

Machinehack-Renew-Power-Hiring-Hackathon

Create a model to get an ideally functioning turbine's expected rotor bearing temperature.

- Basic exploratory data analysis using pandas, matplotlib, seaborn packages.
- Data pre-processing
 - Feature Engineering
 - tip ratio calculation
 - angle to cardinal direction conversion
 - check raw active power greater than converter active power
 - check raw reactive power greater than converter reactive power
 - check nacelle inside temperature greater than nacelle temperature
 - calculate apparent power
 - group by numerical summary for numerical columns
- The final features for the model
 - 0_active_power_calculated_by_converter
 - 1_active_power_raw
 - 2_ambient_temperature

- 3_generator_speed
- 4_generator_winding_temp_max
- 5_grid_power10min_average
- 6_nc1_inside_temp
- 7_nacelle_temp
- 8_reactice_power_calculated_by_converter
- 9_reactive_power
- 10_wind_direction_raw
- 11_wind_speed_raw
- 12_wind_speed_turbulence
- 13_turbine_id
- 14_tip_ratio
- 15_cardinal_direction
- 16_active-convert_pwr_grt_apwer
- 17_reactive-convert_pwr_grt_reapwer
- 18_in_tmp_grt_out_tmp
- 19_apparent_power
- 20_active_power_calculated_by_converter_min
- 21_active_power_calculated_by_converter_median
- 22_active_power_calculated_by_converter_max
- 23_active_power_calculated_by_converter_mean
- 24_active_power_raw_min
- 25_active_power_raw_median
- 26_active_power_raw_max
- 27_active_power_raw_mean

- 28_ambient_temperature_min
- 29_ambient_temperature_median
- 30_ambient_temperature_max
- 31_ambient_temperature_mean
- 32_generator_speed_min
- 33_generator_speed_median
- 34_generator_speed_max
- 35_generator_speed_mean
- 36_generator_winding_temp_max_min
- 37_generator_winding_temp_max_median
- 38_generator_winding_temp_max_max
- 39_generator_winding_temp_max_mean
- 40_grid_power10min_average_min
- 41_grid_power10min_average_median
- 42_grid_power10min_average_max
- 43_grid_power10min_average_mean
- 44_nc1_inside_temp_min
- 45_nc1_inside_temp_median
- 46_nc1_inside_temp_max
- 47_nc1_inside_temp_mean
- 48_nacelle_temp_min
- 49_nacelle_temp_median
- 50_nacelle_temp_max
- 51_nacelle_temp_mean
- 52_reactice_power_calculated_by_converter_min

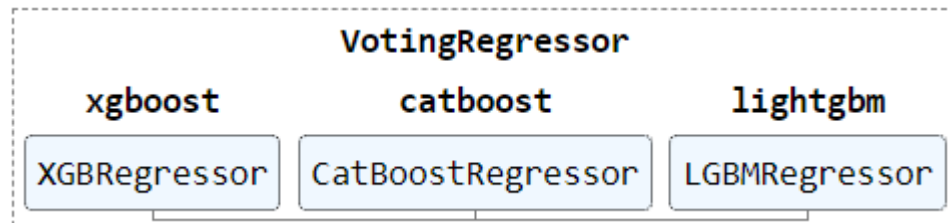
- 53_reactice_power_calculated_by_converter_median
- 54_reactice_power_calculated_by_converter_max
- 55_reactice_power_calculated_by_converter_mean
- 56_reactive_power_min
- 57_reactive_power_median
- 58_reactive_power_max
- 59_reactive_power_mean
- 60_wind_direction_raw_min
- 61_wind_direction_raw_median
- 62_wind_direction_raw_max
- 63_wind_direction_raw_mean
- 64_wind_speed_raw_min
- 65_wind_speed_raw_median
- 66_wind_speed_raw_max
- 67_wind_speed_raw_mean
- 68_wind_speed_turbulence_min
- 69_wind_speed_turbulence_median
- 70_wind_speed_turbulence_max
- 71_wind_speed_turbulence_mean
- 72_tip_ratio_min
- 73_tip_ratio_median
- 74_tip_ratio_max
- 75_tip_ratio_mean
- 76_apparent_power_min
- 77_apparent_power_median

- 78_apparent_power_max
 - 79_apparent_power_mean
 - 80_target
- By using pycaret regressor compared more than one regressor model with 5-fold cross-validation and evaluated by the MAPE.

