Project Summary

# Capstone Project 2

## Bike Sharing Demand Prediction (ML-Regression)

|  |
| --- |
| Details and Contribution:-  Name: LALIT  Email id: nagar9210@gmail.com |

* Made the Google Colab notebook with the help of Google Drive data connectivity, data cleaning, data manipulation, and EDA Visualization.
* Development of the business objective and problem statement along with the relevant questionnaire for the business objective.
* Solved all the questions and also created visualization in the Google Colab.
* Created the design and contents of Technical documentation and ensured that everything is covered in the documentation.
* Created the project summary to ensure all the points were covered.

|  |
| --- |
| GitHub Link: **https://github.com/LALIT9210/ML-Regression-Bike-Sharing-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):**  Seoul city locals can rent bicycles for a charge and for a set period of time. However, the demand for motorbikes varies according to various variables, and they cannot keep up. Our approach aims to estimate demand for rented bikes at  any given time, considering all factors that will assist in regulating the flow of motorbikes. Rental bikes are already available in numerous urban locations to increase transit convenience. It is critical that the rental bikes are accessible and available to the general public at the proper time since this eliminates waiting. Maintaining a consistent supply of rental bikes for the city eventually emerges as a major issue. The ability to estimate the number of bikes needed each hour is essential. We are going to explore the machine learning algorithms which function effectively with the data and the factors that have a significant impact on the demand for rented bikes. In this project, I dig deep into the datasets by doing an exploratory data analysis to see certain patterns that can be helpful for businesses in their profits. After that, I go into the machine learning part and implement the split data which is (70:30) ratio into the different regressors like Linear regressor, Lasso, Ridge, and so on and find out the best model accuracy.  I began by in-depth cleaning the dataset. After that, I did a generalized analysis to get numerous insights. In this dataset, we see several data like Date, Rented Bike Count, Hour, Temperature, Humidity(%), wind speed, visibility, Dew Point, Solar radiation, rainfall, snowfall, seasons, holiday, and Functional days. After analyzing all the insights we find some correlation between several things and causes and also solutions for that which will help in business problems. We also analyze the regression MSE, RMSE, R2, R2 adjusted, etc. values to predict the accuracy of the model and compare it with the help of the chart and we can see that Random Forest Regressor gives the highest accuracy |