

# The use of match-based exact movement activities to classify elite rugby league players into positional groups

Victor Elijah Adeyemo (✉ [v.adeyemo@leedsbeckett.ac.uk](mailto:v.adeyemo@leedsbeckett.ac.uk))

Leeds Beckett University

Anna Palczewska

Leeds Beckett University

Ben Jones

Leeds Beckett University

Dan Weaving

Leeds Beckett University

---

## Research Article

**Keywords:** player movement profiling, movement patterns, rugby league positional groups, classification

**Posted Date:** October 13th, 2023

**DOI:** <https://doi.org/10.21203/rs.3.rs-3424741/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

**Additional Declarations:** No competing interests reported.

---

# The use of match-based exact movement activities to classify elite rugby league players into positional groups

Victor Elijah Adeyemo<sup>1,2,3,4\*</sup>, Anna Palczewska<sup>1</sup>, Ben Jones<sup>2,3,4,5,6</sup>, Dan Weaving<sup>2</sup>

<sup>1</sup>School of Built Environment, Engineering and Computing,, Leeds Beckett University, Leeds, UK.

<sup>2</sup>Carnegie School of Sport, Leeds Beckett University, Leeds, UK.

<sup>3</sup>England Performance Unit, Rugby Football League, Leeds, UK.

<sup>4</sup>Leeds Rhinos Rugby League Club,, Leeds, UK.

<sup>5</sup>School of Science and Technology, University of New England, Australia.

<sup>6</sup>The University of Cape Town and the Sports Science Institute of South Africa, South Africa.

\*Corresponding author(s). E-mail(s): [v.adeyemo@leedsbeckett.ac.uk](mailto:v.adeyemo@leedsbeckett.ac.uk);

Contributing authors: [a.palczewska@leedsbeckett.ac.uk](mailto:a.palczewska@leedsbeckett.ac.uk); [b.jones@leedsbeckett.ac.uk](mailto:b.jones@leedsbeckett.ac.uk);  
[d.a.weaving@leedsbeckett.ac.uk](mailto:d.a.weaving@leedsbeckett.ac.uk);

## Abstract

The cluster analysis of elite rugby league players identified groups of distinct playing positions, that can be referred to as broad positional groups. However, the identified positional groups were based on traditional indicators (physical and technical-tactical) that provided no information about the exact match-based movement activities that led to such similarity grouping and the classification of elite rugby league players into these broad positional groups remains unexplored. Hence, this study finds the best model to classify elite rugby league players into positional groups, using data characterised by movement patterns to uncover the similar movement activities of distinct playing positions within a positional group. Key movement patterns for the positional group classification and differences between the groups were also investigated. A total of 18,173 unique movement patterns were derived from 422 players' GPS data across the 2019 and 2020 seasons, where only 36 were identified as key patterns. The highest classification accuracy of 77.58% using all unique patterns and 74.5% accuracy using the key patterns was achieved, outperforming studies that used traditional indicators. Further analyses, based on key patterns revealed differences between forwards and backs. These findings establish movement patterns as viable indicators to classify rugby league players into positional groups, enabling coaches and trainers to develop position-specific training programs that cater to the unique physical demands of each position, leading to better player development and team performance. Movement patterns are therefore recommended as an alternative approach to quantifying players' external loads and obtaining granular information.

**Keywords:** player movement profiling, movement patterns, rugby league positional groups, classification

# 1 Introduction

In sports science, and in rugby league football, traditional (i.e., physical and technical-tactical) performance indicators that are used to characterize match games [1, 2] are widely used as variables for various performance analysis problems. In [3], cluster analysis was primarily carried out on data characterised by traditional indicators (e.g., high-speed running distance and total tackles) to determine the similarity between playing positions (i.e., fullbacks, centres, half-backs, wingers, hookers, prop-forwards, second-rows and loose-forward) of professional rugby league. For each data characterised by technical, physical and combined indicators [3], various cluster sizes (i.e., 3, 4 and 3) were later identified respectively where players of different playing positions were grouped into the clusters, based on top-ranked important indicators. This phenomenon is in tandem with the ground truth [4] that broad positional groups, i.e. forwards and backs (comprising distinct playing positions), exist in professional rugby and that players of some distinct playing positions perform similar activities. Although broad positional groups are known to exist in professional rugby league, the classification of players into these groups remains unexplored while other classification problems [5–7] such as injury (status or type) prediction and competition levels classification are thoroughly researched.

