PREDICTION OF GLOBAL SOLAR RADIATION USING MACHINE LEARNING

A PROJECT REPORT ON NAAN MUDHALVAN

Submitted by

Jaithra K – 2020509022

Kamesh A – 2020509023

E J Manoj – 2020509026

Piyush M – 2020509033

in partial fulfillment for the award of the degree

of

BACHELOR OF ENGINEERING

IN

MECHANICAL ENGINEERING

MADRAS INSTITUTE OF TECHNOLOGY

ANNA UNIVERSITY: CHENNAI 600 025

NOVEMBER 2022

ANNA UNIVERSITY: CHENNAI 600 025

BONAFIDE CERTIFICATE

Certified that this project report "**Prediction of Global Solar Radiation Using Machine Learning**" is the bonafide work of "**Jaithra K (2020509022), Kamesh A (2020509024), E J Manoj (2020509026) and Piyush M (2020509033)** " who carried out the Naan Mudhalvan project work under my supervision.

SIGNATURE

Dr. A. Siddharthan
Professor and Head
Department of Production Technology
MIT Campus, Anna University
Chennai 600 044

SIGNATURE

Dr. P. Ganesh
Machine Learning In-Charge
Assistant Professor (Sl.Gr.)
Department of Production Technology
MIT Campus, Anna University
Chennai 600 044

SIGNATURE

Mr. S. Mohamed Shazuli
Machine Learning In-Charge
Teaching Fellow
Department of Production Technology
MIT Campus, Anna University

ACKNOWLEDGEMENT

We are very grateful to **Dr. A. SIDDARTHAN**, Professor and Head of the Department for his constant encouragement and suggestions and for his valuable and in-depth experienced advice which helped us in completion of this project successfully. We are highly thankful to **Dr. J. PRAKASH**, Dean, Madras Institute of Technology, Chennai for providing the necessary facilities to carry out the project on Naan Mudhalvan.

It is a pleasure to place on record our profound sense of gratitude to our Faculty Advisor **Dr. P. GANESH**, Assistant Professor, Department of Production Technology, Madras Institute of Technology, Chennai for his valuable guidance, advice, motivation and encouragement which helped us immensely to accomplish this project successfully.

We wish to express our sincere thanks and gratitude to the Machine Learning In-charges **Dr. P. GANESH**, Assistant Professor, Department of Production Technology, Madras Institute of Technology, Chennai and **Mr. S. Mohamed Shazuli**, Teaching Fellow, Department of Production Technology, Madras Institute of Technology, Chennai for their valuable suggestion and constant encouragement.

Jaithra K

Kamesh A

E J Manoj

Piyush M

ABSTRACT

The ever-growing interest in and requirement for green energy have led to an increased focus on research related to forecasting solar irradiance recently. This study aims to develop forecast models based on machine learning (ML) methodologies. Solar radiation prediction is necessary as solar energy is essential not only in electricity generation but in other applications such as solar distillation, water heating, metrology and producing solar conversion energy.

TABLE OF CONTENTS

CHAPTER NO.	TITLE	PAGE NO.
	ABSTRACT	**iv**
	LIST OF TABLES	**vi**
	LIST OF FIGURES	**vi**
1.	**INTRODUCTION**	1
2.	**PROBLEM DEFINITION**	**3**
	2.1 NEED FOR STUDY	3
	2.2 OBJECTIVE OF PROJECT	4
3.	**METHODOLOGY**	**5**
4.	**DATA COLLECTION**	**7**
5.	**GOOGLE COLAB CODE**	**20**
6.	**RESULTS AND DISCUSSION**	**30**
7.	**CONCLUSION**	**33**
	REFERENCES	**34**

LIST OF TABLES

TABLE NO.	TITLE	PAGE NO.
4.1	Data Collection	7

LIST OF FIGURES

FIGURE NO.	TITLE	PAGE NO.
3.1	Methodology	5
6.1	Hourly Solar Radiation	30
6.2	Mean Radiation by Month	31
6.3	Humidity by Hour	31
6.4	Validation curve for polynomial regression	32

CHAPTER 1
INTRODUCTION

Energy is invisible, it is a driving force in the entire universe. It is a fundamental input to any kind of human activity. There are many types of energy exist; these include light energy, sound energy, solar energy, chemical energy, and nuclear energy. In the 19^{th} century, the use of fossil fuels had enhanced rapid industrialization.

However, the excessive exploitation of the fossil fuels has directly and indirectly assisted in global warming and other problems which drive our planet towards a dark future. Today, technological advancements enabled the development of an alternative and renewable energy sources, one of which is the solar energy.

Solar energy is free, it does not create pollution nor destroy our planet, and helped man to become less dependent on the other more costly and damaging forms of energy. Solar energy is one the most abundant renewable energy resource available to mankind and it is available in most parts of the world, and has become increasingly attractive.

Generally, people understand real-world phenomena in a better when they are represented symbolically. As such, modelling should not only be used to illustrate and deepen understanding, but it should also help predict real world phenomena. For an instance, the design and development of various solar energy systems and the estimation of solar radiation available is considered the most crucial issue. The importance of models for solar radiation prediction is further emphasized because, in most cases, the density and number of solar radiations measuring stations cannot describe the necessary variability.

Most of the available models for predicting the global solar radiation mostly use all the following factors as input: sunshine data, solar declination, latitude, longitude, extra-terrestrial radiation, relative humidity, soil temperature, temperature of the air, cloudiness, evaporation, and precipitation among others. These are the factors that require special skills and equipment's to measure. In addition to that models with multiple inputs have been shown to have insignificantly improved in accuracy compared with other variants of the models in which fewer inputs are considered.

Therefore, one of the main objectives of this paper is to develop models, with few input factors that can be used to predict global solar radiation. The new models that are introduced in this study require a little computational effort, and the only information needed to use them is the knowledge of longitude and latitude.

CHAPTER 2
PROBLEM DEFINITION

Variable nature of solar energy poses challenges in its integration to the power grid. Accurate forecasting is required for a techno-economically viable solar energy systems. To pave the way for solar energy to be a major type of green energy.

2.1 NEED FOR STUDY

Climate changes in recent times had shown a high demand for electricity that have led to the requirement of power generation from green and renewable sources. Solar energy being one of them. Solar energy, which is an abundant sustainable energy resource, which causes very less harm to the environment, turning the Sun into a major source of energy for the humans. This solar power can be harnessed either through concentrated power plants or photovoltaic (PV) power plants. Here, we deal with the PV power plants; their performance is mainly related to the factors of electrical parameters of its components (PV panels, inverters), characteristics of the installation (orientation, tilt angle) and meteorological conditions. The meteorological factors that are affecting the power produced by a PV field is mainly the absorbed solar irradiance. There is, in fact, a linear correlation between the PV modules' maximum power and the solar irradiance. The value of the solar irradiance being high or low depends on the geographical location and time along with the orientation of the panel that is relative to both the Sun and the sky. As such, solar power tends to have a chaotic and intermittent behaviour. In this study, the main aim is to forecast irradiance optimally also in a generalized manner, as we face problems like that of solar power forecasts. The solar irradiance forecasting is also performed on historical data from two locations in India for protection and conservation of the environment as well as energy security. The main aim is to achieve an increase in the amount of renewable or green energy contribution to the power generated.

