Capstone Project Submission

Instructions:

i) Please fill in all the required information.

ii) Avoid grammatical errors.

Team Member's Name, Email and Contribution:
Chandrashekhar Awate Email – shekharawate420@gmail.com Contribution – Done Project single Handedly
Please paste the GitHub Repo link.
Github Link: https://github.com/Chandr25/AlmabetterRegressionproject
Please write a short summary of your Capstone project and its components. Describe the problem statement, your approaches and your conclusions. (200-400 words) **Currently Rental bikes are introduced in many urban cities for the enhancement of mobility comfort. It is important to make the rental bike available and accessible to the public at the right time as it lessens the waiting time. Eventually, providing the city with a stable supply of rental bikes becomes a major concern. The crucial part is the prediction of bike count required at each hour for the stable supply of rental bikes. From the Exploratory Data Analysis we got some insights which are listed following. The Demand changes mainly with these variables.** **Season** **In the season column, we are able to understand that the demand is low in the winter season.** **Holiday** **In the Holiday column, The demand is low during holidays, but in no holidays the demand is high, it may be because people use bikes to go to their work.** **Functioning Day** **In the Functioning Day column, If there is no Functioning Day then there is no demand** **Days of week** **In the Days of week column, We can observe from this column that the pattern of weekdays and weekends is different, in the weekend the demand becomes high in the afternoon. While the demand for office**

timings is high during weekdays, we can further change this column to weekdays and weekends.

month

In the month column, We can clearly see that the demand is low in December January & Febuary, It is cold in these months and we have already seen in season column that demand is less in winters.

year

The demand was less in 2017 and higher in 2018, it may be because it was new in 2017 and people did not know much about it.

After EDA the preprocessing done for model development after that the data was exposed to linear regression, polynomial regression and Random forest regressor out of which polynomial became the more suitable one with r-squared score of 88% followed by 77% and 79% for linear regression and Random forest regressor respectively. We also used regularization techniques of regression but the results were more or less similar to linear regression one.