Kaveh Bakhtiyari
PhD Candidate in Artificial Intelligence | Data Scientist (Bakhtiyari.com)
Jul 15 · 6 min read

The official logo of the FIFA World Cup 2018 — Russia

Why Did Artificial Intelligence Fail in the FIFA World Cup 2018?

Here is what we learned from the AI's failure to predict the results of the FIFA World Cup 2018 Russia.

· · ·

The FIFA World Cup 2018 (Russia) ended today on Sunday July 15th, with France as the champion, followed by Croatia and Belgium. Just like the previous World Cup 2014, many researchers tried to predict the outcome of the tournaments in Russia in advance. This year, the FIFA World Cup 2018 was no exception, and researchers and scientists tried to exploit Artificial Intelligence (AI) and statistics to predict the outcomes of all 64 matches in FIFA World Cup.

Artificial Intelligence (AI) has made a lot of noise recently, and it is known as the future technology. Nowadays, AI is becoming a part of every large and medium businesses. But how reliable would it be? In this article, I talk about the performance of AI in predicting the results of the World Cup 2018 as a sample use-case. Either you are an expert in

AI or not, I try to keep this article as simple and understandable as possible.

.   .   .

There are different approaches to predict the results of FIFA World Cup. One approach is to simulate every single match in a paired comparison in terms of team's capabilities and the winning odds. Zeileis, Leitner, and Hornik (2018) used the same technique, and they predicted that Brazil would win the FIFA World Cup 2018 with a probability of 16.6%, and it is followed by Germany (15.8%) and Spain (12.5%) [1].

Swiss Bank UBS also predicted the same three teams as the top 3 teams but in a different order. They predicted Germany (24.0%) as the champion, followed by Brazil (19.80%) and Spain (16.1%). Their generated model was based on four factors: 1) the Elo rating; 2) the teams' performances in the qualifications preceding the World Cup; 3) the teams' success in previous World Cup tournaments, and 4) a home advantage. The model was calibrated by 10,000 Monte Carlo simulations to determine winning probabilities and the results of the last five tournaments [2].

On June 8, 2018, four researchers (A. Groll et al.) from Technical University of Dortmund (Germany), Ghent University (Belgium), and Technical University of Munich (Germany) published a research paper on arXiv predicting the results of the FIFA World Cup 2018 using a well-known algorithm of Artificial Intelligence, *Random Forest,* and *Poisson ranking algorithm* [3]. This paper was published online days before the opening game of the world cup between Russia and Saudi Arabia, on June 14. They used a dataset covering all matches of the last four FIFA World Cups (2002–2014). They predicted Spain as the champion, followed by Germany and Brazil as runner-ups.

These three mentioned research came up with the same top 3 teams of Spain, Germany, and Brazil, in different orders. They used three different methods, data and data features, but they came with almost a similar result. Now, the world cup is over, and we can see that all those models failed to predict the world cup results correctly, and none of the predictions happened.

Among these research, the methodology of A. Groll et al. is my favorite in this area. First of all, they used a good data source. Secondly, they considered many features and parameters for training. Thirdly, they

employed the algorithm of *Random Forest*. In the rest of this article, I discuss its data features, error, and the reason of its failure in this area.

Data Features

A. Groll et al. considered various features related to the team itself, such as *1) Economic factors* (GDP per Capita, Population); *2) Sportive factors* (ODDSET probability, FIFA Ranking); *3) Home advantage* (Host, Continent, Confederation); *4) Team's structure factors* (Maximum number of teammates for each squad, Average Age, Number of Champions League players, Number of Legionnaires); *5) Team's coach factors* (Age, Duration of tenure, Nationality). In total, they had 16 features for each team and each world cup.

Classification Model

As I mentioned earlier, they used "*Random Forest*" that is one of the well-known algorithms in AI and Data Mining. This algorithm works based on "*Decision Tree*", and it has shown high performance in data classification in many cases. They also used *Poisson models* to rank the teams based on their current abilities.

			Round of 16	Quarter finals	Semi finals	Final	World Champion	Oddset
1.		ESP	88.4	73.1	47.9	28.9	17.8	11.8
2.		GER	86.5	58.0	39.8	26.3	17.1	15.0
3.		BRA	83.5	51.6	34.1	21.9	12.3	15.0
4.		FRA	85.5	56.1	36.9	20.8	11.2	11.8
5.		BEL	86.3	64.5	35.7	20.4	10.4	8.3
6.		ARG	81.6	50.5	29.8	15.2	7.3	8.3
7.		ENG	79.8	57.0	29.8	15.6	7.1	4.6
8.		POR	67.5	46.1	19.8	7.3	2.5	3.8
9.		CRO	65.9	30.8	15.6	6.0	2.2	3.0
10.		SUI	58.9	30.6	13.1	5.6	2.2	1.0
11.		COL	79.2	33.1	14.0	5.7	2.1	1.8
12.		DEN	59.0	26.1	12.4	4.8	1.7	1.1
13.		URU	86.6	37.5	13.5	4.4	1.3	2.8
14.		SWE	54.0	21.7	8.0	3.1	1.0	0.8
15.		POL	60.6	18.9	6.8	2.3	0.7	1.5
16.		PER	39.2	15.4	6.6	2.1	0.6	0.4
17.		ICE	36.6	12.9	5.3	1.7	0.5	0.6
18.		SRB	36.2	13.8	4.7	1.5	0.4	0.6
19.		SEN	39.7	10.9	3.7	1.1	0.3	0.6
20.		MOR	30.3	14.8	4.0	1.0	0.3	0.3
21.		TUN	22.8	8.9	2.8	0.8	0.2	0.2
22.		MEX	41.5	13.9	3.7	1.1	0.2	1.0
23.		CRC	21.4	6.4	1.7	0.4	0.1	0.3
24.		EGY	45.5	10.3	2.1	0.4	0.1	0.6
25.		RUS	50.4	10.5	2.4	0.4	0.1	2.2
26.		NGA	15.8	4.0	1.2	0.3	0.1	0.6
27.		AUS	16.2	4.2	1.2	0.3	0.1	0.3
28.		JPN	20.5	4.1	0.9	0.2	0.0	0.6
29.		KOR	17.9	4.0	0.8	0.2	0.0	0.6
30.		IRN	13.8	5.1	0.9	0.1	0.0	0.3
31.		PAN	11.1	2.5	0.5	0.1	0.0	0.1
32.		KSA	17.5	2.6	0.4	0.0	0.0	0.1

