Performance Measurement

On a macro scale, our simulator intends to estimate and predict match outcomes between two teams. On a micro scale, there are several performance measures that actually determine the final outcome of the match. To properly model these measures, we began by asking ourselves several questions:

1. What are the significant predictors for final match outcome?
2. How do we model these predictors based on available data?
3. If data is not available, how then do we ensure the validity of our model?

Upon answering the above questions, we narrowed down the predictors to: distance of ball from either goal, relative strength of teams involved and team tactics. From the data obtained, we were able to derive estimates of the team strength and classify their tactics (as shown in the table below):

buildUpPlaySpeed	buildUpPlayDribbling	buildUpPlayPassing
Balanced	Normal	Mixed
Fast	Lots	Long
Slow	Little	Short
chanceCreationPassing	chanceCreationCrossing	chanceCreationShooting
Normal	Normal	Normal
Risky	Lots	Lots
Safe	Little	Little
defencePressure	defenceAggression	defenceTeamWidth
Medium	Press	Normal
High	Double	Wide
Deep	Contain	Narrow

Fig 1 : Classification of Football Tactics

Using the above classifiers and data retrieved from Sofifa.com, we were able to use Python and Jupyter Notebook to generate a final dataset that includes the average team rating (code attached) and numerical tactics.

home_avg_rating	away_avg_rating	home_tactics	away_tactics
81.090909	73.636364	[66.0, nan, 30.0, 30.0, 45.0, 35.0, 30.0, 40.0...	[58.0, nan, 30.0, 31.0, 70.0, 50.0, 30.0, 70.0...
69.545455	73.909091	[70.0, nan, 70.0, 70.0, 70.0, 60.0, 70.0, 70.0...	[70.0, nan, 59.0, 65.0, 70.0, 50.0, 30.0, 70.0...
83.545455	75.454545	[70.0, nan, 60.0, 56.0, 70.0, 70.0, 30.0, 60.0...	[65.0, nan, 70.0, 70.0, 70.0, 70.0, 70.0, 70.0...

Fig 2 : Example of Team Rating and Tactics

Thus, our performance measure is the number of goals scored by each team in 90 minutes, which is weighed against both the expected results (from real life matches) and the predicted results from the Poisson model generated. More details on how we run the experiments using the simulator are in the Experimental Design and Experimental Runs documents.