SOLAR RADIATION PREDICTION

Pre-Project Report

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENT FOR THE AWARD OF DEGREE OF

BACHELOR OF TECHNOLOGY

IN

ELECTRICAL ENGINEERING

Under the guidance of
DR. HAILIYA AHSAN
CO-SUPERVISOR: DR. SHOEB HUSSAIN

SUBMITTED BY: DHRUV SHARMA (2020BELE024)
ANURUDRA YADAV (2020BELE077)
KAUSINDER (2020BELE099)

DEPARTMENT OF ELECTRICAL ENGINEERING
NATIONAL INSTITUTE OF TECHNOLOGY
SRINAGAR – 190006 (INDIA)
FEBRUARY,2024

ACKNOWLEDGEMENT

We would like to take this opportunity to express our sincere appreciation and gratitude to our esteemed guide, Dr. Hailiya Ahsan and Dr. Shoeb Hussain, for his invaluable support and guidance throughout our project work. His expertise, knowledge and dedication have been instrumental in shaping the direction of our project and ensuring its successful completion. Dr. Hailiya Ahsan and Dr. Shoeb Hussain have played a pivotal role in providing us with a strong foundation and framework for our project. His guidance has not only improved the quality of our project but has also enriched our overall learning experience. Once again, we extend our heartfelt thanks and appreciation to Dr. Hailiya Ahsan and Dr. Shoeb Hussain for their exceptional guidance and support.

We would further like to extend our heartfelt gratitude towards the Department of Electrical Engineering, National Institute of Technology Srinagar without whom this project work would not have been possible. Thus, we thank our beloved Department for its support.

It is noteworthy to mention that during our project work, we came across various new technologies and learnt them which not only helped us in completing the work but also enriched our learning experience.

We feel indeed that our entire project journey was a good tunneling path which would help us in our future endeavors.

Dhruv Sharma (2020BELE024)

Anurudra Yadav (2020BELE077)

Kausinder (2020BELE099)

National Institute of Technology, Srinagar.

DECLARATION

NATIONAL INSTITUTE OF TECHNOLOGY SRINAGAR
(J&K)
CANDIDATE'S DECLARATION

I hereby certify that the work which is being presented in the project titled "**SOLAR RADIATION PREDICTION** " in partial fulfilment of the requirements for the award of the Degree of Bachelor of Technology and submitted in the Department of ELECTRICAL ENGINEERING, National Institute of Technology Srinagar, is an authentic record of our own work carried out during a period from September 2022 to June 2023 under the supervision of **Dr. Hailiya Ahsan** and **Dr. Shoeb Hussain**, Assistant Professor, Department of Electrical Engineering, National Institute of Technology Srinagar.
The matter presented in this project report has not been submitted by us for the award of any other degree of this or any other Institute/University.

Sd/-
(Dhuruv Sharma, Anurudra, Kausinder)

This is to certify that the above statement made by the candidate is correct to the best of my knowledge.

Sd/-
Dr. Hailiya Ahsan

Sd/-
(Dr. Shoeb Husaain)

The project Viva-Voce Examination of Dhruv, Anurudra, Kausinder, has been held on

Signature of Supervisor(s) **Signature of External Examiner**

Date:

ABSTRACT

Solar radiation prediction is crucial for optimizing the efficiency and effectiveness of solar energy systems. Traditional methods of predicting solar radiation often rely on complex physical models and historical data, which can be time-consuming and less accurate. In recent years, artificial intelligence (AI) techniques, such as machine learning (ML) and deep learning (DL), have shown promising results in predicting solar radiation with higher accuracy and efficiency.

This paper presents a comprehensive review of AI-based approaches for solar radiation prediction. It explores the use of ML and DL algorithms, such as support vector machines (SVM), random forests, and convolutional neural networks (CNN), in predicting solar radiation. The paper discusses the advantages and limitations of these algorithms and compares their performance with traditional methods.

Furthermore, the paper examines the various factors that affect solar radiation prediction, such as weather conditions, geographical location, and time of day. It also discusses the importance of feature selection and data preprocessing in improving the accuracy of solar radiation prediction models.

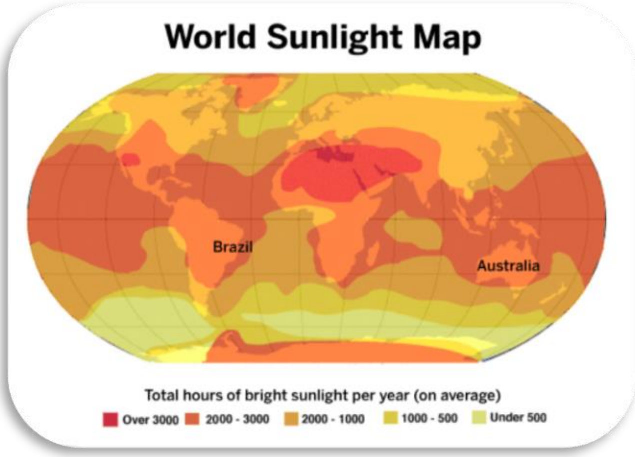
The study evaluates the performance of AI-based models using real-world solar radiation data and compares their accuracy with traditional methods. The results show that AI-based approaches outperform traditional methods in terms of accuracy and computational efficiency. Additionally, the paper discusses the challenges and future directions of AI-based solar radiation prediction, such as the need for more comprehensive and diverse datasets, and the integration of AI models with IoT devices for real-time prediction.

In conclusion, this paper highlights the potential of AI techniques in improving solar radiation prediction and emphasizes the importance of further research in this area to harness the full potential of solar energy.

