**Bike Sharing Prediction**

The goal of this project is to combine the historical bike usage patterns with the weather data to forecast bike rental demand. The data set consists of hourly rental data spanning two years.

Exploratory Data Analysis is done on the dataset and compare the target variable with the other variables to find the distribution of graph. We look for null values which were not found and outliers and appropriately modify them by z-score. We also perform correlation analysis to extract out the important and relevant feature from dataset and later perform train test split to train the model.

The main objective is to build a predictive model, which could help to train a model to predict the number of bike rentals of the year given the weather conditions. This would in turn help to predicting quickly and efficiently.

Bike sharing systems are a means of renting bicycles where the process of obtaining membership, rental, and bike return is automated via a network of kiosk locations throughout a city. Using these systems, people are able rent a bike from a one location and return it to a different place on an as-needed basis.

The goal of this project is to combine the historical bike usage patterns with the weather data to forecast bike rental demand. The data set consists of hourly rental data spanning two years.

**Contributor role** :- **Kartik Pisudde**

* Preview Data
* Check total number of entries and column types
* Check the null values
* Plot distribution of numeric data
* Building the model
* Linear regression
* Lasso regression
* Ridge regression
* Elastic net
* Polynomial regression
* KNN and Tree based model
* Random forest
* Gradient boosting
* Extreme gradient boosting
* Catboost
* LightBGM
* Model interpretation
* Model Explaninbility

**Conclusion**

1. In holiday or non-working days there is demands in rented bikes.

2. There is a surge of high demand in the morning 8AM and in evening 6PM as the people might be going to their work at morning 8AM and returing from their work at the evening 6PM.

3. People prefered more rented bikes in the morning than the evening.

4. When the rainfall was less, people have booked more bikes except some few cases.

5. The Temperature, Hour & Humidity are the most important features that positively drive the total rented bikes count.

6.After performing the various models the lightGBM and Catboost found to be the best model that can be used for the Bike Sharing Demand Prediction since the performance metrics (mse,rmse) shows lower and (r2,adjusted\_r2) shows a higher value for the lightGBM and Catboost models !

7. We can use either lightGBM or catboost model for the bike rental stations.

**GitHub Link:**

[kartik-pisudde/Seoul-Bike-Sharing-Demand-Prediction (github.com)](https://github.com/kartik-pisudde/Seoul-Bike-Sharing-Demand-Prediction)

**Drive Link:**

https://drive.google.com/drive/folders/1h5owrqAk1oRluU7UMnWDRyUh5vdgW2MK?usp=drive\_link