Data characterised by traditional indicators has the capacity to classify players into positional groups. However, insight into similar movement activities performed by players of different playing positions within and between positional groups will be inaccessible. This limitation arises because physical performance indicators [8], primarily derived from wearable Global Positioning Systems (GPS) devices [9], offer little to no context or explanation on how players accumulate measured physical activities. Physical performance indicators [10] (e.g., total distance covered, maximum momentum, and high-speed running distance) are often aggregated or reported in volume and without regard to the order of match activities or events occurrences. Player movement profiling [8, 10] now offers alternative indicators (i.e., movement patterns - also extracted from GPS data) that capture sequential match activities completed by players, which provide the

context and explanation lacking in traditional indicators.

In this study, we propose the use of data characterised by movement patterns for elite rugby league players’ positional group classification as well as for uncovering the movement activities of players within the same positional group. This will assist in identifying sets of match-based sequential movement activities performed by players of both groups (which is useful for talent identification, training design and load management [8, 10]), measure the extent to which exact match-based movement activities can separate players into rugby league positional groups, establish the on-field movement activities of players required to profile the performance of players in each group and help develop predictive models to classify new players into each group among others benefits.

Therefore, using data characterised by movement patterns, this study aims to find the best machine learning model to classify elite rugby league players into forwards and backs, identify the key movement patterns necessary for such classification, uncover the similar movement activities of different playing positions within a positional group and quantify the difference in players’ performance of key movement activities between positional groups.

This paper is organized with Section 2 providing background and related works; Section 3 gives details about GPS data conversion to sets of discrete movement sequences, movement pattern extraction, classification data generation, and data analyses. Section 4 presents the result of GPS data conversion and data analysis. Finally, in Section 5, the results are discussed from the implementation of the proposed framework and Section 6 concludes the study with relevant future works.

## 2 Related Works

Traditional performance indicators are widely used to classify players into different sporting aspects but related studies for rugby positional group classification are scarce. In [2], 82 elite rugby league players of both forward and back positional groups were classified into competition levels (i.e. under-19 Academy and European Super League (ESL)) using 157 physical and technical-tactical performance indicators. Forwards were classified into competition levels at 68% accuracy while

backs were classified at 83% accuracy [2] by the random forest classification models. These physical, technical and tactical performance indicators are limited [8, 24] such that they are aggregated series of completed activities reported in arbitrary volume(s) e.g., total distance covered, average match-speed, line break and defensive play-the-ball loss or win.

In a more related study [28], 211 elite junior Australian footballers were classified into positional groups (i.e., forward, defence, ruck or mid-field) using 12 technical skill performance indicators. Most players were classified as midfielders by the linear discriminant analysis (LDA) model with an accuracy of 56.8% among others. The use of physical and technical-tactical indicators for competition classification [2] and only technical skill indicators for positional group classification [28] produced a sub-optimal classification model, especially for positional group classification.

Besides the classification of sporting aspects, the identification of key independent variables in sports science [2, 14] is often carried out as a solution to overcome the high dimensional and multicollinearity of independent variables. Few key variables (performance indicators in most cases) are often identified as needed for practical utility. For instance, study [17] used a repeated-measures analysis of variance to identify key variables in its assessment of the difference in external loads of two competitive levels (i.e., the National Rugby League and National Youth Competition) of rugby league players. However, this method of multiple univariate analyses of independent variables is computationally expensive as multiple analyses are executed.

In [2], the machine learning algorithm variable importance method was used to identify and select the independent variables that are important to the predictive model for the initial classification of players into competition levels. The limitation of this method includes the identification of independent variables that are suboptimal [25] and biased (specific) towards the underlying machine learning algorithm. The application of feature selection techniques [3], especially the correlation-based filter feature subset-selection method [14], has proven to be an alternative method for identifying key independent variables in sports analytics.

In [14], the correlation-based filter feature subset-selection method was used to optimise

[2] classification analysis and also to identify key physical and technical-tactical indicators for classifying rugby league players into competition levels. The identified key physical and technical-tactical indicators are more useful practically than all indicators and they also produced an increased accuracy for competition levels classification (i.e., 84.55% from 82% for forwards and 77.42% from 67.5% for backs). Additionally, based on the identified key physical and technical-tactical indicators, the difference between competition levels was assessed [14] using Cohen's effect size.