2.2 OBJECTIVE OF OUR PROBLEM

i. To develop forecast models based on ML algorithms to predict solar energy irradiance
ii. To develop multiple regression model based on the historical data collected
iii. To study the accuracy, performance and reliability of the model

CHAPTER 3
METHODOLOGY

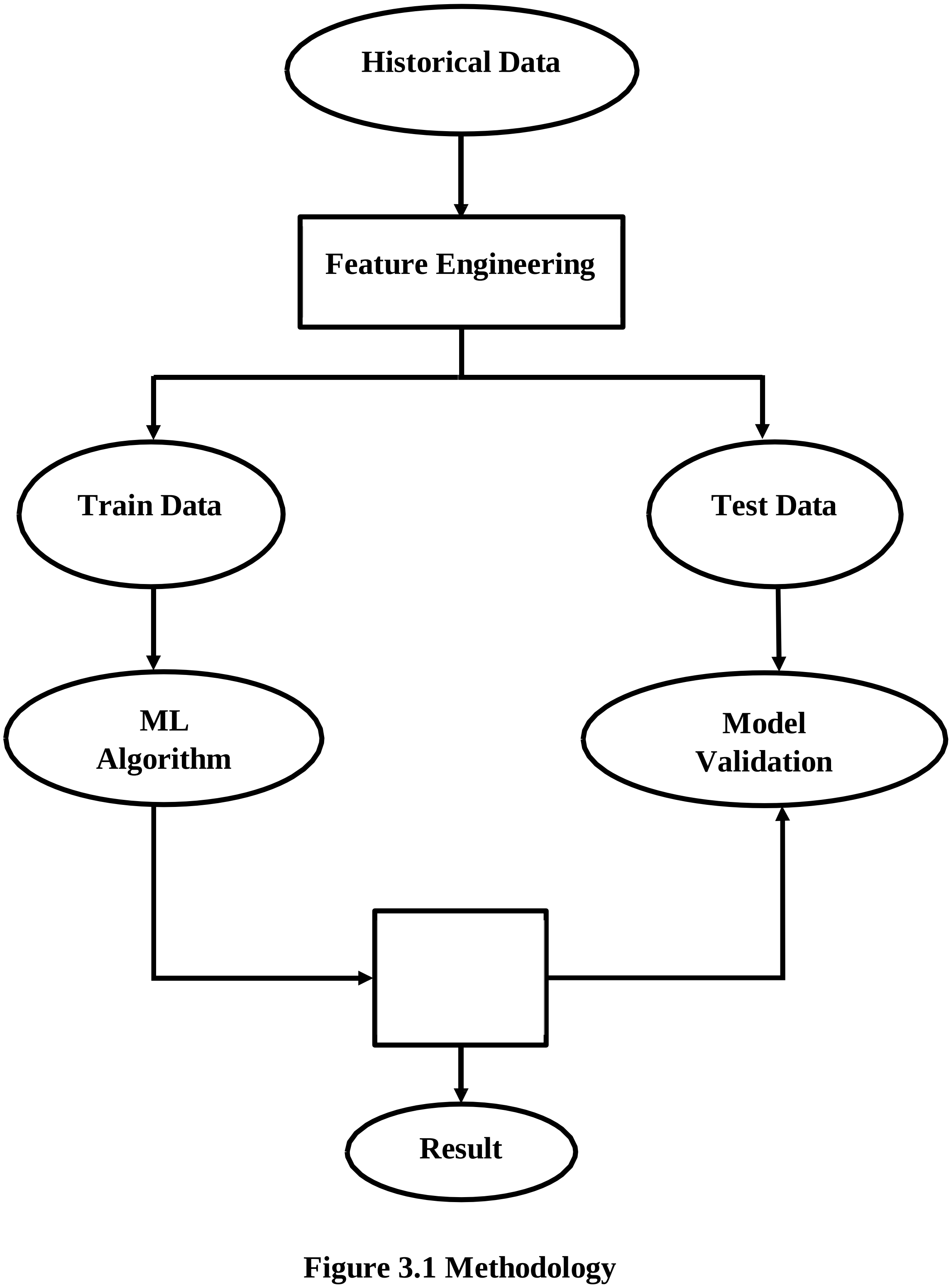


Figure 3.1 Methodology

The methodology of the proposed machine learning framework for the prediction of solar radiation using regression model is shown in Figure 3.1.

Regression is a supervised machine learning that helps in predicting continuous numerical values or quantity.

Regression model can be simple linear or multiple linear regression model. If a linear regression model involves single predictor variable it is called linear regression. If it involves multiple predictor variables it is called multiple regression.

Mathematically, a regression model is represented as y=f(x) where y is the target or dependent variable and x is the set of predictors at independent variables.

CHAPTER 4
DATA COLLECTION

Table 4.2 Data Collection

UNIX Time	Data	Time	Temperature	Pressure	Humidity	Wind Direction (Degrees)	Speed	Time sunrise	Time sunset
1.48E+09	9/29/2016 12:00:00 AM	23:55:26	48	30.46	59	177.39	5.62	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:00 AM	23:50:23	48	30.46	58	176.78	3.37	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:00 AM	23:45:26	48	30.46	57	158.75	3.37	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:00 AM	23:40:21	48	30.46	60	137.71	3.37	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:00 AM	23:35:24	48	30.46	62	104.95	5.62	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:00 AM	23:30:24	48	30.46	64	120.2	5.62	06:13:00	18:13:00

1.48E+09	9/29/2016 12:00:00 AM	23:25:19	49	30.46	72	112.45	6.75	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:00 AM	23:20:22	49	30.46	71	122.97	5.62	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:00 AM	23:15:22	49	30.46	80	101.18	4.5	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:00 AM	23:10:22	49	30.46	85	141.87	4.5	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:00 AM	23:05:23	49	30.47	93	120.55	2.25	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:00 AM	23:00:25	49	30.47	98	144.19	3.37	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:00 AM	22:55:20	49	30.47	99	139.8	6.75	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:00 AM	22:50:19	50	30.47	99	140.92	2.25	06:13:00	18:13:00
1.48E+09	9/29/2016	22:45:31	50	30.47	99	147.61	5.62	06:13:00	18:13:00

	12:00:00 AM								
1.48E+09	9/29/2016 12:00:00 AM	22:40:23	50	30.47	99	113.78	4.5	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:00 AM	22:35:19	50	30.47	99	123.03	10.12	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:00 AM	22:30:22	50	30.47	99	173.73	6.75	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:00 AM	22:25:19	50	30.47	98	91.43	6.75	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:00 AM	22:20:22	50	30.47	98	109.74	6.75	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:00 AM	22:15:22	50	30.47	98	143.53	2.25	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:00 AM	22:10:21	50	30.47	97	146.76	5.62	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:	22:05:20	50	30.47	97	158.35	4.5	06:13:00	18:13:00

