Seoul Bike Demand Pridiction

Candidate Name: Sourav Chowdhury
Email: sourav.20497@gmail.com

Contributor Roles:

- Importing libraries
- Exploratory Data Analysis
- Extracting features from date
- Data Preparation
- Model Building Using Hyper Parameter Tuning
    1. Linear Regression
    2. Lasso Regression
    3. Ridge Regression
    4. ElasticNet Regression
    5. DecisionTree Regressor
    6. RandomForest Regressor
    7. Gradient Boost
    8. Xg Boost
- Model Evaluation through Metrics :
    1. Mean Squared Error
    2. Root Mean Squared Error
    3. R-Squared
    4. Adjusted R-Squared
- Table for test and train Data.
- Uploading on Github
- Making Technical Document
- Presentation Slides.

Seoul Bike Sharing Demand Prediction is a Regression based bike demand prediction project for the rental bikes, the model is build on the data provided consisting of bike rented in each hour during the year 2017 and 2018. The data contains 8760 non null entries with 14 columns.

As the first step, perform date wrangling over the raw data and select those columns which are important for analysis. Once the data is prepared the machine learning algorithms were implemented are Linear Regression, Lasso regression, Ridge Regression, ElasticNet Regression, DecisionTree Regressor, RandomForest Regressor, Gradient Boost, Xg Boost. After analysis on every algorithm the Evaluation metrics was calculated for each one of them with best parameters tuned using hyperparameter Tuning.

In the Analysis the more Focus was on the R- Squared Evaluation metrics , Tried to reach maximum value of it. And also the line Graph for Actual and Predicted Test results were Plotted for each of the Algorithm , along with that, Data contained Heteroscedasticity, So to minimize that , for each Case the Scatter graph was plotted.

At the last a table of Model Evaluation Summary for Train and Test dataset was made. Where the evaluation parameter values were mentioned.
Following is the Table for test dataset.

SL NO	MODEL_NAME	Test MSE	Test RMSE	Test R^2	Test Adjusted R^2
1	Linear Regression	169871.70934024057	412.15495792267325	0.5624656403341544	0.5584307413401177
2	Lasso Regression	169871.72597325977	412.1549781007865	0.5624655974928983	0.5584306981037839
3	Ridge Regression	169871.77465071727	412.1550371531534	0.5624654721155751	0.5584305715702432
4	ElasticNet Regression	183719.51037262668	428.6251396880807	0.5267982017652659	0.5224343811475394
5	DecisionTree Regressor	91626.47632392391	302.6986559664973	0.7639999514779179	0.7618235821543713
6	RandomForest Regressor	52363.04539891553	228.8297301464902	0.8651297992599828	0.863886039483706
7	Gradient Boost	67959.97028150507	260.69133142761973	0.8249571856578451	0.823342957975151
8	Xg Boost	68250.19807949613	261.2473886558412	0.8242096530978654	0.8225885317431483

The best Algorithm found to be RandomForest Regressor which gives 86.51% of R-Squared Value.

Thus with this analysis a Bike prediction model was made to give an accurate no. of Bikes needed as to establish a Rental Bike Store at a given place can be calculated.

GitHub link: https://github.com/SrvPioneer/Supervised-Machine-Learning-Regression-Bike-Sharing-Demand-Prediction-

Drive Link:
https://drive.google.com/drive/u/1/folders/1HXLf_uxalmWN4Lf7_krXQSbP-4q59PNu