Due to traditional indicators' limitation [8], movement patterns emerged as alternative indicators providing the exact match-based sequential locomotive activities of players. Recently, a Sequential Movement Pattern-mining (SMP) framework was developed for player movement profiling [24] and validated on rugby league GPS data. The framework was reported to be the better, more stable and more robust method to discretise GPS data into movement units, form (discrete) movement sequences and identify the longest common movement patterns of elite rugby league players within twenty-five clusters respectively. The SMP framework was used by another study [16] to explore the differences in the movement patterns of elite rugby league players among three competition levels (i.e. Super League regular, Super League (semi-) finals and international rugby league matches). However, the study found no difference in the movement patterns of players among the competition levels as extracted by the SMP framework. But a further analysis, using a three-stepwise Linear Discriminant Analysis process for dimension reduction and classification, revealed an 81% accuracy for classifying elite rugby league players into competition levels.

In [13], three distinct pattern mining algorithms (i.e., the Longest Common Sequence algorithm of the SMP framework [24],  $l$ -length closed contiguous sequential pattern mining algorithm (LCCspm) [10] and AprioriClose [20]) was used to extract different types of obtainable movement patterns to profile and classify rugby league players into two distinct playing positions (i.e., hookers and wingers) as a means to identify and thus validate the best pattern mining algorithm (and type of movement pattern) for player movement profiling. The results of [13] revealed closed contiguous movement patterns extracted by LCCspm

as the best type of movement pattern for profiling and classifying elite rugby league players into selected tactical roles at 91.02% accuracy based on a Multi-Layered Perceptron model.

Overall, existing studies provided various machine learning algorithms useful for classifying sporting aspects and accuracy as a go-to performance evaluation metric. It also revealed the extent movement patterns are being applied within sports analytics, particularly in rugby league, which is for competition level and two distinct playing position classifications. However, the use of movement patterns for classifying elite rugby players into forward and back positional groups has not been previously explored despite its potential usefulness in practice. Also, the identification of the key movement patterns required for classifying players into positional groups is expedient. To identify key independent variables, the use of a feature selection algorithm is most suitable and the use of classification model(s) feature importance should be avoided. Also, Cohen’s effect size can be utilised to assess the difference(s) between groups characterised by the same variables. Based on all findings, an experimental framework that addresses the identified revelations is proposed in Section 3.

### 3 Method

The experimental framework depicted in Fig. 1 illustrates an overview of data collection, processing and analysis. To achieve the aim of this study, this study utilised the SMP framework [24] to convert GPS data to sets of discrete movement sequences, used LCCspm algorithm [10] to extract movement patterns from sets of discrete movement sequences, and conducted data analyses in two phases. The first phase of data analysis involved the creation of classification input data using movement all unique movement patterns as independent variables and finding the most accurate model for positional group classification. The second phase involved identifying key movement patterns for classification, re-training models for positional group classification and comparative analysis of models using all unique or key movement patterns for positional group classification.

#### 3.1 GPS Data Conversion to Sets of Discrete Movement Sequences

Player tracking data of elite rugby league players that participated in The Rugby Football League (RFL) competitions over two seasons (i.e. 2019 and 2020) collected via Global Position System [GPS] micro-sensor units (Optimeye S5, Catapult Sports, Melbourne, Australia) per player per fixture were used. The method published by [24] was followed to process the GPS data of elite rugby league players (i.e., 239 forwards and 203 backs) from 11 teams (i.e. Huddersfield Giants, Hull FC, Warrington Wolves, Leeds Rhinos, Hull Kingston Rovers, Wakefield Trinity, Salford Red Devils, Wigan Warriors, St. Helens, Castleford Tigers and Catalans Dragons) that participated in 349 fixtures into discrete movement sequences. Elite rugby league forwards consist of players that participated in any fixture as a prop-forward, hooker, second-row, and lock-forward/loose-forward. Elite rugby league backs consist of players that played in any fixture as a fullback, centre, five-eighth, half-back, and winger.