	Model	MAE	MSE	RMSE	R2	RMSLE	MAPE	TT (Sec)
xgboost	Extreme Gradient Boosting	0.8855	1.7207	1.3117	0.7486	0.0273	0.0190	3.8520
catboost	CatBoost Regressor	0.8869	1.7129	1.3088	0.7498	0.0272	0.0191	26.2020
lightgbm	Light Gradient Boosting Machine	0.9480	1.9353	1.3911	0.7173	0.0289	0.0204	9.0580
br	Bayesian Ridge	1.1740	2.8389	1.6849	0.5853	0.0347	0.0251	9.8880
omp	Orthogonal Matching Pursuit	1.2666	3.1716	1.7809	0.5367	0.0368	0.0271	0.8300
en	Elastic Net	1.3312	3.6257	1.9041	0.4703	0.0393	0.0285	0.6480
lasso	Lasso Regression	1.3675	3.8313	1.9574	0.4403	0.0404	0.0293	0.6000
knn	K Neighbors Regressor	1.3952	3.7906	1.9469	0.4463	0.0407	0.0301	3.3800
ridge	Ridge Regression	1.5145	4.1408	2.0324	0.3949	0.0423	0.0325	0.4040
lr	Linear Regression	1.5544	4.4717	2.0838	0.3470	0.0434	0.0334	1.0160
llar	Lasso Least Angle Regression	1.9015	6.8458	2.6164	-0.0000	0.0547	0.0411	0.8840
dummy	Dummy Regressor	1.9015	6.8458	2.6164	-0.0000	0.0547	0.0411	0.2740