TABLE OF CONTENTS

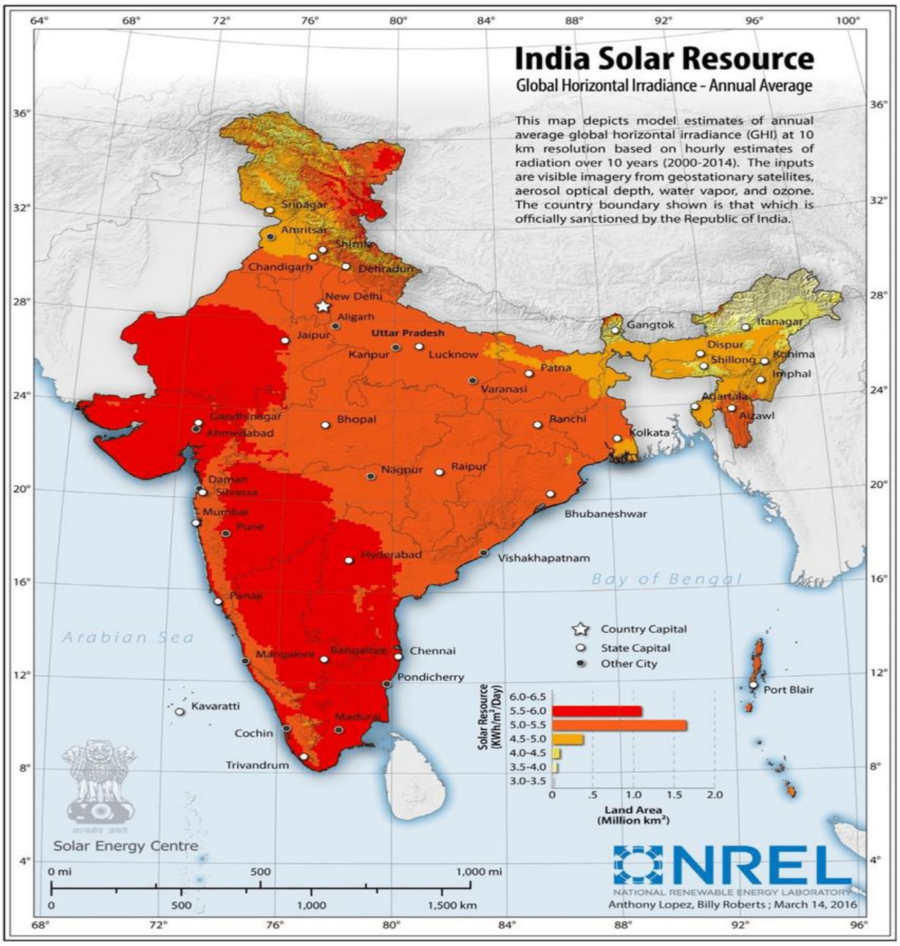
1. INTRODUCTION

The Sun has been worshiped as a life-giver to our planet since ancient times. The industrial ages gave us the understanding of sunlight as an energy source. India is endowed with vast solar energy potential. About 5,000 trillion kWh per year energy is incident over India's land area with most parts receiving 4-7 kWh per sqm per day. Solar photovoltaic power can effectively be harnessed providing huge scalability in India. Solar also provides the ability to generate power on a distributed basis and enables rapid capacity addition with short lead times. Off-grid decentralized and low-temperature applications will be advantageous from a rural application perspective and meeting other energy needs for power, heating and cooling in both rural and urban areas. From an energy security perspective, solar is the most secure of all sources, since it is abundantly available. Theoretically, a small fraction of the total incident solar energy (if captured effectively) can meet the entire country's power requirements.



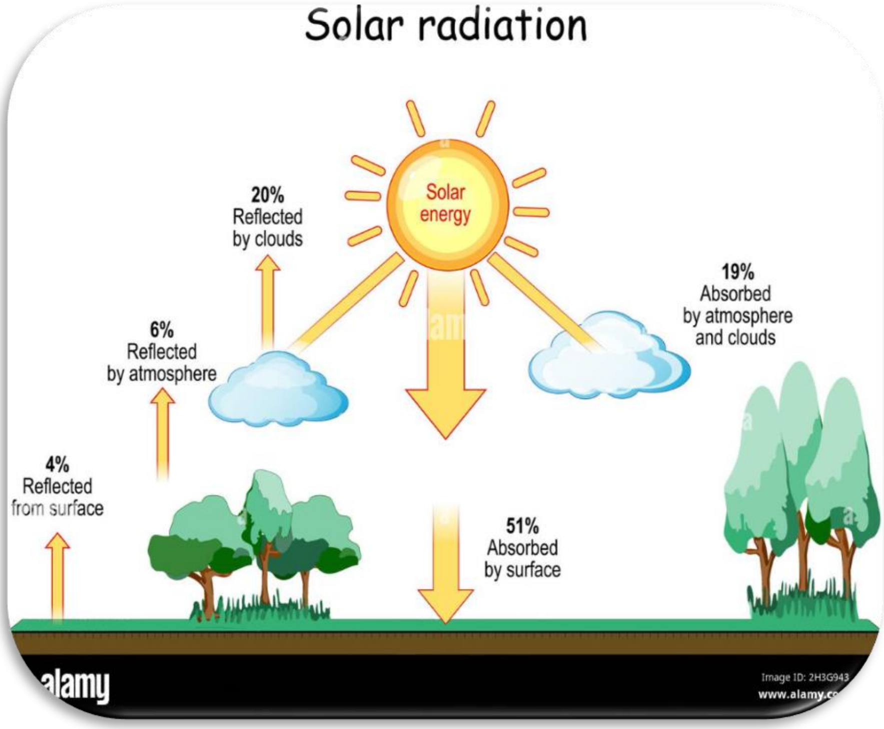
There has been a visible impact of solar energy in the Indian energy scenario during the last few years. Solar energy based decentralized and distributed applications have benefited millions of people in Indian villages by meeting their cooking, lighting and other energy needs in an environment friendly manner.The social and economic benefits include reduction in drudgery among rural women and girls engaged in the collection of fuel wood from long distances and cooking in smoky kitchens, minimization of the risks of contracting lung and eye ailments, employment generation at village level, and ultimately, the improvement in the standard of living and creation of opportunity for economic activities at village level. Further, solar energy sector in India has emerged as a significant player in the grid connected power generation capacity over the years. It supports the government agenda of sustainable growth, while, emerging as an integral part of the solution to meet the nation's energy needs and an essential player for energy security.NationalInstitute of Solar Energy (NISE) has assessed the country's solar potential of about 748 GW assuming 3% of the waste land area to be covered by Solar PV modules. Solar energy has taken a central place in India's National Action Plan on Climate Change with National Solar Mission (NSM) as one of the key Missions. NSM was launched on 11th January, 2010.

NSM is a major initiative of the Government of India with active participation from States to promote ecological sustainable growth while addressing India's energy security challenges. It will also constitute a major contribution by India to the global effort to meet the challenges of climate change. The Mission's objective is to establish India as a global leader in solar energy by creating the policy conditions for solar technology diffusion across the country as quickly as possible. This is line with India's Nationally Determined Contributions (NDCs) target to achieve about 50 percent cumulative electric power installed capacity from non-fossil fuel-based energy resources and to reduce the emission intensity of its GDP by 45 percent from 2005 level by 2030.In order to achieve the above target, Government of India have launched various schemes to encourage generation of solar power in the country like Solar Park Scheme, VGF Schemes, CPSU Scheme, Defence Scheme, Canal bank & Canal top Scheme, Bundling Scheme, Grid Connected Solar Rooftop Scheme etc.



WHAT IS SOLAR RADIATION?

Solar radiation is the energy emitted by the Sun, which is sent in all directions through space as electromagnetic waves. Emitted by the surface of the Sun, this energy influences **atmospheric and climatological processes.** It is also, directly and indirectly, responsible for common phenomena, such as plant photosynthesis, keeping the planet at a temperature compatible with life and wind formation, which is essential for generating wind power.



The **Sun** emits energy in the form of short-wave radiation, which is weakened in the atmosphere by the presence of clouds and absorbed by gas molecules or suspended particles. After passing through the atmosphere, solar radiation reaches the oceanic and continental land surface and is reflected or absorbed. Finally, the surface returns it to outer space in the form of long-wave radiation.

How is solar radiation measured

Solar radiation is measured on a horizontal surface by means of a radiation sensor or pyranometer, which is placed in a south-facing, shadow-free location. Data are collected in units of power, **watts per square metre (W/m^2),** at all weather stations and tend to be taken at ten-minute or 24-hour intervals to establish averages. In the case where it is desired to convert solar radiation from power units to energy units, the data in W/m2 must be multiplied by the number of seconds comprising ten minutes (600) or 24 hours (86,400) and the result will be provided in joules per square metre (J/m^2).

Types of solar radiation

Depending on the form in which it reaches the Earth:

- **Direct solar radiation.** This type of radiation penetrates the atmosphere and reaches the Earth's surface without dispersing at all on the way.

- **Diffuse solar radiation.** This is the radiation that reaches the Earth's surface after having undergone multiple deviations in its trajectory, for example by gases in the atmosphere.

- **Reflected solar radiation.** This is the fraction of solar radiation that is reflected by the earth's surface itself, in a phenomenon known as the *albedo effect*.

