Data Science Project Proposal

Overview: Problem and Background

The major challenges to integrating renewable energies in the grid are instability and intermittent production. Most renewable energy sources are dependent on the weather or other environmental phenomena. This adds an intrinsic uncertainty in their production working with Renewable energy so in this project we will predict the power generation for next couple of days and this could help on better management for solar plants.

Data

This data has been gathered at two solar power plants in India over a 34 day period. It has two pairs of files - each pair has one power generation dataset and one sensor readings dataset. The power generation datasets are gathered at the inverter level - each inverter has multiple lines of solar panels attached to it. The sensor data is gathered at a plant level - single array of sensors optimally placed at the plant. And we can use those features like date time , dc_power, ac_power, ambient_temperature, module_temperature, irradiation which can help to predict power generated on that day.

Tools

The main tool will going to use is Python and Jupyter Notebook were Pandas, numpy, matplotlib, and Scikit-Learn.

Minimum Viable Product (MVP) Template

Train test, run a basic linear regression model, and find the R2