The data comprises three files: match\_history.csv, test\_set.csv, worldcup\_2022.csv.

match\_history.csv contains the data required to train your model. It consists of historical match information of all the teams worldwide. The features include:

* date : the date of the match
* home\_team : the team playing at home
* away\_team : the team playing away
* home\_team\_continent : the continent where the home team is from
* away\_team\_continent : the continent where the away team is from
* home\_team\_fifa\_rank : the fifa rank of home team before the match
* away\_team\_fifa\_rank : the fifa rank of away team before the match
* home\_team\_total\_fifa\_points : the fifa points of home team before the match
* away\_team\_total\_fifa\_points : the fifa points of away team before the match
* home\_team\_score : score of home team in the match
* away\_team\_score : score of away team in the match
* tournament : the type of match (i.e. World Cup Qualification)
* city : the city where the match took place
* country : the country where the match took place
* neutral\_location : True if the match location is not home for both teams
* shoot\_out : whether there is a shoot-out
* home\_team\_result : Win, Draw or Lose for home team
* home\_team\_goalkeeper\_score : score of the goalkeeper of home team during the match
* away\_team\_goalkeeper\_score : score of the goalkeeper of away team during the match
* home\_team\_mean\_defense\_score : score of defence players of home team during the match
* home\_team\_mean\_offense\_score : score of offense players of home team during the match
* home\_team\_mean\_midfield\_score : score of midfield players of home team during the match
* away\_team\_mean\_defense\_score : score of defense players of away team during the match
* away\_team\_mean\_offense\_score : score of offense players of away team during the match
* away\_team\_mean\_midfield\_score : score of midfiled players of away team during the match

test\_set.csv contains the data of all the game happened in 2022 with only the teamnames and the match result. This should be used as a test set to see how your model performs. This does not have the features provided in the training files (because obviously it does not contain the match information before the match) and you should engineer your own features.

worldcup\_2022.csv contains the match information during the coming world cup. You should give your predictions on these matches. The submission format should be the same as test\_set.csv.