

# Footie: Web Simulator of Football Leagues and Championships

Martin Urbanczyk



### Abstract

The goal of this project is to create a web football simulator, which would simulate realistic matches of the football world. The concept of the whole simulator is complex. It generates future results based on real historical match results. However, unpredictable and crazy results are also a part of the fun. Each user has its own simulation and he or she can have completely different results, which are more and more diverse year after year of the simulation. Because of the complexity of the football world, the simulator supports only selected football nations. The main focus of the project is to improve the simulation of matches. Calculation of match results is based on real historical results and it uses neural networks. It is mixed with random factors described in the paper. The project focuses on the group of football fans who would like to only watch simulations of the football world without the need to manage teams and players. This is something that is not available in the current market. Every game focuses on management mainly, not the simulation. The core of the simulator could be also used for trying to predict real football matches because it is based on real data. This paper describes base concepts of predicting a football match. It presents football systems and shows, how the simulator game is different from existing solutions.

Keywords: Football - simulator - machine learning - match prediction

# Supplementary Material: N/A

\*xurban61@stud.fit.vutbr.cz, Faculty of Information Technology, Brno University of Technology

# 1 1. Introduction

The main value of this project is a system predicting 2 match results based on real data (real results). It is con-3 nected to a simulating system which includes random 4 factors and provides a unique evolution of the football 5 world, based on current data. These random factors 6 are not just random functions. They are based on some 7 potential motivating factors, which exist in the real 8 sports world. A user who would like to see scenarios 9 of the next football season can have satisfaction in this 10 game same as the user who likes sci-fi scenarios and 11 would like to see generated football world for example 12

in 2100. The core of the prediction can be also used 13 for prediction of match results in the current season in 14 the real world. 15

The current version of the system simulates the 16 biggest European football nations (England, Spain, 17 France, Italy, and Germany) and the Czech Republic. It 18 also contains the biggest club competitions like Cham-19 pions League or Europa League, which use the UEFA 20 coefficients<sup>[1]</sup> for recalculation of nations' ranking af-21 ter each season. The ranking is used to calculate the 22 number of participating clubs for every nation. Hence, 23 it is possible that some smaller nations might have 24

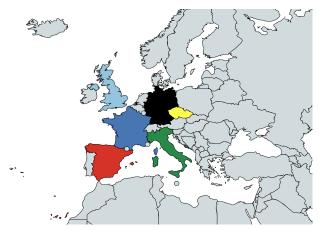


Figure 1. UEFA nations which are included

<sup>25</sup> more teams the competitions in future seasons. The

data of all world leagues are included, but the project
is pretty complex. I chose to focus on the match pre-

diction. Figure 1 present map of UEFA nations, whichare included in detail simulation.

The system needs to be entertaining and playable 30 for football stats maniacs as well as for casual fans 31 who would like to see how their favourite team evolves 32 in the nearby future. The project does not guarantee 33 fun for every user. It focuses on the people who are 34 interested in football and its future. It offers some-35 thing to either casual fans and bigger experts who 36 have experience with playing manager games and who 37 are interested in deeper stats. Who knows, maybe in 38 39 your simulation Zbrojovka Brno will play Champions League final in the year 2100. It is also planned to 40 create a community hub where users can share their 41 results and compare the state of their world. 42

The crucial task of the project was to find the key features of the prediction. I also had to find the *soft* features which might also have an impact on match (e. g., motivation, frustration, form, or match importance).

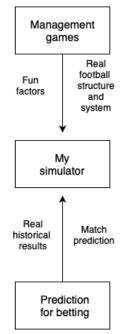
#### 47 2. Analysis

Most of my inspiration came from playing manage-48 ment games and detailed watching of the football 49 world. The biggest football manager is the game from 50 SEGA called Football Manager 2019 [2]. It is a pretty 51 realistic game but it focuses more on the management 52 factor. No matter which team you choose, the world 53 automatically changes due to your actions. I took some 54 55 inspiration from its user interface (Figure 2). I also used to play Czech Soccer Manager [3] in the past. 56 The game is not so huge and complex as FM 2019, but 57 I took some inspiration from it's easy going UI and its 58 simplicity. 59

