

## Capstone Project Submission

### Instructions:

- i) Please fill in all the required information.
- ii) Avoid grammatical errors.

### **Team Member's Name, Email and Contribution:**

**1. Sanjay Ramkishan Verma**

**Email-id:** - [shankyverma1998@gmail.com](mailto:shankyverma1998@gmail.com)

**Contribution-**

- Checking the presence of NAN values using missing values mno
- matrix.
- Feature engineering
- Best fitting line between Target & features.
- Correlation matrix
- Normalizing the features of the dataset.
- One hot coding technique on categorical variables.
- Model developed for Linear Regression, Lasso Regression and
- Random Forest regressors.

**2. Poonam Dattu Shevkar**

**Email-id:** - [dattupoonam@gmail.com](mailto:dattupoonam@gmail.com)

**Contribution -**

- Histogram plot for dependent & independent variables.
- Feature Distribution.
- Box plot and Scatter plot with features and Rented bike count.
- Outliers treatment with imputation of mean.
- Multicollinearity to evaluate the VIF factor of the numerical
- features.
- Model fitting by splitting the dataset into train and test data.
- Model developed for Decision Tree and Gradient Boosting
- Regressors.
- Hyperparameter tuning on Gradient Boosting Regressors.

**GitHub Link:** - <https://github.com/VermaSanjayrk/Seoul-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)**

Bicycle rental or bike hire businesses rent out bicycles for a shorter duration of time, usually for a few hours. Basically, most of these rentals are provided by the bike shops as a sideline to the main business of sales and services. This model allows a customer to get a bike from any station for free for up to half an hour which is enough for most of the customers to travel to a destination where they can drop this bike at the respective stations. After completion of the first half hour tour the actual bike sharing starts in order to encourage the customer to return the bike after the end of each trip and take another bike for the next one.

To enhance the mobility of comfort in the crowded urban cities, the Seoul bike rental data sets a view to predict the demands for rental bikes. For customer satisfaction it is important to make this bike available at the right time at the right place for customer usage. The information contains the data set from the year 2016 to 2017 with 8761 rows and 14 columns respectively.

The main objective of the Seoul bike rental is to protect the number of bikes required at each hour and to maintain the supply of Rental bikes using regression technique.

How to get the useful data without NAN values and observe the summary statistics of the data set, so we perform data wrangling on the given raw data. we have prepared a data set with feature engineering and feature scaling, we also dropped out the unnecessary columns. From the analysis point of view, we have to check outliers and try to overcome it by using outlier treatment using imputation of mean. The data was normalized using the standard deviation.

Moreover, to cross check the multicollinearity in the data set VHF factor is evaluated. Also, one hot coding is used to create a new data frame with dummy variables.

After reviewing all the data, the model is fitted with splitting the data into train and Tests in the ratio 70:30 respectively, standard scalar technique is used to scale the data which helps to sustain the model prediction rate. To train the model we used linear regression, lasso regression, decision tree regressors, random forest regression and gradient

boosting regressor. Cross validation technique is applied for all fitted model We evaluated the gradient boosting regressor with hyperparameter tuning which in returns give the result as best fit and module for or the given data set.

**Drive Link:** [https://drive.google.com/drive/folders/1xaIJKsRTv4dB1Vn7fkq5pAvDa\\_7unn-v?usp=sharing](https://drive.google.com/drive/folders/1xaIJKsRTv4dB1Vn7fkq5pAvDa_7unn-v?usp=sharing)