

Can a team picked by machine learning beat an average player's team in Fantasy Premier League?

By Marko Mylyanyk

Introduction

1

Fantasy Premier League is a game, where participants form virtual team of football players, which prototypes are playing in real world and based on their performance get points for each gameweek. This project is developed with aim to pick a team of footballers, based on their recent performance, which will beat an average players score.



Methodology

2

The methodology implies three steps:

- 1st : Explore and prepare the data.
- 2nd: Create an algorithm that would predict future players performance.
- 3rd: Pick the best team based on results

Working process

3

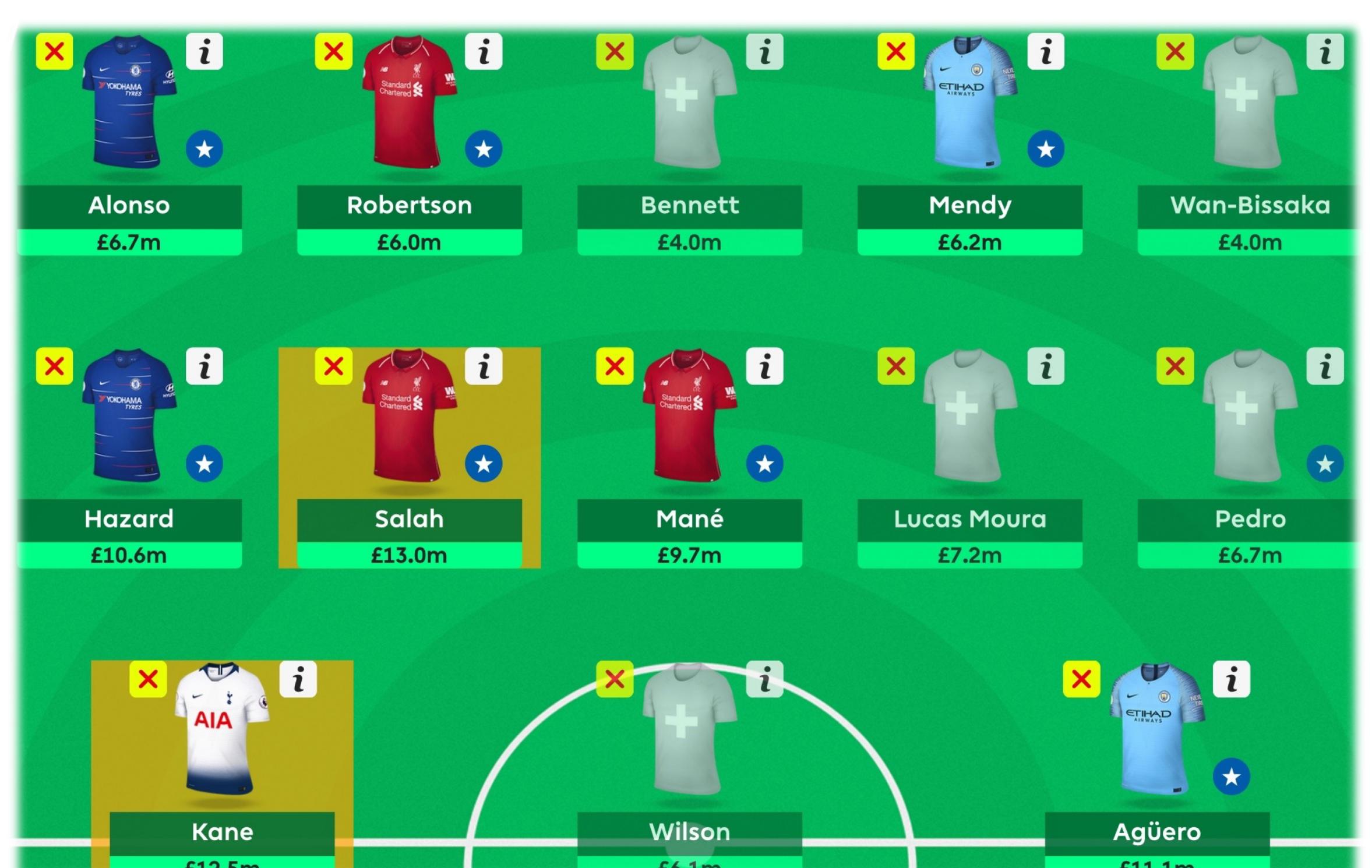
Data exploration:

- Particular dataset if of size 700 rows by 174 columns. It consist all information about each Premier League player for past three years.

Data preparation:

- Understand the data.
- Investigate which parameters are most important.
- Deal with missing values

After this steps, I understood that in this case, outliers is something what I am looking for. Players that has the best correlation between their cost and total points is should be in final team.



The next step was to predict future players results using information from three previous seasons. For this I have not seen another solution, except a model similar to sequence prediction.

First approach – Linear regression:

I understood that the data I used behave similar to random, which is very hard for linear model to approximate.

