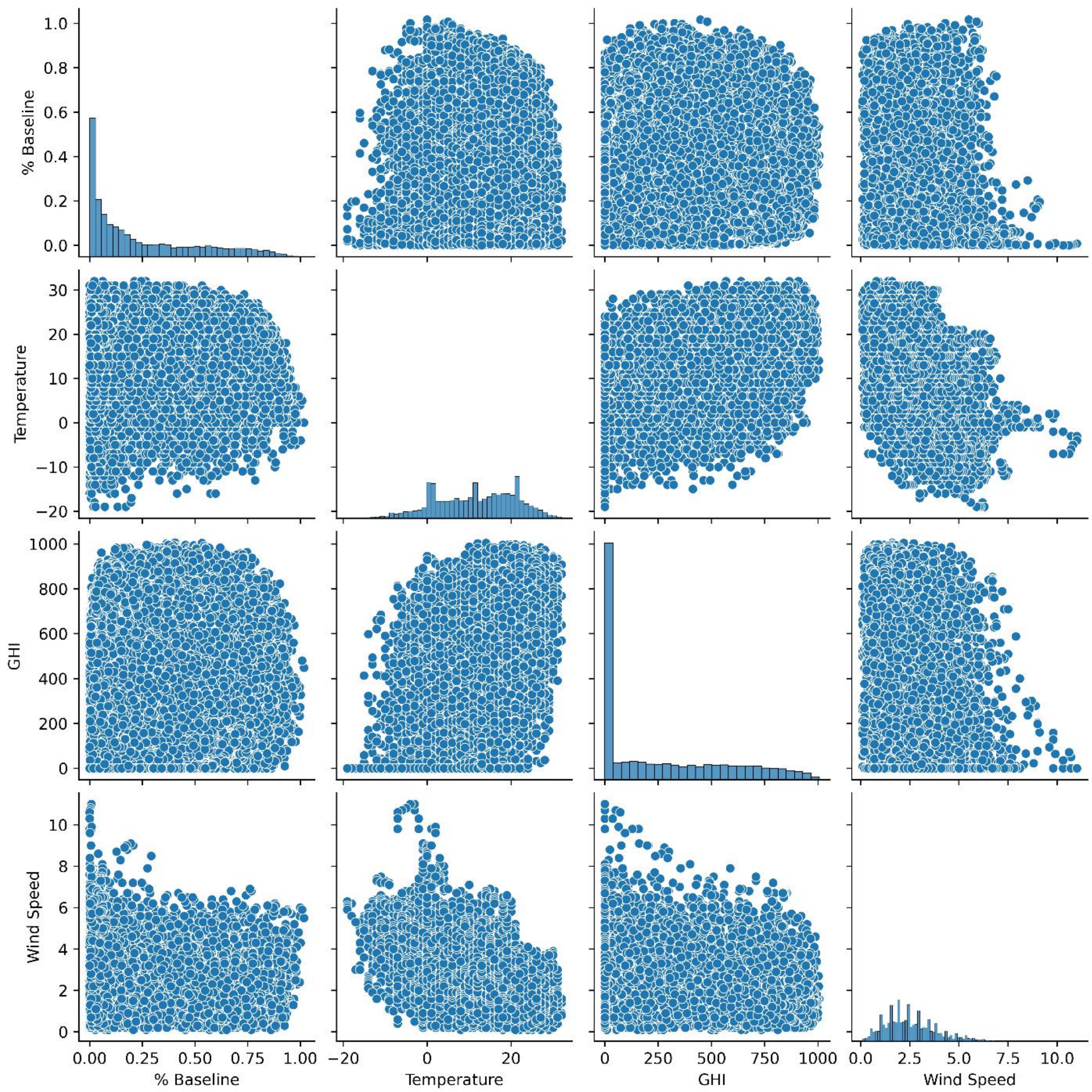


Figure 18: Pairplot of Selected Features

3.3.1.1 Dataset 1 Relationships

Find the most moderate relationship distance with the target variable in the dataset. '%Baseline', 'Temperature', 'GUI', and 'Wind Speed' are selected. Figure 18 shows the relationship between these four variables, which supports the later object selection and model construction.

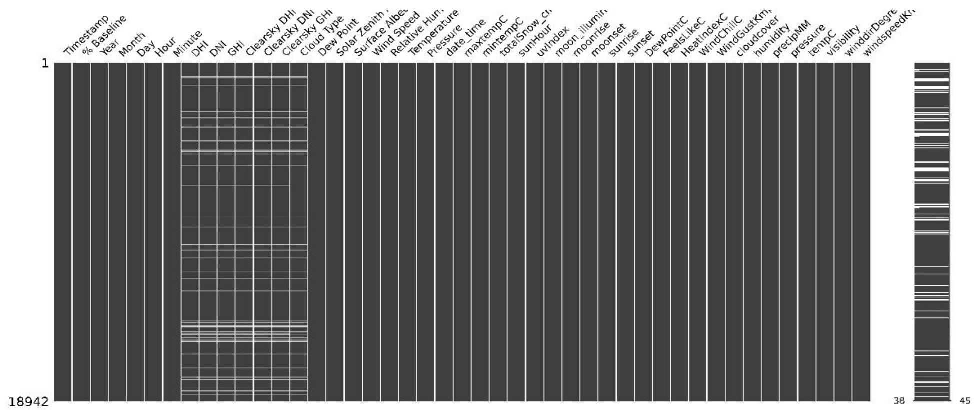


Figure 19: Null Values Visualization (Training data)

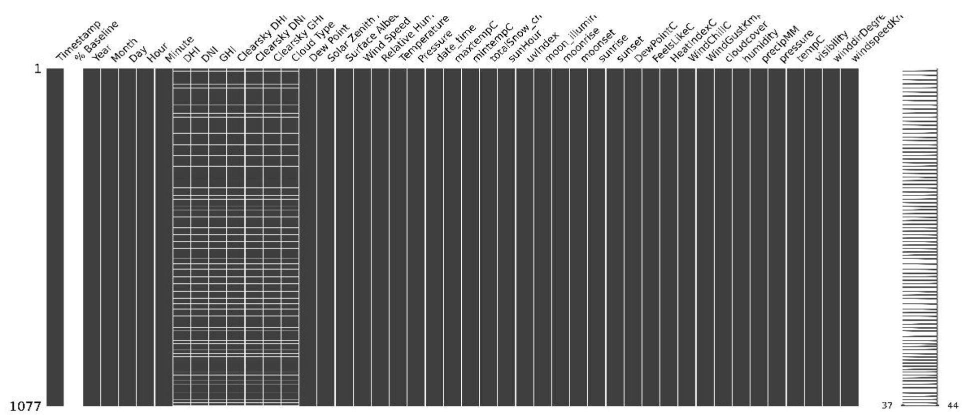


Figure 20: Null Values Visualization (Testing data)

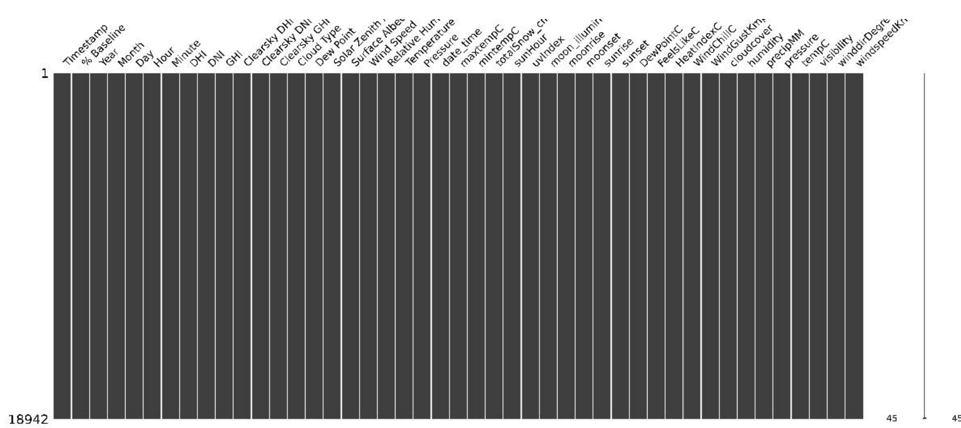


Figure 21: Null Values (After-Imputation-Training Data)

3.3.1.2 Dataset 1 Preprocessing

Data preprocessing, including data cleaning, standardization to ensure data quality and availability. As shown in Figures 19 and 20, there are many empty values in the unprocessed data. In order to ensure the accuracy of subsequent experiments, the empty values will be filled in and the processed data will be finally obtained as shown in Figures 21 and 22. Not only that. As shown in Figure 23, check the outliers in its data and eliminate the data points.

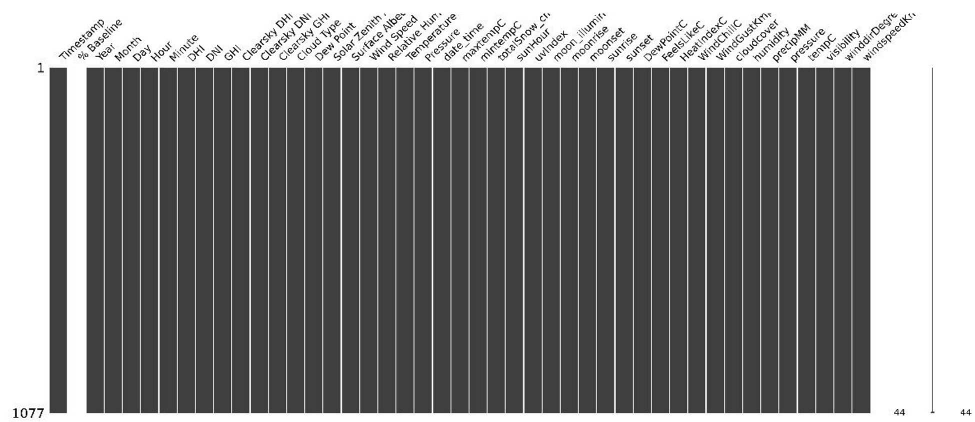


Figure 22: Null Values (After-Imputation-Testing Data)

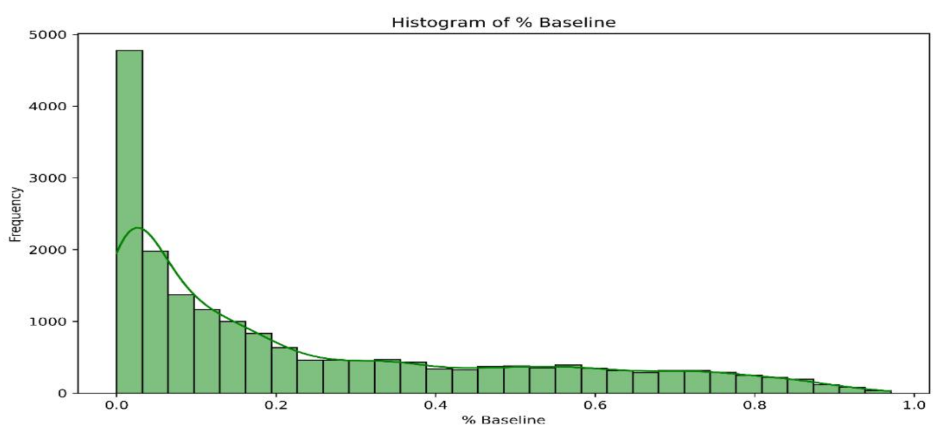


Figure 23: Histogram for '% Baseline' after being cleaned

3.3.1.3 Dataset 1 Split

The experiment uses an automated data partitioning technique to randomly split the data into training and validation sets. In this, we selected 20% of the data as the validation set, which is used to evaluate the performance of the model, while the remaining 80% is used to train the model. In order to ensure that the same result is obtained for each partition, and a fixed random seed is set. The advantage of this division is that it is random and repeatable, which helps to obtain consistent results across different experiments.

3.3.2 Dataset 2

The collection of datasets is from https://github.com/Soumyapro/Solar-Cells-Output-Prediction/blob/main/Solar%20Cells%20output%20prediction.ipynb, which uses time-stamp columns to merge solar irradiance data and weather data. This dataset already contains solar output and the corresponding weather data, so there is no need to merge other datasets

3.3.2.1 Dataset 2 Preprocessing

Since the quality of this dataset is good enough, the data standardization and normalization are simple enough. Figure 24 allows you to verify the effect of normalization and observe the difference between the original and normalized data.

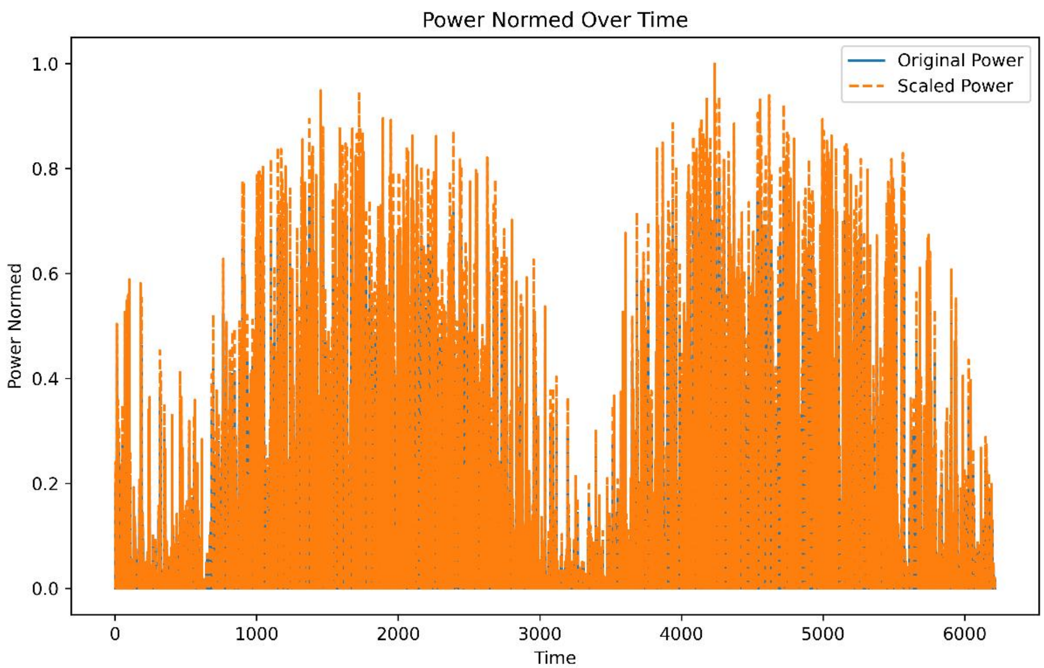


Figure 24: Power Normed Over Time

3.3.2.2 Dataset 2 Split

A more manual partitioning strategy is adopted for the experiment. First, 80% of the whole data set is calculated as the training set and the remaining 20% is used as the validation set shown in Figure 25. This partitioning method does not use a randomization process, so different results may be obtained each time it is executed. Nonetheless, this approach provides more flexibility, allowing us to adjust the size of the training and validation sets according to specific needs.

