Recognition of Patterns of Play in Rugby Union

Name: Kirk Meyer
Student No.: MYRKIR006
Supervisor: Neil Watson

29 May 2020

1 Background

Recognition of patterns of play in rugby union is an interesting and challenging problem. This is due to the nature of the game, which consists of multiple possessions between two teams. A single possession can be broken down into a sequence of phases, with each phase consisting of multiple actions. A pattern of play may be defined within one phase, several phases or even an entire possession. What defines a pattern of play is a sequence of certain actions in a particular order, with the associated field position and identified players.

The recognition of patterns of play or tactical structure is a problem that has been tackled in other sports through various methods. Teich, Lutz, and Kassarnig (2016), investigated predicting the success of plays in the NFL. Their focus was on real-time prediction so that coaches can make decisions in-game. The "accuracy" of each model was based on several measures of error. Overall, their best performing models were Decision trees and Support Vector Machines using a Radial Basis Function kernel. Their exploration into Neural Networks was limited, as they didn't attempt to find an optimal network configuration. They also stated that the lengthy running times, defeated the purpose of real-time in-game predictions. The use of Neural Networks, to recognise plays in basketball was the focus of Wang and Zemel (2016). Using standard neural networks and recurrent neural networks they were able to classify plays in fraction of the time a human can, with relative success. Pfeiffer and Perl (2006), analysed the tactical structures of handball and were able to classify different patterns in offensive plays. An integral part was the use of a process-orientated model. This allowed sequential data to be analysed, providing deeper insight into tactical structures. Recently, Coughlan, Mountifield, Sharpe, and Mara (2019) used clustering to identify which patterns of play lead to tries in rugby.

Previous research has demonstrated that there is potential to recognise patterns of plays in sports. The purpose of this project is to use various modeling techniques to identify patterns of play in rugby union.

2 Objective

The objective of this project is to identify several patterns of play in rugby union, through use of a Recurrent Neural Network (RNN). The use of other modelling methods will also be explored . Furthermore, we will attempt to find which pattern of plays leads to an increase in probability of winning matches by comparing patterns of play of winning teams against that of losing teams.

3 Data

The dataset is large and consists of game data from various competitions (Heineken Cup, Champions Cup, Six Nations, Rugby Championship and Super Rugby). Each observation has multiple variables, each of which is described in the table below:

Variable Name	**Variable Description**
id	Identification Number of Observation
fx_id	Idenfication Number of Fixture
prd	The Half (First or Second)
pl_id	Unique Player Identification Number
tm_id	Unique Team Identification Number
time	Time passed in half (in seconds)
act	The Group Qualifier
act_type	Action Qualifier
act_res	Result of Action Qualifier Description
q3, q4, q5	Further Qualifier Descriptions
m	Meters gained during observation
x_crd and y_crd	X and Y coordinate at start of observation
x_end and y_end	X and Y coordinate at end of observation
score_adv	Score Advantage in terms of Home Team
play_num	Number of Phase in current possession

Table 1: Feature Name and Descriptions

The data is rich - the qualifier variables have over four hundred factor levels or descriptors. The data is spatiotemporal, and thus provides all the key information needed to build a model to recognise patterns of play. Data pre-processing forms an integral part of this project, where the labelling of the data is especially important to facilitate the use of supervised learning methods. Hence, data wrangling will be one of the main focuses of the project.

4 Proposed Methods

Since the data is spatiotemporal in nature, the focus will be on building a recurrent neural network to classify the different patterns of play. Recurrent Neural Networks are designed to model sequential data. While they, are similar to standard feed forward neural networks in that they have an input layer, a specified number of hidden layers and an output layer, they differ in that they have internal memory. The hidden layers, take in a new information as well as, input from the previous observation. This allows for sequences of events to be classified. The classification performance of other supervised, and unsupervised methods will be explored. This comparison between the RNNs and other machine learning methods, will be of interest when considering the spatial nature of the data.

5 Timeline

- 30 June: First draft of Literature review.
- 15 July: Have completed Literature review and continuously update.
- 21 July: Have enough data processed and labelled to train Neural Net.
- 31 July: Have first iteration completed of RNN, with some sort of results.
- 3 August: Hand-in Progress Report
- 11 September: Have final version of RNN, start exploring alternative methods and secondary objective.
- 12 October: First Draft of Project
- 22 October: Project Presentation
- 6 November: Complete Project
- 9 November: Hand-in Final Project
- 1-10 December: VIVA

References

Coughlan, M., Mountifield, C., Sharpe, S., & Mara, J. K. (2019). How they scored the tries: applying cluster analysis to identify playing patterns that lead to tries in super rugby. *International Journal of Performance Analysis in Sport*, *19*(3), 435–451.

Pfeiffer, M., & Perl, J. (2006). Analysis of tactical structures in team handball by means of artificial neural networks. *International Journal of Computer Science in Sport*, *5*(1), 4–14.

Teich, B., Lutz, R., & Kassarnig, V. (2016). Nfl play prediction. *arXiv preprint arXiv:1601.00574*.

Wang, K.-C., & Zemel, R. (2016). Classifying nba offensive plays using neural networks. In *Proceedings of mit sloan sports analytics conference* (Vol. 4).