Depending on the types of light:

- **Infrared rays (IR).** Longer wavelength than visible light, they emit heat and are given off by anybody whose temperature is greater than 0º Kelvin.
- **Visible rays (VI).** They emit light and are those that the human eye perceives in the form of colours (red, orange, yellow, green, cyan, blue and violet).
- **Ultraviolet (UV) rays.** They are invisible to the human eye and have the most serious impact on the skin (burns, spots, wrinkles). They are divided into three subcategories:
    - Ultraviolet A (UVA). Ultraviolet light that passes through the atmosphere easily, most of it reaching the surface of the planet.
    - Ultraviolet B (UVB). This does not penetrate the atmosphere so easily. Even so, it reaches the surface and is responsible for the worst skin damage.
    - Ultraviolet C (UVC). This type of ultraviolet radiation cannot get through the atmosphere, because it is absorbed by the ozone layer.

The UV index and the effects of the sun on our skin

Humans are exposed to UV radiation, especially UVA and UVB radiation, which can be dangerous to their skin. One of the ways we have to measure the negative consequences of this type of radiation on people is the **global solar UV index (UVI).** This index ranges from one to eleven and the higher the index, the greater the likelihood of skin and eye damage.

Among other consequences, it increases the chances of sunburn, premature ageing and even **skin cancer,** especially in people with a lighter phototype. For this reason, the IUV is an important and differential element in raising public awareness of the risks of excessive exposure, warning of the imminent need to adopt protective measures to minimise the risks.

Staying out of the sun in the middle of the day and, if there is no alternative, staying in the shade and drinking plenty of water.

Wearing protective clothing, like hats, caps or carrying parasols to protect the eyes, face and neck, and light garments.

Wear good quality sunglasses, in other words those with certified lenses and, if possible, with a wraparound design and with side panels.

Use sun protection cream with a sun protection factor higher than 15, although it is advisable to choose according to the skin phototype, half an hour before exposure. Apply generously and repeat as often as necessary.

IMPORTANCE OF SOLAR RADIATION PREDICTION?

Solar resource forecasting is very important for the operation and management of solar power plants. Solar radiation is highly variable because it is driven mainly by synoptic and local weather patterns. This high variability presents challenges to meeting power production and demand curves, notably in the case of photovoltaic (PV) power plants, which have little or no storage capacity. For concentrating solar power (CSP) plants, variability issues are partially mitigated by the thermal inertia of the plant, including its heat transfer fluid, heat exchangers, turbines and, potentially, coupling with a heat storage facility; however, temporally and spatially varying irradiance introduces thermal stress in critical system components and plant management issues that can result in the degradation of the overall system's performance and reduction of the plant's lifetime.

Solar radiation prediction is a crucial component of renewable energy planning, grid management, and climate research. Its importance stems from its role in optimizing energy production, improving grid stability, reducing costs, advancing climate understanding, and aiding in urban planning and agriculture

1**. Optimizing Energy Production:** Solar radiation prediction enables solar power plants to optimize their energy production. By forecasting sunlight availability, these plants can adjust their operations, such as tilting solar panels or managing energy storage, to maximize energy output. This optimization leads to increased efficiency and a more stable energy supply.

2. **Grid Management**: Solar radiation prediction is essential for grid operators to manage the integration of solar energy into the grid. Accurate predictions allow

operators to anticipate fluctuations in solar power generation and balance supply and demand accordingly. This helps prevent overloading the grid during periods of high solar generation and ensures a reliable energy supply.

3**. Cost Reduction:** Accurate solar radiation prediction can lead to cost reductions in several ways. By optimizing energy production, solar power plants can reduce their operating costs and improve their profitability. Additionally, grid operators can avoid costly grid upgrades by efficiently managing solar energy integration, leading to overall cost savings for energy consumers.

4. **Climate Research:** Solar radiation data is crucial for climate research and modeling. It provides insights into the Earth's energy balance, which is essential for understanding climate patterns and trends. Solar radiation prediction helps researchers study the impact of solar variability on the climate and improve climate change predictions.

5**. Urban Planning and Agriculture:** Solar radiation prediction has practical applications in urban planning and agriculture. In urban areas, accurate predictions can help architects and city planners design buildings and infrastructure that maximize natural lighting and energy efficiency. In agriculture, solar radiation data can assist farmers in planning crop planting and harvesting times, leading to improved yields and resource management.

Overall, solar radiation prediction is a critical tool for maximizing the benefits of solar energy while minimizing its impact on the environment. By improving our ability to forecast solar radiation, we can increase the reliability and efficiency of solar energy systems, contributing to a more sustainable future.

CURRENT METHOD AND CHALLENGES?

Current Methods:

1**. Satellite Data:** Satellites provide a wide range of data for solar radiation prediction, including direct measurements of solar radiation and indirect measurements such as cloud cover and atmospheric conditions. These data are used in models to estimate solar radiation levels at different locations on Earth. While satellite data offer broad coverage, they can be limited by factors such as cloud cover, which can obscure the view of the sun and affect the accuracy of the predictions.

2**. Ground-Based Sensors:** Ground-based sensors directly measure solar radiation at specific locations. These sensors provide real-time data, allowing for more immediate adjustments to solar energy production and grid management. However, their coverage is limited to the areas where the sensors are installed, making it challenging to obtain comprehensive data for large regions.

3. **Numerical Models:** Numerical models simulate the interactions of sunlight with the Earth's atmosphere and surface to predict solar radiation levels. These models incorporate data from satellite observations, ground-based measurements, and atmospheric models to estimate solar radiation under different conditions. While numerical models can provide valuable insights into solar radiation patterns, they can be computationally intensive and require accurate input data to produce reliable predictions.

Challenges:

1. **Cloud Cover:** Cloud cover is a major challenge in solar radiation prediction, as clouds can block or scatter sunlight, significantly affecting the amount of solar radiation that reaches the Earth's surface. Predicting cloud cover and its impact on solar radiation is complex due to the variability of cloud formations and movements.

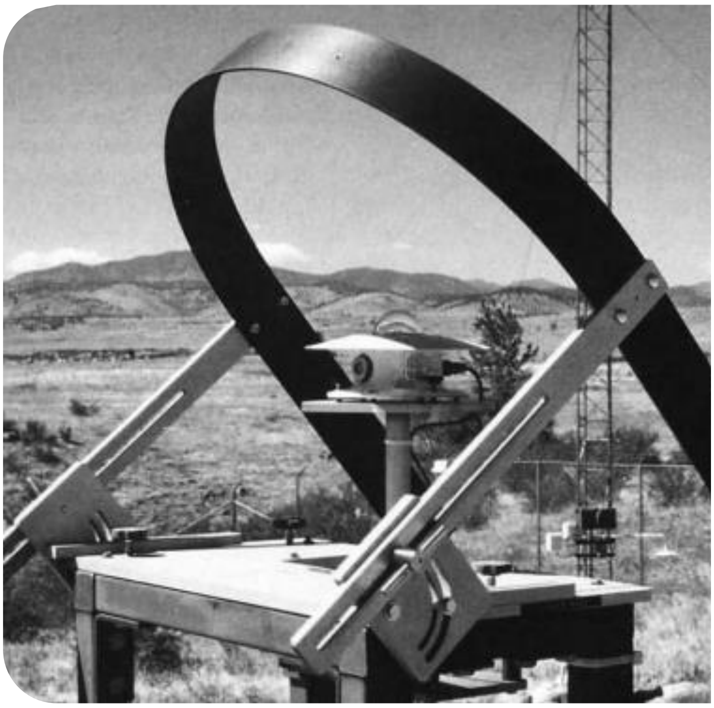
2. **Atmospheric Conditions:** Changes in atmospheric conditions, such as aerosol concentrations and water vapor content, can alter the path of sunlight and affect solar radiation levels. These changes are challenging to predict accurately, particularly in regions with complex terrain or weather patterns.

