edunet foundation

Solar power prediction using Linear Regression

Learning Objectives

- ✓ Understanding the data set
- ✓ Identifying the key features
- ✓ Building a predictive Linear Regression Model
- ✓ Evaluating Model Performance

Source : www.freepik.com/

Tools and Technology used

- Jupyter note book
- Python
- Pandas
- Sea born
- Numpy
- Matplotlib
- Scikit learning

Methodology

1. Data collection:

Description: Gather the solar power generation dataset which contains historical data.

2. Exploratory Data Analysis (EDA):

Description: Explore the dataset to understand its structure, summary statistics, distributions, and relationships between variables.

3.Data Cleaning:

Description: Identify and handle missing values and duplicate records.

4.Data Visualization:

Description: Create visualizations to better understand the data distribution and correlations among variables.

5. Data preprocessing:

Description: Split the data into training and testing sets and standardize the features value

6. Model Building:

Description: Train a machine learning model on the training data

7.Model Evaluation:

Description: Evaluate the models on the test data

Problem Statement:

The project aims to **predict solar power generation** using historical data. By analyzing various factors that influence solar power output, the goal is to develop a machine learning model to make accurate predictions. This involves:

1.Data Understanding

2. **Data Cleaning**

3.Feature Selection

4.Model Building

5.Model Evaluation

Solution:

Data Loading:

•Load the dataset containing historical solar power generation data using Pandas.

Exploratory Data Analysis (EDA):

•Conduct EDA to understand the dataset's structure, summary statistics, and distributions.

•Use Pandas functions like head(), describe(), info().

Data Cleaning:

•Check for and handle missing values and duplicates.

Data Visualization:

•Create visualizations to understand data distributions correlations

Feature Selection:

•Identify and select relevant features that contribute to the prediction of solar power generation.

•Separate features and target variable.

Data Preprocessing:

•Split the dataset into training and testing sets.

•Standardize the feature values.

•**Model Building**:

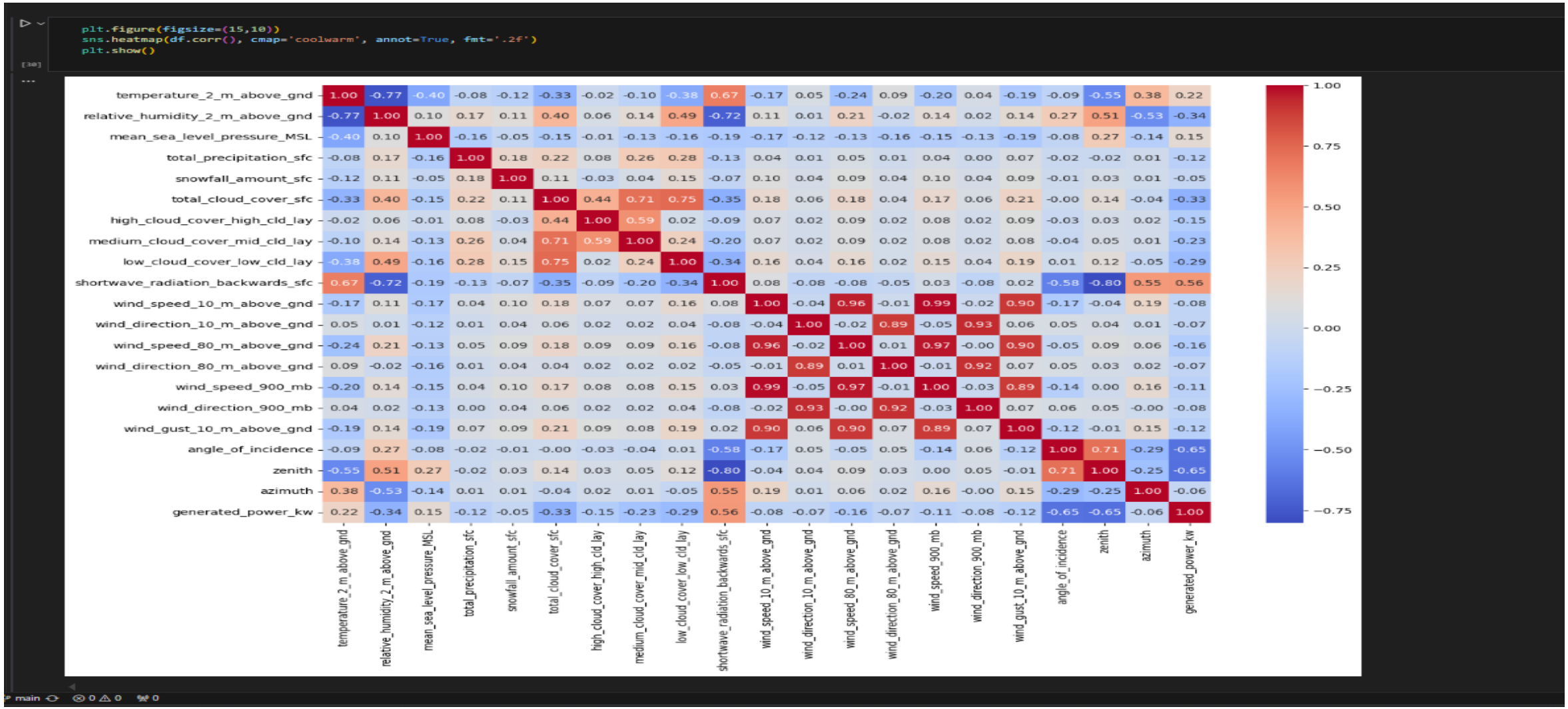
•Train a Linear Regression model using the training data.

•**Model Evaluation**:

•Evaluate the model's performance on the test data using the Mean Absolute Error (MAE) metric

edunet foundation

Screenshot of Output:



Conclusion:

- Evaluating the model on the test dataset using Mean Absolute Error (MAE) provided insights into its accuracy and reliability.
- The model's performance metrics indicated that it can make reasonably accurate predictions.