**Steps for Power Plant Energy Output Prediction Project**

* **Step 1: Problem Understanding**

- Define the problem: Predict energy output based on environmental conditions.

- Identify target variable (PE) and features (Temperature, Pressure, Humidity, Vacuum).

* **Step 2: Data Exploration**

- Load the dataset and check for missing values.

- Perform descriptive statistics to understand the data range.

- Visualize relationships using pairplots and correlation heatmaps.

* **Step 3: Feature Selection**

- Select key features (Temperature, Pressure, Humidity, Vacuum) based on correlations.

* **Step 4: Data Splitting**

- Split the data into training (80%) and test (20%) sets.

* **Step 5: Model Selection**

- Choose two models: Linear Regression and Random Forest.

- Train both models using cross-validation.

* **Step 6: Model Evaluation**

- Use Mean Absolute Error (MAE) to evaluate the models.

- Compare MAE scores of Linear Regression and Random Forest.

* **Step 7: Final Model Selection**

- Select the Random Forest model due to its lower MAE.

* **Step 8: Test Set Evaluation**

- Test the final model on the test set and check MAE performance.

* **Step 9: Visualization**

- Visualize feature importance and plot predictions vs actual values.

- Analyze residuals for model fit.

* **Step 10: Conclusion**

- Summarize the results: Random Forest model performs better with an MAE of 2.47.