To generate sets of movement sequences, per player per fixture, players’ movement units were first derived through the discretization of velocity, acceleration and turning angle values into qualitative locomotive labels based on some thresholds published by White et. al. [24] SMP framework. Velocity locomotive labels are: walk, jog, run, and sprint. Acceleration locomotive labels are: acceleration, neutral and deceleration. Turning angles locomotive labels are: straight, acute-change, large-change, and backwards. Hence, movement units are combinations of all (i.e.,  $4 \times 4 \times 3 = 48$ ) possible locomotive labels encoded by lower and upper-case alphabets (see Table 2 for examples). For each player per fixture, GPS data are processed into a sequence of movement units (e.g., uuvvvuSSHGHopdK) and later split into subsequences (i.e. a set of discrete movement sequences) after removing movement units representing the time of locomotive inactivity on the field (i.e., velocity is less than 1.2 m/s). Each set of discrete movement sequences was generated at player-per-fixture granularity. A total of 10,811 sets of discrete movement sequences were generated for both elite rugby league forwards and league backs.

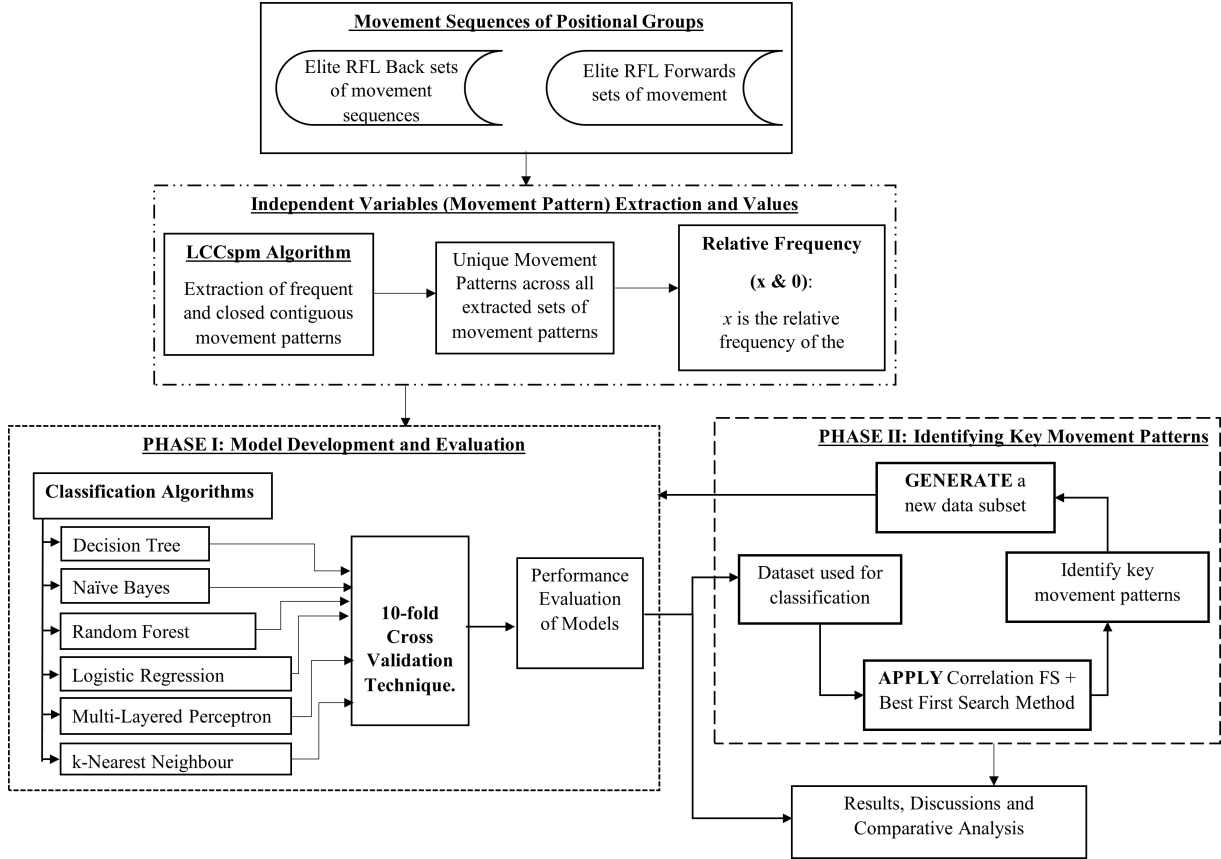


Fig. 1 Experimental Framework

Table 1 Performance metrics extracted from the confusion matrix

| Metric    | Equation  | Description  |
|-----------|---|--|
| Precision | $\frac{TP}{TP+FP}$  | The proportion of correctly identified forwards/backwards based on actual and incorrectly predicted forwards/backwards |
| Recall    | $\frac{TP}{TP+FN}$  | The proportion of correctly identified forwards/backwards based on all forwards/backwards.                             |
| F1-Score  | $2 \times \frac{Precision \times Recall}{Precision + Recall}$ | The weighted harmonic mean of both precision and recall  |
| Accuracy  | $\frac{TP+TN}{TP+FP+TN+FN} \times 100$                        | The percentage of forward and back players that are correctly predicted as forward and back players.                   |

True Positive (TP), False Positive (FP), True Negative (TN) and False Negative (FN).