None of these games and projects focuses on realistic simulation. You can get some crazy results

	111				B. MADRID. CATAOT -
pante			zázna 2013 Goder dní 21	SPRÁVY Z TÝMU VELANDENE	
				Seego Avego nexi regatrován Mara Iberlanes servinezátrován	CO.CO D - Fixel Heppile
	<b>X</b>	Villarreal	(R. Madrid 👷 –		06.03 V E Linepool O
1 jima	(CVA)		- Internet (1990)		17.03 D 🕹 Ellar
			O Marcello (VS)		35.03 V 🎒 Defate
			MINISTRA	Toni Kroce je mine hru japortovnikýtů Alvara Odropala je mine hru jaztačení zadní stetenní ováli	CO.04 V - Valencia
				And a Coll scale of mine the parameters as a service state	07.04 D 💆 Lepente
			STATISTIKY INACE -	- Olytex years -	WEDV A FRANKE
Pep			- Northyllou	Report	£160e4
			Mariana	Shenata a podaržano, vedeni je se spilosbem, jelvjm tým redeli,	£120+1
Z			A 1400	spokojero.	all
***			C National Accession	Witked vivaced bits Chain Warld Champsonship tryle bez posttyk referenceared been.	
1		21 20 3 3 43 63	Sector Sector		
3.		24 17 6 3 25 67		Yedení však je zklamánovýkonem týmu v souhlží Spenish Cup.	eve all in ted
2.		21 16 5 5 21 53	Toni Knos	Southa	CELINARY FOR WERE DOM, CTRICK & UTOBASHED
٤.			Carlos Carlos	26949	00222606.075 0016.013,7040
	V A. Biber	28 18 4 6 15 62	- O surviviantiane risesion	Postavani tizma	#817,998,120
	Espanyol 1 Yest	26 13 5 6 3 44	% Marcas Lhawrite	Pertury	Planadni subpodet był zw zyta 677,32mil
	Vies Sovila	21 13 4 5 5 43	and the second s	Kideni france	20% z přípnu z photopů k dopceno
· .	V Sevila	20 10 4 9 9 40	A Structure of the second state of the second	Protecte Kabo	#20047#02001ar #225.128.3555 miles
	de Levante	21 10 1 12 1 40	Marcelo Transmitter and finalized	Shinebe e sou	Mitcalini ryskeji 625.333.588 mitc.
15	V Leganda	20 8 7 11 12 21			belaveré výdes tot. NY. Kristeke. Bestapní natový szpole rs.M.
12	S Williamood	28 7 9 10 -10 50	A Contraction becausing a seaso	Robours assurbara	
12	Red Sandroattin	2111112	STATISTICS THE		· Tolesa ·
14.		28 7 4 18 -18 28	Sparable Line Division	Covenino	TRÉNINGONÝ STVL - TECHNIKA
15.			Potet of ehrenisch zeigandis 20	habitrout novos anticore	
18.			a tanak	Zhaina Anic B O Mexispach	
17.	💧 Gatalia		A Verbievé beavley	Theast Coarbin	
15			V 50 - 3. repybli	Vic hearly pamin tyme	<u> </u>
12.				Zhipa de b 2 Dissipatels	
20.	Statecare .		Children branky		NUMERIA TRANSPORTATION AND A PROVIDENCE TONIS
			SP 10-2 rejesti		
					o Masana men
			Addataty 40 - 5 retrolli		D.Min. 20 M
			🗸 40 - 1. rajvjili		NUMBER OF THE PROPERTY AND A POST TO THE
			A	1	
			Cervenikarty		n Karin Bergana miti Soberi nine, hine da

Figure 2. FM 2019 main dashboard [2]



**Figure 3.** Difference between existing solutions and my simulator

practically immediately. It is alright when you play a game just for fun. On the other hand, there are also some projects which focus just on match prediction 64 mainly for betting [4]. These projects do not focus 65 on the simulation of a hypothetical world. Figure 3 66 presents how I mixed my project from these types of 67 existing solutions. 68

69

#### 3. Design

The solution is on the halfway to both analyzed cat-70 egories. It uses real prediction of the results. Then, 71 it adds the random factor to make the game more en-72 tertaining. Results of the matches are close to the 73 predicted results but they might be different for every 74 player. This helps to make the evolution of the football 75 world diverse and crazy in the future. The simulator, 76 however, does not force players to plan tactics, buy or 77 sell proper players or negotiate about their salaries. It 78 focuses more on the simulation of the games than any 79

80 other game does.

