Capstone Project Submission

Team Member's Name, Email and Contribution:

Team	Name	Email	Contribution
Individual	Deepak	Dshekhawat30@gmail.com	From EDA to model selection and implementation and final conclusion

Please paste the GitHub Repo link.

Github Link:-

https://github.com/deepakshekhawat1209/BikeSharingDemand/

Please write a short summary of your Capstone project and its components. Describe the problem statement, your approaches and your conclusions. (200-400 words)

Problem Statement

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.

Conclusions:

● We started with loading the data, then we did Exploratory Data Analysis (EDA), null values treatment, feature selection, encoding of categorical columns, and then model building. In all of these models, our accuracy ranges from 77% to 93%, which can be said to be good.

From the exploratory analysis, we saw how clear and sunny weather attracts more riders as compared to rainy and snow weather. We also found that the demand is maximum during morning (7-9 AM) and evening (4-7 PM) travel hours. We also saw how the demand of bikes drop in winter as compared to other seasons and how it has raised in the span of just one year. With the developed model using all these variables, we could have strategically maintained the supply depending on the 30-day weather forecast. And this could have helped our company to reduce the gap between the demand and supply leading to a higher revenue and a better profit margin. This also gives the company scope to change the revenue model by introducing price surge during peak hours based on the forecast of demand.

● After fitting so many regressor models into the dataset, if I consider low RMSE and high R2 score then I got the best results from XGB Regressor and then from random forest and Gradient Booster regressor.

Therefore, We can use either XGB regressor or random forest model for the bike rental stations.