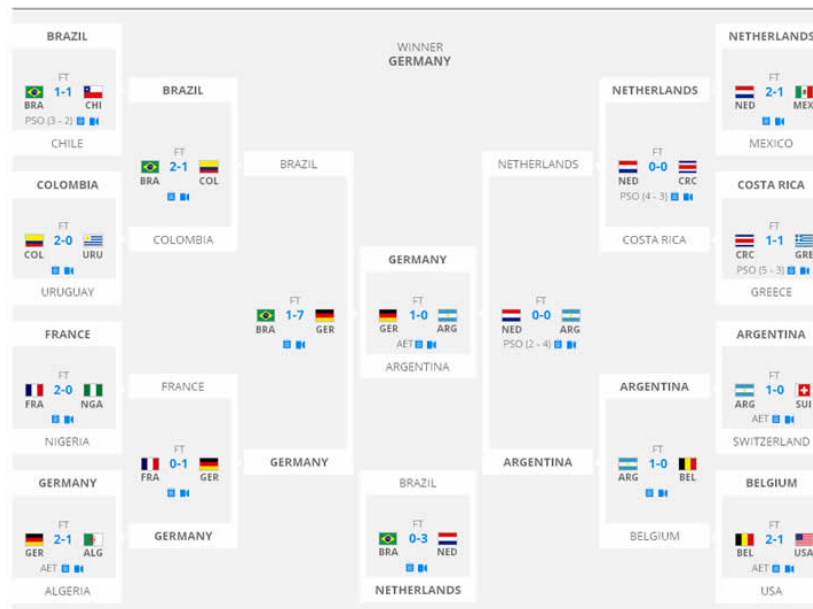


World Cup Recap

Group Members: Wen Bo, Weichen Ning

1. Subject

We choose to build a model on World Cup match results prediction, especially focus on elimination games. We plan to use the final-16 soccer teams in 2014 Brazil World Cup to build our model and make the prediction, then compare with the real results to test the accuracy of our model. To be mentioned, we can also use other tournament elimination games to test our model(e.g. 2010 South Africa World Cup).



2. Method

By building Bayesian linear regression model, we select historical records containing the variables of weather, home or away game, having super star or not, ranking in FIFA, passing success rate, shots on target and possession percentage. Moreover, we will also perform variable selection on the measurements to adjust the weight of variables.

3. Data

We plan to collect data from the website Opta(<http://www.optasports.com/>) which holds the historical records of all kinds of match results. We can use some parsing tools such as JSON to parse the data to make it more readable and editable. If we cannot find the actual parsing tools to parse, we will type in by hand.

4. Further Development

We can use our model on any other tournaments in a specific time period such as future European Cup and World Cup. If possible, we can improve our model to predict the final champion before the tournament starts.