Capstone Project Submission

Instructions:

i) Please fill in all the required information.

ii) Avoid grammatical errors.

Team Member's Name, Email and Contribution:

- **Jai Harish S (ppjairam2@gmail.com):**
  - Handling Null Values
  - EDA
  - Finding Skewed Data
  - Transformation
  - Removing Outliers
  - Removing Multicollinearity
  - Conclusion

- **Harish Patil (harishpatil0172@gmail.com):**
  - Relationship b/w dependent and Categorical Value
  - Label Encoding
  - Feature Engineering
  - Conclusion

- **Pranil Satish Thorat (pranilthorat@gmail.com):**
  - Feature Selection
  - Implementing Linear Regression
  - Regression Evaluation Metrics for Linear Regression
  - Implementing Polynomial Regression
  - Conclusion

- **Saransh Srivastava (saranshoffice@gmail.com):**
  - Regression Evaluation Metrics for Polynomial Regression
  - Implementing Random Forest Regressor
  - Regression Evaluation Metrics for RFR
  - Hyperparameter Tuning using Grid Search for RFR & its Regression Evaluation Metrics
  - Comparing Evaluation Metrics for all Models used
  - Conclusion

Please paste the GitHub Repo link.

GitHub Link:
https://github.com/saransh2396/Seoul-Bike-Sharing-Demand-Prediction
Drive link:
https://drive.google.com/drive/u/2/folders/1UL2coEKkv5OStGddU7KNvcHP6DNrepyQ

Please write a short summary of your Capstone project and its components. Describe the problem statement, your approaches and your conclusions. (200-400 words)

Currently Rental Bikes are introduced in many urban cities for the enhancement of mobility comfort. It is important to make the rental bike available & accessible to the public at the right time as it lessens the waiting time. So, our problem statement was to find the Rental Bike Count required at each hour for the stable supply of rental bikes. So, in our dataset there were no null values and after handling outliers we observed that Dew Point Temperature was correlated so after removing that performing Label Encoding, Feature Engineering, selecting the correct features and splitting them for training & testing datasets (80-20), we applied 3 models Linear Regression, Polynomial Regression and Random Forest Regressor. After using these models we selected Random Forest Regressor and performed Tuning in it as it was giving the best results we observed that Rental Bike Count was dependent on Hour, Temperature, Humidity, Functioning Day and if it is Raining outside or not.