	00 AM								
1.48E+09	9/29/2016 12:00:00 AM	22:00:26	50	30.47	97	166.05	5.62	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:00 AM	21:55:20	50	30.47	96	151.82	5.62	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:00 AM	21:50:32	50	30.47	96	152.52	6.75	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:00 AM	21:45:28	50	30.47	95	127.82	7.87	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:00 AM	21:40:20	50	30.47	95	157.58	6.75	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:00 AM	21:35:21	50	30.47	95	172.96	7.87	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:00 AM	21:30:21	50	30.47	95	143.82	6.75	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:00 AM	21:25:22	50	30.47	94	154.03	5.62	06:13:00	18:13:00

1.48E+09	9/29/2016 12:00:00 AM	21:20:21	50	30.47	93	165.93	4.5	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:00 AM	21:15:23	50	30.47	93	139.82	7.87	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:00 AM	21:10:20	49	30.47	93	166.07	9	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:00 AM	21:05:26	50	30.47	93	116.16	5.62	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:00 AM	21:00:20	50	30.47	92	157.48	4.5	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:00 AM	20:55:18	50	30.47	91	112.27	7.87	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:00 AM	20:50:20	50	30.47	90	133.69	6.75	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:00 AM	20:45:24	50	30.47	88	140.64	6.75	06:13:00	18:13:00
1.48E+09	9/29/2016	20:40:19	50	30.47	87	52.89	1.12	06:13:00	18:13:00

	12:00:00 AM								
1.48E+09	9/29/2016 12:00:00 AM	20:35:19	50	30.47	86	132.66	4.5	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:00 AM	20:30:18	50	30.46	85	138.35	5.62	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:00 AM	20:25:20	50	30.46	84	109.87	3.37	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:00 AM	20:20:23	50	30.47	81	123.25	3.37	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:00 AM	20:15:20	50	30.46	75	158.64	4.5	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:00 AM	20:10:22	50	30.46	70	165.22	5.62	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:00 AM	20:05:22	50	30.46	63	160.51	3.37	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:	20:00:18	50	30.46	51	155.36	4.5	06:13:00	18:13:00

	00 AM								
1.48E+09	9/29/2016 12:00:00 AM	19:55:23	50	30.46	47	157.78	6.75	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:00 AM	19:50:22	50	30.46	46	151.6	3.37	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:00 AM	19:45:25	50	30.46	47	105.78	4.5	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:00 AM	19:40:22	50	30.46	48	135.55	4.5	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:00 AM	19:35:22	50	30.46	56	140.92	4.5	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:00 AM	19:30:20	50	30.46	57	165	4.5	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:00 AM	19:25:21	51	30.46	55	126.63	3.37	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:00 AM	19:20:20	51	30.46	54	119.97	4.5	06:13:00	18:13:00

1.48E+09	9/29/2016 12:00:00 AM	19:15:22	51	30.45	53	95.57	2.25	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:00 AM	19:10:24	51	30.45	51	100.02	4.5	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:00 AM	19:05:20	51	30.45	51	111.73	2.25	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:00 AM	19:00:24	51	30.45	52	145.77	5.62	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:00 AM	18:55:19	51	30.45	51	105.89	6.75	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:00 AM	18:50:24	51	30.45	52	120.86	5.62	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:00 AM	18:45:19	51	30.44	55	170.11	5.62	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:00 AM	18:40:21	52	30.44	55	107.15	5.62	06:13:00	18:13:00
1.48E+09	9/29/2016	18:35:21	52	30.44	55	129.29	6.75	06:13:00	18:13:00

	12:00:00 AM								
1.48E+09	9/29/2016 12:00:00 AM	18:30:23	52	30.45	59	158.73	3.37	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:00 AM	18:25:19	52	30.44	59	211.38	2.25	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:00 AM	18:20:23	53	30.44	61	174.97	4.5	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:00 AM	18:15:21	53	30.44	62	120.78	7.87	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:00 AM	18:10:52	53	30.44	59	118.82	5.62	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:00 AM	18:05:22	54	30.44	59	154.16	4.5	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:00 AM	18:00:22	55	30.44	57	58.42	6.75	06:13:00	18:13:00
1.48E+09	9/29/2016 12:	17:55:22	55	30.44	53	47.86	4.5	06:13:00	18:13:00

	00 AM								
1.48E +09	9/29/2 016 12:00: 00 AM	17:5 0:19	56	30.4 4	52	74.56	2.2 5	06:1 3:00	18:1 3:00
1.48E +09	9/29/2 016 12:00: 00 AM	17:4 5:22	56	30.4 4	49	25.72	6.7 5	06:1 3:00	18:1 3:00
1.48E +09	9/29/2 016 12:00: 00 AM	17:4 0:22	56	30.4 4	50	102.1 1	2.2 5	06:1 3:00	18:1 3:00
1.48E +09	9/29/2 016 12:00: 00 AM	17:3 5:38	56	30.4 4	48	67.49	3.3 7	06:1 3:00	18:1 3:00
1.48E +09	9/29/2 016 12:00: 00 AM	17:3 0:25	56	30.4 3	49	180.2 9	2.2 5	06:1 3:00	18:1 3:00
1.48E +09	9/29/2 016 12:00: 00 AM	17:2 5:25	57	30.4 3	57	150.2	2.2 5	06:1 3:00	18:1 3:00
1.48E +09	9/29/2 016 12:00: 00 AM	17:2 0:25	57	30.4 3	55	47.3	5.6 2	06:1 3:00	18:1 3:00
1.48E +09	9/29/2 016 12:00: 00 AM	17:1 5:20	58	30.4 3	52	333.8 1	1.1 2	06:1 3:00	18:1 3:00

1.48E+09	9/29/2016 12:00:00 AM	17:10:20	58	30.43	49	41.62	4.5	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:00 AM	17:05:22	58	30.43	45	132.36	5.62	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:00 AM	17:00:24	58	30.43	41	332.09	4.5	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:00 AM	16:55:25	58	30.43	55	45.88	3.37	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:00 AM	16:50:22	58	30.43	56	94.24	7.87	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:00 AM	16:45:38	58	30.43	50	81.57	6.75	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:00 AM	16:40:20	58	30.43	49	351.03	7.87	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:00 AM	16:35:41	58	30.43	51	78.81	5.62	06:13:00	18:13:00
1.48E+09	9/29/2016	16:30:22	58	30.43	46	85.96	5.62	06:13:00	18:13:00

	12:00:00 AM								
1.48E+09	9/29/2016 12:00:00 AM	16:25:21	59	30.43	48	19.78	4.5	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:00 AM	16:20:25	58	30.43	59	18.26	4.5	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:00 AM	16:15:23	58	30.43	52	69.08	3.37	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:00 AM	16:10:22	58	30.43	52	40.87	4.5	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:00 AM	16:05:20	58	30.43	49	0.13	4.5	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:00 AM	16:00:20	59	30.43	58	27.34	9	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:00 AM	15:55:22	59	30.43	46	359.93	3.37	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:	15:50:21	60	30.43	46	22.18	7.87	06:13:00	18:13:00