### 3.2 Movement Pattern Extraction and Classification Input Data

The  $l$ -length closed contiguous pattern mining algorithm (i.e., LCCspm) [10] was applied to extract closed contiguous movement patterns from each set of generated movement sequences for both rugby league positional groups. The  $l$  parameter of the LCCspm algorithm was set to 20 (i.e. 2 seconds timeframe) to ensure long closed contiguous movement activities of players are extracted. The *support* parameter of the LCCspm algorithm was set to 5% to ensure a large number of frequent

movement patterns can be identified as a higher support value will identify few frequent movement patterns.

To generate the dataset for classification modelling, the unique movement patterns extracted across all sets of discrete movement sequences were identified by finding the union of all extracted movement patterns. This set of unique movement patterns was used as independent variables for the classification modelling input data. The values of each unique movement pattern in the dataset are "x" or "0", where "x" is the relative frequency of



the performed movement pattern within a fixture and 0 if the movement pattern was not performed. This classification input data is referred to as “full dataset”.

### 3.3 Data Analyses Phases

The framework shown in Figure 1 captures the two phases of data analysis conducted. In phase I, machine learning classification algorithms were utilised to develop models for classifying rugby league players into groups, to measure how well movement patterns can classify players into groups and find the best model. Six classification algorithms (i.e., Decision Tree, Naive Bayes, Random Forest, Logistic Regression, Multilayered Perceptron, and k-Nearest Neighbours) were chosen based on their distinct learning method as well as for identifying the best model through comparison of performance. The default parameters of each algorithm were used except for the  $k$  parameter of the k-Nearest Neighbour algorithm was set to 5 and MLP whose hidden layer sizes were set to 1000, 500, 250, 125, 50 and 25.

The classification modelling task is to predict whether an elite rugby league player (per fixture) is a forward or back. Using the data generated in Section 3.2, classification models were developed via 10-fold cross-validation (Figure 1). The 10-fold cross-validation technique [15] splits data into ten partitions, nine partitions are used for training models while the remaining partition is utilized for testing. It is repeated 10 times until each partition is used for testing the models. The result of all cross-validated models is aggregated during performance evaluation. Classification models are evaluated [11] using the following evaluation metrics: accuracy, precision, recall and f1-score (Table 1).

In phase II, key movement patterns necessary for the classification of players into positional groups were identified. From Section 2, the ‘Correlation-based feature subset’ (Cfs) feature selection method was identified as a useful feature selection method in sports analytics (See [14] for details of the algorithm). It was selected and applied to the dataset containing all unique movement patterns as independent variables to identify key (movement patterns) variables. Another dataset was extracted based on the outputted movement patterns subset, identified through this

process and referred to as the ‘reduced dataset’, which was used for re-classification modelling.

Upon identifying the key movement patterns, further analysis that assessed if there were significant differences between players of both positional groups when performing those patterns across matches was executed. The analysis was carried out using the following Cohen’s effect size thresholds: trivial ( $\leq 0.1$ ), small ( $\leq 0.2$ ), moderate ( $\leq 0.6$ ), large ( $\leq 1.2$ ), very large ( $\leq 2.0$ ), nearly perfect ( $\leq 4.0$ ) and perfect 4.0). Also, the percentage average of identified key movement patterns performed by players per positional group was assessed and presented.

The analysis for the application of LCCspm for extraction of movement patterns, finding the unique set of patterns, classification modelling, and statistical analysis were carried out using the Python programming language on a private cloud Linux-based server. The source can be found on GitHub [12]. The libraries (i.e., Scikit Learn and LCCspm) were imported and utilized. Feature selection was conducted using Waikato Environment for Knowledge Analysis (WEKA) GUI software version 3.8.5 [26] on a private cloud Linux-based server.

## 4 Results

The collected GPS data are converted into 10,811 sets of movement sequences for both elite rugby league forwards and backs players. An example of the result of such conversion is presented in the appendix (Table A). The movement sequence derived from Table A is “jjjvvuuuuuGGGHH-Hzmnnnmmononbabfbrvuvvwm” after removing periods of inactivity (i.e. “eeee”).