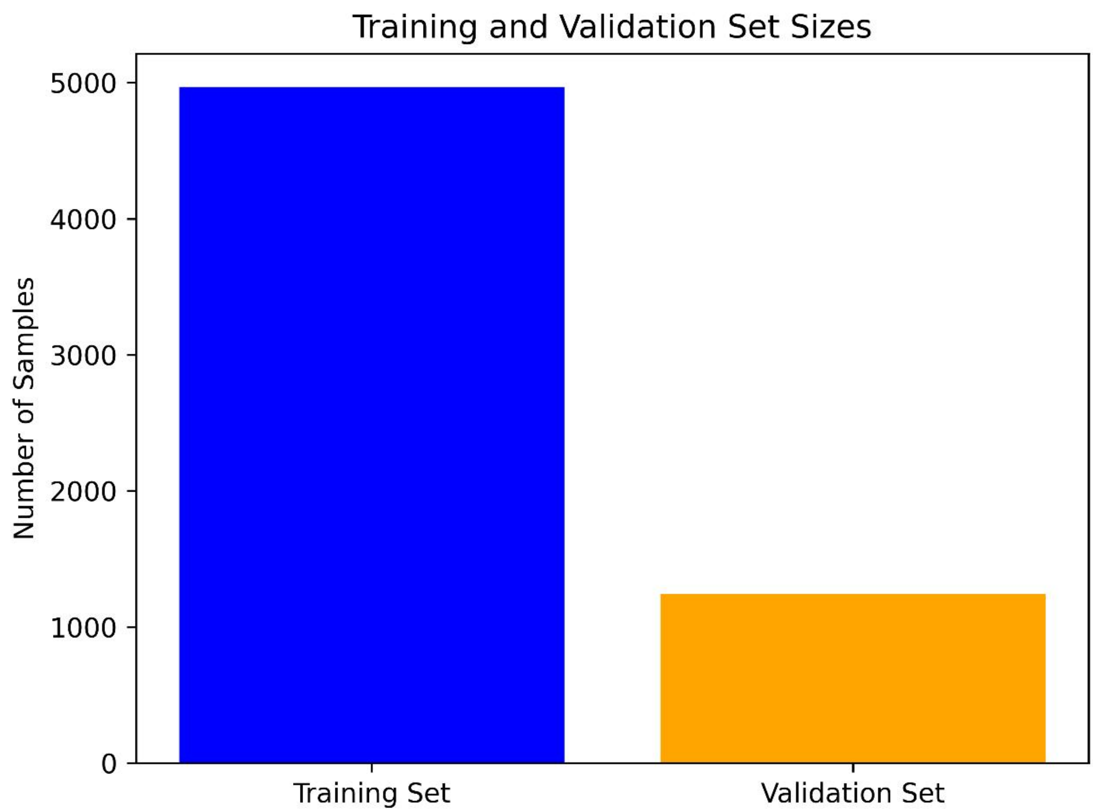


Figure 25: Training and Validation Set Sizes

3.3.3 Dataset Annotation

The data is annotated to evaluate the model's recognition performance for different actions. At the same time, appropriate evaluation indices were selected, such as accuracy, loss value, MAE, MSE, RMSE, R^2, etc.

3.4 Experimental Setup & Technology

In the experimental setup, we carefully selected several key parameters to optimize model performance as shown in Table 2, including setting 60 training epochs to ensure that the model has enough time to learn data features, using a batch size of 32 to balance memory usage and model update frequency. Adam optimizer is chosen because of its adaptive learning rate feature, which helps the model converge quickly and effectively. In total, 28,407 parameters are occupying approximately 110.96 kilobytes

Table 2: Experimental setup

Parameter	Value
Epoch	60
Batch size	32
Optimizer	Adam
Total parameters	28407 (110.96 kB)

The technology used in this project is displayed in Table 3

Table 3: Summary of Technology Utilization

Software	Framework	Tensorflow
	Language	Python
	Libraries	Numpy, Keras
	Version tool	Matplotlib
Hardware	CPU	Modern multi-core processors, such as Intel Core i7 or higher
	GPU	NVIDIA GPUs (such as Tesla, Quadro, or GeForce series)

3.5 Project Version Management

To manage the different versions of code modifications, I plan to use GitHub as the version management tool for keeping code updated and secure.

URL is as follow: https://github.com/Duan-wenhui/project

3.6 Evaluation Metrics

This project will evaluate the performance of the model through the following standards.

Loss: This criterion measures the gap between the results of prediction labels and actual labels. Below shows the equation for classification and with the expression presented in equation (1).

$$\text{Loss} = -\sum_{c=1}^{M} y_{ic} \, log\,(\mathrm{P_{ic}}) \quad (1)$$

MSE: MSE calculates the average of the squared difference between the predicted value and the true value for all samples. MSE can amplify large errors, so it is very sensitive to outliers and with the expression shown in equation (2).

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

RMSE: RMSE is the square root of MSE. It is similar to the MSE, but measures errors in the same units as the original data. RMSE similarly gives greater weight to large prediction errors because it is the square root of the sum of squares of the errors and the expression is presented in equation (3).

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (y_i - \hat{y}_i)^2} \quad (3)$$

MAE: MAE calculates the average of the absolute values of the difference between the predicted and true values of all samples. Unlike MSE and RMSE, MAE does not square large errors, so it is less sensitive to outliers and the expression is shown in equation 4.

$$MAE = \frac{1}{n}\sum_{i=1}^{n} |y_i - \hat{y}_i| \quad (4)$$

Chapter 4 Experimental Result and Analysis

This chapter shows the results of this experiment by training different models as follows.

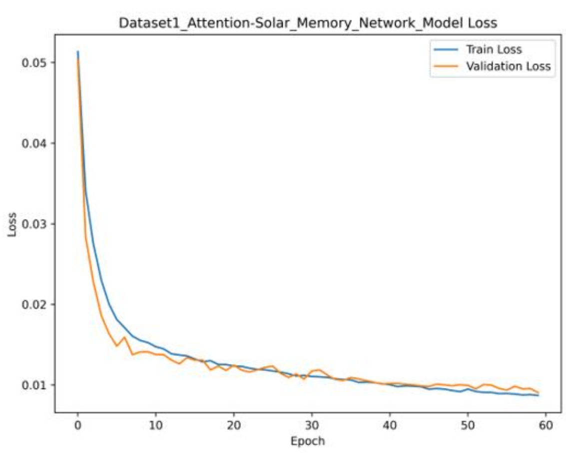
4.1 Design and Implementation

The main experimental model of this project is the Attention-SolarMeNet model. At the same time, to further prove the feasibility and accuracy of the model. Use different data sets to compare in different models. Finally, it is proven that different data sets are the best in the Attention-SolarMeNet model.

4.1.1 Attention-SolarMeNet Model Using Dataset 1

Put dataset 1 into the Attention-SolarMeNet model. The loss value and the MAE as shown in Figure 26 are finally obtained. The comprehensive figure of loss value and MAE value is shown in Figure 27.

Figure 26 shows the loss values during training and validation. As you can see from the graph, the loss value drops rapidly over the first few epochs, suggesting that the model learns quickly and ADAPTS to the data at an early stage. For example, at epoch 10, the training loss was reduced from the initial 0.048 to 0.015, and the validation loss was also reduced from 0.035 to 0.015, showing the rapid learning ability of the model in the early stages. Figure 26 also shows MAE during training and validation. Similar to the loss value, MAE drops rapidly in the initial period and remains stable in subsequent epochs. This further confirms the validity of the model in terms of prediction accuracy.



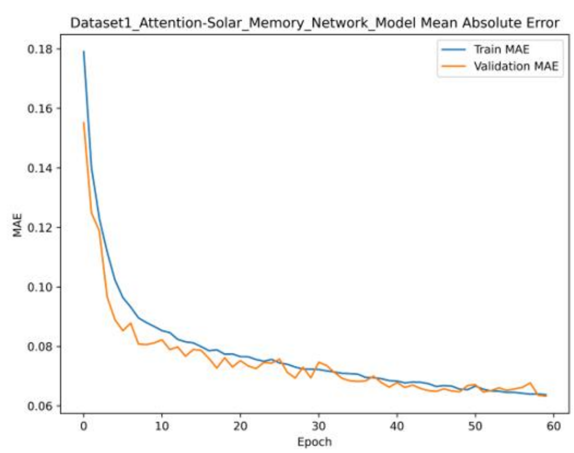


Figure 26: Dataset 1 Attention-SolarMeNet Model Loss & MAE

At the 10th epoch, the training MAE decreased from 0.17 to 0.08, and the validation MAE also decreased from 0.14 to 0.08, indicating a significant improvement in the model's prediction accuracy. Figure 27 provides a comprehensive view of the loss value and MAE, allowing us to look at these two key metrics simultaneously. As shown in Figure 27, both the training and validation loss values and MAE maintained a low and stable level throughout the training process, indicating that the model showed good performance throughout the training cycle. In particular, at the 60th epoch, both the training loss and validation loss remained around 0.01, while the training MAE and validation MAE remained around 0.06, which further demonstrated the stability and prediction accuracy of the model

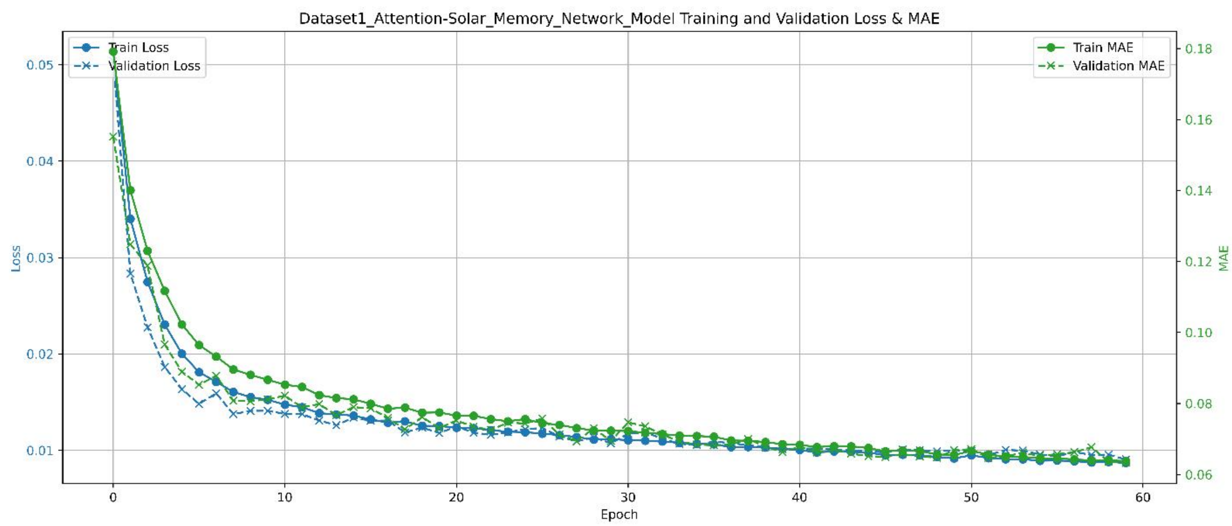


Figure 27: Dataset1 Attention-SolarMeNet Model Training and Validation Loss & MAE

4.1.2 RadiMeNet Model Using Dataset 1

Put dataset 1 into the RadiMeNet model. The loss value and the MAE as shown in Figure 28 are finally obtained. The comprehensive figure of loss value and MAE value is shown in Figure 29.

Figure 28 shows the loss values during training and validation. In the initial stage, the loss value drops rapidly, from 0.050 to about 0.015, indicating that the model learns well in the early stage. However, compared with the Attention-SolarMeNet model, the loss value of the RadiMeNet model declined more gently in the late training period, and finally stabilized at a level slightly higher than 0.010. Figure 28 also shows MAE during training and validation. MAE also declined rapidly in the initial period, from 0.18 to around 0.08, showing the effectiveness of the model in terms of prediction accuracy. However, compared with the Attention-SolarMeNet model, the MAE of the RadiMeNet model fluctuates more in the late training period and finally stabilizes at a level slightly higher than 0.07. Figure 29 provides a comprehensive view of the loss value and MAE, allowing us to look at these two key metrics simultaneously. As shown in Figure 29, both the training and verified loss values and MAE maintain a lower level during the whole training process, but compared with the Attention-SolarMeNet model, the curve of the RadiMeNet model is more volatile in the late training period, and the final loss values and MAE are slightly higher.

In summary, although the performance of the RadiMeNet model on dataset 1 is effective, its performance is slightly inferior to that of the Attention-SolarMeNet model. The RadiMeNet model has a faster learning speed in the early stage of training, but it is inferior to the Attention-SolarMeNet model in terms of stability and prediction accuracy in the later stage of training. This suggests that the Attention-SolarMeNet model may be a better choice for dataset 1, as it exhibits better stability and a lower error rate during training

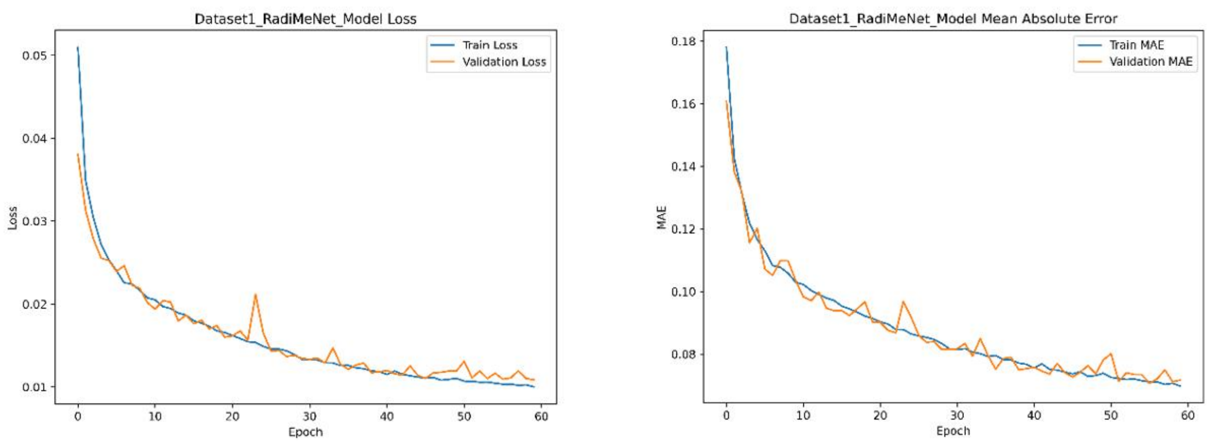


Figure 28: Dataset 1 RadiMeNet Model Loss & MAE

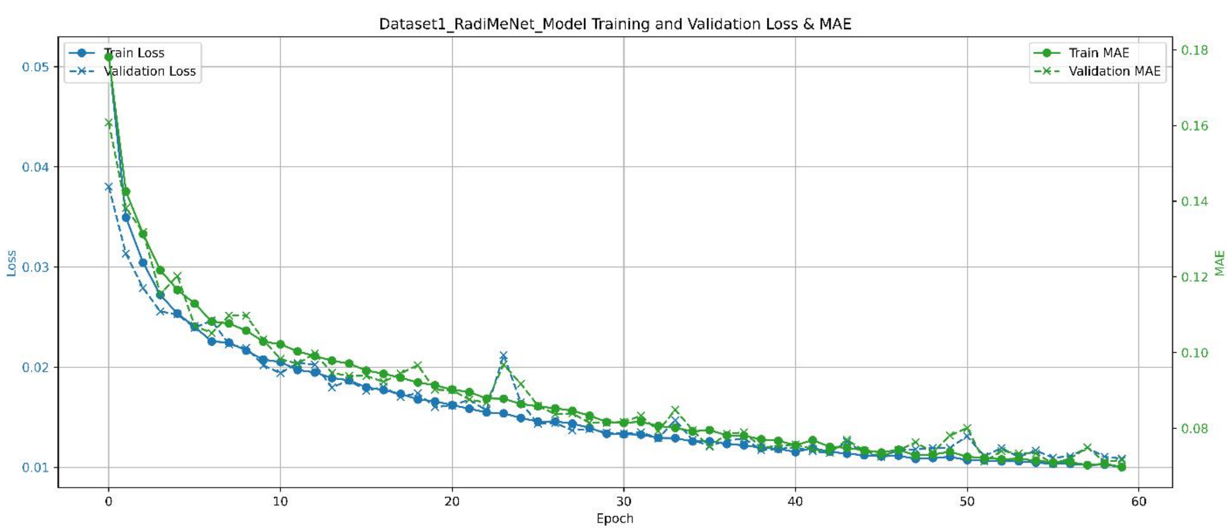


Figure 29: Dataset1 RadiMeNet Model Training and Validation Loss & MAE

4.1.3 Attention-SolarNet Model Using Dataset 1

Put dataset 1 into the Attention-SolarNet model. The loss value and MAE as shown in Figure 30 are finally obtained. The comprehensive figure of loss value and MAE value is shown in Figure 31.

