**Capstone Project Submission**

**Instructions:**

i) Please fill in all the required information.

ii) Avoid grammatical errors.

|  |
| --- |
| **Team Member’s Name, Email and Contribution:** |
| 1. Name – Rushikesh Arjun Mane   Email -  [rushikesham9094@gmail.com](mailto:rushikesham9094@gmail.com)  Role –   * Understand the data * Manipulate and clean the data * Data Wrangling and Feature Engineering * Analyse the data using various techniques of EDA * Preparation of Data for Model Building * Model Selection and Evaluation * Find insight form the output for writing effective conclusions. * Research for the whole project. |
| **Please paste the GitHub Repo link.** |
| Github Link:- |
| **Please write a short summary of your Capstone project and its components. Describe the problem statement, your approaches and your conclusions. (200-400 words)**         In recent days, public rental bike sharing is becoming popular because of is increased comfortableness and environmental sustainability. A bike sharing system provides peoples a beneficial and effective way of transportation. So, 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. Because bike use is rising constantly, understanding the system demand in prediction is important to boost the operating system readiness. Data used include Seoul Bike and Capital Bikeshare program data. Data have weather data associated with it for each hour. The crucial part is the prediction of bike count required at each hour for the stable supply of rental bikes. This project discusses the models for hourly rental bike demand prediction. Data used include weather information (Temperature, Humidity, Windspeed, Visibility, Dewpoint, Solar radiation, Snowfall, Rainfall), the number of bikes rented per hour and date information. We understand the problem statement then we decided the work flow. After collecting and understanding the data we done data wrangling and feature engineering for further EDA. In Exploratory Data Analysis we visualize the data and get the important insights for further modelling. In this EDA we understand the relation between the 21 attributes and our target variable.  For this dataset we are using various regression (Linear Regression, Lasso Regression (regularized regression), Ridge Regression (regularized regression),Decision Tree regression, Random Forest regression, Gradient Boosting regression) models were train with optimize hyper-parameters using a repeated cross validation approach and testing set is used for evaluation. . Multiple evaluation indices such as 𝑅2 , Root Mean Square error are used to measure the prediction performance of the regression models. The performance of the model is varied with the time interval used in transforming data.  The Gradient Boosting regression (GridsearchCV) and random Forest (GridsearchCV) are both fitted models with Good R2. And this is very convincing insight form the coclusions.    . |
|  |