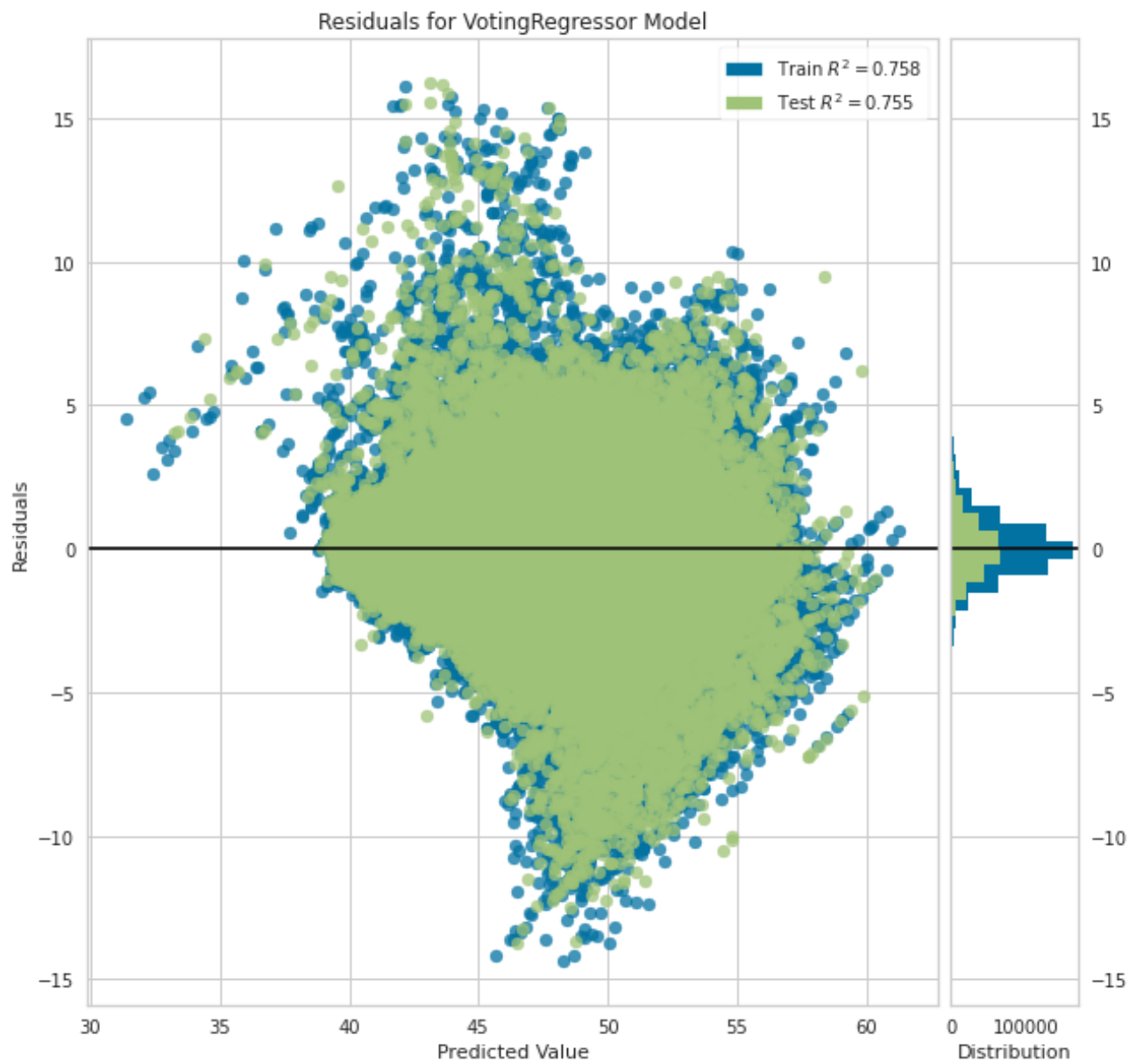
	Model	MAE	MSE	RMSE	R2	RMSLE	MAPE	TT (Sec)
lar	Least Angle Regression	1360.8	457.00	3022.116	-66.000	9.6135	29.3006	0.8780