Figure 30 illustrates the loss values during training and validation. At the beginning of training, the loss decreases rapidly, from around 0.06 to around 0.02, indicating that the model is learning well in the early stages. However, compared with the Attention-SolarMeNet model, the loss values of the Attention-SolarNet model level off later in training and eventually stabilize at a level slightly higher than 0.02, which is higher than the loss values of the Attention-SolarMeNet model. Figure 30 also presents the Mean absolute error (MAE) during training and validation. MAE also decreases rapidly at the beginning, from about 0.18 to about 0.10, showing the effectiveness of the model in terms of prediction accuracy. However, compared with the Attention-SolarMeNet model, the MAE of the Attention-SolarNet model fluctuates greatly in the later stage of training and finally stabilizes at a level slightly higher than 0.10, which is higher than the MAE of the Attention-SolarMeNet model. Figure 31 provides a comprehensive view of the loss values and MAE, allowing us to look at these two key metrics simultaneously. As shown in Figure 31, the loss values and MAE of both training and validation remain low throughout the training process, but the curve of the Attention-SolarNet model shows larger fluctuations in the late training period and the final loss values and MAE are slightly higher compared to the Attention-SolarMeNet model.

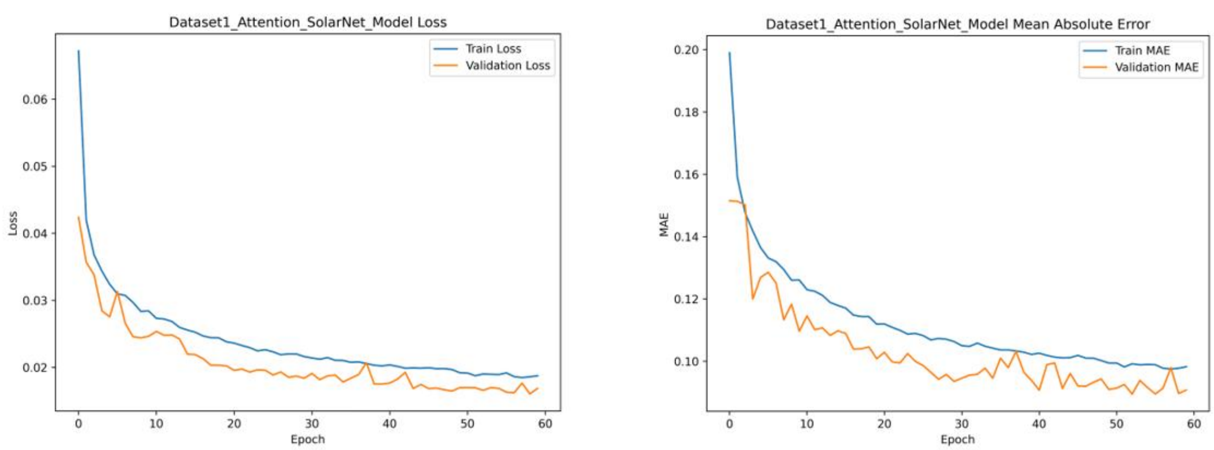


Figure 30: Dataset 1 Attention-SolarNet Model Loss & MAE

In summary, although the Attention-SolarNet model is effective on dataset 1, its performance is slightly inferior compared with the Attention-SolarMeNet model. The Attention-SolarNet model has a faster learning speed in the early stage of training, but it is inferior to the Attention-SolarMeNet model in terms of stability and prediction accuracy in the later stage of training. In addition, the Attention-SolarNet model also has a higher final loss value and MAE compared to the RadiMeNet model, which indicates that the Attention-SolarMeNet model may be a better choice when processing dataset 1 because it exhibits better stability and lower error rate during training.

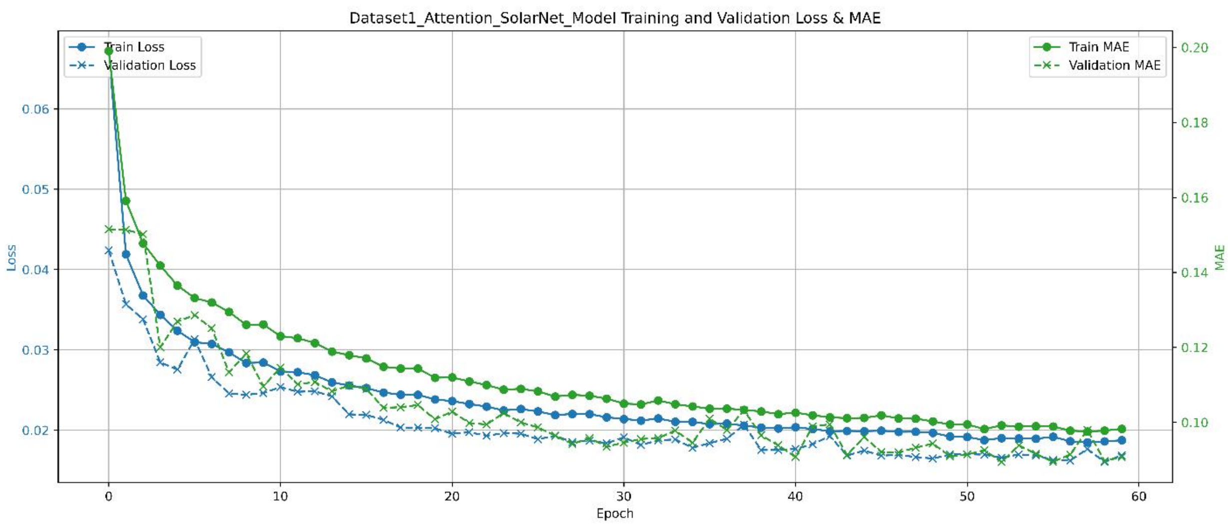


Figure 31: Dataset1 Attention-SolarNet Model Training and Validation Loss & MAE

4.1.4 No-Attention-SolarNet Model Using Dataset 1

Put dataset 1 into the No-Attention-SolarNet model. The loss value and MAE as shown in Figure 32 are finally obtained. The comprehensive figure of loss value and MAE value is shown in Figure 33.

Figure 32 shows the loss values during training and validation. In the initial stage, the loss value dropped rapidly, from 0.18 to about 0.07, indicating that the model learned significantly in the early stage. However, compared with the Attention-SolarMeNet model, the loss value of the Attention-SolarMeNet model without the attention mechanism tends to level off at the later stage of training, and finally stabilizes at a level slightly higher than 0.06, which is higher than the loss value of the Attention-SolarMeNet model. Figure 32 also shows MAE during training and validation. MAE also declined rapidly in the initial period, from 0.30 to around 0.22, showing the effectiveness of the model in terms of prediction accuracy. However, compared with the Attention-SolarMeNet model, the MAE of the Attention-SolarMeNet model without attention mechanism fluctuates more in the late training period, and finally stabilizes at a level slightly higher than 0.21, which is higher than the MAE of the Attention-SolarMeNet model. Figure 33 provides a comprehensive view of the loss value and MAE, allowing us to look at these two key metrics simultaneously. As shown in Figure 33, both the training and verified loss values and MAE maintain a lower level during the whole training process, but compared with the Attention-SolarMeNet model, the curve of the No-Attention-SolarNet model has greater volatility in the late training period, and the final loss values and MAE are slightly higher.

In summary, although the performance of the No-Attention-SolarNet is effective on dataset 1, its performance is significantly worse than that of the Attention-SolarMeNet model. The Attention-SolarMeNet model has a faster learning speed in the early training period, and shows better stability and lower error rate in the later training period. In addition, compared with the RadiMeNet model and the Attention-SolarNet model, the Attention-SolarMeNet model also has a lower final loss value and MAE, which indicates that the Attention-SolarMeNet model has better performance when processing dataset 1, and is the best choice among the three models.

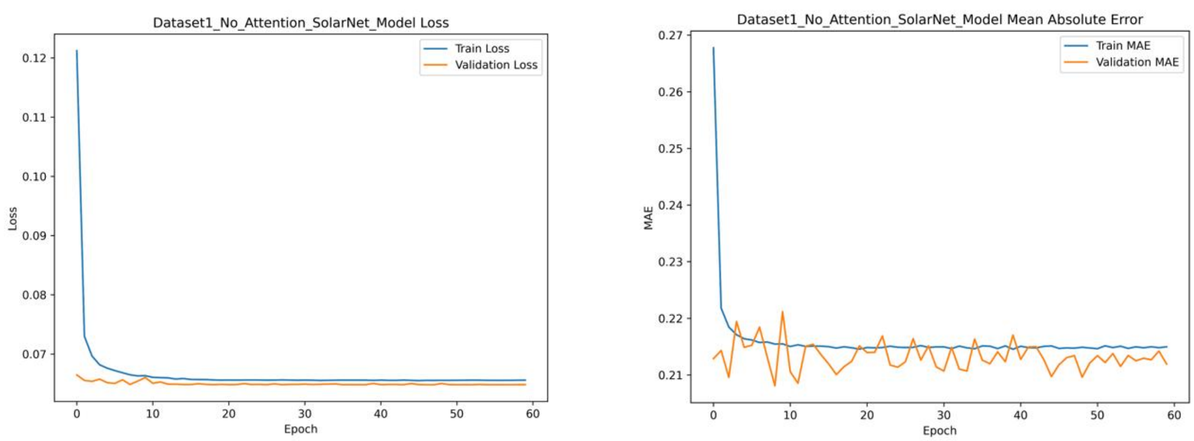


Figure 32: Dataset 1 No- Attention-SolarNet Model Loss

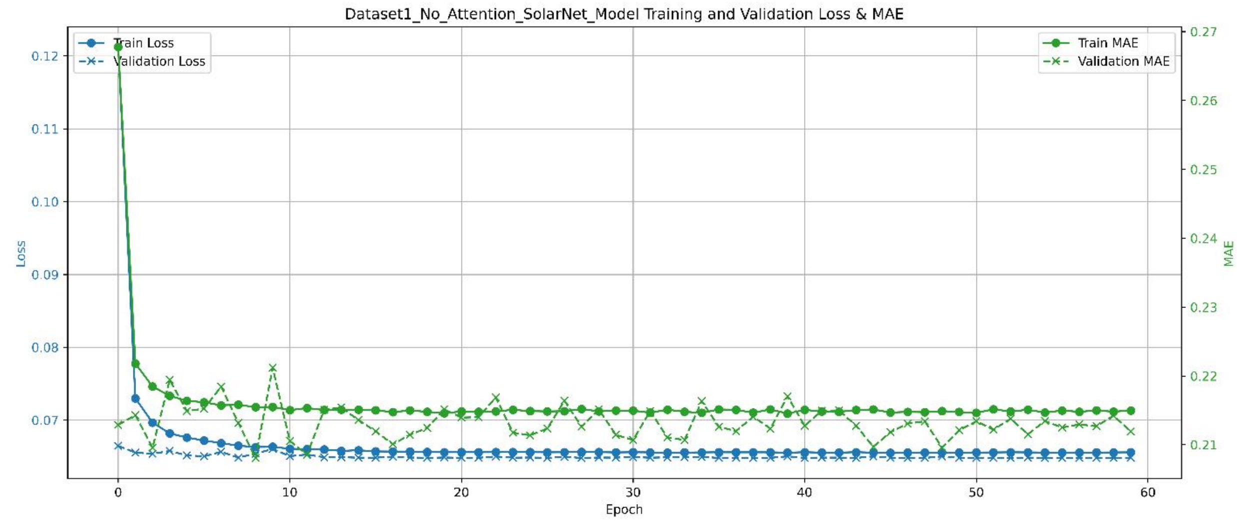


Figure 33: Dataset1 No- Attention-SolarNet Model Training and Validation Loss & MAE

4.1.5 Attention-SolarMeNet Model Using Dataset 2

Put dataset 2 into the Attention-SolarMeNet model. The loss value and MAE as shown in Figure 34 are finally obtained. The comprehensive figure of loss value and MAE value is shown in Figure 35.

Figure 34 shows the loss values during training and validation. In the initial stage, the loss value drops rapidly, from 0.0250 to about 0.0075, indicating that the model learns well in the early stage. As the training progressed, the loss value continued to decline steadily and reached a level close to 0.006 by the end of the training, which shows the stability and effectiveness of the model throughout the training process. Figure 34 also shows MAE during training and validation. MAE also declined rapidly in the initial period, from 0.10 to about 0.05, showing the effectiveness of the model in terms of prediction accuracy. At the later stage of training, MAE tends to be stable, and the MAE curves of training and verification are very close, which indicates that the model has good generalization ability. Figure 35 provides a comprehensive view of the loss value and MAE, allowing us to look at these two key metrics simultaneously. As shown in Figure 35, both the training and verified loss values and MAE maintained a low and stable level during the whole training process, which further proved the stability and prediction accuracy of the model.

In summary, the Attention-SolarMeNet model performs very well on dataset 2. The model not only performs well on the training set, but also shows stable performance on the validation set, which indicates that the model has high generalization ability. The low level and stability of the loss value and MAE further demonstrate the prediction accuracy and robustness of the model. So, this is a pretty good model.

