**Big Data Analytics- Proposal**

**Sport Analytics**

**--Project title: Sport Analytics**

**--Group Member : He NI et Bader GUETARI**

**--URL: https://github.com/openfootball/england**

In current assumption the dataset need to be further preprocessed.

**--The goals of the project:**

Given two teams: analyse the data from the last 10 games between the two teams (scores, shots on goal, points scored in the first and second halves, etc.) through the calendar.

Make a classifier to predict which of these two teams will win the next match.

A dataset will be created with examples of matches where X beat Y, Y beat X, or X and Y drew. Report the accuracy of my classifier

Optional: the predicted outcome varies week by week

**--Project timeline:**

**-December 27: Finish ETL get usable dataset , and also the train and test data.**

**-January 3: Make the classifier and let the prediction correct accuracy as higher as possible.**

**-January 15: Do the prediction changes week by week during the matches. And Upload on Moodle a Project Presentation.**

**-January 17, 2020: Upload on Moodle the final project report.**

**--Role of team members:**

**He NI: Data processing**

**Bader GUETARI: Classifier and Prediction**