

Predicting good strategies for picking players in Fantasy premier league

PRANEET NANDAN AND ADARSHA MOHIT SAHU

Experiment

First we implemented a simple SVM to the model, as a baseline, and also to get insight into the direction we should move in

Taking just the Game-week data, the team selected by percentage and the ict index, and asking the machine to predict which players would score 2 or more points (the player actually played, even if he didn't score or do something amazing), we got

Accuracy: 0.8478260869565217

Precision: 0.6595744680851063

Recall: 0.7209302325581395

We don't bother about the recall but the precision is important for us and it is low, we tried it with the rbf kernel

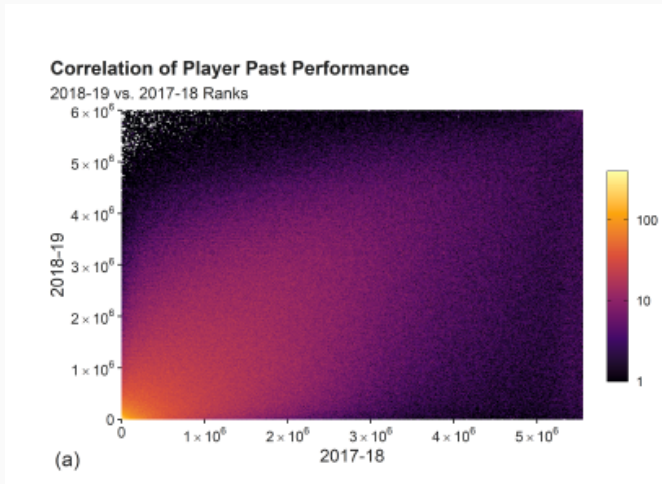
Accuracy: 0.875

Precision: 0.7288135593220338

Recall: 0.86

In search of additional features that could help train the model better, we landed on

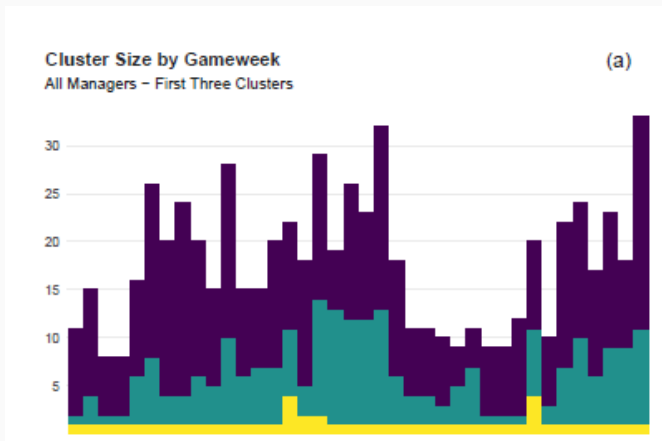
- O'Brien JD, Gleeson JP, O'Sullivan DJP (2021) Identification of skill in an online game: The case of Fantasy Premier League. PLoS ONE 16(3): e0246698. <https://doi.org/10.1371/journal.pone.0246698>



The managers who performed well in past seasons are much more likely to perform well this season

Literature

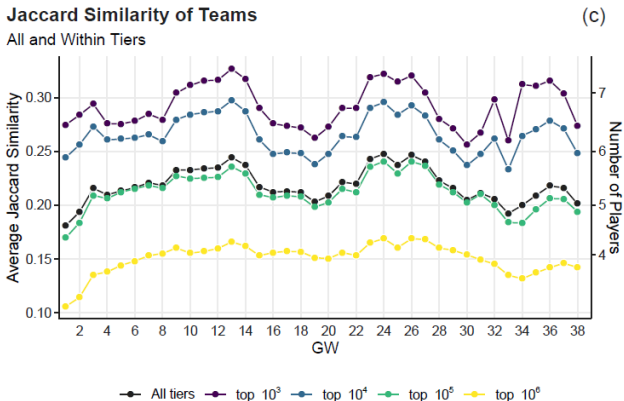
In the paper they created a network of players, each edge having weight equal to the number of managers who have both the players in squad
Then they applied hierarchical clustering with 4 clusters



Literature

Then they calculated the jaccard similarity between managers in different tiers

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|} \quad (1)$$



Experiment

So then we scraped data about last year's winners, and this year's leading managers and the composition of their teams, getting a new team selected by percentage that would not be affected by noobs,

Accuracy: 0.8804347826086957

Precision: 0.7413793103448276

Recall: 0.86

That is a very slight improvement, but not too much.

