Smith School of Business, Queen's University | Master of Management Analytics Toronto

MMA 861
Analytical Decision Making

Professor Guang Li

Project Report: Fantasy Premium League
Due Date: April 11, 2022

Team Adelaide

Student Name	Student Number
Ishaan Sinha	20290167
Jiali Bai	20316342
Juefei Chen	20311141
Kripa Shanker Nayak	20315710
Oluwaseun Adelana	20310041
Yichen Yuan	20315046

Order of files:

Filename	Pages	Comments and/or Instructions
Team Adelaide_Final Report	12	
FPL_Project Model_Team Adelaide	1	

Additional Comments:

Executive Summary

Fantasy Premier League[1] is a virtual competition that involves multiple organization to compete and win the final prize. The ultimate objective of our model is to maximize the points over all games, while considering the player's position, total budget, and other variables. To initiate our modelling process, we start with the input data. As the input for our model, we believe the last year's performance points would accurately reflect the player's recent physical condition due to the nature of sports, thus, we have extracted past year player performance data from Fantasy Premier League's statistics website[2]. Based on our initial analysis of the data, our team decide to use an optimization model from Evolver, would be valuable as it produces informational output which helps to identity the best mix of players that would help in maximizing the overall point for our team during the tournament, thus, to beat other game organizations in FPL within the budget. This system could help us defining all the assumption, constraints, based on the rule of football, and the commercial guideline of the league, thus creating information value. After performing the modeling and its analysis, the optimization model helps the competing organization Raging Bull score the maximum possible score of 2,332 points thereby winning the overall tournament. The team is made up of 15 players, among them 11 are starters and 4 are bench players. After looking at the result from last season, the maximum point of a team could reach up to 2600 points. Due to the limitation of model, the best result obtained is solely based on the performance data from last season and our assumption to the model, any changes during the season such as penalty, injury and decline of performance will be updated in the dataset as our next steps, which could lead to a re-modeling process, and possibly a completely different solution.

[1] Fantasy Premier League official website, https://fantasy.premierleague.com/

[2] Fantasy Premier League, Statistics, https://fantasy.premierleague.com/api/bootstrap-static/

Introduction

Virtual Fantasy Premier League (FPL) is an elite league in which multiple organizations create a 15-squad team to compete among themselves to win the prize of €3 million. FPL is co-owned by 20 Club Members which provide pool of players who can be purchased by each competing organization. Player selection should follow the rules highlighted in the league playbook[3] and player cost depends on the points they scored in the previous season. Also, FPL limits the purchasing of each competing organization by capping the overall squad budget. Players are awarded points for goal scored, assists, saves and clean sheets. Player is also eligible for bonus points as defined in the rulebook. FPL rulebook is to be adhered to by all competing organizations and any serious breach would result in punishment which can range from points deduction to expulsion from the competition.

Raging Bulls, one of the competing organizations which is participating in FPL this year, has hired us as consultants to provide them guidance on how to select the best mix of players which would help achieve the maximum points thereby winning the prize. We are going to create a winning combination of players by utilizing an optimization model.

Data and Limitation

To gather data for our study, we used the FPL statistics website available at *fantasy.premierleague.com*. This is an official site hosted by FPL itself and contains all relevant FPL data relating to players and clubs. Thus, all historical FPL player statistics were available on this site. From all this, we chose to work with the FPL player data of last season. The reason for not selecting a longer window of time in the scope of our report is that the older a statistic is, the lesser its reliability is to predict a player's current form.

[3] Fantasy Premier League, Help, Rules, https://fantasy.premierleague.com/help/rules

Many factors affecting a player's performance, such as age and fitness level, can change over time. So, it's best to use data for last season.

We have identified four critical parameters to compare the benefits of including different individual players in our team, namely, performance, contribution, value, and position. Performance indicates the productivity of a player during a match and shows how well a player utilizes their game time (the amount of time they are involved in the game). But just looking at the performance of a player based on statistics can be misleading as players who play more often will have better statistics, irrespective of the quality of play. Thus, we are using 'contribution' to identify how much of the time is a player contributes to the team. Some players are not able to play all the matches in a season due to injury or fitness, or any other issues. Thus, the contribution is just a measure of the amount of time played out of the total playtime available in the season.

