**Capstone Project Submission**

**Seoul Bike Sharing Demand Prediction**

**Instructions:**

i) Please fill in all the required information.

ii) Avoid grammatical errors.

|  |
| --- |
| **Team Member’s Name, Email and Contribution:** |
| Sachin Yallapurkar  [sachin.almabetter@gmail.com](mailto:sachin.almabetter@gmail.com)  Individual project |
| **Please paste the GitHub Repo link.** |
| Github Link:- <https://github.com/SachinYallapurkar/-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)** |
| **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.  **Approach:**  Here first we imported a data set and performed EDA where we got valuable insights and further we Encoded the Categorical Columns, Feature scaling and fitting into the models. At first we tried with basic linear regression and also with Lasso regularization technique but soon realized we will need a much more complex model and so we then used a Decision tree Regressor, XGB Model, Random Forest Regressor and LightGBM compared the results.  **Conclusion:**  The analysis is done with Seoul Bike data. Four regression techniques Linear Regression, Decision Tree, XG Boosting and Random Forest are used to predict the trip duration. This statistical data analysis shows interesting outcomes in prediction methods and also in an exploratory analysis.  **The experimental results show that:**   * Most numbers of Bikes were rented in summer, followed by autumn, spring, and winter. May-July is the peak Bike renting Season, and Dec-Feb is the least preferred month for bike renting. * Majority of the client in the bike rental sector belongs to the Working class. This is evident from EDA analysis where bike demand is more on weekdays, working days in Seoul. * Temperature of 20-30 Degrees, evening time 4 pm- 8 pm, Humidity between 40%-60% are the most favourable parameters where the Bike demand is at its peak. * Temperature, Hour of the day, solar radiation, and Humidity are major driving factors for the Bike rent demand.   Feature and Labels had a weak linear relationship; hence the prediction from the linear model was very low. Best predictions are obtained with a **LightGBM** model with an r2\_score of **0.894** and RMSE of **203.91 .** |