Capstone Project Submission

Instructions:

i) Please fll in all the required information.

ii) Avoid grammatical errors.

Team Member's Name, Email and Contribution:
1. **ANAS MUSTAFA:** **Email: Mustafaanas84464@gmail.com** **Contribution:** • **Data Wrangling** • **Data Cleaning** • **EDA** • **Total rented bikes in 2017 and 2018** • **Effects of numerical features in rented bike count** • **Finding correlation between different variables** • **Linear Regression** • **Random Forest Regression** • **XGboost Regression** • **Important Features** 2. **SARTHAK RASTOGI:** **Email:sartakrastogi1@gmail.com** **Contrinution:** • **Data Cleaning** • **Checking distribution of our dependent variable** • **Exploring numerical Variables** • **Average bike count per hour** • **Ridge Regression** • **Elastic Net Regression** • **Gradient Boosting Regression** 3. **CHETAN RAJPUT:** • **Data Wrangling** • **Data Cleaning** • **Total rented bikes on monthly basis** • **Rented bikes count in different seasons** • **Exploring categorical Variables** • **Finding Correlation of dependent variables with other variables** • **Lasso Regression** • **Decision Tree Regression** • **Hyperparameter Tuning**
Please paste the GitHub Repo link.
Github Link:- https://github.com/sarthakR1/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)

We had finished our capstone project on seoul bicycle share expectation. As more number of leased bicycles are being utilized in the urban communities these days, it becomes significant for the organization to foresee the quantity of required rental bicycles expected across a day so that no interest supply hole would be created for rental bicycles. This undertaking targets giving fundamental answer for anticipate the rental bicycles request utilizing AI calculations with the goal that every one of the partners of the business can be fulfilled.

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.

The present scenario is about how good is the customer service is in any industry as the number of options at the customer's disposal are unlimited. So, it becomes extremely important to make sure that the customers will not be made to wait for the rental bikes. It would also not be practical to keep a lot of bikes even when the demand is low. Hence , with the help of machine learning, this project aims at predicting the rental bike demand so that no problems arise.

Steps:

The initial step of our venture is playing out the EDA interaction on the dataset with the goal that we can get the thought regarding the dataset for example the quantity of factors, the information sort of the factors envision the dataset for better comprehension and conclude the appropriate techniques and calculations that could deliver wanted outcome.In EDA process we track down the kind of dataset and choose the methodology, in this venture the preprocessing steps would eliminating the accentuations, stopwords , create count vectorizer and report term

lattice which would help in developing the model. After the information preprocessing is done then the information will be fit to be squeezed into AI models.For current issue articulation theme displaying approach would be appropriate .In point demonstrating, a subject is characterized by a group of words with each word in the bunch having a likelihood of event for the given subject, and various subjects have their particular groups of words alongside comparing probabilities.

Conclusion:

The undertaking reaches a conclusion now. Starting with stacking the dataset, up to this point we have done EDA, pre-handling the information, Label encoding, Scaling the information, dividing the information into train and test information, applying different AI calculations followed by hyper boundary tuning. We executed 8 M.L. models. Subsequent to contrasting the mean square mistake and mean root square blunder of the multitude of models, XGBoost has least mean square mistake and root mean square mistake. XGBoost has the most elevated exactness of 91.9% among all calculations. In this way, We can reason that XGBoost is the best model to anticipate leased bicycle count. The quantity of business hours of the day and the interest for leased bicycles were generally related and It checks out moreover. Largest number of bicycles leased at the eighteenth hour of day. All out number of bicycle counts expanded when there was ideal temperature. Along these lines, this can be a significant variable in foreseeing hidden examples of leased bicycle count.