Blended the top 3 model

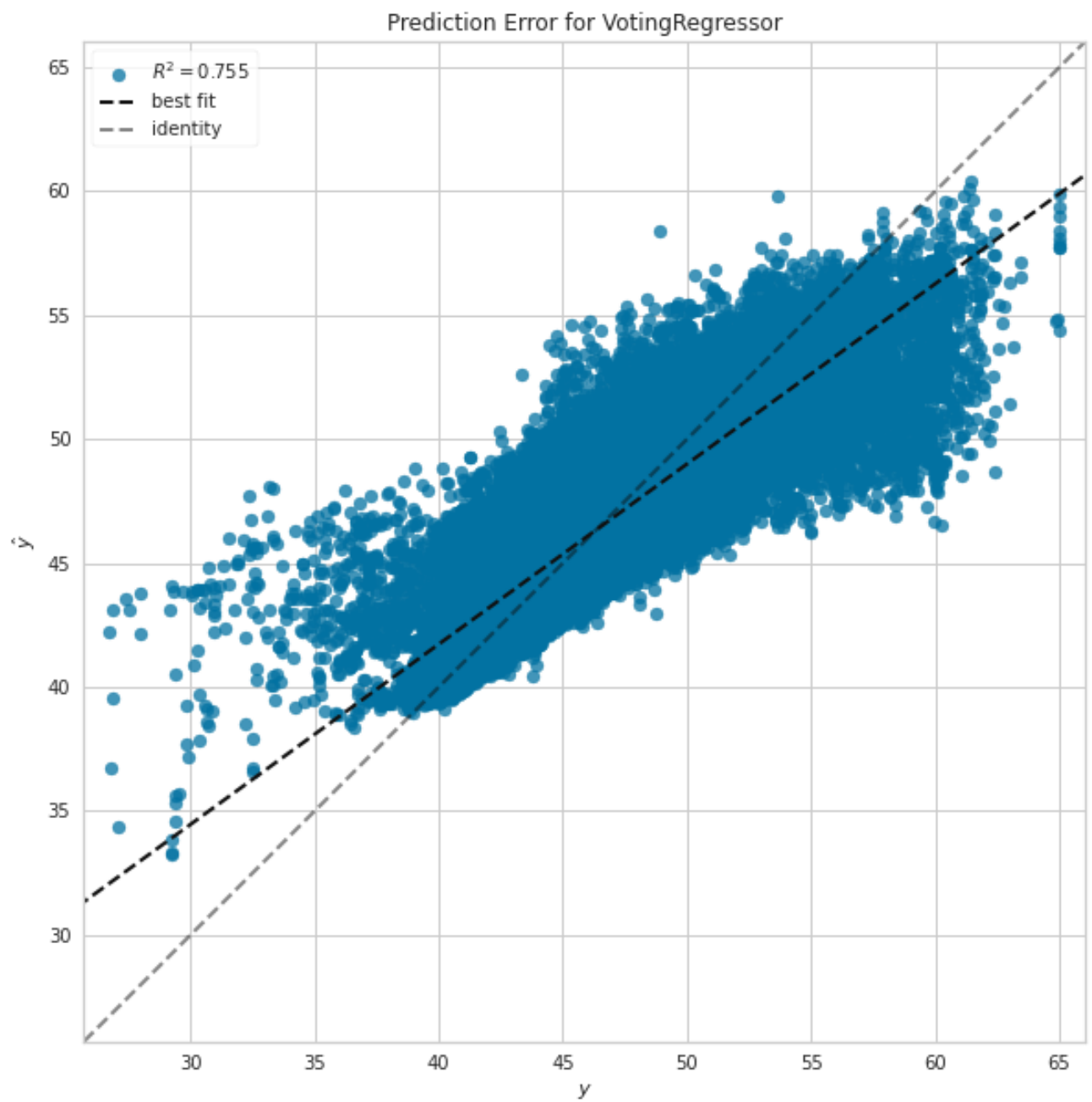


	MAE	MSE	RMSE	R2	RMSLE	MAPE
Fold						
0	0.8891	1.7142	1.3093	0.7482	0.0272	0.0191
1	0.8901	1.7258	1.3137	0.7478	0.0274	0.0191
2	0.8913	1.7385	1.3185	0.7440	0.0275	0.0192
3	0.8884	1.7293	1.3150	0.7480	0.0273	0.0191
4	0.8964	1.7481	1.3222	0.7476	0.0275	0.0192
Mean	0.8910	1.7312	1.3157	0.7471	0.0274	0.0191
Std	0.0029	0.0115	0.0044	0.0016	0.0001	0.0001

