Asad Khan

November 28, 2022

Dr. Barber

Data606: Data Science Capstone

Personal Contribution Write-up

To whom it may concern,

I hope you are doing well. I am a one of the teammates of *Baltimore Arrest Data* capstone project. These are following part where I contributed myself and co-collaborated with my team specially with Saivijay.

I performed some of the EDA and data visualization while discussing other aspect “how to represent the data” such cleansing, designing, and extracting meaningful information. The point of performing EDA and data visualization was to understand the trends and pattern to get better insight of the data. After discussing with Saivijay we decided that some of the columns were not as relevant as we thought it would be, so we decided to get rid of it.

I worked closely with Sai when it came to perform the task of data modeling. Collaborated with him in performing different task of doing the SparkML. The goal was to find highest accuracy of the model using sklearn library. I believe we used few different pre-made models such as Linear and Logistic regression to see which one fits our data. We came to realize that logistic regression was the best fit for our data, so the team decided to go with that

I believe we all contribution to planning of the project. All my teammate went above and beyond whether it was meeting up before or after our assigned meetings. Regardless how busy were they with their role we kept encouraging and helping each other to finish the task despot the fact how.

Personally, I would like to thank everyone one of my teammates for their hard work for keeping the team together and stepping up to help each other when needed.

Best regards,

Asad Khan