3. **Data Accuracy:** Ensuring the accuracy of input data is crucial for reliable solar radiation prediction. Errors or inaccuracies in satellite data, ground-based measurements, or atmospheric models can lead to inaccuracies in the predictions. Calibration and validation of data sources are essential to improving prediction accuracy.

4. **Resolution:** Achieving high spatial and temporal resolution in solar radiation prediction is important for applications such as solar energy planning and grid management. However, obtaining high-resolution data can be challenging, particularly in remote or inaccessible areas.

5. **Model Complexity:** Modeling sunlight interactions with the Earth's atmosphere and surface requires complex algorithms and computational resources. Improving the accuracy and efficiency of these models is an ongoing challenge in solar radiation prediction.

Addressing these challenges requires ongoing research and development in areas such as remote sensing, atmospheric modeling, and data assimilation. Advances in technology and data availability are essential for improving the accuracy and reliability of solar radiation prediction models.



OUR APPROACH

Our approach to solar radiation prediction involves a detailed methodology that integrates advanced machine learning algorithms with high-resolution satellite data. Here's a comprehensive overview of our approach

1**. Data Collection: -** We collect a diverse range of data sources, including: Historical solar radiation data from ground-based stations or satellite observations. High-resolution satellite imagery to capture cloud cover, atmospheric conditions, and surface reflectance. Weather data from ground-based stations, including temperature, humidity, wind speed, and precipitation. Atmospheric conditions data, such as air pressure, moisture content, and aerosol levels. This comprehensive dataset provides the foundation for our solar radiation prediction model.

2**. Data Preprocessing:** The collected data undergoes preprocessing to ensure its quality and suitability for analysis. This includes cleaning the data to remove noise, correcting errors, and standardizing formats. We also perform data normalization and transformation to prepare it for input into our machine learning models.

3**. Feature Selection:** We use feature selection techniques to identify the most relevant features for predicting solar radiation levels. These features may include variables such as time of day, day of year, solar zenith angle, cloud cover, and atmospheric conditions. Feature selection helps reduce the dimensionality of the data and improve the efficiency of our models.

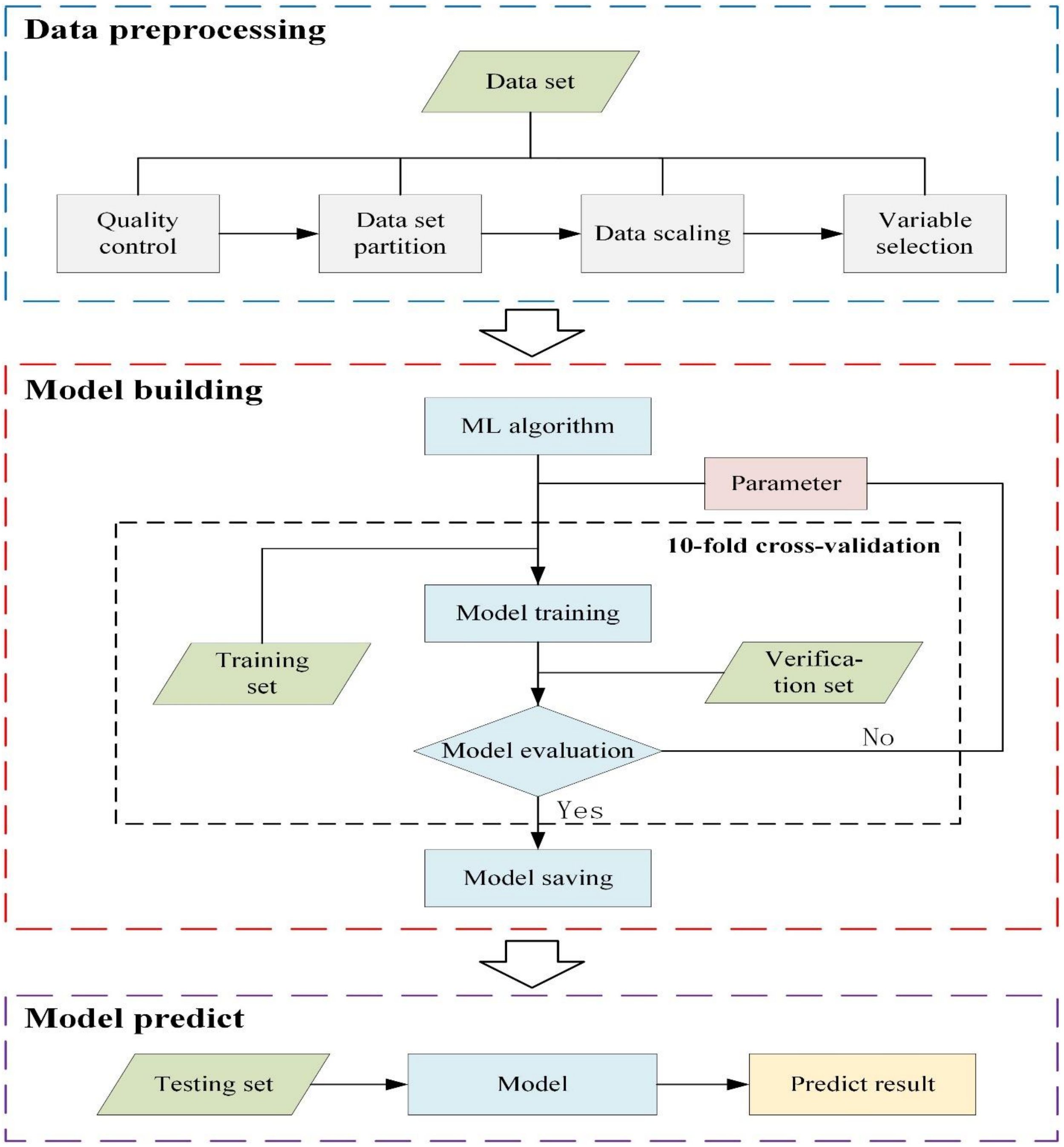
4. **Model Selection:** We select machine learning algorithms based on their suitability for solar radiation prediction. Commonly used algorithms include regression models (e.g., linear regression, support vector regression), decision trees (e.g., random forests), and neural networks. We also consider ensemble methods, such as bagging and boosting, to combine multiple models for improved accuracy.

5**. Model Training:** The selected models are trained using the pre-processed data and selected features. We use techniques like cross-validation to split the data into training and validation sets, optimizing the model hyperparameters to prevent overfitting. Training the model involves iteratively adjusting the model parameters to minimize the prediction error.

6. **Model Evaluation:** We evaluate the trained models using validation data to assess their performance. Common evaluation metrics include mean absolute error (MAE), root mean square error (RMSE), and coefficient of determination (R-squared). We compare the performance of our models with baseline models or existing methods to validate their effectiveness.

7. **Prediction Generation:** Once trained and validated, the models are used to generate solar radiation predictions. These predictions are based on current weather conditions, satellite imagery, and other relevant factors, providing real-time or near-real-time estimates of solar radiation levels.

8**. Model Deployment:** The trained models are deployed for real-world use, such as in solar energy planning, grid management, or climate research. We continuously monitor the performance of the models and update them as needed to ensure their accuracy and reliability.