For example, if players of Football Manager go for 81 one year lasting holiday (the active management of the 82 team is interrupted), they will see more diverse results 83 in the league they chose. Just because they chose it. 84 For example, you took the manager seat in England. 85 Even if you do nothing, there will be some shocking 86 results that are not so realistic. However, in a league 87 where you are not a manager (for example Germany), 88 results are still going to be similar (for example Bayern 89 Munchen will win 10 titles in 10 next seasons). There 90 will be some unpredictable results but not more in 91 some region just because you choose it. 92

If the player starts a new game in Football Man-93 ager, they will see more diverse results in the league 94 they chose. Just because they chose it. Other leagues 95 evolve too, but the change is far slower. So it is not 96 quite realistic. But it is mainly a game, so it is un-97 derstandable because players want to go higher fast. 98 For example, when a player begins a game in England, 99 weaker teams can finish higher in England. Although, 100 in Germany or Spain it is different. Every time it takes 101 more than a decade when some weaker team wins the 102 league. 103

However, I plan to add some extensions in the future, that will allow players to add their teams, players to real leagues, or make some transfers to see how things might evolve (e.g., What would happen if Messi and Ronaldo play in the same team that is on the bottom of the table).

All of the biggest football manager games are available for the only desktop platform. I decided to make a cloud solution, so you can connect everywhere and it play it on (almost) every device, that has an internet browser.

### 115 4. Implementation

#### 116 4.1 Architecture

The architecture of the system is displayed in Figure 117 4. The system is structured into modules which are 118 not dependent on each other. Each functionality has it 119 separate control (i.e. team screen - controls for play-120 ers, fixtures, stats). This makes the system extensible 121 and scalable. In the future, it is planned to provide 122 subsets of the modules to users according to their pur-123 chase plans. They will be able to choose the modules 124 according to their interests (e. g. leagues, nations). 125

I implemented the system as a web application in PHP 7.2 (Nette framework) and JavaScript mainly for asynchronous loading and user actions. It has a lightweight user interface. It is mobile friendly and it does not overwhelm users with stats and data (the detailed data are accessible in specific pages). The system uses the MVC architecture and it supports REST API written in node.js for the connection with the database. It is expected that native mobile applications will be implemented in the future.

The system uses the standard MySQL database. 136 The database has two parts. The first part consists of 137 "hard" data which will not be changed in the future 138 and are same for all users. For example, it contains 139 example names of players and teams, the number of 140 teams in each league etc. The second part differs for 141 each user. It consists of actual data of a simulation 142 (e. g., player goals, team history etc.). The whole 143 database contains a lot of data (for example, around 144 80 000 players). 145

Also, I needed to do a lot of parsing and merging of146data, because of different sources. It was necessary to147merge the Football Manager database, my system and148database and api-football data. I used mainly Python149scripts for this.150

151

#### 4.2 Match Prediction

The match prediction is provided by using the Python 152 scikit-learn library [5]. It uses the neural network 153 which was trained with the data gathered from the apifootball.com service [6]. The service provides a REST 155 API for getting real results. The results are mixed 156 with the results of the simulation (in the first years of 157 simulation). The main features of machine learning 158 are: 159

- 1. **Team rating**: the actual rating from the database; 160 it is calculated after each season from the results. 161
- 2. Actual team form: the last 10 matches played 162 by each team. 163
- Form at home/away stadium: the last 5 matches 164 played at home and away from fields, depending 165 on which team is the home team and which is 166 away team.
- 4. Last matches between the two teams: the matches between the two teams during last five years (not the 5 last matches, because it is not relevant if these two teams played together 15 171 years ago under completely different circumstances).