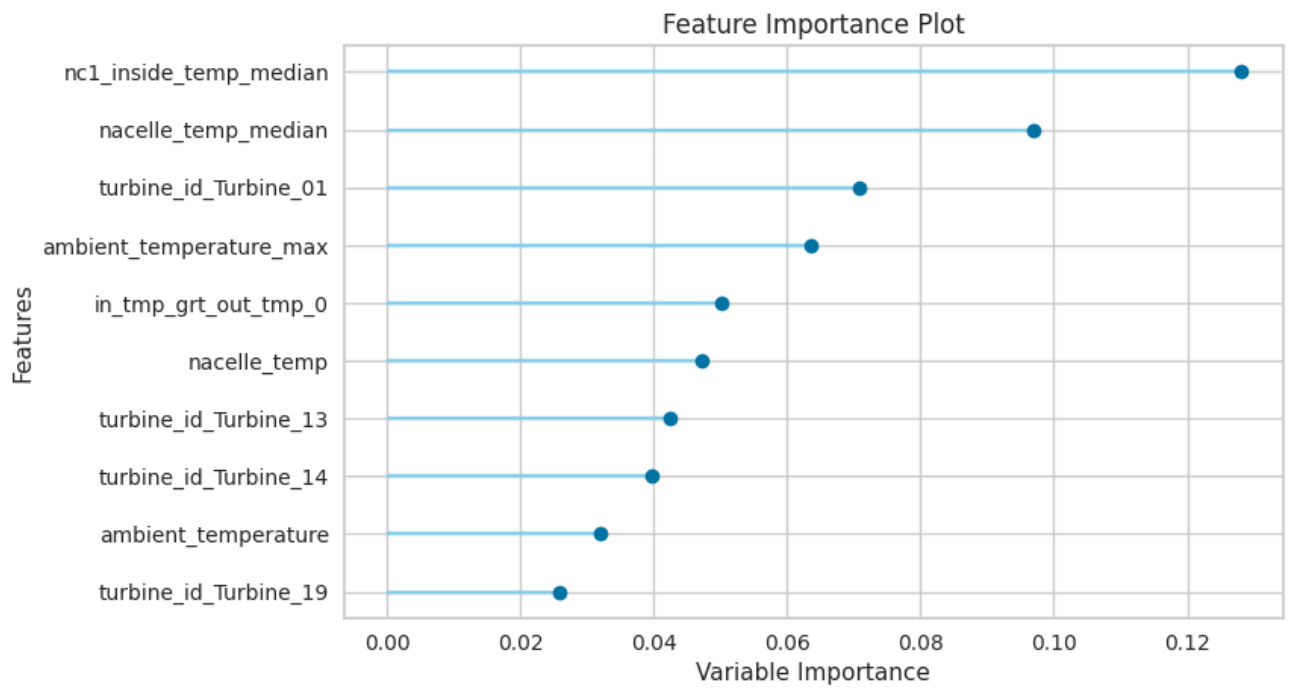
- Xgboost Regressor Residual Plot



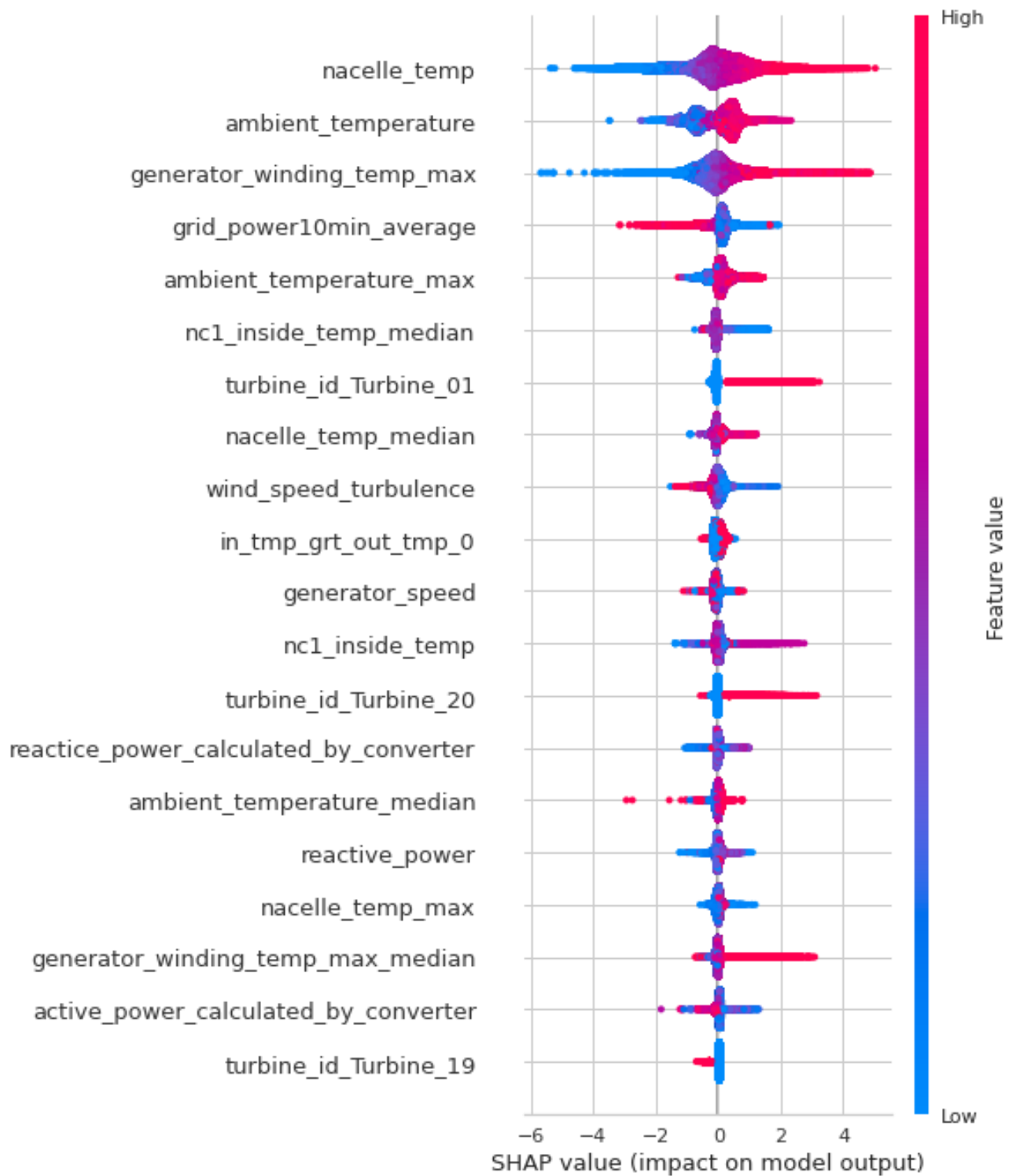
- Xgboost Regressor Prediction Error Plot