P
R

OUR OBJECTIVES

1**. Develop a Robust Prediction Model:** Utilize advanced machine learning algorithms, such as deep learning and ensemble methods, to develop a prediction model. Incorporate diverse datasets, including historical solar radiation data, satellite imagery, weather data, and atmospheric conditions, to improve prediction accuracy. Implement feature selection techniques to identify the most relevant features for predicting solar radiation levels.

2. **Improve Renewable Energy Planning:** Provide accurate and reliable solar radiation predictions to support the planning and operation of solar power plants. Optimize energy production forecasts to maximize the efficiency of solar power generation. Enhance grid integration of solar energy by providing real-time or near-real-time solar radiation predictions to grid operators.

3**. Enhance Grid Management:** Develop tools and models to help grid operators manage the variability of solar energy generation. Provide insights and recommendations for balancing supply and demand, maintaining grid stability, and optimizing energy distribution.

4**. Advance Climate Research:** Contribute to climate research by providing high-quality solar radiation data for studying climate patterns and trends. Improve understanding of solar variability and its impact on the Earth's climate system.

5. **Support Sustainable Development:** Promote the use of renewable energy sources, such as solar energy, to reduce reliance on fossil fuels and mitigate climate change. Support sustainable development goals by providing reliable and actionable information for energy planning and decision-making.

6. **Enable Data-Driven Decision Making:** Provide stakeholders with reliable and up-to-date information based on data-driven predictions. Enable informed decision-making regarding energy planning, grid management, and climate adaptation.

7**. Facilitate Integration of Renewable Energy:** Support the integration of renewable energy sources into existing energy systems by providing accurate solar radiation predictions. Optimize the use of solar energy and reduce dependence on non-renewable energy sources.

By achieving these objectives, our project aims to contribute to the advancement of renewable energy technologies, improve the efficiency and reliability of solar energy systems, and support sustainable development practices.

DATA COLLECTION

Our approach to data collection for solar radiation prediction involves gathering diverse datasets from multiple sources. Here's a detailed outline of our data collection process:

1. **Historical Solar Radiation Data:** Collect historical solar radiation data from ground-based stations or satellite observations. Include data on solar radiation levels at regular intervals (e.g., hourly or daily averages) over an extended period (e.g., several years). Ensure data quality by performing quality control checks and correcting any anomalies or errors.

2**. Satellite Imagery:** Obtain high-resolution satellite imagery from sources such as NASA or commercial satellite providers. Use satellite imagery to capture cloud cover, atmospheric conditions, and surface reflectance, which are critical factors influencing solar radiation levels. Ensure timely access to satellite data to provide up-to-date information for solar radiation prediction.

3. **Weather Data:** Gather weather data from ground-based stations, including temperature, humidity, wind speed, and precipitation. Use weather data to account for local weather conditions that can impact solar radiation levels, such as cloud cover and atmospheric stability. Ensure the accuracy and reliability of weather data by calibrating and validating against other sources.

4**. Atmospheric Conditions Data:** Collect data on atmospheric conditions, such as air pressure, moisture content, and aerosol levels. Use atmospheric data to account for the effects of atmospheric composition and conditions on solar radiation levels. Ensure the availability of atmospheric data for the target region and time period of interest.

5. **Data Preprocessing:** Preprocess the collected data to remove outliers, correct errors, and standardize formats. Perform data normalization and transformation to prepare the data for input into machine learning models. Ensure that the pre-processed data is clean and suitable for analysis.

6. **Data Integration:** Integrate the various datasets (e.g., solar radiation data, satellite imagery, weather data) to create a comprehensive dataset for solar radiation prediction. Use data integration techniques to combine data from different sources and formats into a unified dataset.By following this detailed data collection process, we aim to gather the necessary data for training our machine learning models and improving the accuracy of solar radiation prediction.

TECHNICAL APPROACH

1. Reading the dataset

```
import numpy as np
import pandas as pd
import seaborn as sns
from datetime import date
import matplotlib.pyplot as plt
from collections import defaultdict
```

```
df = pd.read_csv('SolarPrediction.csv')
display(df)
```

	UNIXTime	Date	Time	Radiation	Temperature	Pressure	Humidity	WindDirection	Speed	TimeSunRise	TimeSunSet
0	1475229326	9/29/2016 12:00:00 AM	23:55:26	1.21	48	30.46	59	177.39	5.62	06:13:00	18:13:00
1	1475229023	9/29/2016 12:00:00 AM	23:50:23	1.21	48	30.46	58	176.78	3.37	06:13:00	18:13:00
2	1475228726	9/29/2016 12:00:00 AM	23:45:26	1.23	48	30.46	57	158.75	3.37	06:13:00	18:13:00
3	1475228421	9/29/2016 12:00:00 AM	23:40:21	1.21	48	30.46	60	137.71	3.37	06:13:00	18:13:00
4	1475228124	9/29/2016 12:00:00 AM	23:35:24	1.17	48	30.46	62	104.95	5.62	06:13:00	18:13:00
...	...	...	...	...	...	...	...	...	...	...	...
32681	1480587604	12/1/2016 12:00:00 AM	00:20:04	1.22	44	30.43	102	145.42	6.75	06:41:00	17:42:00
32682	1480587301	12/1/2016 12:00:00 AM	00:15:01	1.17	44	30.42	102	117.78	6.75	06:41:00	17:42:00
32683	1480587001	12/1/2016 12:00:00 AM	00:10:01	1.20	44	30.42	102	145.19	9.00	06:41:00	17:42:00
32684	1480586702	12/1/2016 12:00:00 AM	00:05:02	1.23	44	30.42	101	164.19	7.87	06:41:00	17:42:00
32685	1480586402	12/1/2016 12:00:00 AM	00:00:02	1.20	44	30.43	101	83.59	3.37	06:41:00	17:42:00

32686 rows × 11 columns

2. Data Cleaning

```
def func_date(data):

    data = data.split()
    data = data[0]
    data = data.split('/')

    day = int(data[1])
    month = int(data[0])
    year = int(data[2])

    date1 = date(year-1, 12, 31)
    date2 = date(year, month, day)
    diff = date2 - date1
    diff = str(diff)
    diff = diff.split(' ')

    return int(diff[0])
```

```python

def func_time(data):
    data = data.split(':')
    time = int(data[0])*3600 + int(data[1])*60 + int(data[2])
    return time
df['Date'] = df['Date'].apply(func_date)
df['Time'] = df['Time'].apply(func_time)
df['TimeSunRise'] = df['TimeSunRise'].apply(func_time)
df['TimeSunSet'] = df['TimeSunSet'].apply(func_time)
display(df)
```

	Date	Time	Radiation	Temperature	Pressure	Humidity	WindDirection	Speed	TimeSunRise	TimeSunSet
0	273	86126	1.21	48	30.46	59	177.39	5.62	22380	65580
1	273	85823	1.21	48	30.46	58	176.78	3.37	22380	65580
2	273	85526	1.23	48	30.46	57	158.75	3.37	22380	65580
3	273	85221	1.21	48	30.46	60	137.71	3.37	22380	65580
4	273	84924	1.17	48	30.46	62	104.95	5.62	22380	65580
...	...	...	...	...	...	...	...	...	...	...
32681	336	1204	1.22	44	30.43	102	145.42	6.75	24060	63720
32682	336	901	1.17	44	30.42	102	117.78	6.75	24060	63720
32683	336	601	1.20	44	30.42	102	145.19	9.00	24060	63720
32684	336	302	1.23	44	30.42	101	164.19	7.87	24060	63720
32685	336	2	1.20	44	30.43	101	83.59	3.37	24060	63720