I made a dataset of historical results and tried to predict some older games. I have a lot of different data from all over the world. I did a lot of experimentation with a different type of matches. I took also some official odds for comparison. It was provided by the api-football API, but I also did a comparison with Fortuna's odds [7]. I had around 85 percent of 180

successfully predicted results of the matches contain-181 ing a strong favorite. The success of prediction means 182 the right prediction of win/draw/loses. It was worse 183 for mixed datasets, where it was around 65 percent of 184 successful predictions. Then, I did a competition with 185 Fortuna on near planned football matches through the 186 whole weekend in the biggest football leagues, which 187 are mentioned upper. The most important comparison 188 here was between my prediction and odd prediction 189 (my prediction versus the lowest odd on each match). 190 Betting company is still a winner, but not too heavily. 191 In each iteration, they beat me just by a few percents 192 (average is 6 percent, which is 3 matches on average-193 from 49 weekend matches). I chose a polynomial 194 regression model and the rate between training and 195 testing data is 0.8/0.2. The final model is extracted, 196 but it requires preprocessed data from the system and 197 api-football API. 198

#### 199 4.3 Prediction with Randomness

The simulation of future matches uses mostly histor-200 ical simulator's data (it calculates new results from 201 the results of the last 10 matches and last 5 years). 202 Some random factors can have an impact in later years. 203 These main random factors are used mainly to simu-204 late extra motivation, or frustration, or importance of 205 upcoming matches which can produce an unexpected 206 207 score. It generates the rank of "randomness".

Then, the final result is mixed with the result gener-208 ated by the neural network. The algorithm works with 209 a tape of numbers representing the random difference 210 from the predicted results. In the beginning, there are 211 mostly the numbers 0 and 1 (but there are also higher 212 numbers). If the level of randomness should be higher 213 than I add higher numbers (2, 3, 4, ...). Some of the 214 numbers of this tape is chosen for each team. The cho-215 sen values can add or remove some goals which were 216 predicted by the neural network. Mostly, the result 217 will be similar to the prediction, but there can be a 218 situation (which happens also in real football world) 219 220 when a result is shocking. The goal is not to have still the same results, but it is not going to be completely 221 random. Hence, I take the result of the prediction, gen-222 erate the random value and use the value to change the 223 originally predicted result. 224

#### 225 5. Testing

#### 226 5.1 Target Groups

Last year, before the beginning of the work, I spoke
with a group of different people who are involved and
deeply interested in football or football management
games. I asked them for help with testing the sys-

tem. I tried to ask both experts and casuals. As an 231 expert, I mean a person who deeply understands foot-232 ball leagues structure and watches more than just his 233 favourite teams. As a casual, I mean a person who is 234 just a fan of some teams and is not really interested in 235 complex statistics etc. The website is going to be free 236 for all, however, if users would like to edit the date for 237 their own there will be a small fee for this addition. 238

#### 5.2 Actual testing

Firstly, when I finished the UI templates, I asked the 240 users for feedback and they gave me a lot of hints. 241 Currently, they are giving me feedback about the actual 242 functionality of system and prediction. It helped me to 243 fix some bugs. 244

Many of the users made long-lasting simulation 245 so they gave me feedback on how the football world 246 can evolve in the future. Such testing would be time- 247 consuming for one person. Their simulations com- 248 pletely differ in later years but had some common 249 results in the nearby future. After decades of sim-250 ulation, some football giants have fallen into lower 251 leagues and new giants were born. It is quite different 252 for everybody, but this is part of the fun. The system is 253 currently not publicly accessible. It is planned before 254 the deadline of the thesis (end of May). So I gave users 255 my laptop for a few days to check the simulator. 256

#### 6. Plans for the Future

The project is developed as part of a master thesis but 258 I plan to continue with the development afterwards. 259 Firstly, I plan to add all European leagues (at least first 260 and second level). Secondly, other football competi- 261 tions will be included—e. g., international football 262 (World Cups, European Championships and its qual- 263 ifiers). And also, as I said before, I plan to add some 264 customizers for users. It will be possible to edit/add 265 their own teams, players or make some transfers so 266 the users can check how the world changes after. In 267 the year 2020, I am going to move the application to 268 other platforms (Windows: desktop application, iOS 269 and Android). It is possible because I made universal 270 back-end using REST API. However, users need to be 271 connected to the Internet on all platforms. 272