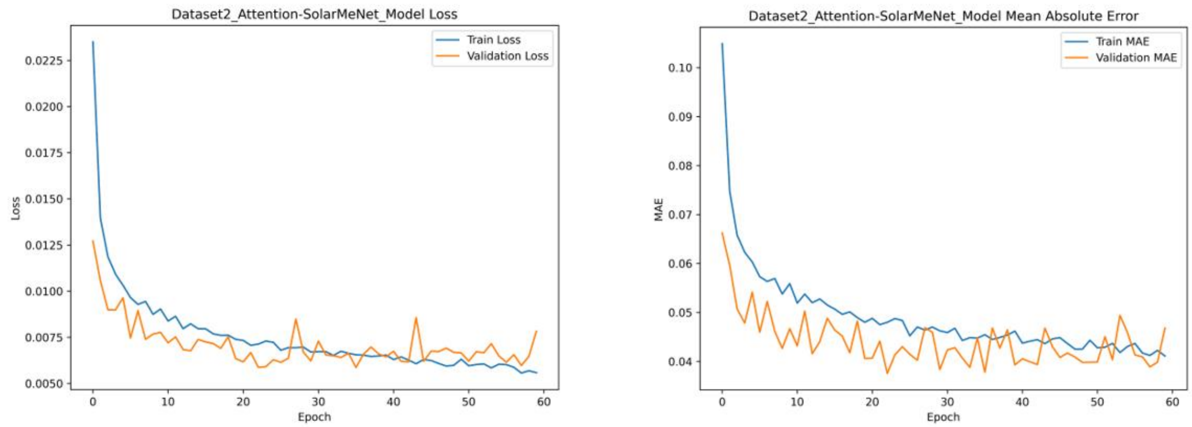


Figure 34: Dataset 2 Attention-SolarMeNet Model Loss & MAE

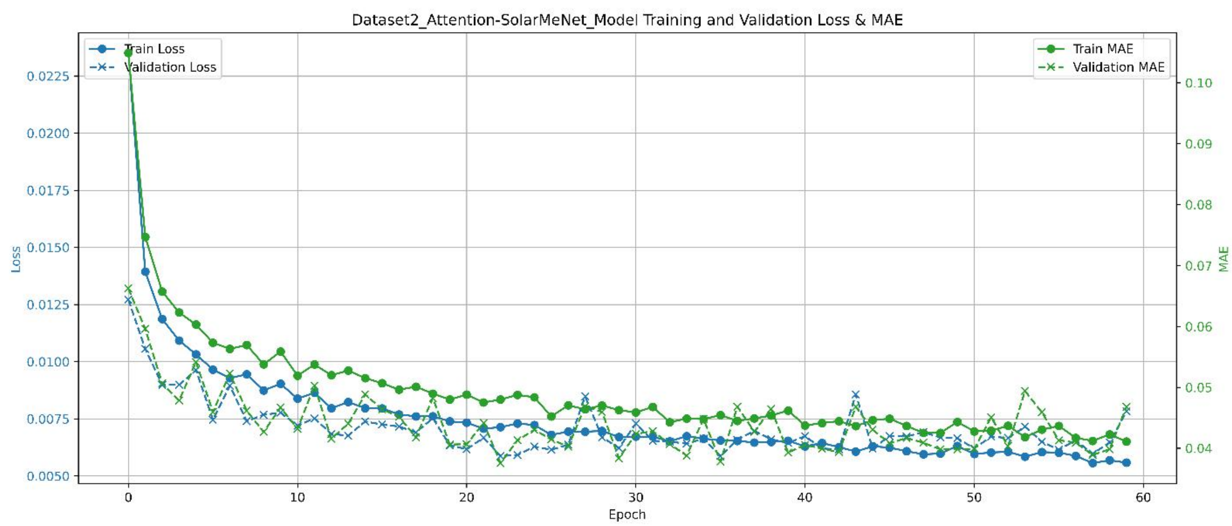


Figure 35: Dataset 2 Attention-SolarMeNet Model Training and Validation Loss & MAE

4.1.6 RadiMeNet Model Using Dataset 2

Put dataset 2 into the RadiMeNet model. The loss value and MAE as shown in Figure 36 are finally obtained. The comprehensive figure of loss value and MAE value is shown in Figure 37.

As shown in Figure 36, training and validation loss values decline rapidly in the initial phase, with training loss dropping from approximately 0.025 to near 0.006 and validation loss dropping from approximately 0.012 to near 0.006. This shows that the model learns significantly in the early stage and maintains stability and effectiveness throughout the training process. Also shown in Figure 36, MAE for training and verification also declined rapidly in the initial stage, with MAE for training decreasing from 0.12 to close to 0.05 and MAE for verification decreasing from 0.08 to close to 0.05. This shows the effectiveness of the model in terms of prediction accuracy, and in the later stage of training, MAE tends to be stable, and the training and verification MAE curves are very close, indicating that the model has good generalization ability. As shown in Figure 37, a comprehensive view of loss values and MAE provides an opportunity to observe these two key metrics simultaneously. As shown in Figure 37, both the training and verified loss values and MAE maintained a low and stable level during the whole training process, which further proved the stability and prediction accuracy of the model.

Although the RadiMeNet model may perform slightly worse on dataset 2 than the Attention-SolarMeNet model, it still shows good performance and generalization ability. The low level and stability of the loss value and MAE further demonstrate the prediction accuracy and robustness of the model.

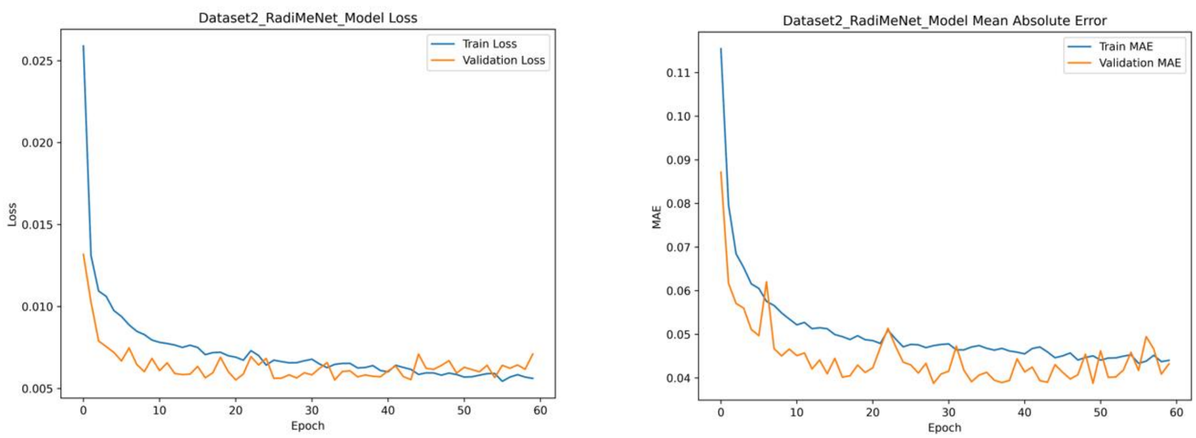


Figure 36: Dataset 2 RadiMeNet Model Loss & MAE

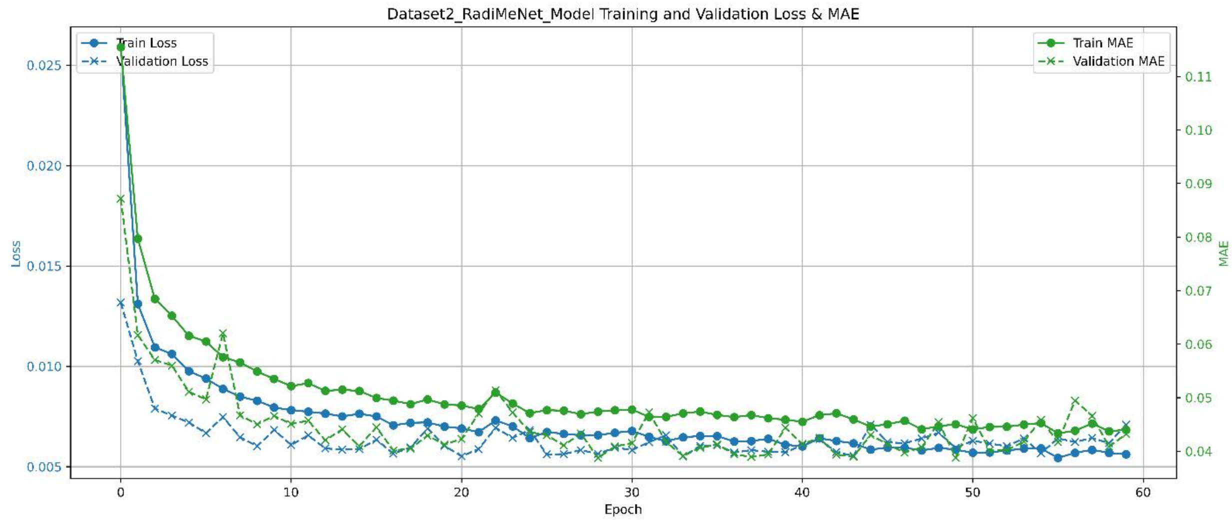


Figure 37: Dataset 2 RadiMeNet Model Training and Validation Loss & MAE

4.1.7 Attention-SolarNet Model Using Dataset 2

Put dataset 2 into the Attention-SolarNet model. The loss value and the MAE as shown in Figure 38 are finally obtained. The comprehensive figure of loss value and MAE value is shown in Figure 39.

Figure 38 illustrates the loss values during training and validation. At the beginning of training, the loss decreases rapidly, from around 0.09 to close to 0.02 for training loss and from around 0.04 to just over 0.02 for validation loss. This indicates that the model learns significantly in the early stage, but the final loss value is slightly higher than that of the Attention-SolarMeNet and RadiMeNet models, which may indicate that the model has some shortcomings in data fitting. Figure 38 also illustrates the Mean Absolute error (MAE) during training and validation. MAE also decreases rapidly at the beginning, from approximately 0.22 for training MAE to near 0.10, and from approximately 0.14 for validation MAE to near 0.08. Despite the improvement in MAE during training, the final MAE value is still higher than that of the Attention-SolarMeNet and RadiMeNet models, which may mean that the model does not perform as well as other models in terms of prediction accuracy. Figure 39 provides a comprehensive view of the loss values and MAE, allowing us to look at these two key metrics simultaneously. As shown in the Figure 39, although the loss values and MAE of training and validation decrease throughout the training process, the final index values are higher than those of Attention-SolarMeNet and RadiMeNet models, which further proves that the stability and prediction accuracy of Attention-SolarNet model on dataset 2 may not be as good as other models.

To summarize, although the Attention-SolarNet model shows some learning ability on Dataset 2, it performs worse than the Attention-SolarMeNet and RadiMeNet models. This may be due to limitations in the model structure or parameter Settings.

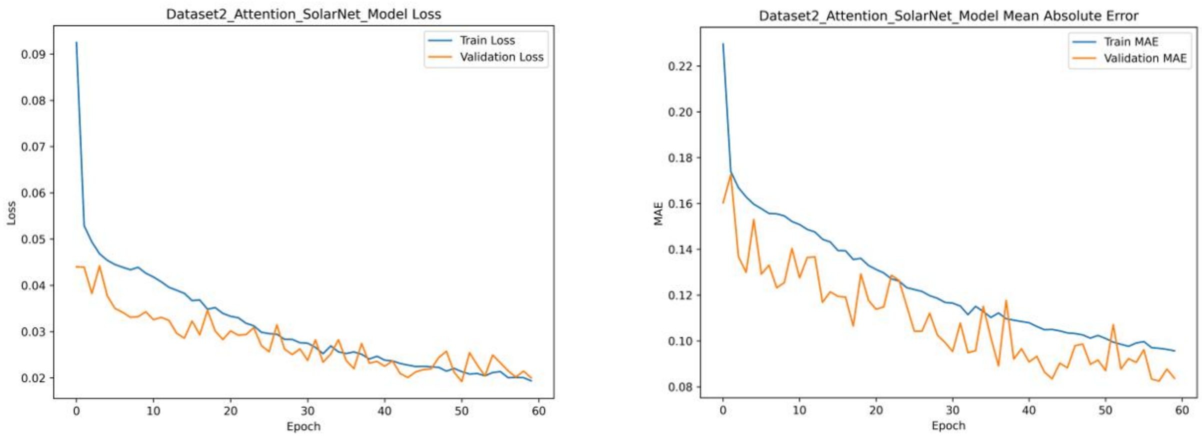


Figure 38: Dataset 2 Attention-SolarNet Model Loss & MAE

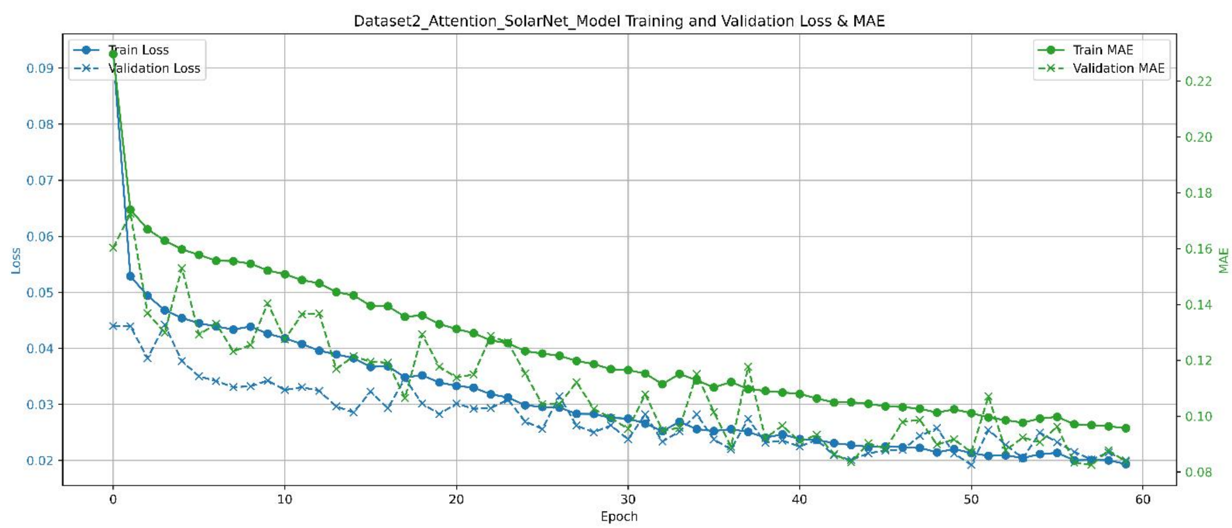


Figure 39: Dataset 2 Attention-SolarNet Model Training and Validation Loss & MAE

4.1.8 No-Attention-SolarNet Model Using Dataset 2

Put dataset 2 into the No-Attention-SolarNet model. The loss value and the MAE as shown in Figure 40 are finally obtained. The comprehensive figure of loss value and MAE value is shown in Figure 41.

As shown in Figure 40, training and validation loss values decline rapidly during the initial phase, with training loss dropping from approximately 0.14 to close to 0.05 and validation loss dropping from approximately 0.05 to just under 0.04. Although the loss value decreased during the training process, the final loss value was higher than that of other models, indicating that the model may have certain shortcomings in fitting the data. Also shown in Figure 40, training and validation MAE also declined rapidly in the initial stage, with training MAE dropping from 0.28 to nearly 0.16 and validation MAE dropping from 0.14 to slightly above 0.14. However, the final MAE value is higher than that of Attention-SolarMeNet, RadiMeNet, and Attention-SolarNet models, indicating that the No-Attention-SolarNet model does not perform as well as other models in terms of prediction accuracy. As shown in Figure 41, a comprehensive view of loss values and MAE provides an opportunity to observe these two key metrics simultaneously. As shown in the Figure 41, although the training and validation loss values and MAE decreased throughout the training process, the final index values were higher than those of the Attention-SolarMeNet, RadiMeNet and Attention-SolarNet models, which further proved that on dataset 2, The stability and prediction accuracy of the No-Attention-SolarNet may not be as good as other models.

In summary, the No-Attention-SolarNet model has the worst performance of the three models on dataset 2. In contrast, the Attention-SolarMeNet model exhibits lower loss values and MAE, demonstrating its advantages in terms of stability and prediction accuracy.