32686 rows × 10 columns

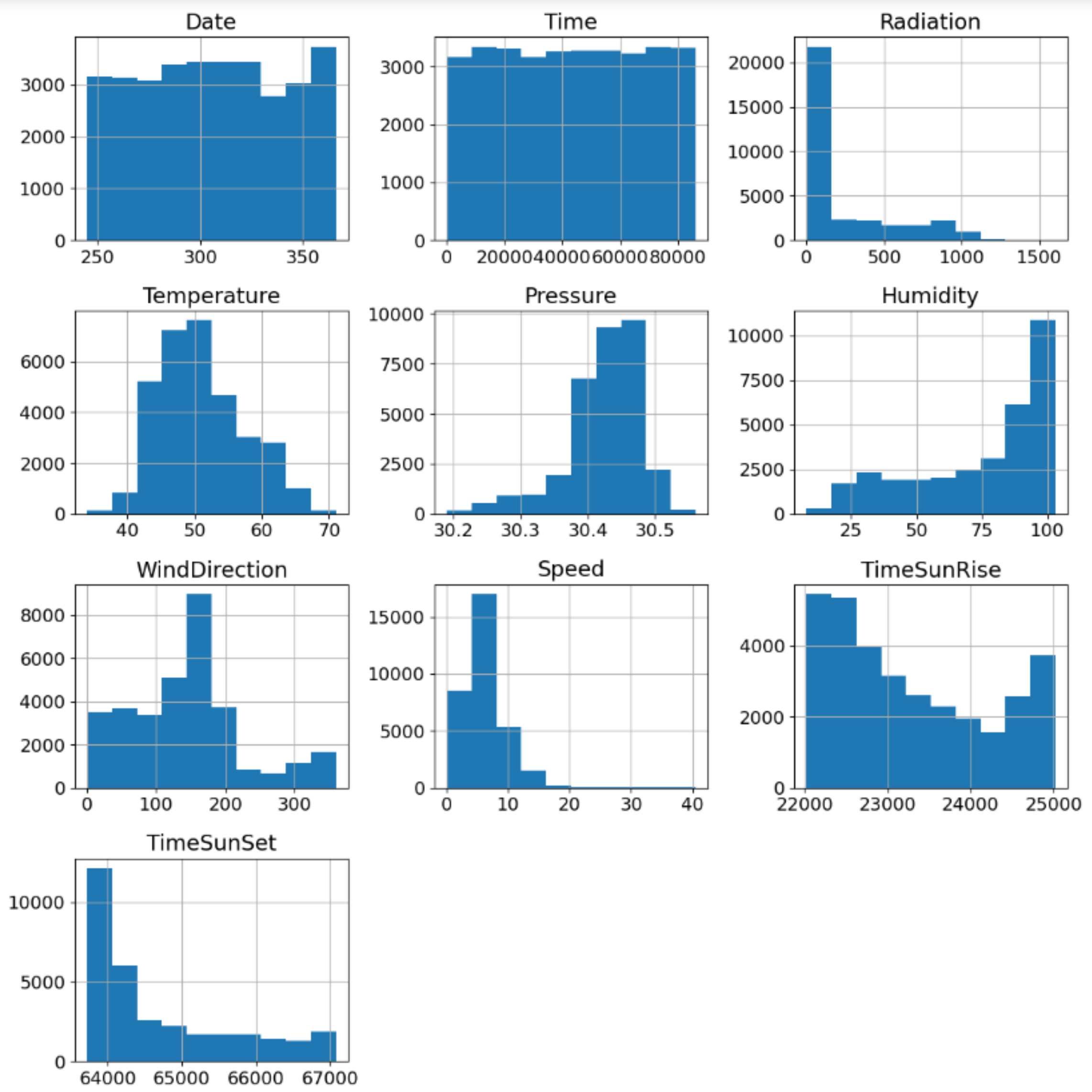
3. Summarizing Dataset

```python
df.describe()
```

9]:

	Date	Time	Radiation	Temperature	Pressure	Humidity	WindDirection	Speed	TimeSunRise	TimeSunSet
count	32686.000000	32686.000000	32686.000000	32686.000000	32686.000000	32686.000000	32686.000000	32686.000000	32686.000000	32686.000000
mean	306.110965	43277.574068	207.124697	51.103255	30.422879	75.016307	143.489821	6.243869	23258.431133	64691.463624
std	34.781367	24900.749819	315.916387	6.201157	0.054673	25.990219	83.167500	3.490474	931.122823	995.053346
min	245.000000	1.000000	1.110000	34.000000	30.190000	8.000000	0.090000	0.000000	22020.000000	63720.000000
25%	277.000000	21617.000000	1.230000	46.000000	30.400000	56.000000	82.227500	3.370000	22440.000000	63900.000000
50%	306.000000	43230.000000	2.660000	50.000000	30.430000	85.000000	147.700000	5.620000	23040.000000	64260.000000
75%	334.000000	64849.000000	354.235000	55.000000	30.460000	97.000000	179.310000	7.870000	24000.000000	65340.000000
max	366.000000	86185.000000	1601.260000	71.000000	30.560000	103.000000	359.950000	40.500000	25020.000000	67080.000000

```python
plt.rcParams["font.size"] = 14
df.hist(figsize=(12,12))
plt.tight_layout()
plt.show()
```


Date
Time
Radiation
Temperature
Pressure
Humidity
WindDirection
Speed
TimeSunRise
TimeSunSet

4. Feature selection

Correlation Matrix

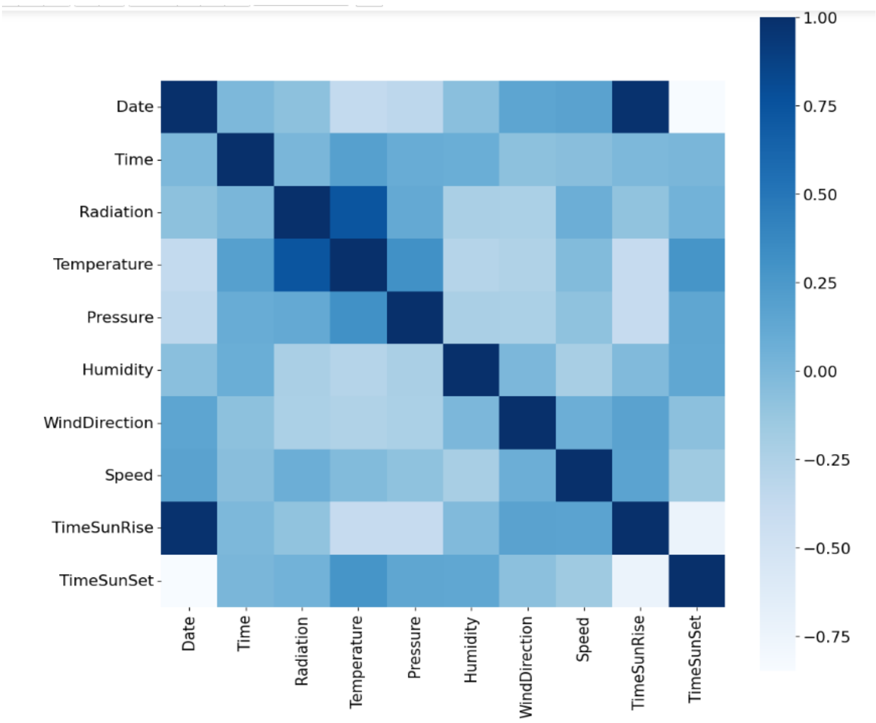
```
display(df.corr(method='pearson'))
```