#### 7. Conclusion

This paper provides information about Footie—the 274 web football simulator. The simulator generates match, 275 league and season results and is not limited by a number of seasons. It is practically limited just by the 277 amount of data which can be stored. 278

257

239

273

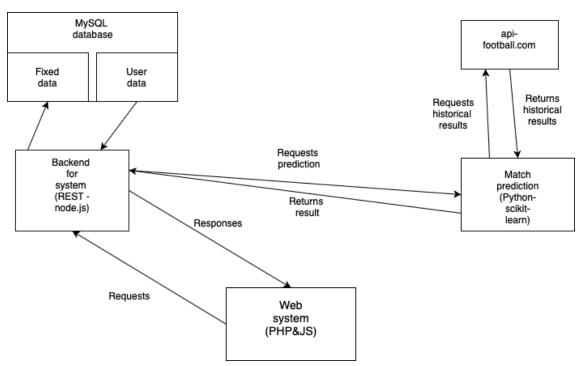


Figure 4. Architecture of the system

In the beginning, it was important to define the main idea of the system. I had a lot of concepts. I tried to compare my concepts with existing solutions (mainly football management games and match prediction for making bets). Then, I designed a which is probably on the halfway to both categories.

Then, I made an architecture that includes multi-285 ple modules. The main modules are the web system, 286 back-end API, neural network and API with real his-287 torical results. The main goal was to create match 288 predictor which combines the predicted results with 289 regulated random factors in order to bring both realis-290 tic but sometimes unpredictable and crazy results (as 291 real football world does). It was tested on both his-292 torical and real-time matches with the help of betting 293 odds. 294

At the current stage, the system includes basic sim-295 ulation of the biggest European leagues and European 296 club cups. The system is composed of modules, so it is 297 not a problem to add or remove any module. The web 298 application is designed responsively so it is runnable 299 on all devices which has a modern internet browser. 300 Its database is optimized as much as possible because 301 of the amount of data. It included around 280 leagues 302 and 70 000 players. All these leagues are going to 303 be playable in the future. Currently, the app is still in 304 development so it can be only seen locally, but at least 305 its first phase prototype is going to be released by the 306 end of May 2019. 307

#### References

[1]	Bert Kassies. UEFA Coefficients calcula-	309
	<pre>tion method. https://kassiesa.home.</pre>	310
	xs4all.nl/bert/uefa/calc.html.	311
[2]	Sports Interactive. Football manager official web-	312
	<pre>site. https://www.footballmanager.</pre>	313
	com.	314
[3]	Czech Soccer Manager official website. http:	315
	//csmweb.net.	316
[4]	Jaroslav Bezděk. Aplikace metod strojového učení	317
	na predikci výsledků fotbalových zápasů. 2015.	318

- [5] SciKit-Learn Developers. *scikit-learn user guide*. 319 [Online; 16.3.2019]. 320
- [6] Api football REST API. https://www.321 api-football.com. 322
- [7] Fortuna offical website. https://www.323 ifortuna.cz. 324
- 8. Screenshots of the application
- 325

308

# 2017/2018

CZECH R	EPUBLIC	ENGL	AND
Het Liga	Narodni Liga	English Premier Division	Sky Bet Championship
1. Banik Ostrava 0	1. Ceske Budejovice 0	1. AFC Bournemouth 0	1 Aston Villa 0
2. Viktoria Plzen 0	2. Vlasim 0	2. Manchester United 0	2. Middlesbrough 0
3. Teplice 0	3. Vitkovice 0	3. Newcastle United 0	3. Milwall 0
4. Sparta Prague 0	4. Viktoria Zizkov 0	4. Southampton 0	4. Norwich City 0
5. Slovan Liberec 0	5. Varnsdorf 0	5. Stoke City 0	5. Nottingham Forest 0
6. Slovacko 0	6. Usti nad Labern 0	6. Swansea City 0	6. Preston North End 0
FRAM	NCE	GERM	IANY
French Under 17 League	Ligue 1 Conforama	German First Division	German Second Division
1. Association Sportive Orly 0	1. Amiens SC 0	1. Augsburg 0	1 Aue 0
2. Cavigal Nice Sports Football 0	2. Montpellier Hérault SC 0	2. Stuttgart 0	2. SG Dynamo Dresden 0
	3. OGC Nice 0	3. RBL 0	3. Sandhausen 0
	4. Olympique de Marseille 0	4. Mainz 0	4. Nürnberg 0

