**Capstone Project Submission**

**Instructions:**

i) Please fill in all the required information.

ii) Avoid grammatical errors.

| **Team Member’s Name, Email and Contribution:** |
| --- |
| Manmohan Madhu - [mmanmohan1992@gmail.com](mailto:mmanmohan1992@gmail.com) |
| **Please paste the GitHub Repo link.** |
| Github Link:- <https://github.com/manm0han/NYC-Taxi-Trip-Time-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)** |
| **The aim of this project was to use Linear Regression to predict the trip duration for a taxi ride in New York City.**  **The dataset provided contained details of a particular trip, such as pickup and dropoff datetime, coordinates, trip duration etc. The pickup and dropoff datetime were stored as string datatype. After converting this to datetime format, day of the week and month were extracted and analysis done on the rest of the features. Id, passenger count and store and forward features were dropped after learning that they had no effect from the analysis. An additional distance feature was added after calculating the same from the pickup and dropoff coordinates. Outliers were treated and the models were trained.**  **Linear Regression : RMSE - 275.83**  **XGB Regressor : RMSE - 262.45**  **Ridge Regression : RMSE - 275.83**  **The lowest RMSE obtained using the above mentioned three models is 260 seconds which is negligible compared to our range of nearly 2000 seconds. Therefore, we can confidently use this model for predictions which can be used for improving operations.** |