	00 AM								
1.48E+09	9/29/2016 12:00:00 AM	15:45:22	60	30.43	48	4.38	10.12	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:00 AM	15:40:24	60	30.43	44	355.03	7.87	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:00 AM	15:35:19	60	30.42	47	0.4	7.87	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:00 AM	15:30:20	60	30.42	51	53.16	4.5	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:00 AM	15:25:21	60	30.42	48	353.87	9	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:00 AM	15:20:24	60	30.43	54	2.04	4.5	06:13:00	18:13:00
1.48E+09	9/29/2016 12:00:00 AM	15:15:18	61	30.42	53	34.85	11.25	06:13:00	18:13:00

CHAPTER 5
GOOGLE COLAB CODE

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.preprocessing import PolynomialFeatures
from sklearn.pipeline import Pipeline
x=pd.read_csv('/content/SolarPrediction.csv')
x.head()
```

	UNIXTime	Data	Time	Radiation	Temperature	Pressure	Humidity	WindDirection(Degrees)	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

```
x.isnull().sum()
UNIXTime        0
Data            0
Time            0
Radiation       0
Temperature     0
Pressure        0
Humidity        0
WindDirection   0
```

Speed	0
TimeSunRise	0
TimeSunSet	0

dtype: int64

x.shape

(32686, 11)

x.info()

<class 'pandas.core.frame.DataFrame'>RangeIndex: 32686 entries, 0 to 32685 Data columns (total 11 columns):

#	Column	Non-Null	Count
	Dtype		
0	UNIXTime int64	32686	non-null
1	Data object	32686	non-null
2	Time object	32686	non-null
3	Radiation float64	32686	non-null
4	Temperature int64	32686	non-null
5	Pressure float64	32686	non-null
6	Humidity int64	32686	non-null
7	WindDirection(Degrees) float64	32686	non-null
8	Speed float64	32686	non-null
9	TimeSunRise object	32686	non-null

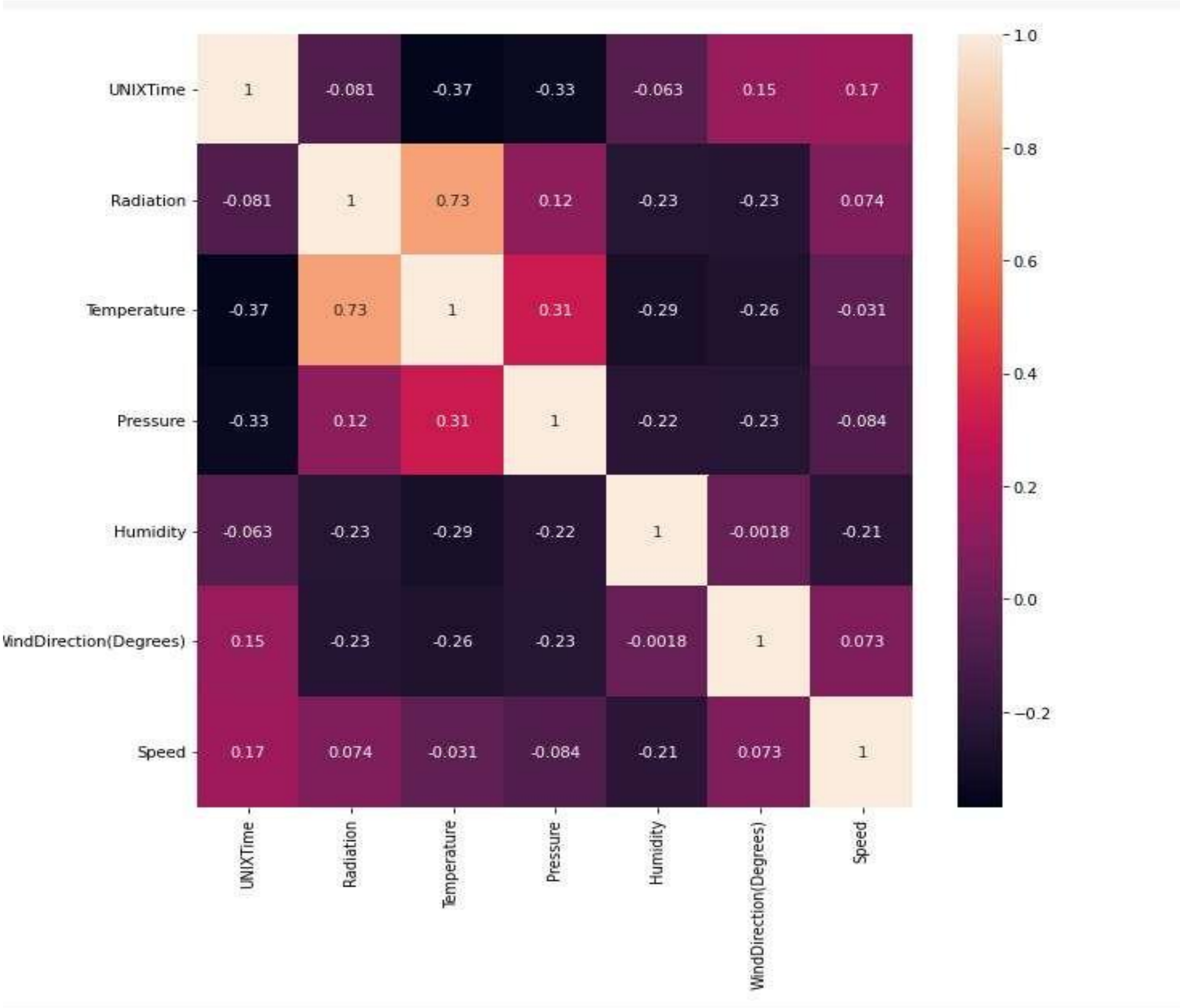
10 TimeSunSet 32686 non-null object

dtypes: float64(4), int64(3), object(4)

x.describe()

	UNIX Time	Radiation	Temperature	Pressure	Humidity	WindDirection(Degrees)	Speed
count	3.27E+04	32686	32686	32686	32686	32686	32686
mean	1.48E+09	207.1247	51.10326	30.42288	75.01631	143.4898	6.243869
std	3.01E+06	315.916	6.201157	0.054673	25.99022	83.1675	3.490474
min	1.47E+09	1.11	34	30.19	8	0.09	0
25%	1.48E+09	1.23	46	30.4	56	82.2275	3.37
50%	1.48E+09	2.66	50	30.43	85	147.7	5.62
75%	1.48E+09	354.235	55	30.46	97	179.31	7.87
max	1.48E+09	1601.26	71	30.56	103	359.95	40.5