The real magic happened if we change the threshold to 5, to find real difference makers

While earlier, in this setting the model was predicting that all the players in the testing set will score below 5, this time it did predict some players and half of those actually did do great

Accuracy: 0.9456521739130435

Precision: 0.5

Recall: 0.2

Experiment

Another aspect of the scrapped data from FPL site was the team strength.

FPL gives a rating to each team at the start of the season that is supposed to signify how strong it is. But the problem with this metric is that it doesn't change with the season, even if the team is performing very bad, having problems with players or managers and other things. Therefore we scrapped oddsportal.com and got betting odds for the gameweek, something that is very dynamic and changes every week and replaced that metric with this.

Threshold 2

Accuracy: 0.8913043478260869

Precision: 0.75

Recall: 0.8478260869565217

Threshold 5

Accuracy: 0.9130434782608695

Precision: 0.75

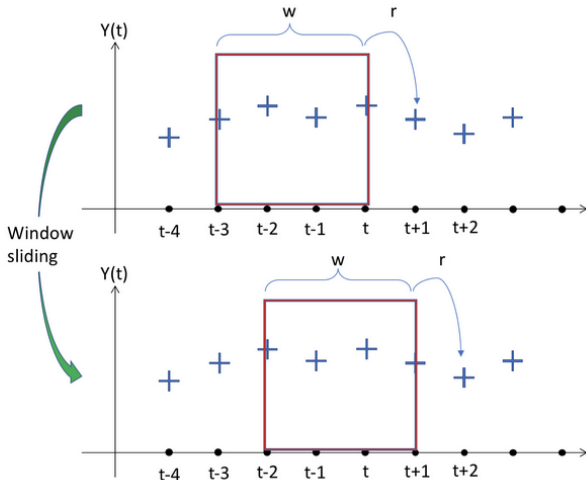
Recall: 0.16666666666666666

Problems

- The features and the data are in a time-series like fashion, even though there are fluctuations, there is some correlation between performance in different game weeks that SVM is not able to capture
- A lot of the data points are players who don't get picked, being 0. The players who get picked and the players who score are who we should be focusing on

Plan A for improvement

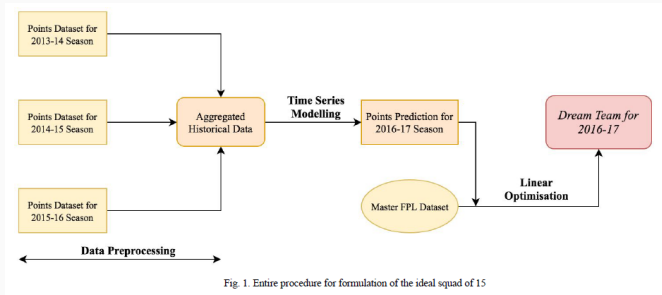
We plan to make windows of data, so that the time series nature can be captured well



Back to Literature

- Gupta A (2019) Time Series Modeling for Dream Team in Fantasy Premier League. International Conference on Sports Engineering (ICSE'17) CoRR abs/1909.12938

They have used LSTM neural networks to study performance over different seasons, and to predict for one season. We want to do that for different game-weeks



References

- Bonello N, Beel J, Lawless S, Debattista J. (2019). Multi-stream Data Analytics for Enhanced Performance Prediction in Fantasy Football.
- Gupta A (2019) Time Series Modeling for Dream Team in Fantasy Premier League. International Conference on Sports Engineering (ICSE'17) CoRR abs/1909.12938
- O'Brien JD, Gleeson JP, O'Sullivan DJP (2021) Identification of skill in an online game: The case of Fantasy Premier League. PLoS ONE 16(3): e0246698. <https://doi.org/10.1371/journal.pone.0246698>