Solar Irradiance Prediction

Solar irradiance is a measure of how much energy the sun sends to the Earth. Imagine that the sun is like a big light bulb in the sky, and it sends out light and heat to the Earth. The amount of light and heat that the sun sends to the Earth is called solar irradiance. We use special tools to measure how much solar irradiance there is at different places on the Earth. Solar irradiance is important because it helps us to understand how much energy we can get from the sun. Example of how solar irradiance can be used:

Imagine that you are planning to install solar panels on the roof of your house. You want to make sure that you will get enough energy from the sun to power your house. To do this, you can use solar irradiance data to help you predict how much energy the solar panels will be able to produce.

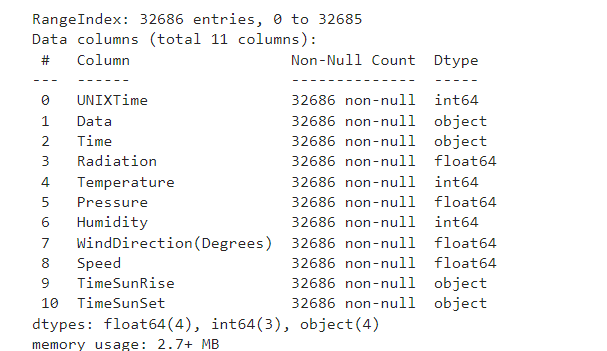
The dataset used in this project is meteorological data from the HI-SEAS weather station from four months (September through December 2016) between Mission IV and Mission V.

For each dataset, the fields are:

A row number (1-n) useful in sorting this export's results The UNIX time\_t date (seconds since Jan 1, 1970). Useful in sorting this export's results with other export's results the date in yyyy-mm-dd format the local time of day in hh:mm:ss 24-hour format The numeric data, if any (may be an empty string) The text data, if any (may be an empty string). Using this data, we will extract features and using those features we can predict the radiations.

The units of each dataset are:

* Solar radiation: watts per meter^2
* Temperature: degrees Fahrenheit
* Humidity: percent
* Barometric pressure: Hg
* Wind direction: degrees
* Wind speed: miles per hour
* Sunrise/sunset: Hawaii time



***Data Info***

***Data Wrangling:***

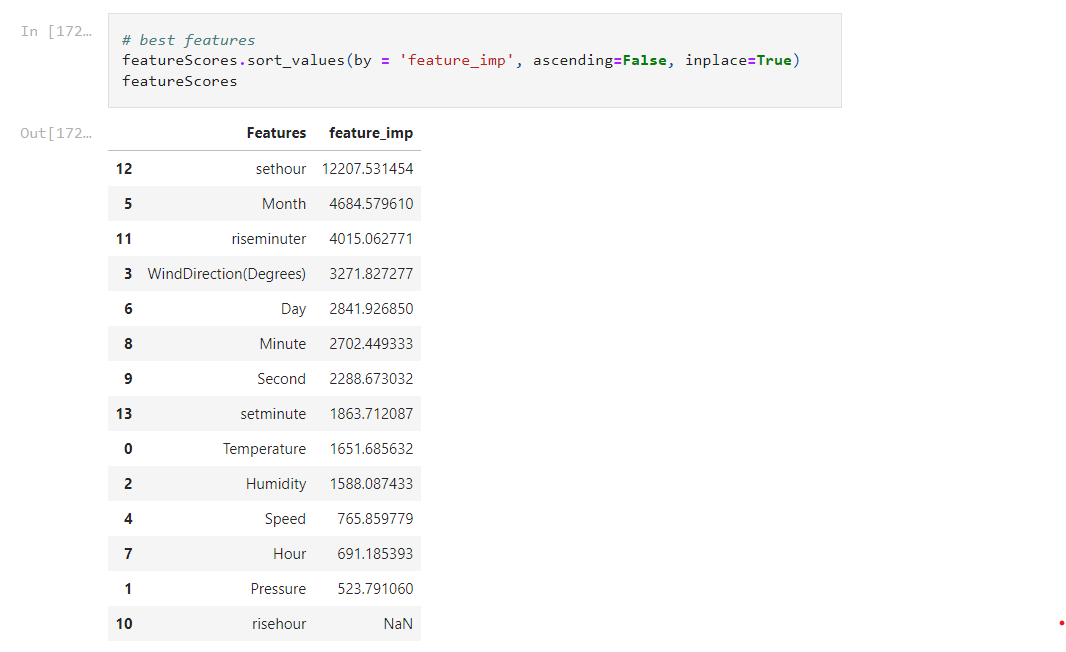
The first step that we can perform is the data wrangling:

* We can split data
* We can split month day year seconds minutes
* Extract Risehour, setminute etc.
* After extracting relevant features, we can drop their parent features to reduce the complexity of the dataset
* We can create the target dataset which will only contain the radiation column