```
plt.figure(figsize=(10,10))
sns.heatmap(x.corr(),annot=True)
plt.show()
```



x['Time_conv'] =pd.to_datetime(x['Time'], format='%H:%M:%S’)

x['hour'] = pd.to_datetime(x['Time_conv'], format='%H:%M:%S').dt.hour

x['month'] = pd.to_datetime(x['UNIXTime'].astype(int), unit='s').dt.month

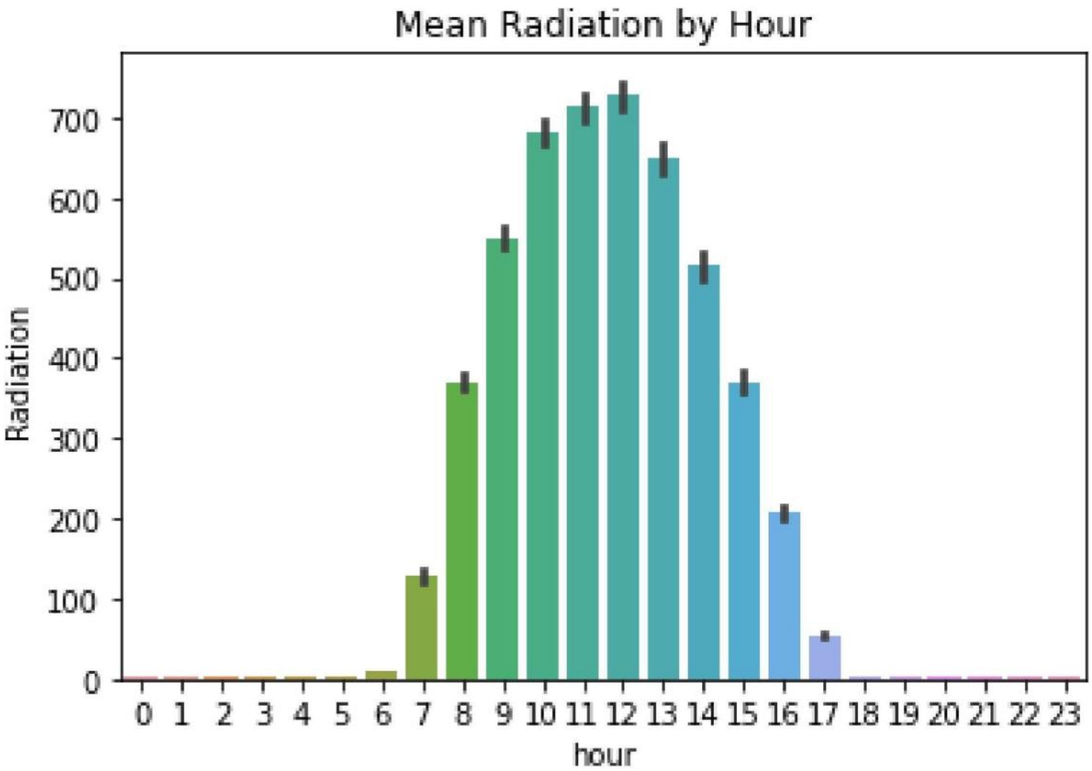
x['year'] = pd.to_datetime(x['UNIXTime'].astype(int), unit='s').dt.year

x.head()

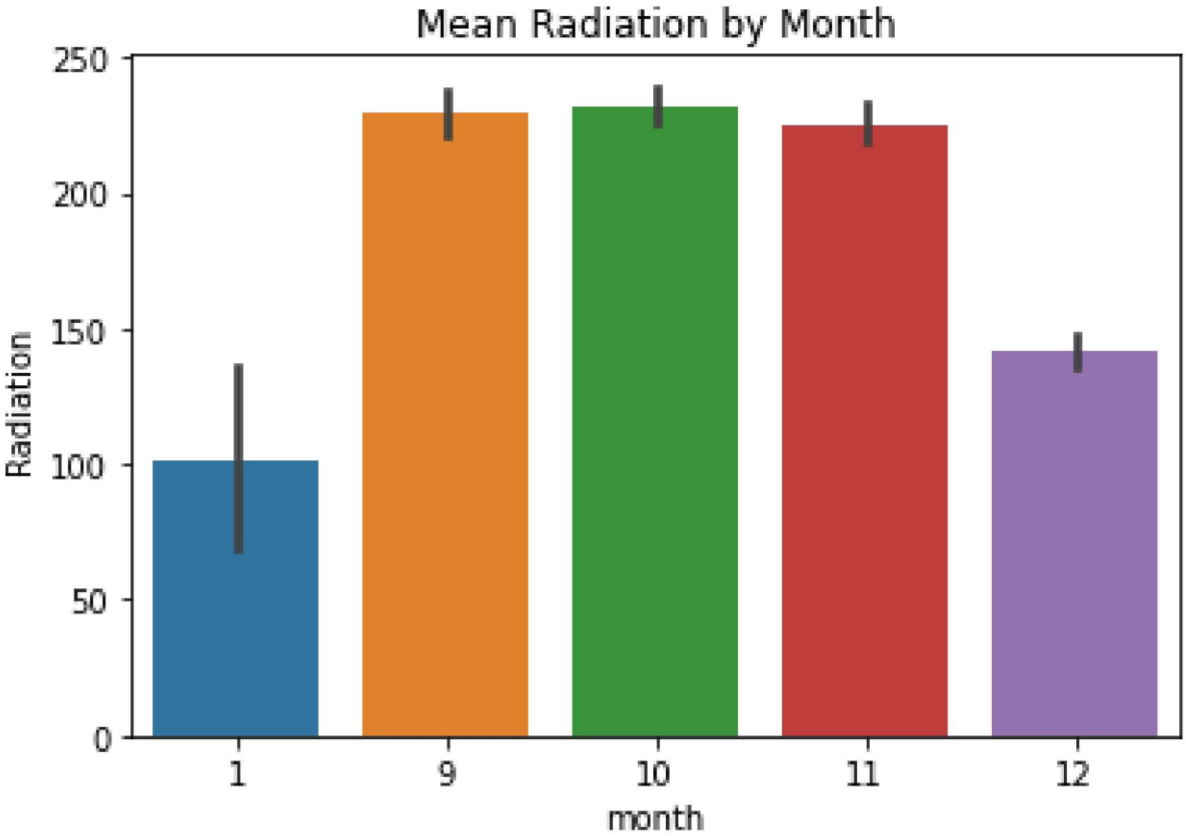
	UNIX Time	Data	Time	Radiation	Temperature	Pressure	Humidity	WindDirection (Degrees)	Speed	TimeSunSet	Time_conv	hour	month	year
0	1.48 E+09	9/29/ 2016 12:00 :00 AM	23:55:26	1.21	48	30.46	59	177.39	5.62	18:13:00	1900-01-01 23:55 :26	23	9	2016

