

# AI for Football (Soccer) Analysis

Cristiano Gaudino - 117434292

Supervisor - Alejandro Arbelaez

## Short Analysis

With football being the most popular sport in the world, there is a large number of extensive datasets available as a result of match recording tools, intelligent video analytics technologies. With the growing interest for football in data analytics the datasets are expected to only get richer in the future.

The aim of this project is to explore how Artificial Intelligence can be utilised for soccer analytics. As outlined by the project supervisor, Alejandro Arbalaez, the project will aim to achieve the following goals:

- Identifying the relevant attributes that will influence the outcome of a given match.
- Studying the correlation between expert ratings vs. (Machine Learning) predictions.
- Using the outcome of the models to optimize betting strategies.

## Broad Plan

The main goals of the development of the project are as follows:

- Build a model to predict the outcome of a given football match before it starts.
- Build a gambling system to make fake investments on bookmakers and observe how this can be used to create better betting strategies.
- Make live predictions based on events within games.

There's a large number of datasets available that all contain different information, once one has been chosen that contains all the required information, we can focus on one season and build the model for predicting the outcome of matches. After scaling up to include more data and once an adequate model has been found, the latter goals can be worked on.

The model can then be used in conjunction with a gambling system. Once a prediction has been made on a game, the gambling system will use this prediction to make fake investments based on the odds offered by bookmakers. Observations can then be made on how better betting strategies can be created for the system.

Finally, using more advanced datasets, make live predictions based on events within games. This will be harder to achieve as it will require using datasets that are updated within games, the model will also have to change for this as these live stats need to be accounted for.