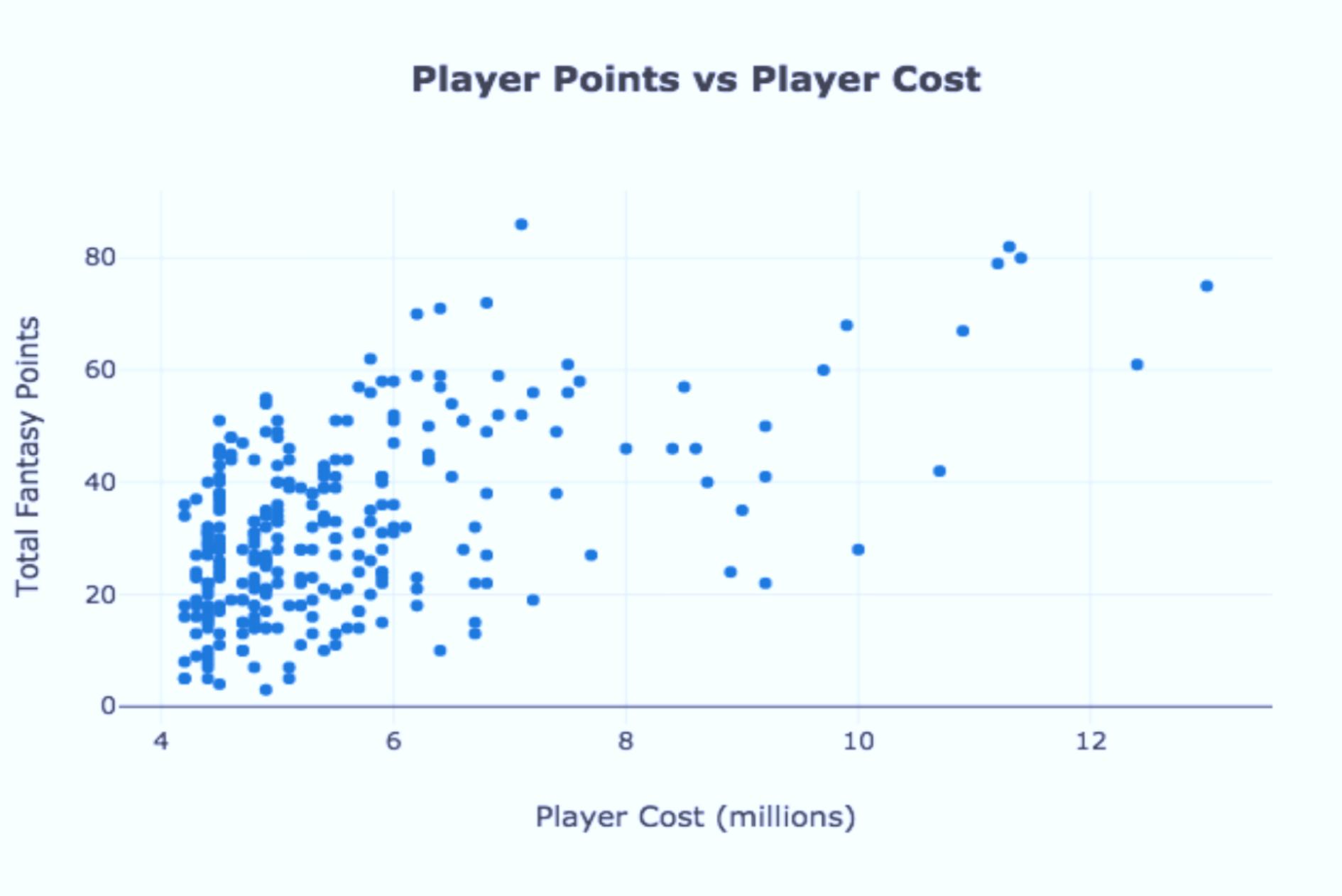
Second one - Long Short Term Memory network:

I have created a simple LSTM using Adam optimizer and Mean Squared Loss function. This one worked fine.

Final step was to pick a team of 2 goalkeepers, 5 defenders, 5 midfielders and 3 forwards, not exceeding budget of \$100.

Results

4



Conclusions

5

Real generated squad in a FPL game



As now is only the middle of a football season, the exact results of my project will be received only in few month. In the end of this season, I will find out if picked by my program team could beat at least half of all 7 millions active users.

By now, I could only estimate this team by my own standards, and for me it looks like it could get in top 25% of all teams.

```
/usr/local/bin/python3.7 /Users/marko/PycharmProjects/FPL_prediction/model.py
[('Joshua King', 'Danny Ings', 'Raúl Jiménez')
 ('Ryan Fraser', 'James McArthur', 'Christian Eriksen', 'Raheem Sterling', 'Heung-Min Son')
 ('César Azpilicueta', 'Nathan Ake', 'Andrew Robertson', 'Patrick van Aanholt', 'John Lundstram')
 ('Ben Foster', 'Jordan Pickford')]

Process finished with exit code 0
```