1	1.48 E+09	9/29/ 2016 12:00 :00 AM	23:5 0:23	1.21	48	30.4 6	58	176.7 8	3.3 7	18:13: 00	1900- 01-01 23:50 :23	23	9	20 16
2	1.48 E+09	9/29/ 2016 12:00 :00 AM	23:4 5:26	1.23	48	30.4 6	57	158.75	3.3 7	18:13: 00	01- 01- 1900 23:45	23	9	20 16
3	1.48 E+09	9/29/ 2016 12:00 :00 AM	23:4 0:21	1.21	48	30.4 6	60	137.71	3.3 7	18:13: 00	01- 01- 1900 23:40	23	9	20 16
4	1.48 E+09	9/29/ 2016 12:00 :00 AM	23:3 5:24	1.17	48	30.4 6	62	104.95	5.6 2	18:13: 00	01- 01- 1900 23:35	23	9	20 16

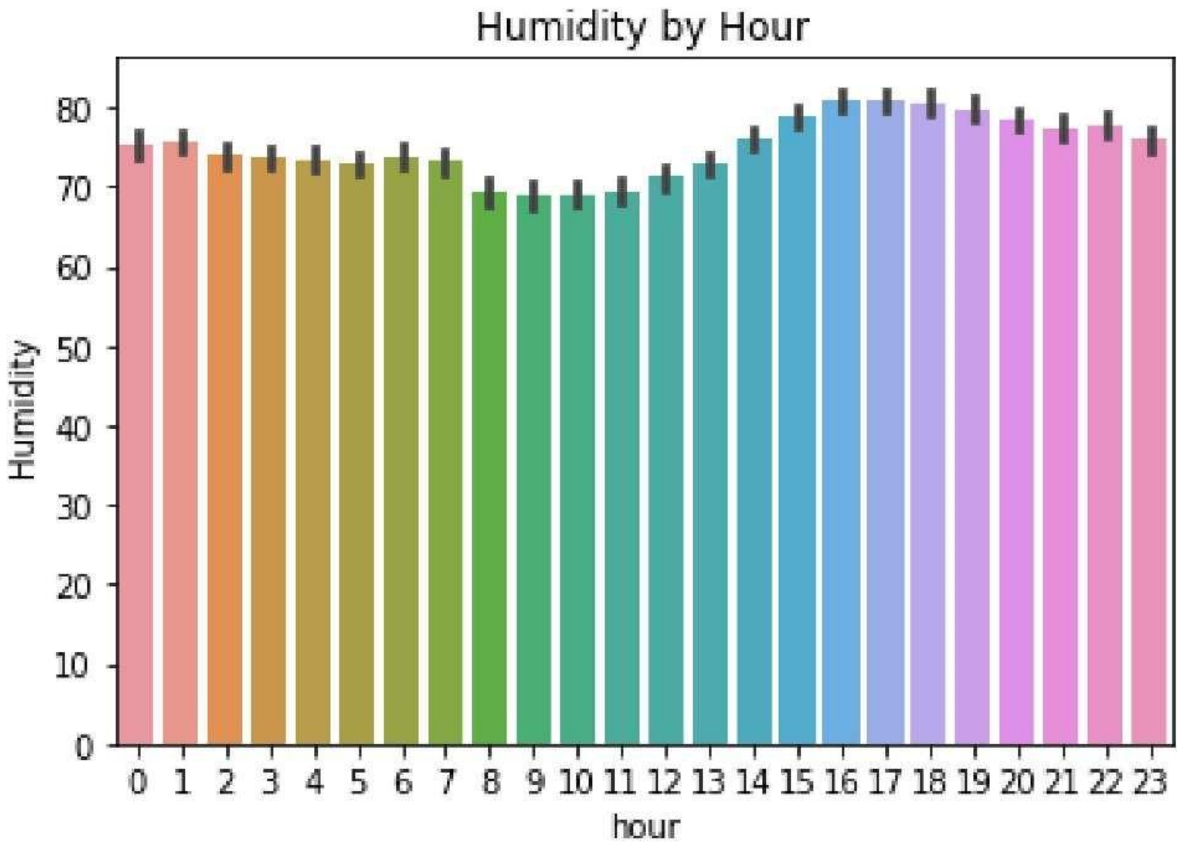
```
ax = plt.axes()
sns.barplot(x="hour", y='Radiation', data=x)
ax.set_title('Mean Radiation by Hour’)
plt.show()
```



