A Capstone Project Submission Summary

Instructions:

i) Please fill in all the required information.

ii) Avoid grammatical errors.

Team Member's Name, Email and Contribution:
AKASH GAWANDE gawandeakash555@gmail.com **BIKE SHARING DEMAND PRIDICTION**
Please paste the GitHub Repo link. https://github.com/gawandeakash/Bike-Sharing-Demand-Pridiction-Almabetter.git
Please write a short summary of your Capstone project and its components. Describe the problem statement, your approaches and your conclusions. (200-400 words)

BIKE SHARING DEMAND PRIDICTION_:-

i. Currently Rental bikes are introduced in many urban cities for the enhancement of mobility comfort. The client is Seoul Bike, which participates in a bike share program in Seoul, South Korea. An accurate prediction of bike count is critical to the success of the Seoul bike share program. 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
ii. In the comparison between "hour vs rented count of bikes" we can clearly notice a high demand in the rush hour of 8:00 am to 9:00 pm.
iii. In the comparison between "holiday non-holiday vs rented count of bikes" we get the notion of high demand of bikes during non-holiday i.e., working days compared to holidays i.e., non-working days Demand of Rented Bike gradually decreases with increase in rainfall.
iv. Same pattern of decrease in demand is observed with the increase in snowfall. In the Winter Months the Demand decreases i.e, December, January, February, however demand spikes up in summer months i,e May, June, July. Modeling Conclusions We used 8 Regression Models to predict the bike rental count at any hour of the day - 'Linear','Lasso','Ridge','Elasticnet','Decision_ Tree','Random_Forest','Gradient_Boosting',' Xtreme_GB'.
v. Using the predictions made by these level 1 individual models as features, we trained 4 level 2 stacking algorithms (Linear Regression, Random Forest Gradient Boost and Xtreme Gradient Boosting) to make more refined predictions.
vi. Below is a summary of the model performances Of all the models, we found a simple XGBoost Model providing the best/lowest RMSE score and the adjusted_r2 of 99% which made the model deployable