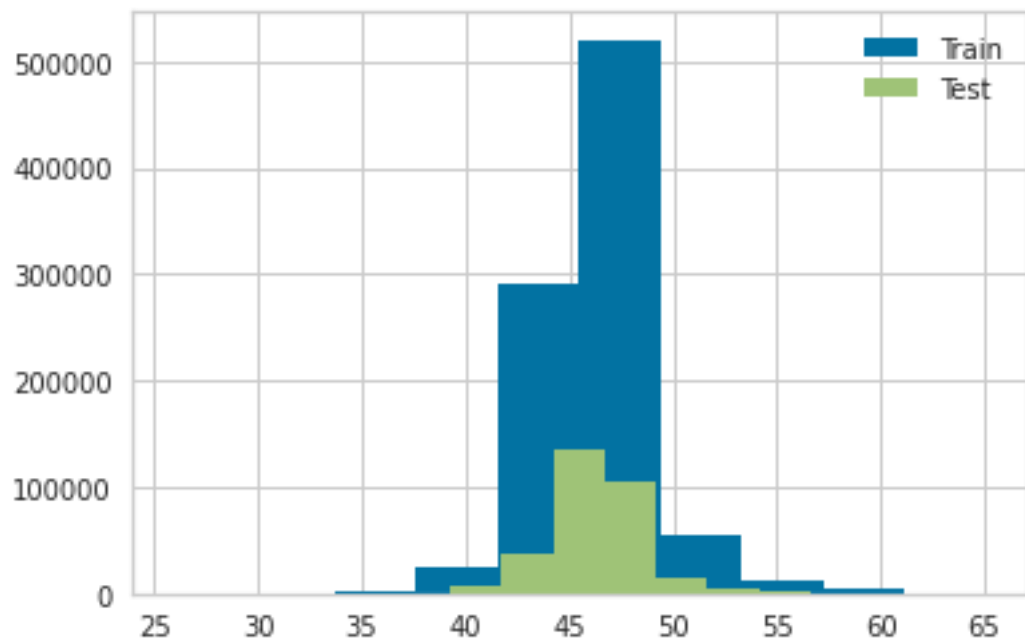
- Xgboost Model Feature Importance Plot



- SHAP - Xgboost Model Feature Importance Plot



- Rotor bearing temperature of train and test data



- Final score is 0.01852