Figure 5. Main screen - list of nations

≡					Э
	<b>K</b> English Premier Division			_	
	Tottenham H	lotspur			
				TOTTENHAN	
	PLAYERS RESU	ILTS FIXTURES	TOP SCORERS HISTO	RY	
	Matchday 1	Matchday 2	Matchday 3	Matchday 4	
			$\overline{\mathbf{e}}$	<b>Ø</b>	
	EVERTON	CHELSEA	BRIGHTON AND HOVE ALBION	LIVERPOOL	
	AWAY	НОМЕ	НОМЕ	HOME	
	Matchday 5	Matchday 6	Matchday 7	Matchday 8	
	<b>M</b>	U.	<b>I</b>	ALBON	
	NEWCASTLE UNITED	STOKE CITY	AFC BOURNEMOUTH	WEST BROMWICH ALBION	
	НОМЕ	НОМЕ	AWAY	AWAY	

Figure 6. Team screen - upcoming fixtures

€

# C English Premier Division Tottenham Hotspur 0-0-0



€

PLAYERS	RESULTS	FIXTURES	TOP SCORERS	н	STORY		
Name and Nation	Position	Value	Games	Goals	Assists	Yellow cards	Red cards
ALFIE WHITEMAN (England)	GK	£92k	0	0	0	0	0
ANTHONY GEORGIOU (England)	AM L	£350k	0	0	0	0	0
ARMANDO SHASHOUA (England)	AM C	£5k	0	0	0	0	0
BEN DAVIES (Wales)	D/WBL	£18m	0	0	0	0	0
BRANDON AUSTIN (England)	бк	£17k	0	0	0	0	0
BROOKLYN (Lyons)	D RLC	£22k	0	0	0	0	0
CHARLIE FREEMAN (England)	БК	£11k	0	0	0	0	0
CHRISTIAN ERIKSEN (Denmark)	AM LC	£58m	0	0	0	0	0
CHRISTIAN MAGHOMA (DR Congo)	DC	£52k	0	0	0	0	0
CY GODDARD (Japan)	MC	£117k	0	0	0	0	0
DANNY ROSE (England)	D/WB/M L	£26m	0	0	0	0	0
DAVINSON SÁNCHEZ (Colombia)	DC	£24m	0	0	0	0	0
DELE ALLI (England)	AM RLC	£48m	0	0	0	0	0

Figure 7. Team screen - tables with the players

Kostons     ENGLISH PREMIER DIVISION						
ACTUAL ROUND	RESULTS	TABLE	STATS			
Matches schedule						
Matchday 1						
Home		Αωαγ	Score			
Liverpool	VS.	AFC Bournemouth	5:1			
Arsenal	VS.	Manchester United	Simulate!			
Brighton and Hove Albion	VS.	Newcastle United	Simulate!			
Burnley	vs.	Southampton	Simulate!			
Chelsea	VS.	Stoke City	Simulate!			
Crystal Palace	VS.	Swansea City	Simulate			
Everton	VS.	Tottenham Hotspur	Simulate!			
Huddersfield Town	V8.	Watford	Simulate!			

Figure 8. Competition screen - actual round results/fixtures



Figure 9. Player detail screen

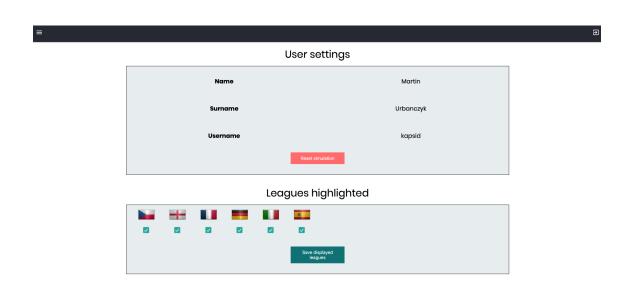


Figure 10. Settings of simulated leagues