**Self Assessment**

**Self Assessment**

Over the course of the final project, I had plenty of involvement in group discussions. From the initial group meeting on zoom, I started noting down whatever ideas we were brainstorming as a group, while sharing my screen to increase clarity and visibility. I had developed visualizations and dashboard on Tableau for the purpose of exploratory analysis. Some of the visualizations were utilized in the final project. Moreover, I took part in data cleansing. There were plenty of unnecessary columns and rows in the huge dataset that we had chosen, which were negatively impacting the accuracy of our Machine Learning model, so these needed to be dropped. One of the greatest challenge for me was the revision of the module that contained Machine Learning models. Machine learning module was the one I found least interesting, it was a challenge for me to go back and review the entire content of this module in order to apply it in the project.

I really endeavored to find a suitable dataset for the project but could not succeed. The dataset we utilized was found by a different group member. However, I provided good suggestions in group discussions, in terms of what we can use this dataset for, such as, what we can predict from this dataset and how to classify the severity of the crash using the given options. Moreover, in terms of the application of the Machine Learning model, I initiated the application of the M.L model in Jupyter Notebook which led our group towards the right direction. However, the model I selected initially were not the ones that our group decided to use.

**Project and Team Summary**

Our project was based on the bike riding crash data for the state of North Carolina. The dataset contained approximately 1200 rows and 60+ columns, where each row represented a crash. Plenty of unnecessary columns and rows were deleted as part of the cleansing process. Using graphs and visualizations for our exploratory analysis, we were able to find key information from the dataset. Our goal was to predict the level of severity of the crash: whether serious or non-serious based on the given variables. The Machine Learning models we utilized were Random Forest Classifier and Adaboost Classifier. Our accuracy was above 90% for both the models.

The medium of our group communication was zoom while we also created a WhatsApp group to keep each other updated. Other than providing inputs for Machine Learning Model, I had created visualizations on Tableau and shared them on Google slides. Google slides was used for the compilation of the entire presentation. One of our group’s strengths was good collaboration and willing to help each other by providing explanation if someone was having hard time understanding a concept. One thing that was really challenging for me was to attend the group meeting on Sunday mornings. Because I arrive home at 5 am on Sundays, after doing Night shift, attending group meeting at 9 am was an extremely difficult task for me. I should not have agreed upon that timing initially, but I survived. One advice I would like to give to the students of future cohort is to make sure to study the Machine Learning Module well so when time comes for the project, they won’t be spending extra time studying the module.