

Predicting Wins and Losses in Women's Rugby

The provided data contained a large amount of information about the wellness of each player throughout the course of the rugby season. We wanted to use this data to predict wins and losses in the team's games. To do this, we supplemented the wellness data with outside data representing various conditions at the time the games were played. This included environmental factors such as temperature, humidity, precipitation, and the stadium's elevation. It also included other relevant factors to the game, such as the opponent's relative strength and the tournament location. We hoped this data, together with the information about fatigue, hydration, and sleep patterns, would give a more complete method of predicting wins and losses. The weather data was collected via NOAA, and elevation data comes from USGS. Information about other rugby teams was obtained via World Rugby's organization website.

Our first model was a binary classification to predict whether the team would win or lose that game based on the variables above. We used a ridge GLM due to the low number of game observations. This model correctly predicted 30 of the 38 outcomes provided.

The second model, instead of using a binary W/L variable as the response, used the continuous point differential between the two teams. Thus, a negative result indicates a loss, as well as by how many points. We modeled this differential using a ridge GLM, and it correctly predicted the sign of the point differential for 25 of the 38 games.

Below is a summary for the coefficients of models 1 and 2, respectively. Below that is a confusion matrix to describe the performance of the binary classification model.

(Intercept)	-1.021143e-01	(Intercept)	.
TournamentCommonwealth	1.861984e-02	Altitude	5.216346e-04
TournamentDubai	-1.069813e-01	OppStrength	6.605774e-02
TournamentKitakyushu	3.882187e-02	TeamPoints	1.421650e-01
TournamentLangford	3.947653e-02	TeamPointsAllowed	-1.569618e-01
TournamentParis	1.257454e-01	Average.Sleep.Hours.night.before.game	4.234827e-01
TournamentSydney	-6.083361e-02	Team.Fatigue.Level	9.623709e-01
TournamentWorld Cup	-4.946172e-05	Temperature	7.329486e-03
Altitude	1.841708e-01	Rainno	1.876342e-01
OpponentBrazil	-1.095614e-01	Rainyes	-1.854550e-01
OpponentEngland	1.764393e-01	game_time	-2.764725e-09
OpponentFiji	8.546161e-02		
OpponentFrance	1.030195e-01		
OpponentIreland	2.082133e-01		
OpponentJapan	1.701458e-01		
OpponentKenya	-2.698784e-01		
OpponentNew Zealand	2.828827e-02		
OpponentRussia	1.799606e-01		
OpponentSouth Africa	3.018589e-02		
OpponentSpain	-3.824551e-02		
OpponentUSA	8.178061e-03		
OppStrength	4.755040e-02		
Average.Sleep.Hours.night.before.game	8.697154e-02		
Team.Fatigue.Level	-3.269687e-01		
Average.Team.USG	1.865366e-03		
Temperature	-3.169154e-01		
Humidity	-1.55562e-02		
Rainyes			

	0	1
0	15	0
1	8	15