Capstone Project Submission

Team Member's Name, Email and Contribution:

1) Chandan Kumar Raxit
E-mail: chandanraxit2468@gmail.com
   - Data Understanding
   - Feature Analysis
   - Feature Engineering
   - Linear Regression modeling
   - Random forest
   - Gradient Boosting
   - Hyperparameter tuning
   - Decision Tree
   - Elastic Net Regressor

2) Deepak Kumar Jena
E-mail: jdeepakumar986@gmail.com
   - Multivariate Analysis
   - Decision Tree
   - Elastic Net Regressor
   - Research Analytics
   - Lasso Regression
   - Ridge Regression
   - Research Analytics
   - Research Analytics

The contents of the data came from a city called Seoul. A bike-sharing system is a service in which bikes are made available for shared use to individuals on a short-term basis for a price or free. Many bike share systems allow people to borrow a bike from a "dock" which is usually computer-controlled wherein the user enters the payment information, and the system unlocks it. This bike can then be returned to another dock belonging to the same system. The data had variables such as date, hour, temperature, humidity, wind-speed, visibility, dew point temperature, solar radiation, rainfall, snowfall, seasons, holiday, functioning day and rented bike count.

Currently rental bikes are introduced in many urban cities for the enhancement of mobility comfort. The purpose of this movement is to modernize cities and encourage people to head to a green world. Let's take the examples of Paris in 2007, where "veils" were introduced and Amsterdam, where there are more bikes than cars. The goal is to facilitate the commute in the Seoul and reduce the amount of cars and the pollution. Indeed, the development of the way to commute reduced the use of cars to go to work and visit the city

It is important to make the rental bike available and accessible to the public, as it provides many alternatives to commuters in metropolises. There are a lot of advantages to bike rents, it is convenient because it permits people not to keep the bike all day long, whether it is at work or at school. Furthermore it is the healthiest way to travel and it has many environmental benefits..

The model performed well in this The problem statement was to build a machine learning model that could predict the rented bikes count required for an hour, given other variables. The first step in the exercise involved exploratory data analysis where we tried to dig insights from the data in hand. It included univariate and multivariate analysis in which we identified certain trends, relationships, correlation and found out the features that had some impact on our dependent variable. The second step was to clean the data and perform modifications. We checked for missing values and outliers and removed irrelevant features. We also encoded the categorical variables. The third step was to try various machine learning algorithms on our split and standardized data. We tried different algorithms namely; Linear regression, Random forest and

XGBoost. We did hyperparameter tuning and evaluated the performance of each model using various metrics. The best performance was given by the Gradient boosting and Random Forest model where the R2_score for training and test set was 0.95 and 0.92 respectively. case but as the data is time dependent, values of temperature, wind-speed, solar radiation etc. will not always be consistent. Therefore, there will be scenarios where the model might not perform well. As Machine learning is an exponentially evolving field, we will have to be prepared for all contingencies and also keep checking our model from time to time

Please paste the GitHub Repo link.
GitHub Link: - https://github.com/Chandanraxit/Bike_Sharing_Demand_Prediction#bike_sharing_demand_prediction
Please write a short summary of your Capstone project and its components. Describe the problem statement, your approaches and your conclusions. (200-400 words)