```
ax  =  plt.axes()
sns.barplot(x="month", y='Radiation', data=x)
ax.set_title('Mean Radiation by Month’)
plt.show()
```




```
ax  =  plt.axes()
sns.barplot(x="hour", y='Humidity', data=x)
ax.set_title('Humidity by Hour’)
plt.show()
```



x['Time_conv'] =pd.to_datetime(x['Time'], format='%H:%M:%S')

x['hour'] = pd.to_datetime(x['Time_conv'], format='%H:%M:%S').dt.hour

x['month'] = pd.to_datetime(x['UNIXTime'].astype(int), unit='s').dt.month

x['year'] = pd.to_datetime(x['UNIXTime'].astype(int), unit='s').dt.year

x.head()

```
y = x['Radiation']
x = x.drop(['Radiation', 'Data', 'Time', 'TimeSunRise', 'TimeSunSet','Time_conv'], axis=1)
```

```
x.head()
```

	UNIXTime	Temperature	Pressure	Humidity	WindDirection(Degrees)	Speed	hour
0	1475229326	48	30.46	59	177.39	5.62	23
1	1475229023	48	30.46	58	176.78	3.37	23
2	1475228726	48	30.46	57	158.75	3.37	23
3	1475228421	48	30.46	60	137.71	3.37	23
4	1475228124	48	30.46	62	104.95	5.62	23

from sklearn.model_selection import train_test_split

from sklearn.linear_model import LinearRegression

from sklearn.preprocessing import StandardScalerx_train, x_test, y_train, y_test = train_test_split(x, y, test_size = 0.2, random_state =poly_reg_pipeline= Pipeline([('std',StandardScaler()),("poly", PolynomialFeatures(degree=('lin',LinearRegression())]) poly_reg_pipeline.fit(x_train,y_train)

Pipeline(steps=[('std', StandardScaler()),

('poly', PolynomialFeatures(degree=3)), ('lin', LinearRegression())])

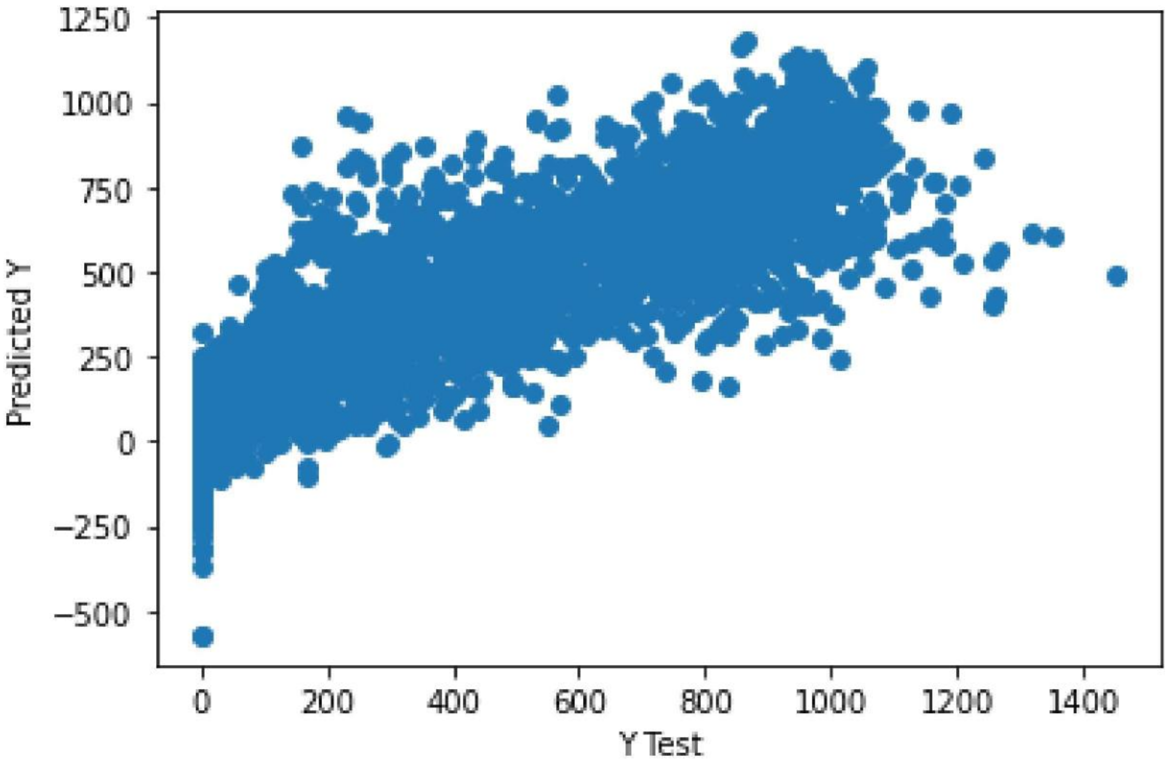
predict=poly_reg_pipeline.predict(x_test)

plt.scatter(y_test,predict)

plt.xlabel('Y Test’)

plt.ylabel('Predicted Y’)

Text(0, 0.5, 'Predicted Y’)



from sklearn import metrics

print('MAE:', metrics.mean_absolute_error(y_test, predict)) print('MSE:’, metrics.mean_squared_error(y_test, predict))

print('RMSE:', np.sqrt(metrics.mean_squared_error(y_test, predict)))

MAE: 88.45172769746836

MSE: 17546.505445750725

RMSE: 132.46322299321696

from sklearn.metrics import r2_score

print(r2_score(y_test, predict))

0.8267034477535906

from sklearn.model_selection import validation_curve

degree = [1, 2, 3, 4, 5]

train_scores, test_scores = validation_curve(

poly_reg_pipeline, x_train, y_train, param_nam param_range=degree, cv=10, scoring="neg_mean_a)

train_errors, test_errors = -train_scores, -test_scores

plt.plot(degree, train_errors.mean(axis=1), 'b-x', label='Training error’)

plt.plot(degree, test_errors.mean(axis=1), 'r-x', label='Test error’)

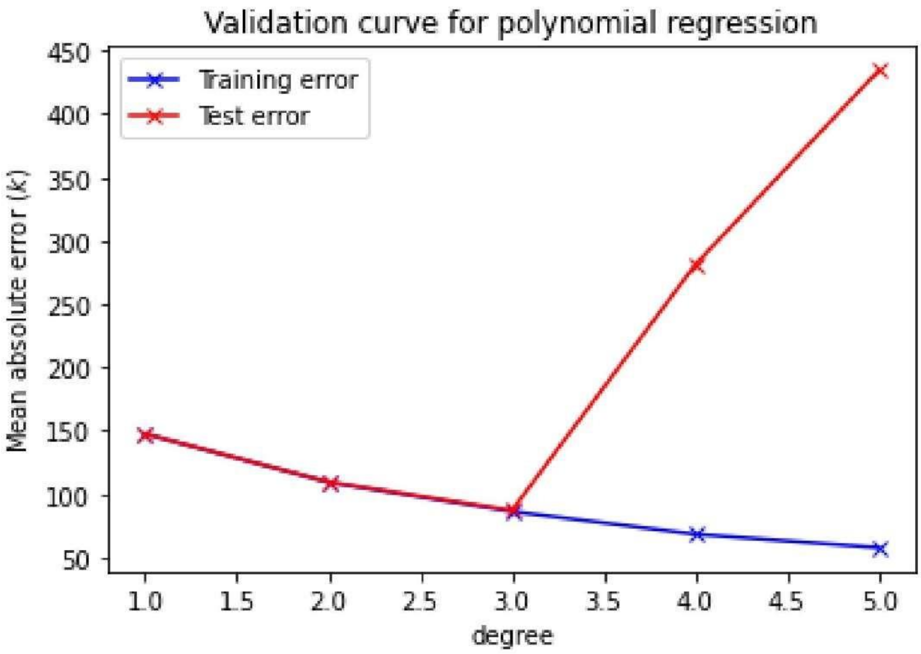