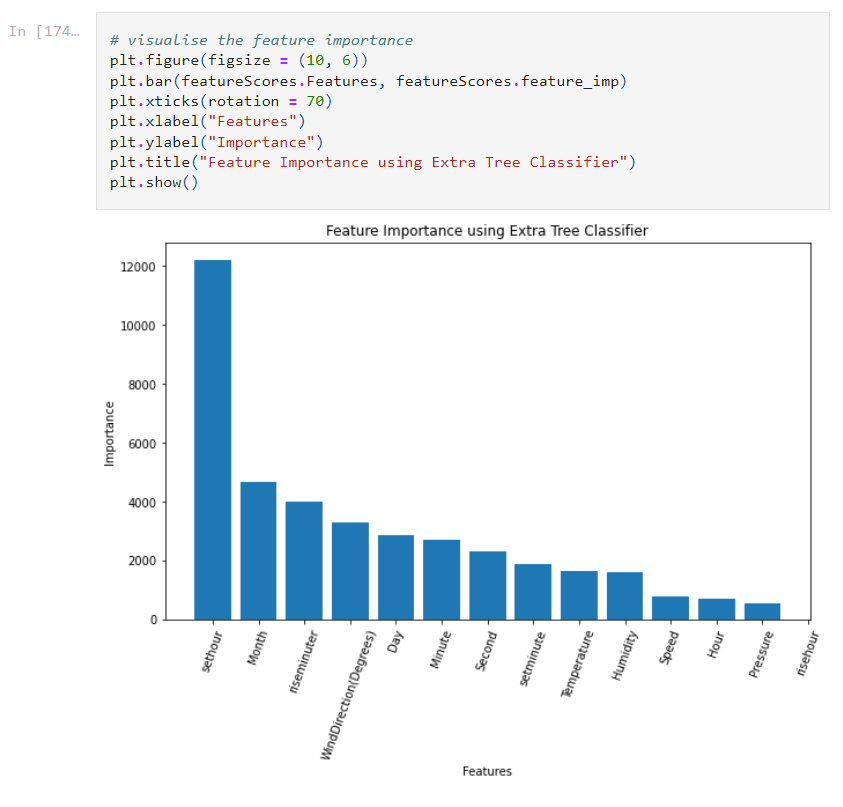
***Feature Selection***

* Feature Selection using Correlation Matrix
* Feature Selection using SelectKBest Method
* Feature Selection using Extra Tree Classifier

We can perform Feature selection using these 3 techniques



**Using SelectKBest Method**



**Feature Engineering**

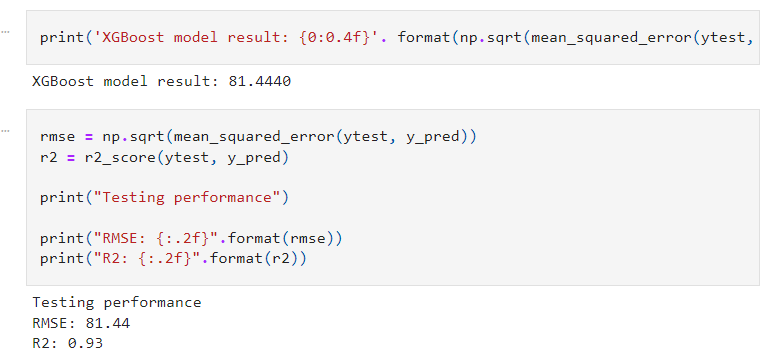
After Feature Selection we’ll perform Feature Engineering on the Dataset using **BoxCox, Log, MinMax and Standard Transformations.**

Then we’ll prepare the data:

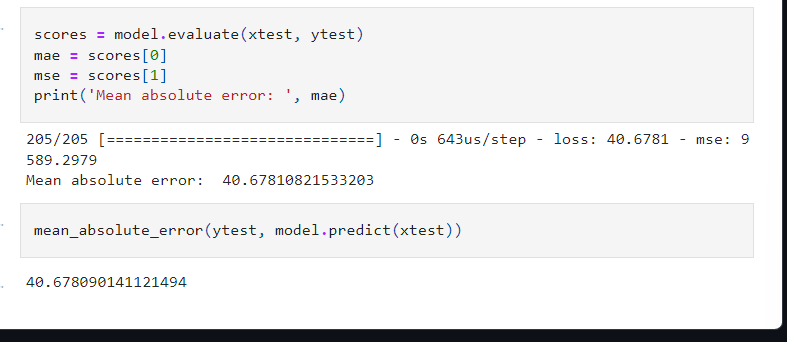
* Standardization
* Train Test Split

**Modelling**

The first prediction will be done using XGBoost:



We can also make a multilayer perceptron model:



Data: <https://github.com/AshishJangra27/Machine-Learning-with-Python-GFG/blob/main/Solar%20Irradiance%20Prediction/SolarPrediction.csv>