Date: 9-Mar-22

Data BootCamp ITESM

Individual Self-Assesstment

Sergio Roman Gaytan Tanguma

Self-Assestment

* On this project I played two main roles. First, I was the leader of the team, and my principal tasks were creating the meetings, the github, communication protocols, set the rhythm of the discussions, get everyone aligned with the objective, validate what is needed and assign tasks to the rest of the team, choose the topic of the project, and evaluate the results. As my secondary role, I was the data architect where I locate the information, create the repository in AWS, made ETL Process with Google colab and postgres, create the schema of the database with PGAdmin and finally load it to the RDS Database in AWS, all the accounts are my own user. Also, I helped to create a bunch of regressions and ML algorithms as Gradient three boosters and Classifications.
* To the rest of the roles, I helped creating the environment (meetings) for discussions, try to align the team, help on how to create the dashboard and ML algorithms and lastly, help on the wording of the readme and presentation regarding data architecture.
* As a personal challenge I observed the team was lack of involvement, they were not attending the meetings or delivering the tasks assigned (in particular, Ricardo Barba and Raciel Tavitas). Jorge and I talked about this after second delivery were almost all the work were performed by us. What we did to overcame this was to talk with our TA, explain the issue and ask him to talk with the rest of the team. After this, the team was focused and working aligned for the rest of the project duration.

Team Assessment.

* Our communication protocol consists of 4 ways, Whatsapp mainly, Slack, email, and meetings, we had some challenges like when we couldn’t make the team to meet, they were not attending to the meetings and what we did was to set and send the next meeting during the current meeting after a verbal approval from all the team and give us a reminder via Whatsapp prior to the meeting time.
* Our strengths as a team we could say that the roles fall to each person easily, since we had a particular set of skills already and we complement each other to form what this project needed

Summary of project

* The project intention was to analyze Formula 1 sport data thru time, our main goal was to analyze what is more significant to win a race, the car, or the driver. What we did was to define a set of variables (17) and define which of those variables are dependent to a driver or car. Here we validate the correlation between the variables and the dependent variable to see how related they were. Also, we sort the variables in importance level to define a race win. Once we have our best machine model validated, we used to evaluate defined data altering the driver variables with different car values and then car variables with different driver values and see if they really got affected altering these values. The result was that Driver and Car are tightly connected and both affects the results, but we were able to see that the driver affect a bit more than the car since a bad driver with a good car don’t improve as much as a good driver with a bad car.