plt.legend()

plt.xlabel("degree")

plt.ylabel("Mean absolute error (k)")

plt.title("Validation curve for polynomial regression")

Text(0.5, 1.0, 'Validation curve for polynomial regression')



Validation curve for polynomial regression
Training error
Test error
Mean absolute error (k)
450
400
350
300
250
200
150
100
50
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
5.0
degree

CHAPTER 6
RESULT AND DISCUSSION

Using the dataset obtained the multiple regression model to predict the solar radiation was successfully built. The regression model enables the prediction of solar radiation in minutes values over a period of few days. The Regression model presents a relationship between solar irradiance, air temperature and relative humidity.

This graph (Figure 6.1) shows that the hourly solar irradiance follows an increasing trend up to a particular point after which it shows decreasing trend.

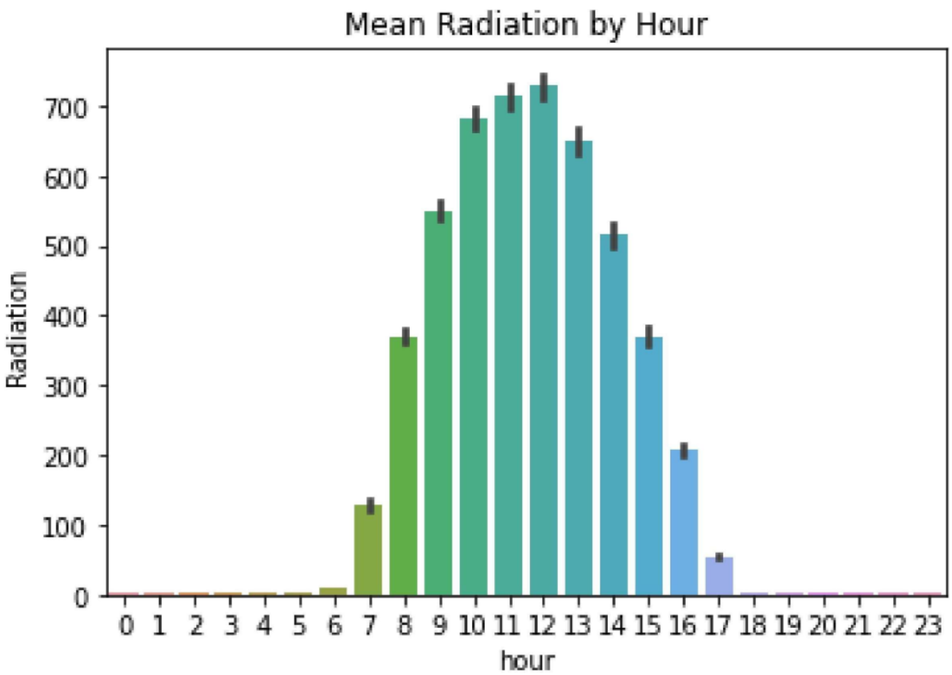


Figure 6.1 Hourly Solar Radiation

This graph (Figure 6.2) shows that the maximum of the global solar radiation is observed in October while the minimum values are appearing in January.

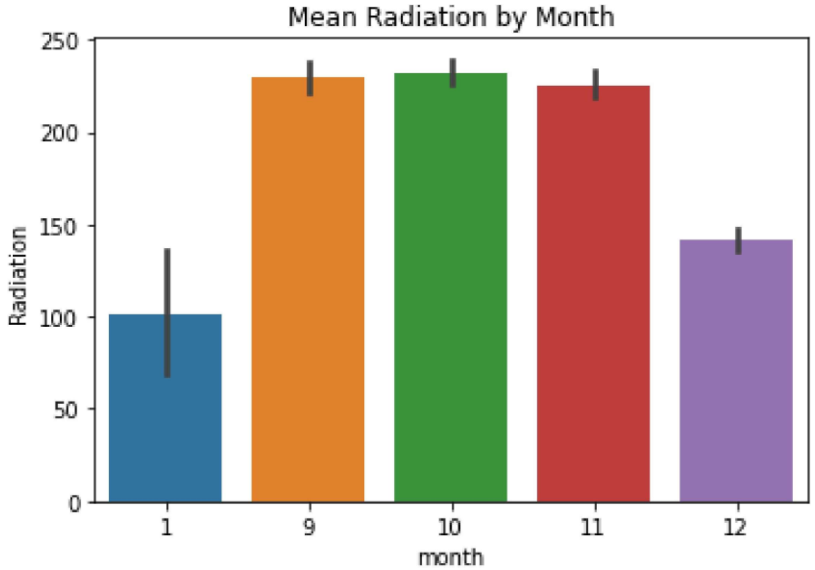


Figure 6.2 Mean Radiation by Month

The graph in Figure 6.3 shows the difference in the amount of humidity by hour.

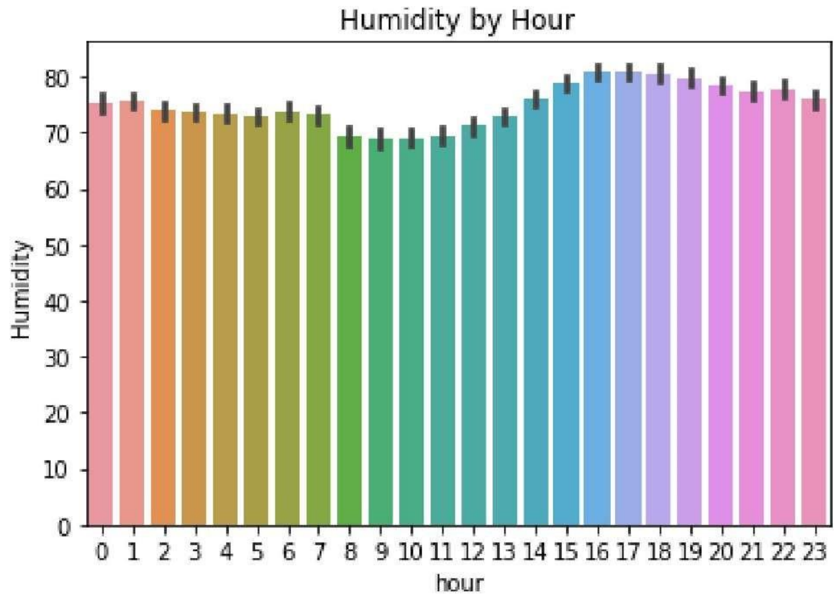


Figure 6.3 Humidity by Hour

The validation curve (Figure 6.4) shows a lower training error and a higher test error which indicates overfitting of the model.

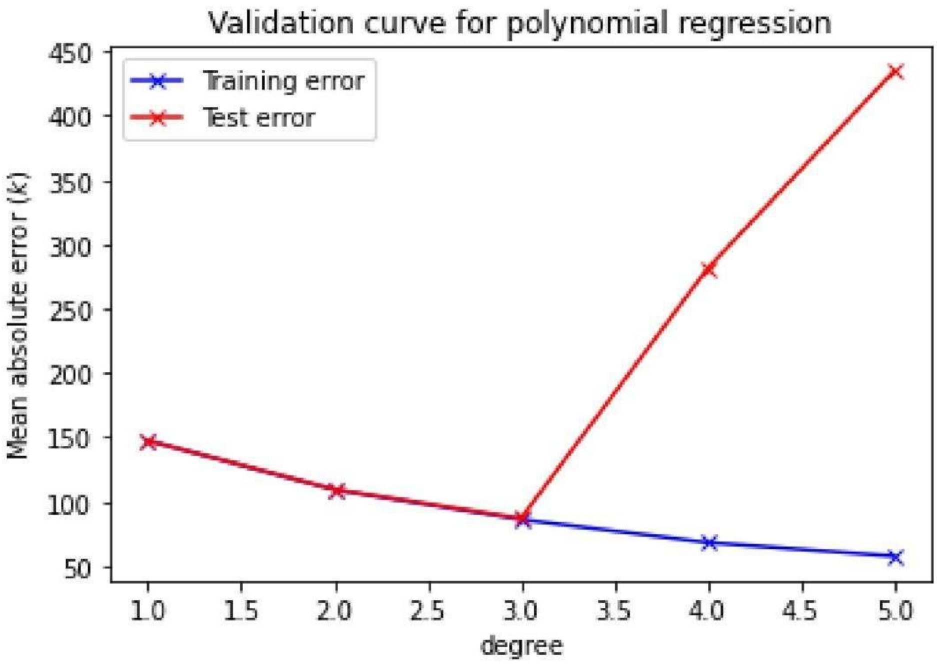


Figure 6.4 Validation curve for polynomial regression

CHAPTER 7
CONCLUSION

Solar radiation forecast has drawn the attention of many research due to the requirement of renewable and green energy. In general, promising models with forecast precision, estimation of tapping potential solar energy in particular locations and advance the sustainable planning of solar power applications.

With the advent of machine learning, we are entering a new world where we are reading a nexus of capability. Machine learning will quietly but meaningfully improve the core operation.

REFERENCES

- Thapar V. A revisit to solar radiation estimation using sunshine duration
- Srivastav, S Lessmann. A comparative study of neural networks in forecasting global irradiance with satellite data.
- Alex G Greg. Neural turing machines.
- Thapar V, Agnihotri G, Sethi VK. Estimation of hourly temperature at a site and its impact on energy yield of PV module.