Additionally, to include a player in our FPL team, we will have to pay the sign-on price of the player. Since our budget is limited, we need to look at value-for-money signings instead of buying the most famous player. Finally, we also need to consider the position of a player. Players belonging to different positions have different roles and, thereby, have other systems of accumulating points. For example, a striker scores more points by scoring goals, and a goalkeeper scores more points by saving goals. Moreover, there are some constraints as to the number of players we can have at each position.

The total points accumulated by a player in the previous season are used as the parameter to measure performance. Points are awarded for positive accomplishments such as goals and assists, and points are deducted for negative accomplishments such as own goals and red cards. Thus, points perfectly represent the level of performance of a player. These tell us how a player utilizes their game time.

We are looking at the total minutes played as a starter and bench players

throughout the previous season to measure a player's contribution. Through this data, we can determine the fraction of the matches played by a player in the previous season. We are also able to determine the nature of the contribution of a player, i.e., the proportion of times played as a starter and bench players.

Position describes what the role of a player in the team is and where he plays. All the available positions are Striker, Midfielder, Defender, and Goalkeeper. Cost is the sign-on price of the player.

There still are some limitations in our data due to which we are forced to deal with some uncertainties. Firstly, we have no way of determining a player's availability to play on any given day in the future. We cannot determine if a player will get injured or for some other reason might not be available to play on the day of a match. Additionally, we do not have the information about player transfers from individual clubs. Thus, we have no info about new players just joining the clubs or any players that have left or retired since last season. Finally, we also do not have the schedule of matches for the upcoming season of the Premier League. It would have been beneficial to be able to decide on our team based on the difficulty of the opponents that individual players are going to face in their matches, but we cannot know which club will play which club when.

Approach

With years of experience or practice, Raging Bull can try to make a good guess on the selection of players to gain points for a single game. However, the level of complexities involved in making a prediction on the highest points possible or selecting a team for an entire season, given budget constraints, uncertainties around the availability of players, and a large player selection pool requires a more sophisticated approach.

FPL optimization project is modeled using the integer programming optimization methodology majorly because the output of the objective and

decision variables are required to be integers at its optimal solution.

This approach takes the guesswork out of the process to meet the objective which is to maximize player points from the immediate past season by summing up total points for each participated player. The point for each player is calculated by their contributed minutes for the game times total points scored from last season, thus each player's points is scored proportionally based on their contribution duration.

Parameters of the objective function, decision variables, and constraints are formulated with excel and optimized by using Recipe method in Evolver; an application that works as an extension of Excel to find optimal solutions for decision models.

Assumptions

Squad selection is a complex problem to optimize with a lot of uncertainties. It replicates the real-life performances of professional football players to earn points and win against their opponents in the game. In order to reduce the limitations of available data and the level of uncertainties, we have made some assumptions that have aided the development of a base and effective model. Assumptions are as follows:

- Performance from the immediate past season is an indication of performance for the current season.
- Cost of the player remains the same so that end cost of last season can be applied to select an optimal solution given constraint on cost.
- Last season's player contribution has been leveraged for the current season's contribution.
- Bench order and substitution – players with the highest points from the previous season are given priority for substitution and team formation may change upon substitution.
- All players are available to be selected throughout the current season.

Decision Variables

In the designed optimization model, all decision variables are binary variables, that for 462 players from the pool, each of them is applied with three sets of binary selections: starter, bench and captain selection. Starter and bench players together build the 15-player squad. First, each of them is nominated as 1 for being selected as starter player, or 0 for not being selected as starter players. Among the ones chosen as 0 for starter players, another binary calculation is applied as 1 for selecting the player as bench, or 0 for not selecting the player as bench. Lastly, among the players selected as starter team, binary variable is applied for 1 for selecting them as the captain, or 0 for not selecting them as the captain. Only 1 player is selected as captain as the captain will score double points during the game. The Evolver then calculates the best mix of all these binary variables to build the 15-player squad in order to achieve maximum points for the game.