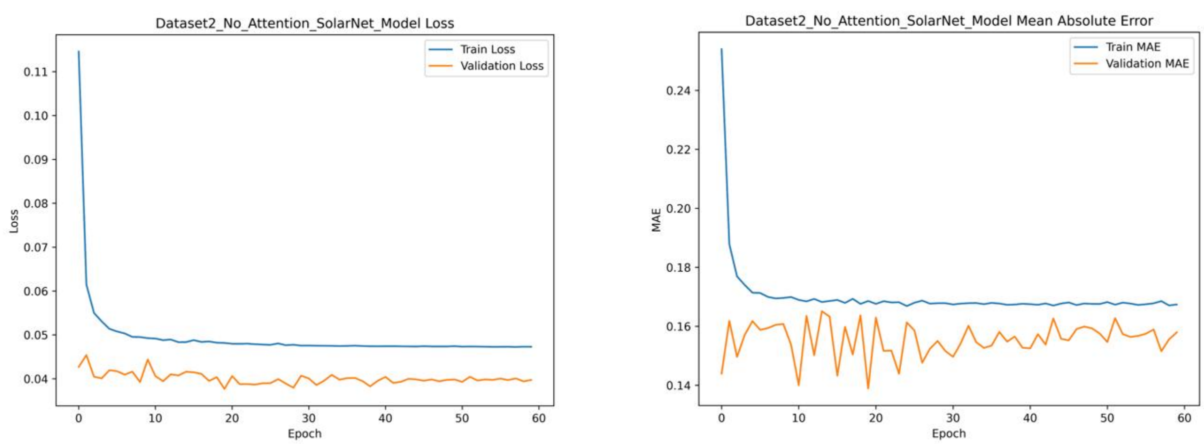


Figure 40: Dataset 2 No-Attention-SolarNet Model Loss

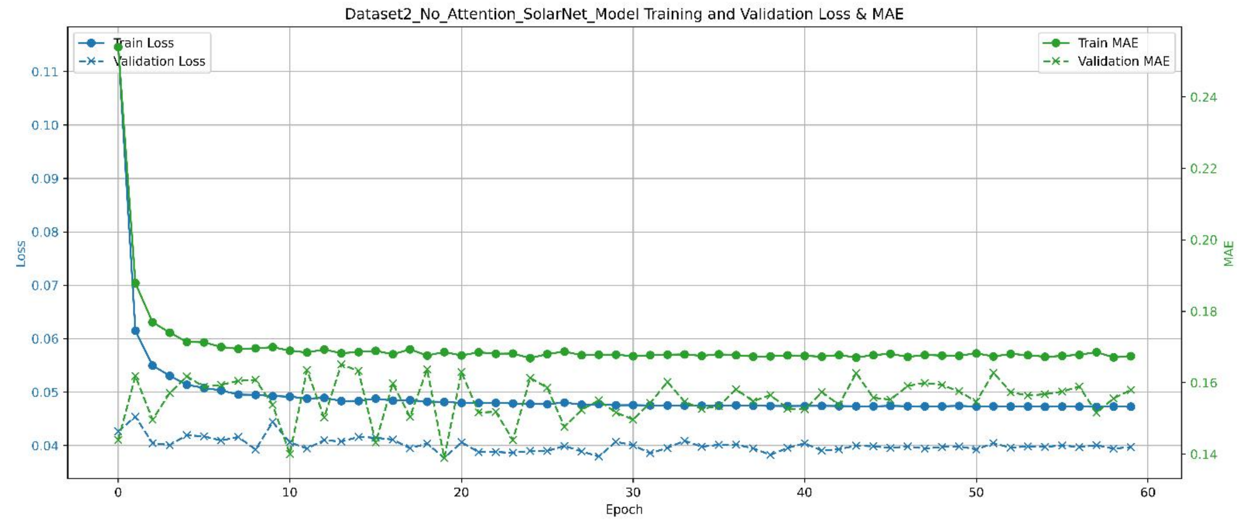


Figure 41: Dataset 2 No-Attention-SolarNet Model Training and Validation Loss & MAE

4.1.9 Comparison Chart

By comparing the results of the two data sets in different models, as shown in Table 4, it can be concluded that the Attention-SolarMeNet model is the best.

Table 4: Comparison chart

	Model	MAE	MSE	RMSE	R^2
Dataset1	Attention-SolarMeNet	0.0653	0.0096	0.0980	0.8521
	RadiMeNet	0.0766	0.0121	0.1102	0.8127
	Attention-SolarNet	0.0900	0.0160	0.1267	0.7525
	No-Attention-SolarNet	0.2149	0.0648	0.2547	-0.0009
Dataset2	Attention-SolarMeNet	0.0399	0.0067	0.0821	0.8293
	RadiMeNet	0.0419	0.0066	0.0811	0.8334
	Attention-SolarNet	0.0904	0.0224	0.1497	0.4329
	No-Attention-	0.1561	0.0401	0.2001	-0.0133

	SolarNet				

4.2 Model Explainability

Figures 42,43,44,45,46 demonstrate five different SHAP plots to visualize the model explainability. SHAP is a way of interpreting the predictions of machine learning models. In machine learning, SHAP regards features as participants, and the prediction result of the model as the total profit of cooperation. The Shapley value of each feature is calculated to quantify its contribution to the prediction result. This approach is not only applicable to any type of machine learning model, but also provides local and global explanations that help us understand the model's decision process, as well as intuitive visualizations that show the importance and impact of features [23].

4.2.1 Force Plot

As Figure 42 shows the Attention-SolarMeNet model of force type SHAP diagram about dataset1. The force diagram presents the contribution of each feature to the model prediction via intuitive bars: red bars indicate positive contributions and blue bars indicate negative contributions. The attempt is to start with the baseline prediction of the model, which is usually the average of all the predictions, and then accumulate the SHAP values for each feature to get the prediction for that sample.

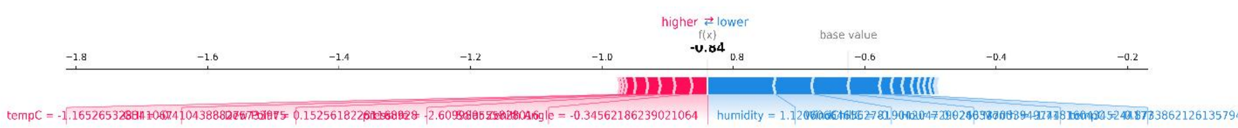


Figure 42: Force Plot

4.2.2 Bar Plot

As Figure 43 shows the Attention-SolarMeNet model of the SHAP-Bar diagram about dataset1. The SHAP-Bar diagram visually shows the average contribution of each feature to the model prediction by the length of the monochromatic bar. The longer the bar is, the more influential the feature is on the prediction result. The simplicity of the plot makes it easy to quickly identify which features contribute the most to a model's prediction, which helps users understand the model's decision-making process.

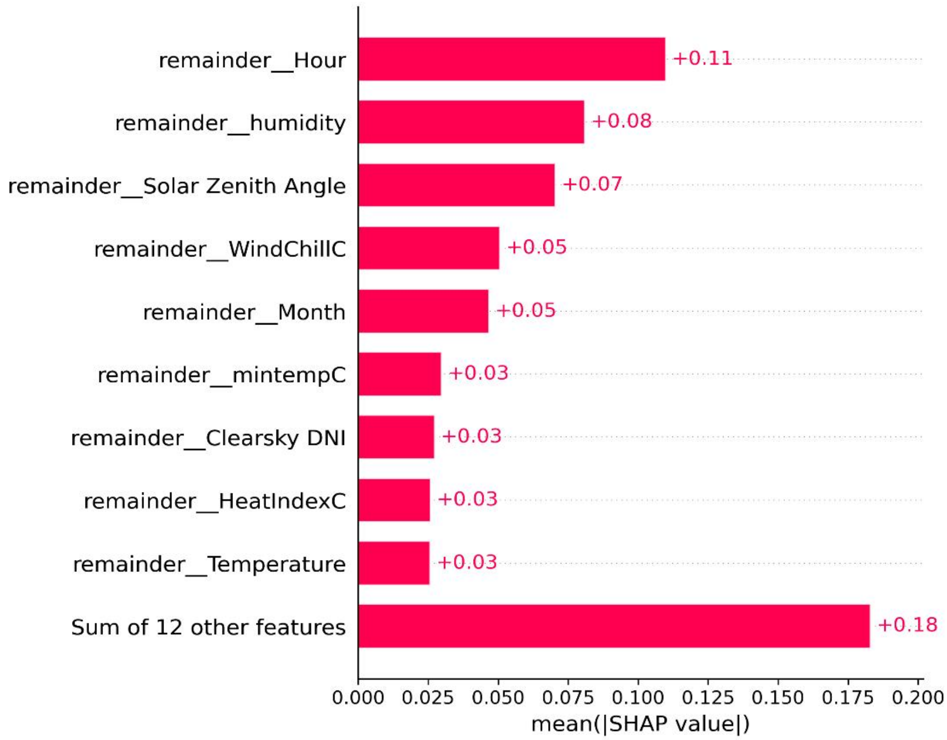


Figure 43: Bar Plot

4.2.3 Summary-Bar Plot

As Figure 44 shows the Attention-SolarMeNet model of the SHAP-Summary-Bar diagram about dataset1. The SHAP-summary-bar diagram visually shows the global importance of each feature by means of monochromatic bars. The length of the bar represents the average absolute SHAP value of the feature, which reflects the overall contribution of the feature to the model prediction, and the features are sorted from highest to lowest importance. Such plots are concise and easy to quickly identify the features that have the most impact on model predictions, and are suitable for globally explaining model behavior.

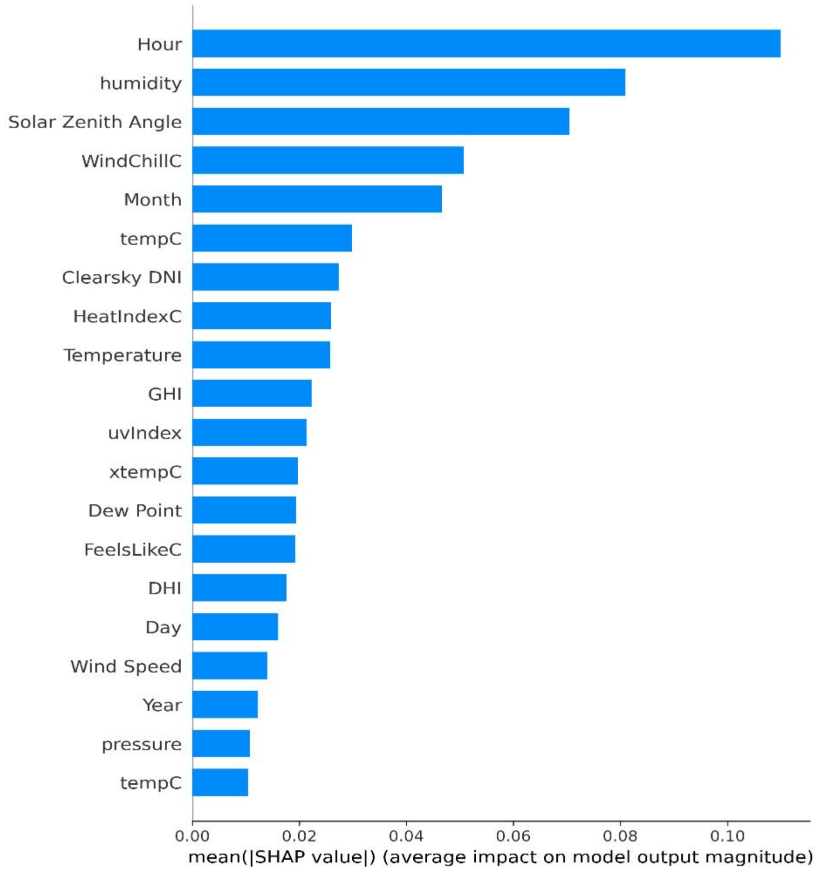


Figure 44: Summary-Bar Plot

4.2.4 Summary-Dot Plot

As Figure 45 shows the Attention-SolarMeNet model of the SHAP-Summary-Dot diagram about dataset1. The SHAP-Summary-Dot diagram visually shows the global importance of each feature by the position and color of the points. Features are ordered by their mean absolute SHAP values from highest to lowest, with important features at the top of the plot. Each point represents the value of a feature and its SHAP value, and the position of the point indicates the contribution of the feature to the prediction, with positive contributions on the right and negative contributions on the left. The color distinguishes the value of the feature, usually, red indicates high values, and blue indicates low values.

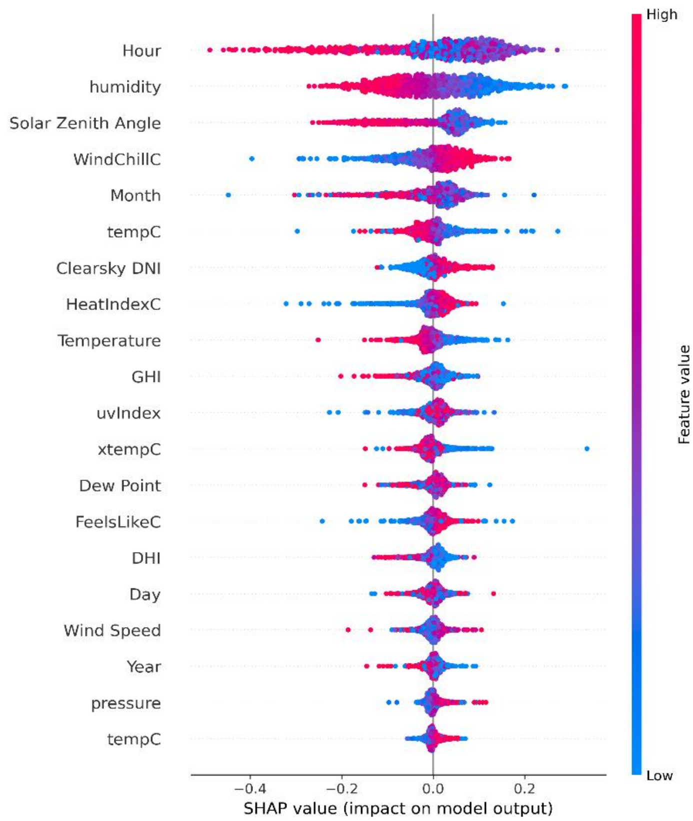


Figure 45: Summary-Dot Plot

4.2.5 Waterfall Plot

As Figure 46 shows, the Attention-SolarMeNet model of the SHAP-Summary-Dot diagram about dataset1. The waterfall diagram provides an intuitive visual explanation of the prediction results for a single sample. It starts with the baseline value of the model (usually the average prediction of the background dataset). It gradually shows each feature in a waterfall that is positive to the prediction in red or negative to the prediction in blue, eventually accumulating to the prediction value for that sample. The features are sorted by their contribution, with the important ones first, so the user can see which features have the most impact on the prediction and how the feature values drive the change in the model output.

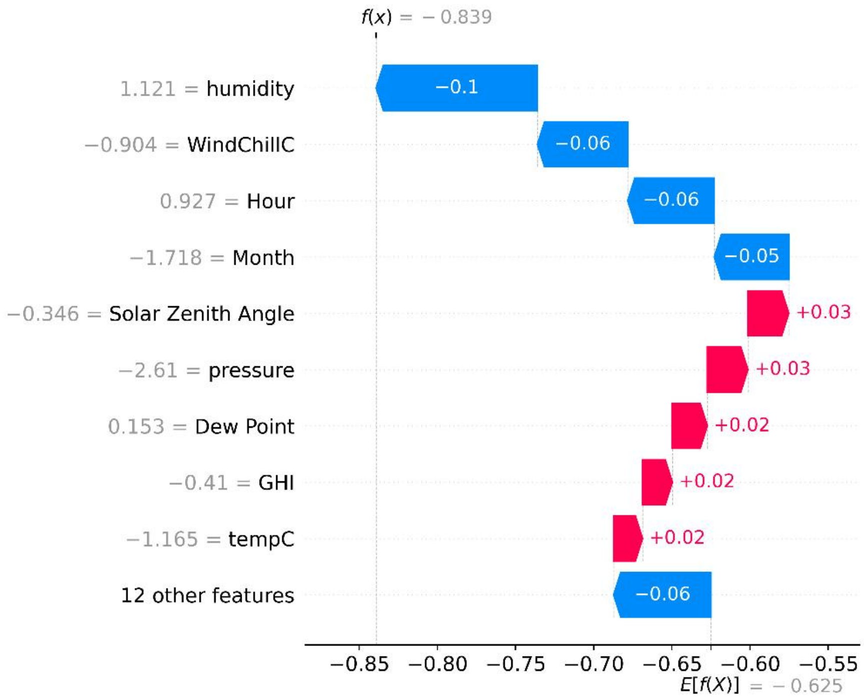


Figure 46: Waterfall Plot

4.3 Model Deployment

The provided Figure 47 presents the deployment website of our proposed model. On the home page, users can learn how to use the site. And understand the types of models used on the website and the advantages of the model. Not only that, one can upload the dataset in the left navigation bar and use the model on the website to make predictions.