	Date	Time	Radiation	Temperature	Pressure	Humidity	WindDirection	Speed	TimeSunRise	TimeSunSet
Date	1.000000	-0.007094	-0.081320	-0.370794	-0.332762	-0.063760	0.153255	0.174336	0.981939	-0.847401
Time	-0.007094	1.000000	0.004348	0.197227	0.091066	0.077851	-0.077956	-0.057908	-0.006639	0.008038
Radiation	-0.081320	0.004348	1.000000	0.734955	0.119016	-0.226171	-0.230324	0.073627	-0.092850	0.045688
Temperature	-0.370794	0.197227	0.734955	1.000000	0.311173	-0.285055	-0.259421	-0.031458	-0.380968	0.285131
Pressure	-0.332762	0.091066	0.119016	0.311173	1.000000	-0.223973	-0.229010	-0.083639	-0.380399	0.146884
Humidity	-0.063760	0.077851	-0.226171	-0.285055	-0.223973	1.000000	-0.001833	-0.211624	-0.023955	0.135243
WindDirection	0.153255	-0.077956	-0.230324	-0.259421	-0.229010	-0.001833	1.000000	0.073092	0.176929	-0.068040
Speed	0.174336	-0.057908	0.073627	-0.031458	-0.083639	-0.211624	0.073092	1.000000	0.167075	-0.159400
TimeSunRise	0.981939	-0.006639	-0.092850	-0.380968	-0.380399	-0.023955	0.176929	0.167075	1.000000	-0.738271
TimeSunSet	-0.847401	0.008038	0.045688	0.285131	0.146884	0.135243	-0.068040	-0.159400	-0.738271	1.000000

Correlation Heatmap

```
plt.rcParams["figure.figsize"] = (12,12)
plt.rcParams["font.size"] = 15

sns.heatmap(df.corr(), square=True, cmap='Blues')
plt.show()
```



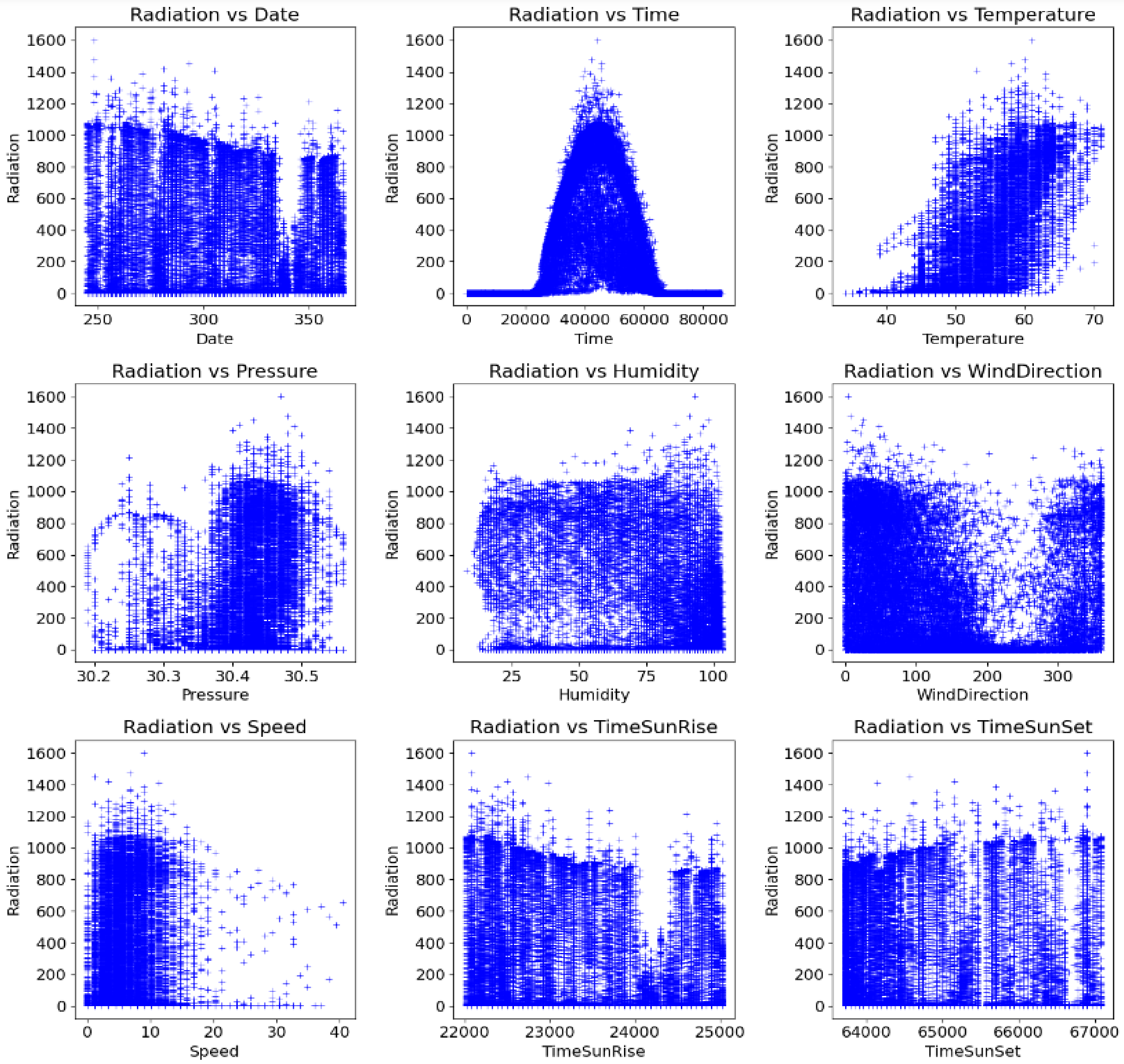
Scatter Plot

```
plt.rcParams["figure.figsize"] = (14, 14)
plt.rcParams["font.size"] = 14

ylabel = 'Radiation'
columns = ['Date', 'Time', 'Temperature', 'Pressure', 'Humidity',
'WindDirection',  'Speed',  'TimeSunRise',  'TimeSunSet']

for index, xlabel in enumerate(columns):
  plt.subplot(3, 3, index+1)
  plt.scatter(df[xlabel], df[ylabel],color='blue',marker='+',linewidth=0.5)
  plt.xlabel(xlabel)
  plt.ylabel(ylabel)
  plt.title(ylabel + ' vs ' + xlabel)

plt.tight_layout()
```



5. Training and Testing the Models

```
!pip install xgboost==1.5.0
from sklearn.svm import SVR
from xgboost import XGBRegressor
from sklearn.tree import DecisionTreeRegressor
from sklearn.neural_network import MLPRegressor
from sklearn.linear_model import LinearRegression
from sklearn.ensemble import RandomForestRegressor
from sklearn.ensemble import GradientBoostingRegressor

from sklearn.model_selection import train_test_split
```

Train-Test split

```
Y = df['Radiation'].values
df.drop(['Radiation'], axis=1, inplace=True)

X = df.values

RS = 1811
X_train, X_test, Y_train, Y_test = train_test_split(X,Y,test_size=0.2,
random_state=RS)
```

Predictions using Linear Regression Model

```
lr = LinearRegression().fit(X_train, Y_train)
print('train', lr.score(X_train, Y_train))
print('test', lr.score(X_test, Y_test))
```

