INTRODUCTION TO DATA SCIENCE [CAP4770-3672, CERT(23937,23940)]

GRADING SHEET

PROJECT GROUP- 10

Group Members- Gonzalo Prats, Avneesh Atrey, Ayati Maitra, Hugo Chapado, Ignacio Gómez Fernández

Overall Hypothesis:

Get information from both races and quali and extract features from that data that explains a correlation between those two variables and, from that, come to conclusions about the importance of a good quali for each race.

Originality 5 __
Difficulty 4 __
Clearly stated and clear outcome 5 __

Data:

- Dataset 1- Race data, https://www.kaggle.com/cjgdev/formula-1-race-data-19502017/
  - races.csv
  Difficulty- (Streaming) 4
  Quality- (Of use) 4
  Completeness 5

- Dataset 1- Race data, https://www.kaggle.com/cjgdev/formula-1-race-data-19502017/
  - results.csv
  Difficulty- (Streaming) 5
  Quality- (Of use) 5
  Completeness 4

Technique:

Technique 1- Gaussian Naïve Bayes

Difficulty 3
Quality 5
Originality 5
Completeness 5

Technique 2- Linear Regression

Difficulty 4

Quality 3
Originality 4
Completeness 4

Technique 3 – Decision Tree

Difficulty 4
Quality 4
Originality 5
Completeness 5