Figure 47: Home Page of Website

The site is divided into three functions. The first of these features, shown in Figure 48 and Figure 49, is Model Training Performance. By uploading the dataset in the navigation bar of the main interface, the data visualization of the loss value, MAE value, and the combination of the two results is obtained by training the model. In this way, it is more intuitive to see whether the training results are good.

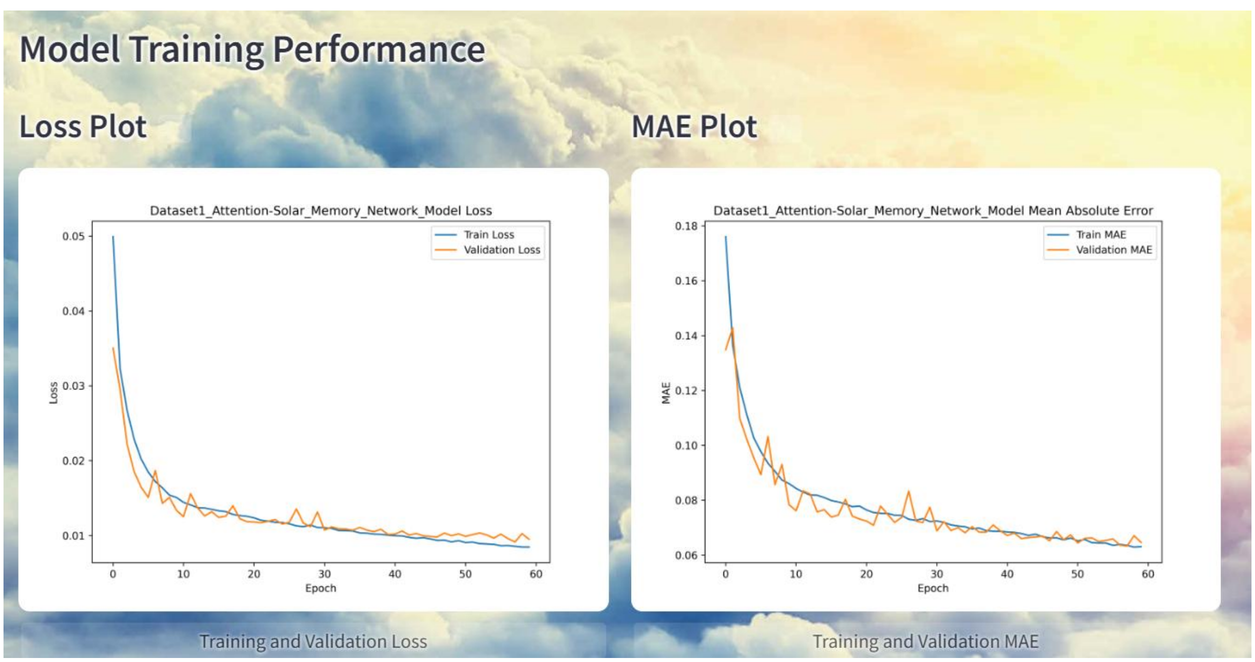


Figure 48: Model Training Performance (Loss value and MAE value) of Website

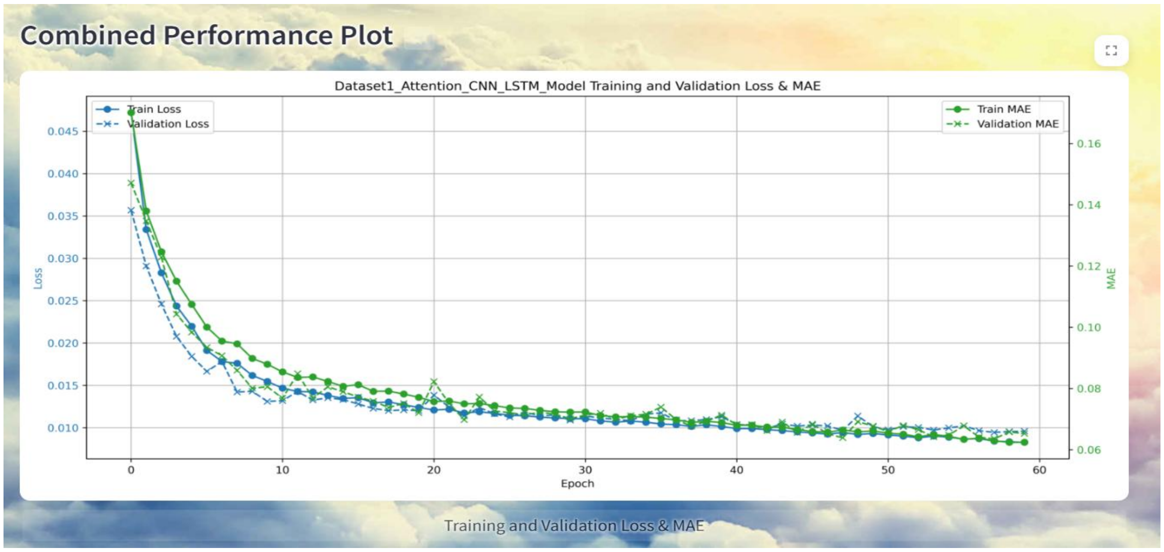


Figure 49: Model Training Performance (Combined value) of Website

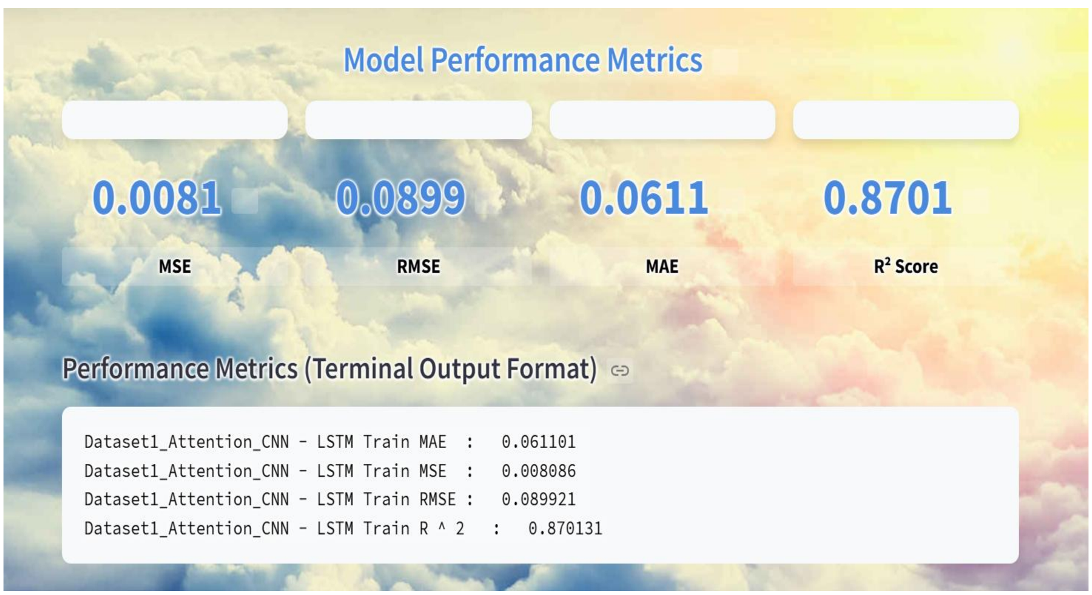


Figure 50: Performance Metric (1)

The second function is to show the performance metrics of the model in Figure 50. This feature gives you a comprehensive view of how your model performs on different evaluation criteria, including but not limited to mean square Error (MSE), Root Mean Square error (RMSE), Mean Absolute error (MAE), and coefficient of determination (R^2). These metrics can help users determine the effectiveness of the model in a specific task and whether the model can meet the needs of the actual application scenario. In addition, by comparing the performance indicators of different models, users can also select the most suitable deployment model to improve the practicality and reliability of the model.

Not only that, but also the distribution map of the error is shown. Used to show the distribution of differences between the predicted values and the actual values. Figures 51, 52 show that the error between the actual value and the predicted value is not large. Also, it provides a comparison chart, as Figure 53 showing between the predicted value and the real value to see the predicted result more intuitively.