Constraints

The integer optimization model is restricted by certain rules from the League's rule book. First, the total budget for the 15-squad team should not exceed €100 million. Each player is labelled for a cost for trading, and no doubt that a player with outstanding performance is labelled at higher cost. In reality, it is not possible to select the best 15 players as the budget is always limited, thus how to allocate the €100 million budget become critical to the problem.

Next, according to League's rule book, total number of players for starter team should be exactly 11 players, while number of players for bench is 4, that to build up the 15 person squad. Among the starter team, only one person is selected as captain to score double points for the team. One other constraint is that number of players selected from the same club cannot exceed 3 for diversification purpose.

There are also several restrictions on the team formation that need to be

taken into consideration in the model. First, for the final 15-player squad, the formation should be exactly built by 5 defenders, 3 forwards, 5 midfielders and 2 goalkeepers. For the 11-player starter team, number of defenders can be from 3 to 5; number of forwards can be from 1 to 3, number of midfielders can be from 3 to 5, and there should be only one goalkeeper.

After all conditions are captured by optimization model, the Evolver tool calculates the best possible outcome for our model.

Output and Conclusion

The model selects a team that could maximize the total points while following the rules. The model suggests a 3-4-3 starting formation, with three defenders, Andrew Robertson, Aaron Cresswell and Matt Targett; four Midfielder Bruno Fernandes, Son Heung-Min, Stuart Dallas and James Ward-Prowse; and three forwards, Harry Kane, Patrick Bamford and Ollie Watkins. Emiliano Martínez is selected as a goalkeeper in the starter lineup. The average cost is €7.3 million for each starter, which could potentially receive an average of 186 points each. The bench players consists of goalkeeper John Stones, defender Antonio Rüdiger, defender Martin Dúbravka and midfielder Hélder Costa. The average cost is €4.9 million for each bench player and could receive 68 points. The Midfielder Bruno Fernandes from Manchester United, who scored the highest points, 244 last season, is also selected as the team captain. He would cost the company €11.3 million and potentially contribute double points 488 to the final score. Fifteen players chosen from 462 are from 10 different clubs among the 20 clubs in the premium league. The model picks two players each from Aston Villa, Newcastle, Spurs and one player each from Chelsea, Liverpool, Man City, Man Utd, Southampton and West Ham.

The model had automatically allotted the most budget on the starter lineup to aim for higher points. All players in the staters' team are ranked top 10

in their position. As a result, the squad costs precisely €100 million, and it could potentially score 2,332 points. Overall the model performs exceptionally well to use all the budget and select the best team with limited resources.

Next Steps

In conclusion, the virtual Fantasy Premier League environment is a highly dynamic system that is rapidly changing. The current model is still facing some challenges and limitations. The output is based on historical data from the last season. It is difficult to predict the players' performance based on limited information; the budget is low for the first team selection of the season, and the model is not yet considering the rules of transfers for future designation. There are a few more steps to furthermore developing the model. Firstly, feed the real-time data as input to the model to keep it up to date. Secondly, increase the budget by trading players to purchase players with higher points. Thirdly, keep up with players' conditions such as penalties, health and injuries. Finally, build a model to predict players' performance in this season.

Appendix

Mathematical Model

Decision Variable:

S_i = 1 if selecting the player for starter,

0 if not selecting the player for starter;

B_i = 1 if selecting the player for bench,

0 if not selecting the player for bench;

C_i = 1 if selecting the player for captain,

0 if not selecting the player for captain;

Where i is the name for each individual players from the pool.

Objective: to maximize total points

Total Points = $\sum$(Starter_min_i * S_i + Bench_min_i *B_i + Starter_min_i*C_i)

Where i is the name for each individual players from the pool.

