**Data Preprocessing:**

While data preprocessing have cleaned the data and removed unwanted columns from the data.

Have joined both whether data and solar power data (as label), The final data created is of 15 mins interval for each day.

Also as the power can be generated only during day time data for night is not necessary and make the prediction more biased towards 0 power generation, hence have selected only the day time data, which is from 6am morning till 6pm evening.

Have handled the NAN values.

Have one hot encoded the categorical variables summary, icon and precip\_type.

Have generated a feature pahar to weight is segregate the data based on time of day.

Have generated a week\_of\_year, to capture the seasonality of the data.

**Gradient boosting regressor:**

Have implemented GBR to predict the power generation based on the provide whether data.

Have generated a feature Hour and minute to categorise the data based on hour and minute of the day.

Couldn’t hyper tune the model due to having low configuration machine.

Train error and test error of the model is as below:

Measuring criteria is sqrt of mean squared error:

Train Error 156.12428115665003 Test Error 262.13128953380277

**Time series analysis:**

Provided data is meeting most of the assumptions of VAR model,

But wasn’t able to complete the model.