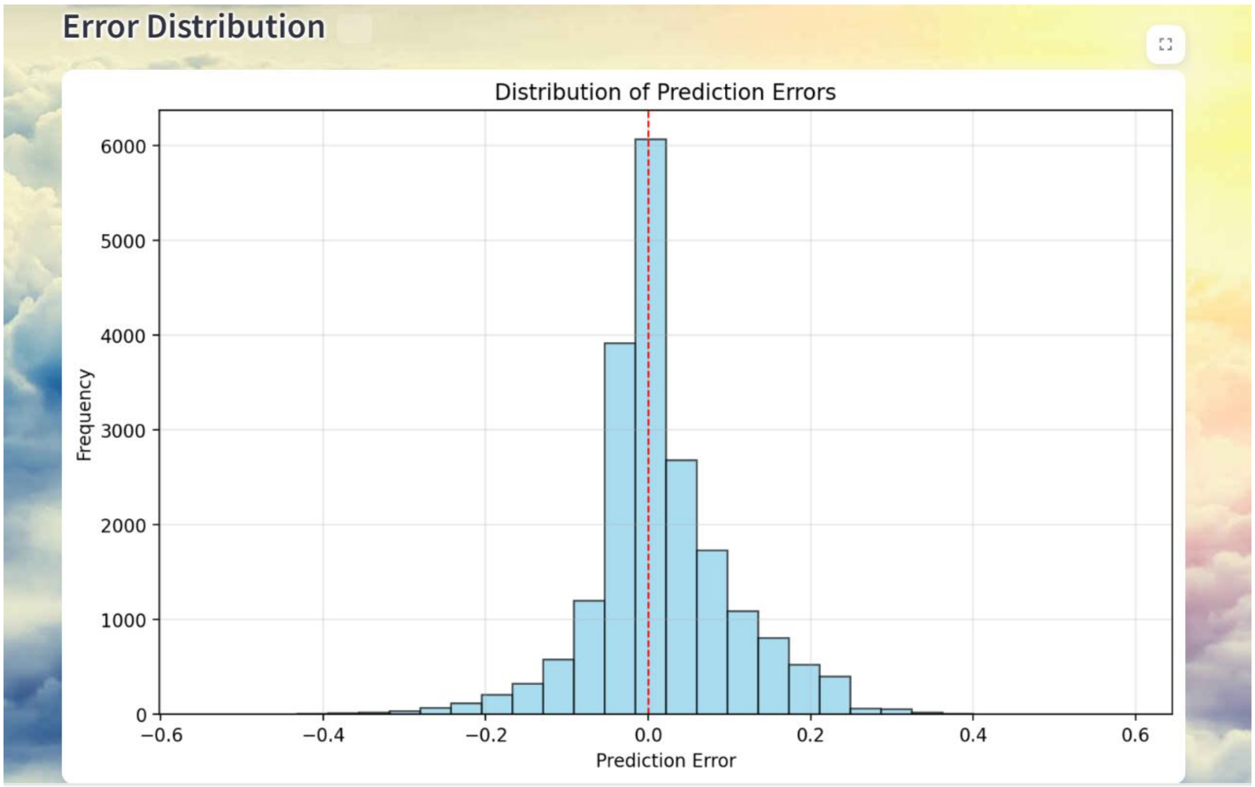


Figure 51: Error Distribution

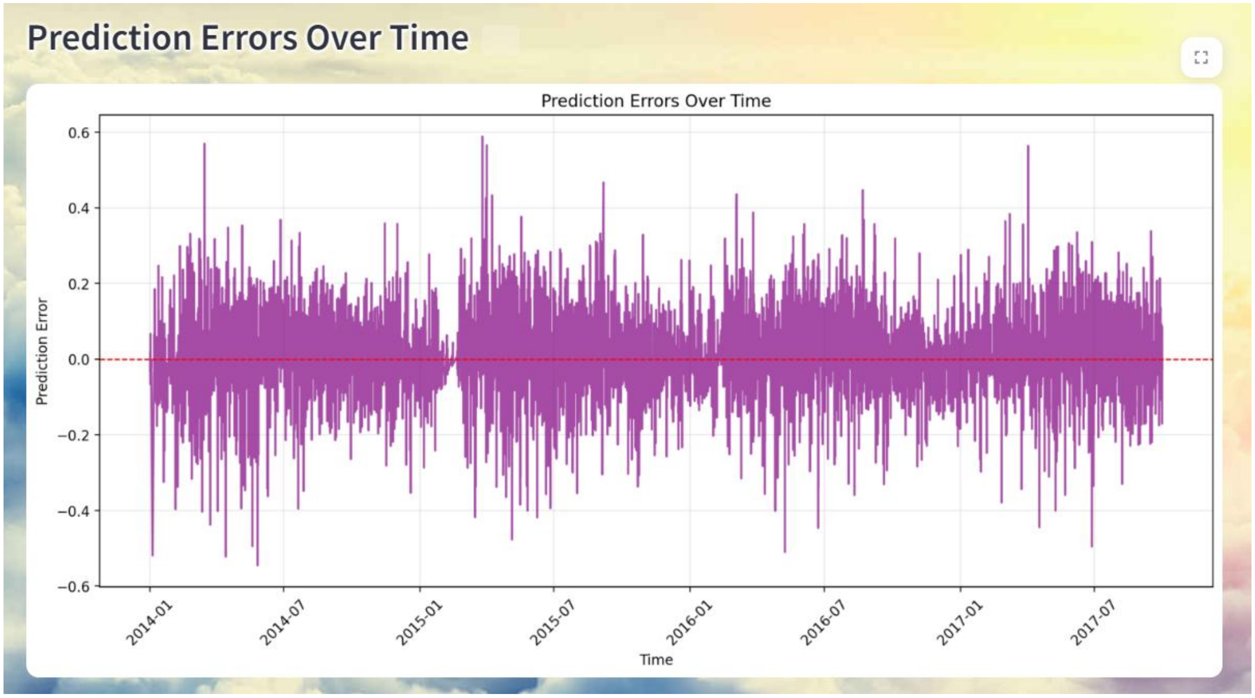


Figure 52: Prediction Errors Over Time

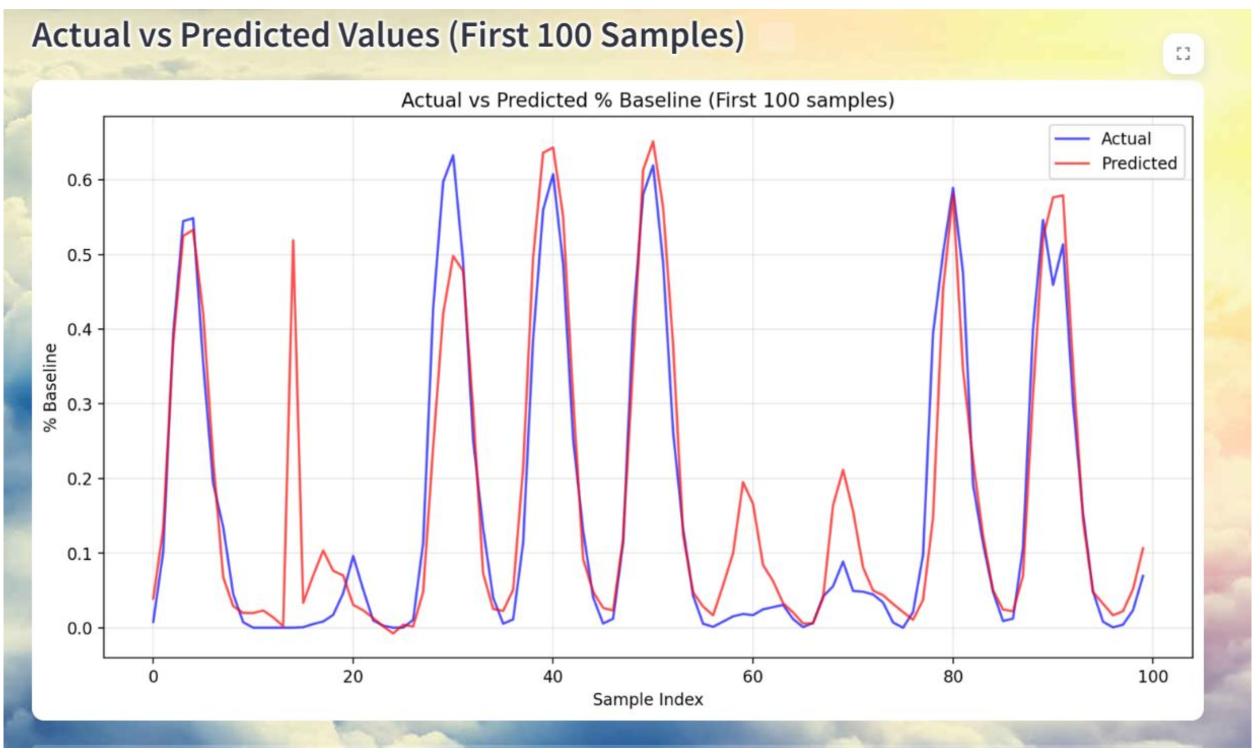


Figure 53: Actual vs Predicted Values

The third feature is to display a SHAP plot of the model as Figures 54, 55, 56, and 57. SHAP plots graphically reveal the degree and direction of influence of each feature on the model output. We also gain a deeper understanding of the inner workings of our models, identify key features, and evaluate the fairness and interpretability of our models to make more informed decisions during model development and deployment.

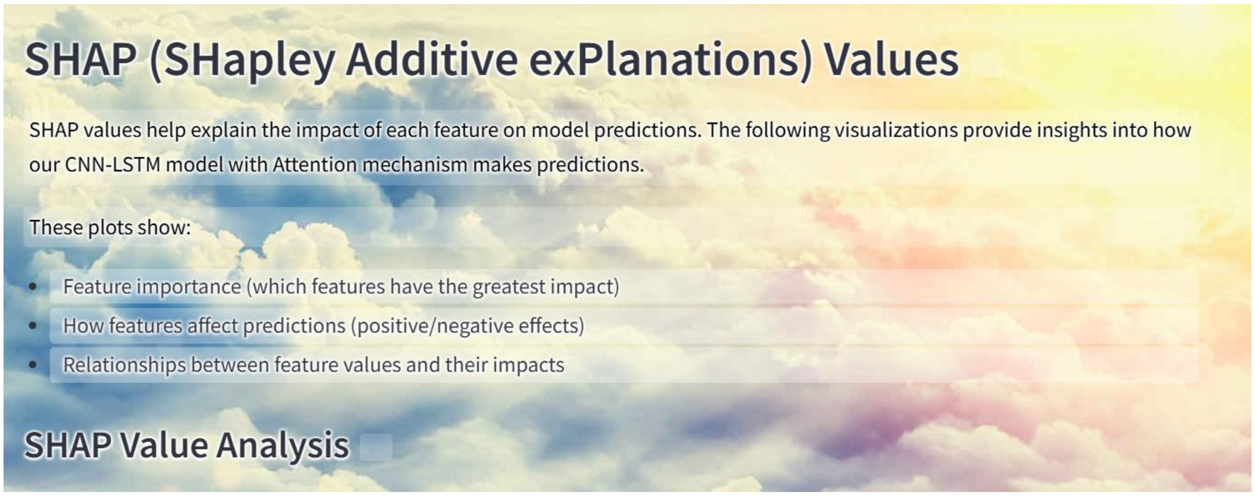


Figure 54: SHAP Value (1)

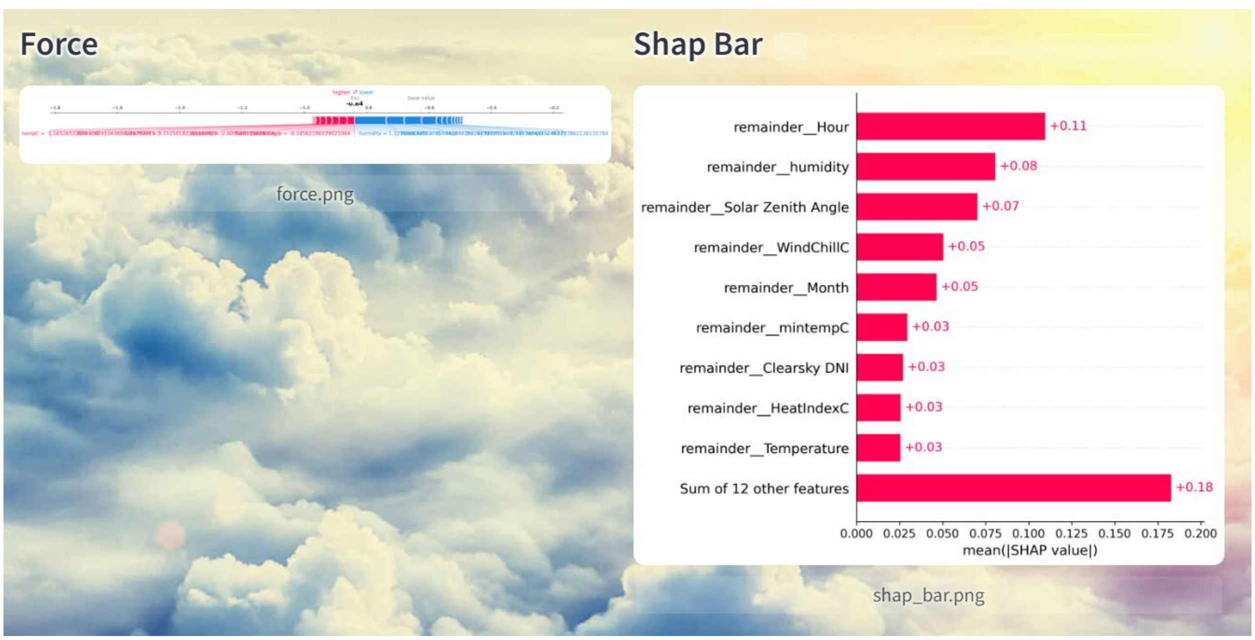


Figure 55: SHAP Value (2)

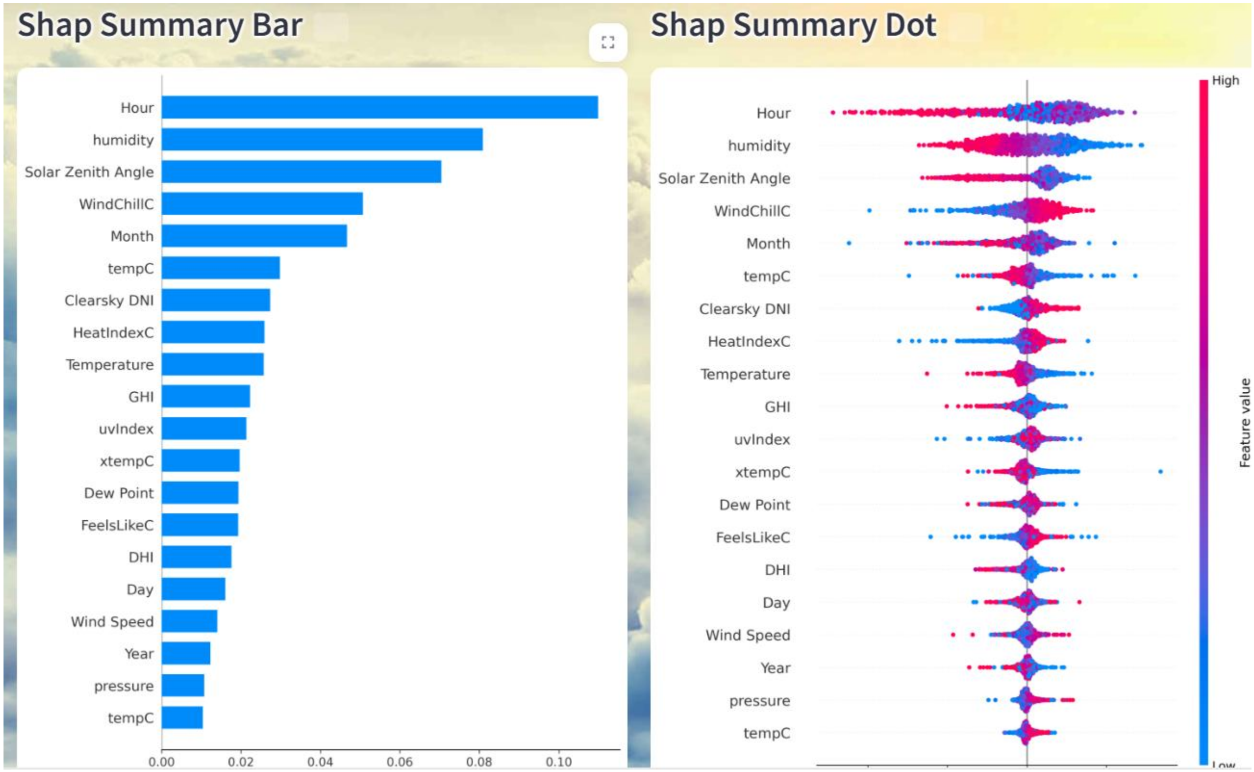


Figure 56: SHAP Value (3)

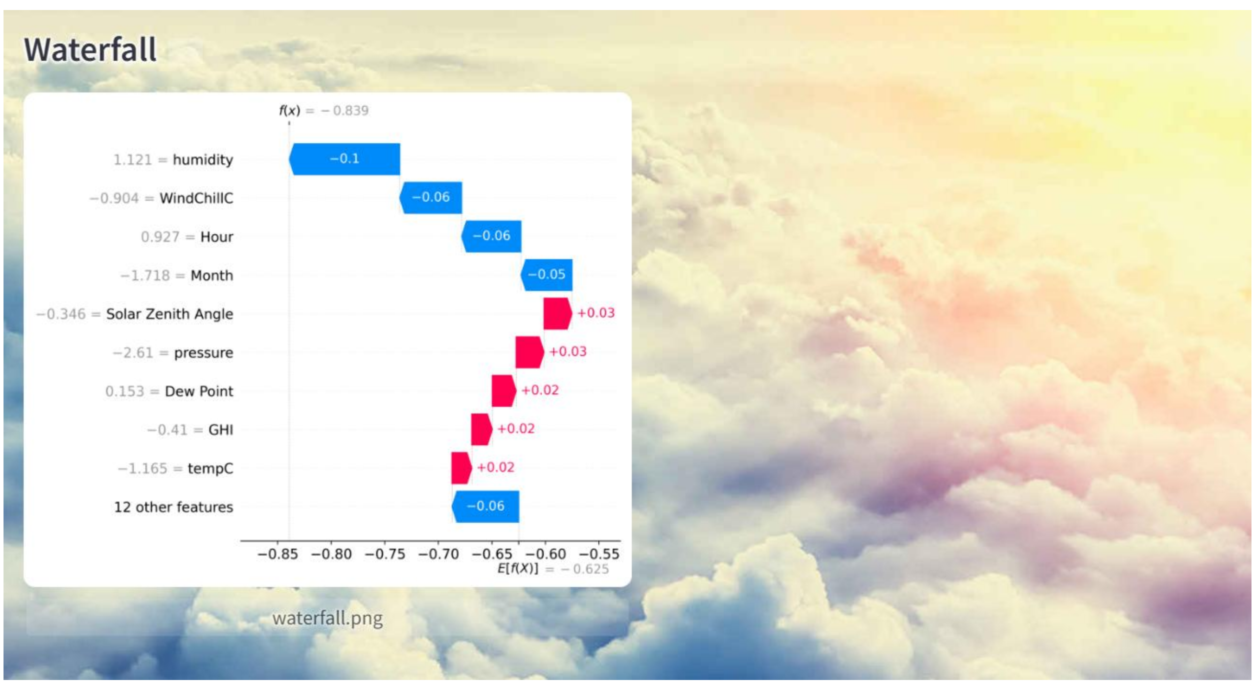


Figure 57: SHAP Value (4)

Chapter 5 Professional Issues

The following elaborates on all aspects of project management, including activity arrangement, time planning, data management and project deliverables, providing a systematic framework for the smooth implementation of the project

5.1 Project Management

5.1.1 Activities

The task list is shown in Table 5

Table 5: Objectives at each phase

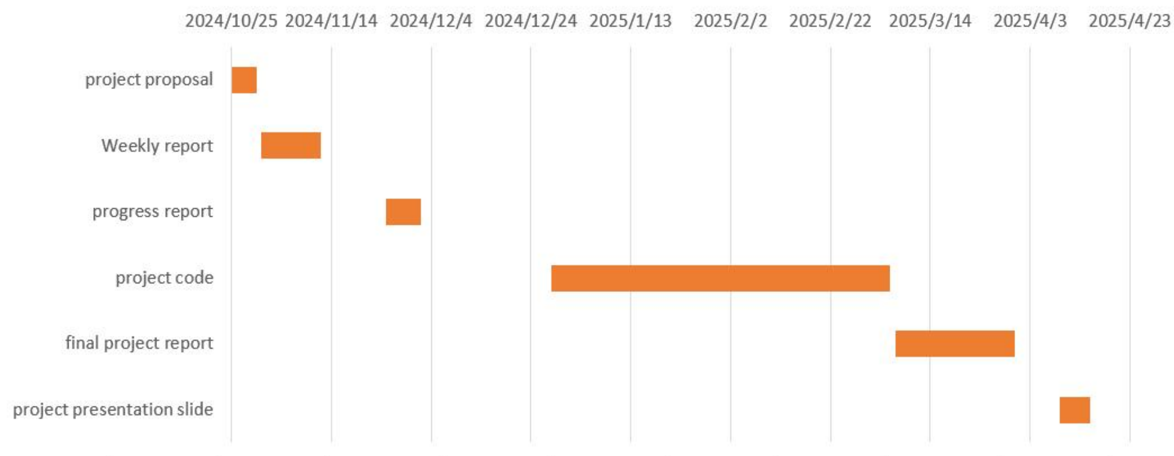
Phase	Objective
Preparation	♦ Review deep learning for solar capacity forecasting. ♦ For some background. ♦ Bring up existing issues. ♦ Try to find what works.
Learn about deep learning knowledge	♦ Learning applies to recognizing image CNN models. ♦ Learn the LSTM model ♦ Master loss functions, optimizers, model building, and optimization. ♦ Master and skillfully use model formulas.
Collect data	♦ Find 2-3 datasets from GitHub. ♦ Divide the data set into training and test sets, and scale them.
Implement experimental ideas	♦ Establish the Attention-SolarMeNet model. ♦ Train, analyze, and compare models. ♦ Adjust hyperparameters to improve model accuracy. ♦ Build the Attention-SolarNet model,

	RadiMeNet model, and No-Attention-SolarNet model to compare with each other.
Test the model and complete the experiment	♦ Check whether the model works properly and the generalization ability of the model. ♦ Check if the Attention-SolarMeNet model is the best among other models.

