Capstone Project II Submission

Bike Sharing Demand Prediction

Name, Email and Contribution:

NAME- Rajakumaran S

EMAIL – rajakumaransrs@gmail.com

CONTRIBUTION -
Data Analysis, Data Visualization, Feature Engineering, Fitting Models, Model Explainability and Report Writing.

GitHub Repo link:

https://github.com/Rajakumaran-S/Seoul_Bike_Sharing_Demand_Prediction_Capstone_Project-II

Short Summary of Capstone project- Bike Sharing Demand Prediction

This study focused on predicting the bike sharing demand using given dataset. Regression techniques Linear Regression, Lasso Regression Ridge Regression, Elastic Net, Decision Tree, Random Forest, Gradient Boosting Regressor, XGB Regressor are used to predict. This statistical data analysis shows interesting outcomes in prediction method and also in an exploratory analysis. First of all, we do EDA on the dataset and observe the following.

- Most number of bikes are rented on time, morning at **8** O'clock and Evening 5-6 o'clock.
- Most numbers of Bikes were rented in **Summer**, followed by **Autumn**, **Spring** and **Winter**.
- Most number of bikes are rented on **Working day** instead of holiday.
- Heat map shows Temperature and Dew point temperature is highly correlated.

After the EDA, we do Feature Engineering then after we simultaneously creating 8 different Regression models.

- Splitting Training and Testing Data set.
- Scaling Training and Testing Dataset with Standard Scaling.
- Applying all Regression models with hyperparameter tuning.
- Analysing Importance Features.

Conclusion:

Comparing to all regression model Gradient Boosting Machine and Random Forest Regression Gives Best Results with r2 score.

Gradient Boosting Machine model, with hyperparameter tuning we got r2 score as 91% on training data and 89% on testing dataset.

Random Forest Regressor model, with hyperparameter tuning we got r2 score as 91% on training data and 86% on testing dataset.