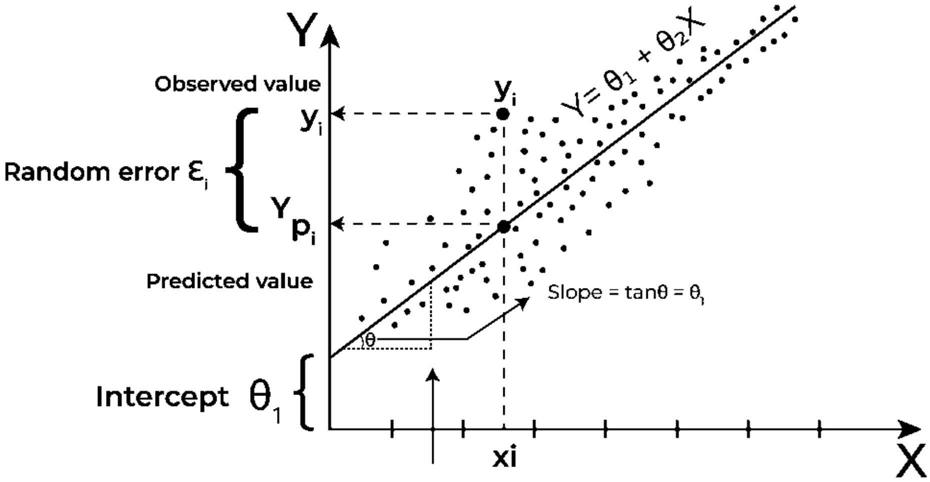
Output

```
train 0.619041454379224

test 0.6259189940799947
```

Mathematical approach Behind Linear Regression

Linear Regression fits a linear model with coefficients w = (w1, …, wp) to minimize the residual sum of squares between the observed targets in the dataset, and the targets predicted by the linear approximation.



R^2 is a number that explains the amount of variation/accuracy that is explained by the developed model.

It always ranges between 0 & 1.

Formula

$$R^2 = 1 - \frac{RSS}{TSS}$$

R^2 = coefficient of determination
RSS = sum of squares of residuals
TSS = total sum of squares

$$\text{RSS} = \Sigma\left(y_i - \widehat{y_i}\right)^2$$

Where: y_i is the actual value and, $\widehat{y_i}$ is the predicted value.

$$\text{TSS} = \Sigma\left(y_i - \overline{y}\right)^2$$

Where: y_i is the actual value and $\overline{y}$ is the mean value of the variable/feature

APPLICATIONS OF SOLAR RADIATION PREDICTION?

Solar radiation prediction has numerous applications across various industries and fields. Here are some key applications:

1**. Renewable Energy Planning:** Solar radiation prediction is crucial for planning and optimizing the operation of solar power plants. By forecasting solar radiation levels, energy producers can schedule the generation of solar energy more efficiently, thereby maximizing energy output and reducing costs.

2. **Grid Integration:** Solar radiation prediction helps grid operators manage the integration of solar energy into the grid. By accurately predicting solar radiation levels, grid operators can anticipate fluctuations in solar power generation and adjust grid operations accordingly to maintain grid stability.

3. **Energy Storage Management:** Solar radiation prediction is essential for managing energy storage systems, such as batteries. By forecasting solar radiation levels, energy storage systems can be charged or discharged at optimal times, maximizing the use of stored energy.

4**. Climate Research:** Solar radiation prediction plays a crucial role in climate research. By providing data on solar radiation levels, researchers can better understand the Earth's energy balance, climate patterns, and the impact of solar variability on the climate.

5. **Agriculture:** Solar radiation prediction is valuable for agricultural planning. By forecasting solar radiation levels, farmers can optimize crop planting and harvesting times, leading to improved yields and resource management.

6. **Urban Planning:** Solar radiation prediction can inform urban planning efforts, particularly in designing energy-efficient buildings and infrastructure. By considering solar radiation levels, planners can optimize building orientation and design to maximize natural lighting and reduce energy consumption.

7. **Disaster Management:** Solar radiation prediction can be useful in disaster management scenarios, such as predicting solar radiation levels in areas affected by natural disasters. This information can help emergency responders plan relief efforts and mitigate the impact of disasters.

Overall, solar radiation prediction has a wide range of applications that contribute to energy efficiency, sustainability, and climate resilience.

FUTURE WORK OF SOLAR RADIATION PREDICTION

The future of solar radiation prediction is promising, with ongoing advancements in technology and research contributing to improved accuracy and reliability. Here are some key trends and developments shaping the future of solar radiation prediction:

1. **Advanced Machine Learning Techniques:** The use of advanced machine learning techniques, such as deep learning and ensemble methods, is expected to improve the accuracy of solar radiation prediction models. These techniques can better capture complex patterns in solar radiation data and enhance prediction capabilities.

2. **Integration of Satellite Data:** Continued advancements in satellite technology are enabling the collection of high-resolution data on solar radiation and atmospheric conditions. Integrating this data into prediction models can improve the spatial and temporal resolution of predictions.

3. **Big Data Analytics:** The increasing availability of big data analytics tools and platforms allows for the analysis of large and diverse datasets. This enables more comprehensive and accurate solar radiation predictions by incorporating a wide range of relevant data sources.

4. **Integration with Renewable Energy Systems:** Solar radiation prediction models are increasingly being integrated into renewable energy systems, such as solar power plants and grid management systems. This integration enables real-time optimization of energy production and grid operations based on predicted solar radiation levels.

5. **Climate Change Adaptation:** Solar radiation prediction is becoming increasingly important for climate change adaptation efforts. By providing insights into future solar radiation patterns, these predictions can help policymakers and planners develop strategies to mitigate the impacts of climate change.

6. **Cross-Disciplinary Collaboration:** Collaboration between researchers from different disciplines, such as meteorology, climatology, and data science, is expected to drive innovation in solar radiation prediction. This collaboration can lead to the development of more holistic and accurate prediction models.

7. **Increased Accessibility:** As solar energy continues to play a larger role in the global energy mix, there is a growing demand for accessible and reliable solar radiation prediction tools. Efforts to make these tools more user-friendly and accessible to a wider audience are expected to continue.

Overall, the future of solar radiation prediction is characterized by advancements in technology, increased integration with renewable energy systems, and a focus on addressing the challenges of climate change. These developments are expected to improve the accuracy and reliability of solar radiation predictions, ultimately supporting the growth of solar energy and sustainable development.

CONCLUSION

In detailed conclusion, solar radiation prediction stands as a crucial pillar supporting the transition towards sustainable energy practices and climate resilience. Through the accurate estimation of solar radiation levels, this field enables optimized energy production, grid stability, cost reduction, climate research, and sustainable urban planning.

The future of solar radiation prediction is marked by advancements in technology, particularly in machine learning and satellite imagery, which promise to enhance prediction accuracy and spatial-temporal resolution. Moreover, the integration of prediction models with renewable energy systems and climate change adaptation strategies will further solidify the importance of this field in addressing global challenges.

Collaboration across disciplines and sectors will continue to drive innovation in solar radiation prediction, ensuring that it remains at the forefront of renewable energy research and application. By leveraging these advancements, solar radiation prediction will play an increasingly pivotal role in shaping a sustainable and climate-resilient future for generations to come.

REFERENCES AND BIBLIOGRAPHY

1. https://www.kaggle.com/datasets/anikannal/solar-power-generation-data

2. https://en.wikipedia.org/wiki/Solar_irradiance

3 https://google.com

4. Yang, H., Yan, J., Li, H., & Zhang, H. (2008). Review on forecasting of solar radiation. Renewable and Sustainable Energy Reviews, 12(1), 112-141.

5. Zhang, J., Wu, Y., & Hong, T. (2016). A review of building energy modelling for the fast urbanizing and industrializing countries. Building and Environment, 97, 196-210.