5.1.2 Schedule

The schedule is shown in Table 6 below

Table 6: Gantt Chart



5.1.3 Project Data Management

A. All files, including datasets, model codes, references, weekly reports, and all sorts, will be replicated into three copies for fail-safe, one on the local computer, one on the hard drive, and one on GitHub.

B. Upload the project to GitHub for every modification, and synchronize the latest version.

The following are documents for the Project for uploading and synchronization:

Reports (Weekly, Proposal, Progress, Final) & Presentation PPT

Attention-SolarMeNet model diagram

References

Datasets Link

Model evaluation documents

Attention-SolarMeNet model codes

Model Deployment Codes

5.1.4 Project Deliverables

Each activity has a corresponding deliverable; they are listed as follows， and as detailed in Table 7.

(1) Detailed Project Proposal and Ethics Forms

(2) Interim Deliverables

(3) Progress Report

(4) Project Log

(5) Project Presentation and Demonstration

(6) Final Report

(7) Legal, Social, Ethical, Environmental, and Professional Issues

(8) Dataset

(9) Project code

(10) Evaluation of the models

Table 7: Deliverables

Deliverables table		
	Activities	Deliverables
1	Collection of documents	Read at least 300 literature methods.
2	Project Proposal	Complete thesis proposal with logical structure and content.

3	Method to study	Research and learn paper methods, finish thesis research.
4	Method implementation	According to the research methodology, complete the code imp lementation.
5	Methods to improve	Improve existing methods, improve video accuracy.
6	Experimental data	The data is authentic and reliable.
7	Experiment 1	Various neural networks are experimented with to find the most suitable network.
8	Experiment 3	Find the most suitable model for predicting solar output based on deep learning.
9	Experiment 5	Find the optimization result.
10	Thesis writing	Complete essay writing
11	Paper Modify	Complete the thesis logic and other improvements in the format.
12	Presentation Prepare	Prepare the PPT required for presentation, review the thesis, and complete the defense preparation.

5.2 Risk Analysis

The research of solar energy prediction based on deep learning faces many challenges, such as technology, resources, and scheduling. When doing a risk analysis, you need to consider the following: Table 8.

Model training problems: Model training problems often arise from data quality issues, such as inaccurate or noisy data, and imbalanced data (i.e., some classes have far more data points than others), which can cause the model to learn wrong patterns and affect the prediction accuracy and generalization ability of the model.

Memory leaks: Memory leaks usually occur when the amount of model training data is too large to be processed by the available hardware resources, which may lead to the exhaustion of system resources, affect the stability of the model training process, or even lead to training failure.

Software version incompatibility: Software version incompatibility may occur when using different versions of libraries or frameworks, due to interface or functionality changes, which may cause the model to not run correctly or produce unexpected errors, affecting the development schedule and model performance.

Table 8: Risk Analysis

Risk ID	Potential Risk	Potential Causes	Severity	Likelihood	Risk	Mitigation
R 1.1	Model training issues	Low data quality	4	1	4	Reduce the risk of overfitting
		Data imbalance	2	4	8	Increase & decrease
R 1.2	Memory leakage	Model training data exceeds the hardware's ability	4	3	12	Use a cloud service & use a high configuration computer
R 1.3	Software version	Version incompatibility	4	3	12	Find compatible versions
R 1.4	Miss deadline	Poor management of time	4	1	4	Follow the schedule strictly

5.3 Professional Issues

A wide range of legal, social, ethical, and environmental issues were critical in the development of the project, including compliance with professional standards set by bodies such as ACM (Association for Computing Machinery) and BCS (British Computer Society).

Legal issues: In the process of collecting and processing solar energy data and weather data, it is essential to strictly abide by relevant laws and regulations, especially the legal requirements regarding personal privacy protection and data security, to ensure the legal acquisition, storage and use of data, and prevent data leakage or abuse. Meanwhile, as solar output detection technology involves innovative methods and models, its application process may encounter intellectual property and patent legal issues, such as the ownership of technological achievements, the compliance of patent applications, and the authorization of technology usage, etc. All these issues need to be fully paid attention to and properly resolved during the project's advancement

Social issues: In practical applications, solar detection technology may involve monitoring and data collection in public areas. This not only may have an impact on public safety but may also be misunderstood as excessive monitoring, thereby raising public concerns about privacy infringement and the expansion of monitoring scope. For instance, during the installation and operation of solar detection equipment, images or data of the surrounding environment may be captured. If these data are improperly used or leaked, they may pose a threat to personal privacy. Therefore, when developing and promoting this technology, researchers must fully consider its possible social impacts, conduct comprehensive risk assessments in advance, formulate corresponding privacy protection measures and data management norms, ensure that the application of the technology complies with social ethics and public interests, and avoid unnecessary social disputes and public resistance

Ethical issues: In the process of collecting and processing solar energy data and weather data, great importance must be attached to the informed consent rights and other legitimate rights and interests of the test subjects. This not only involves clearly informing the test subjects of the purpose, usage method and scope of the data collection to ensure they fully understand and voluntarily participate, but also involves strictly adhering to the data protection principle during the data processing and analysis to prevent unauthorized use or leakage of the data. At the same time, it is necessary to avoid discriminatory applications against specific groups or individuals due to improper processing of data, ensure the fair and impartial use of data, comply with ethical and moral standards, and maintain social fairness and inclusiveness.

Environmental issues: During the training of deep learning models and data processing, a large amount of computing resources is usually consumed. This not only leads to a significant increase in energy consumption but may also harm the environment. For example, large-scale neural network training often requires high-performance computing devices to operate for a long time. During the operation of these devices, a large amount of electricity is consumed, and the production of electricity is often accompanied by environmental issues such as carbon emissions. Therefore, researchers and developers need to fully consider the efficient utilization of energy consumption and computing resources during the model design and training process. They should take measures such as optimizing algorithms, rationally allocating computing tasks, and choosing energy-saving hardware to minimize unnecessary resource waste and reduce the impact on the environment. Meanwhile, attention should also be paid to resource optimization in the data processing process to avoid redundant computing and data storage, so as to achieve sustainable deep learning applications.

Chapter 6 Conclusion

In the whole project, a new deep learning model, Attention-SolarMeNet, has been successfully used and verified on the GitHub shared dataset. It has been proved that the solar energy value can be predicted with high accuracy through this model. This model combines the advantages of extracting features in the Attention-SolarNet model and the advantages of predicting according to time series in the RadiMeNet model and using Attention to make its feature points more prominent and enhance the prediction accuracy. Impressive performance metrics were produced. Dataset 1 includes an MAE of 0.065, an MSE of 0.009, an RMSE of 0.098, and an R2 of 0.852. Dataset 2 includes an MAE of 0.039, MSE of 0.006, RMSE of 0.081, and R2 of 0.829. This shows that the model has great potential to be an important tool for solar energy prediction, and also shows that most of the data applies to the model. By reducing the number of trainable layers, the efficiency and accessibility of the model are improved, especially in real-time diagnostic applications. The project's comprehensive methodology, including dataset preparation such as data cleaning, model construction, and various experiments, demonstrated its effectiveness and utility. The comparative analysis further supports the uniqueness of the proposed model.

Although Attention-Solarmenet performs well on datasets related to solar energy utilization, one of its limitations is that issues such as the need for Attention, continuous learning ability. The continuous learning ability of this model needs to be improved. With the continuous development of solar energy utilization technology and the continuous accumulation of data, models need to be able to dynamically learn new knowledge and patterns to adapt to the constantly changing environment and demands. However, when the current model is faced with new data, it often needs to retrain the entire model, which is not only time-consuming and labor-intensive, but also may lead to the decline of model performance.

In the future, to better meet the practical application requirements, people need to start from multiple aspects. Strengthen data preprocessing and feature engineering to ensure the quality and diversity of the input data, thereby providing a better foundation for model training. In addition, in combination with the demands of actual application scenarios, targeted adjustments and optimizations should be made to the model, such as enhancing its real-time performance, interpretability and adaptability, so that it can better cope with the dynamically changing environment and data. Meanwhile, through continuous learning and model update mechanisms, it is ensured that the model can constantly adapt to new data and task requirements, thereby playing a greater role in key areas such as solar energy prediction.

References

[1] S.-C. Lim, J.-H. Huh, S.-H. Hong, C.-Y. Park, and J.-C. Kim, "Solar Power Forecasting Using CNN-LSTM Hybrid Model," *Energies*, vol. 15, no. 21, Art. no. 21, Jan. 2022, doi: 10.3390/en15218233.

[2] X.-Y. Li, X.-Y. Dong, S. Chen, and Y.-M. Ye, "The promising future of developing large-scale PV solar farms in China: A three-stage framework for site selection," *Renew. Energy*, vol. 220, p. 119638, Jan. 2024, doi: 10.1016/j.renene.2023.119638.

[3] "Automated Deep CNN-LSTM Architecture Design for Solar Irradiance Forecasting | IEEE Journals & Magazine | IEEE Xplore." Accessed: Oct. 27, 2024. [Online]. Available: https://ieeexplore.ieee.org/document/9481345

[4] "Modelling and Experimental Validation of Aging Factors of Photovoltaic Solar Cells | IEEE Journals & Magazine | IEEE Xplore." Accessed: Mar. 16, 2025. [Online]. Available: https://ieeexplore.ieee.org/document/9475857

[5] P. C. Okonkwo *et al.*, "Solar PV systems under weather extremes: Case studies, classification, vulnerability assessment, and adaptation pathways," *Energy Rep.*, vol. 13, pp. 929–959, Jun. 2025, doi: 10.1016/j.egyr.2024.12.067.

[6] S. Impram, S. Varbak Nese, and B. Oral, "Challenges of renewable energy penetration on power system flexibility: A survey," *Energy Strategy Rev..*, vol. 31, p. 100539, Sep. 2020, doi: 10.1016/j.esr.2020.100539.

[7] S. Indolia, A. K. Goswami, S. P. Mishra, and P. Asopa, "Conceptual Understanding of Convolutional Neural Network- A Deep Learning Approach," *Procedia Comput. Sci.*, vol. 132, pp. 679–688, Jan. 2018, doi: 10.1016/j.procs.2018.05.069.

[8] H. Shi, A. Wei, X. Xu, Y. Zhu, H. Hu, and S. Tang, "A CNN-LSTM based deep learning model with high accuracy and robustness for carbon price forecasting: A case of Shenzhen's carbon market in China," *J. Environ. Manage.*, vol. 352, p. 120131, Feb. 2024, doi: 10.1016/j.jenvman.2024.120131.

[9] A. Agga, A. Abbou, M. Labbadi, Y. E. Houm, and I. H. Ou Ali, "CNN-LSTM: An efficient hybrid deep learning architecture for predicting short-term photovoltaic power production," *Electr. Power Syst. Res.*, vol. 208, p. 107908, Jul. 2022, doi: 10.1016/j.epsr.2022.107908.

[10] H. Hayakawa, H. Tamazawa, A. D. Kawamura, and H. Isobe, "Records of sunspot and aurora during CE 960–1279 in the Chinese chronicle of the Sòng dynasty," *Earth Planets Space*, vol. 67, no. 1, p. 82, May 2015, doi: 10.1186/s40623-015-0250-y.

[11] Y. Ledmaoui, A. El Maghraoui, M. El Aroussi, R. Saadane, A. Chebak, and A. Chehri, "Forecasting solar energy production: A comparative study of machine learning algorithms," *Energy Rep.*, vol. 10, pp. 1004–1012, Nov. 2023, doi: 10.1016/j.egyr.2023.07.042.

[12] D. Gottwald, M. Parmar, and A. Zureck, "Forecasting Solar Power Generation: A Comparative Analysis of Machine Learning Models," in *2024 International Conference on Renewable Energies and Smart Technologies (REST)*, Jun. 2024, pp. 1–5. doi: 10.1109/REST59987.2024.10645372.

[13] S. Theocharides, G. Makrides, G. E. Georghiou, and A. Kyprianou, "Machine learning algorithms for photovoltaic system power output prediction," in *2018 IEEE International Energy Conference (ENERGYCON)*, Jun. 2018, pp. 1–6. doi: 10.1109/ENERGYCON.2018.8398737.

[14] M. K. Boutahir, Y. Farhaoui, M. Azrour, A. Sedik, and M. M. Nasralla, "Advancing Solar Power Forecasting: Integrating Boosting Cascade Forest and Multi-Class-Grained Scanning for Enhanced Precision," *Sustainability*, vol. 16, no. 17, Art. no. 17, Jan. 2024, doi: 10.3390/su16177462.

[15] H. K. Adli, K. A. K. Husin, N. H. M. Hanafiah, M. A. Remli, F. Ernawan, and P. W. Wirawan, "Forecasting and Analysis of Solar Power Output from Integrated Solar Energy and IoT System," in *2021 5th International Conference on Informatics and Computational Sciences (ICICoS)*, Nov. 2021, pp. 222–226. doi: 10.1109/ICICoS53627.2021.9651831.

[16] M. Abdoos, H. Rashidi, P. Esmaeili, H. Yousefi, and M. H. Jahangir, "Forecasting solar energy generation in the mediterranean region up to 2030–2050 using convolutional neural networks (CNN)," *Clean. Energy Syst.*, vol. 10, p. 100167, Jun. 2025, doi: 10.1016/j.cles.2024.100167.

[17] J. Zheng *et al.*, "Time series prediction for output of multi-region solar power plants," *Appl. Energy*, vol. 257, p. 114001, Jan. 2020, doi: 10.1016/j.apenergy.2019.114001.

[18] F. D. Campos, T. C. Sousa, and R. S. Barbosa, "Short-Term Forecast of Photovoltaic Solar Energy Production Using LSTM," *Energies*, vol. 17, no. 11, Art. no. 11, Jan. 2024, doi: 10.3390/en17112582.

[19] H. Alharkan, S. Habib, and M. Islam, "Solar Power Prediction Using Dual Stream CNN-LSTM Architecture," *Sensors*, vol. 23, no. 2, Art. no. 2, Jan. 2023, doi: 10.3390/s23020945.

[20] E. M. Al-Ali *et al.*, "Solar Energy Production Forecasting Based on a Hybrid CNN-LSTM-Transformer Model," *Mathematics*, vol. 11, no. 3, Art. no. 3, Jan. 2023, doi: 10.3390/math11030676.

[21] G. Van Houdt, C. Mosquera, and G. Nápoles, “A review on the long short-term memory model,” *Artif. Intell. Rev.*, vol. 53, no. 8, pp. 5929–5955, Dec. 2020, doi: 10.1007/s10462-020-09838-1.

[22] “What Is Attention? | SpringerLink.” Accessed: Dec. 20, 2024. [Online]. Available: https://link.springer.com/chapter/10.1007/978-1-4939-3435-5_2

[23] C. Wang, X. Miao, Y. Xi, M. Bi, L. Li, and W. Hu, “Interpretable and visualized SHAP-based equalizer with feature selection in IMDD system,” in *Optical Fiber Communication Conference (OFC) 2021*, Washington, DC: Optica Publishing Group, 2021, p. Th1A.1. doi: 10.1364/OFC.2021.Th1A.1.