Predicted results of the FIFA World Cup 2018 by the Random Forest algorithm — Source: [3]

Prediction

After running the tournaments' simulations for 100,000 times, Spain was predicted to be the champion reaching final with 28.9% of chance, followed by Germany (26.3%), and Brazil (21.9%).

Error

As we observed in the FIFA World Cup 2018, none of the predicted top 2 teams could reach to quarter-finals, let alone the final games (Brazil reached quarter-finals). Based on the actual results of the world cup, and the predictions, *Root Mean Square Error (RMSE)* and *Mean Absolute Error (MAE)* of the model are calculated as below:

```
RMSE: 8.052
MAE:  6.468
```

These two metrics show the error of the model, and how far it could predict the team rankings accurately in overall. Both RMSE and MAE values are so high, which make the model unreliable, and despite using 16 features and large data set (past four world cups), AI machine learning (Random Forest in particular) still failed to predict the results reliably. In this World Cup, Russia, Japan and Iran played significantly better than what was predicted, and on the other hand, Germany failed to qualify.

	Country	Predicted	Actual	Difference
1				
2	France	4	1	3
3	Croatia	9	2	7
4	Belgium	5	3	2
5	England	7	4	3
6	Russia	25	5	20
7	Brazil	3	6	-3
8	Uruguay	13	7	6
9	Sweden	14	8	6
10	Spain	1	9	-8
11	Colombia	11	9	2
12	Denmark	12	11	1
13	Argentina	6	12	-6
14	Japan	28	13	15
15	Portugal	8	14	-6
16	Switzerland	10	15	-5
17	Mexico	22	16	6
18	Senegal	19	17	2
19	Iran	30	18	12
20	South Korea	29	19	10
21	Peru	16	20	-4
22	Nigeria	26	21	5

The predicted ranking, the actual ranking, and the difference (error) of prediction for each team in the FIFA World Cup 2018.

Why did AI fail?

In AI and Machine Learning, it is very important to have proper data for training and modeling. But in this case, despite having proper data (16

features, cleaned), relatively large data (past four world cups), and good algorithms with right parameters, the trained model failed terribly. The reason for this failure relies on the nature of what we are predicting.

FIFA World Cup like any other human-based incidents is dependent on too many factors (not only 16) before and during the match (for minimum 90 minutes), which are known as *confounding variables*. In order to predict the results correctly, every single minute of each match should be simulated. The result of each state (every minute/second) of the match depends on the preceding states. This is also known as *Markov Chain Process*. An incorrectly simulated state can easily result in unreliable outcomes for the proceeding states of the game.

Besides the internal factors, the results of a football match may also be significantly influenced by some external factors as well, such as an unfair referee, weather, political situation, even personal problems of players, etc. These important features are usually very difficult to be measured and collected. In addition, there is always some chance of exploration, and uncertainty, for instance having a critical mistake or scoring an own goal, which are not easily predictable.

In a nutshell, *stochastic* and *dynamic environments* such as FIFA World Cup or human activities are those areas that the today's technology of AI cannot perform very well. This is a very good example to note that we have to be very careful about the applicability of AI in the similar dynamic fields. Also, by having a very complex data structure, it might be very difficult to audit the trained models for any potential bias. The existence of bias in AI can simply lead to discriminative decisions against a particular group. The implementation of such systems responsible as the sole decision maker may cause huge problems for both individuals and companies. Governments and companies are recommended to use AI for stochastic and dynamic environments only as a supplementary decision-making platform.

. . .

References

[1] Zeileis, A., C. Leitner, and K. Hornik (2018): "*Probabilistic forecasts for the 2018 FIFA World Cup based on the bookmaker consensus model,*" Working Paper
2018–09, Working Papers in Economics and Statistics, Research

Platform Empirical and Experimental Economics, Universität Innsbruck.

[2] Audran, J., M. Bolliger, T. Kolb, J. Mariscal, and Q. Pilloud (2018): "*Investing and football—Special edition: 2018 World Cup in Russia,*" Working paper, UBS.

[3] Groll, A., C. Ley, G. Schauberger, and H. Van Eetvelde (2018): "*Prediction of the FIFA World Cup 2018—A random forest approach with an emphasis on estimated team ability parameters,*" Working Paper.

7/15/2018

Why Did Artificial Intelligence Fail in the FIFA World Cup 2018?