Constraints:

Budget: $\sum$(Cost_i*S_i + Cost_i*B_i) <= 100

Number of players in starter lineup: $\sum$(S_i) = 11

Number of Captain: $\sum$(C_i) = 1

Number of players from each club: $\sum$(club*S_i + club*B_i) <= 3 for each club

Number of defender for the squad: $\sum$(defender*S_i + defender*B_i) = 5

Number of forward for the squad: $\sum$(forward*S_i + forward*B_i) = 3

Number of midfield for the squad: $\sum$(midfield*S_i + midfield*B_i) = 5

Number of goalkeeper for the squad: $\sum$(goalkeeper*S_i +

goalkeeper*B_i) = 2

Number of defender for starter lineup: 3<= ∑(defender*S_i) <= 5

Number of forward for starter lineup: 1<= ∑(forward*S_i) <= 3

Number of midfield for starter lineup: 3<= ∑(midfield*S_i) <= 5

Number of goalkeeper for starter lineup: ∑(goalkeeper*S_i) = 1

Evolver Model:

Decision Variable Output:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	id_player	first_name	second_name	club	position	end_cost	Total Min	Total_Game_Min	Starter_Min	Bench_Min	total_points	Starter(S_i)	Bench(B_i)	captain	costs	points
2	277	Bruno Miguel	Borges Fernandes	Man Utd	Midfielder	113	3,101	450	408	42	244	1	0	1	113	442
3	357	Harry	Kane	Spurs	Forward	119	3,083	450	406	44	242	1	0	0	119	218
5	359	Heung-Min	Son	Spurs	Midfielder	96	3,119	450	410	40	228	1	0	0	96	208
6	189	Patrick	Bamford	Leeds	Forward	66	3,052	450	402	48	194	1	0	0	66	173
8	30	Emiliano	Mart√≠nez	Aston Villa	Goalkeeper	53	3,420	450	450	-	186	1	0	0	53	186
11	188	Stuart	Dallas	Leeds	Midfielder	55	3,410	450	449	1	171	1	0	0	55	171
12	40	Ollie	Watkins	Aston Villa	Forward	63	3,328	450	438	12	168	1	0	0	63	163
14	234	Andrew	Robertson	Liverpool	Defender	73	3,384	450	445	5	161	1	0	0	73	159
19	341	James	Ward-Prowse	Southampton	Midfielder	59	3,420	450	450	-	156	1	0	0	59	156
22	411	Aaron	Cresswell	West Ham	Defender	57	3,170	450	417	33	153	1	0	0	57	142
36	39	Matt	Targett	Newcastle	Defender	50	3,404	450	448	2	138	1	0	0	50	137

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	id_player	first_name	second_name	club	position	end_cost	Total Min	Total_Game_Min	Starter_Min	Bench_Min	total_points	Starter(S_i)	Bench(B_i)	captain	costs	points
57	252	John	Stones	Man City	Defender	51	1,933	450	254	196	128	0	1	0	51	56
130	127	Antonio	R√ºdiger	Chelsea	Defender	47	1,710	450	225	225	93	0	1	0	47	47
204	192	H√©lder Wander	Sousa de Azevedo e Costa	Leeds	Midfielder	50	1,150	450	151	299	63	0	1	0	50	42
238	295	Martin	Dubravka	Newcastle	Goalkeeper	48	1,170	450	154	296	48	0	1	0	48	32

Objective & Constraints Output:

				Constraint		
	Objective	**Total Points**	2,332			
	Constraint on buddget	**Total Costs**	1000	1000		
	Constraint on Captain	**Number_Captain**	1	1		
Constraint on Position	**starter**	**bencher**	**squad**	**min starters**	**max starters**	**squad**
Defender	3	2	5	3	5	5
Forward	3	0	3	1	3	3
Goalkeeper	1	1	2	1	1	2
Midfielder	4	1	5	3	5	5
total	11	4	15			
Limit	11		15			

team	starter	bencher	squad	limit
Arsenal	0	0	0	3
Aston Villa	2	0	2	3
Brentford	0	0	0	3
Brighton	0	0	0	3
Burnley	0	0	0	3
Chelsea	0	1	1	3
Crystal Palace	0	0	0	3
Everton	0	0	0	3
Leeds	2	1	3	3
Leicester	0	0	0	3
Liverpool	1	0	1	3
Man City	0	1	1	3
Man Utd	1	0	1	3
Newcastle	1	1	2	3
Norwich	0	0	0	3
Southampton	1	0	1	3
Spurs	2	0	2	3
Watford	0	0	0	3
West Ham	1	0	1	3
Wolves	0	0	0	3