LCCspm extracted a total of 10,149 movement patterns performed by forward rugby players of which only 6,476 were uniquely performed by players of the group. An example of such movement patterns is “SKKKKLLyyyzTSPTMyzyC” (denoted as Sprint-Acceleration-Straight, [Sprint-Deceleration-Straight]x4, [Sprint-Deceleration-Acute-Change]x2, [Run-Deceleration-Straight]x3, Run-Deceleration-Acute-Change, Sprint-Acceleration-Acute-Change, Sprint-Acceleration-Straight, Sprint-Neutral-Acute-Change, Sprint-Acceleration-Acute-Change, Sprint-Deceleration-Large-Change, Run-Deceleration-Straight, Run-Deceleration-Acute-Change,

**Table 2** Performance of all classification models for Forwards and Backs using all unique movement patterns

| Models                   | Accuracy (%) | Precision   | Recall      | F1-Score    | No of Features | Dataset     |
|--------------------------|--------------|-------------|-------------|-------------|----------------|-------------|
| Decision Tree            | 65.53        | 0.66        | 0.66        | 0.65        | 18,173         | Full        |
| Naïve Bayes              | 52.33        | 0.58        | 0.51        | 0.54        | 18,173         | Full        |
| Random Forest            | 76.31        | 0.76        | 0.76        | 0.76        | 18,173         | Full        |
| Logistic Regression      | <b>77.58</b> | <b>0.78</b> | <b>0.77</b> | <b>0.78</b> | <b>18,173</b>  | <b>Full</b> |
| Multi-Layered Perceptron | 77.23        | 0.77        | 0.77        | 0.77        | 18,173         | Full        |
| 5-Nearest Neighbour      | 72.62        | 0.73        | 0.73        | 0.73        | 18,173         | Full        |

Run-Deceleration-Straight, Run-Neutral-Straight) as uniquely performed by elite rugby league forwards.

LCCspm extracted another total of 11,636 movement patterns performed by elite rugby league backs of which only 7,963 movement patterns were uniquely performed by players of this group. An example of such movement patterns is “BoyyzDonnpppnmabbnqu” (denoted as Run-Deceleration-Backwards, Jog-Deceleration-Large-Change, [Run-Deceleration-Acute-Change]x2, Run-Neutral-Acute-Change, Jog-Deceleration-Large-Change, [Jog-Deceleration-Acute-Change]x2, [Jog-Deceleration-Backwards]x3, Jog-Deceleration-Acute-Change, Jog-Deceleration-Straight, Walk-Deceleration-Straight, [Walk-Deceleration-Acute-Change]x2, Jog-Deceleration-Acute-Change, Jog-Neutral-Straight, and Jog-Acceleration-Straight) as uniquely performed by elite rugby league backs.

Furthermore, a total of 18,173 unique movement patterns were identified across all extracted movement patterns across both positional groups.

#### 4.1 Phase I

The set of 18,173 unique movement patterns was used as the independent variables of the data for classification modelling in this phase. The performances of the six classification models developed on this dataset revealed the Naïve Bayes classification model as the least-performing model (Table 2), with an accuracy of 52.33%, 0.58 precision score, 0.51 recall score and F1-Score of 0.37. The Logistic regression model achieved the highest accuracy of 77.58% and f1-score, recall and precision scores of 0.78 respectively. Meanwhile, the Multi-Layered Perceptron model of 77.23% accuracy and precision, f1-score and recall scores of

0.77 respectively is not significantly lower than the Logistic regression classifier.

#### 4.2 Phase II

The correlation-based feature subset filter method identified 36 key movement patterns out of the 18,173 unique movement patterns used as independent variables for classifying elite RFL players into forward or back (Table 3). This subset of 36 key movement patterns (Table 3) had the highest score (0.097) out of all 74,4281 subsets.

There are significant differences between players of both positional groups when performing the key movement patterns as assessed by Cohen’s effect size. Considering only correctly predicted players that performed each key movement pattern (i.e., non-zeros observations), three (3) of the 36 key movement patterns had small effect sizes between forwards and backs (Table 3). Twenty (20) and eleven (11) of the 36 key movement patterns had moderate and large effect sizes (Table 3) respectively. Two (2) of the 36 key movement patterns had very large effect sizes. Differences were also seen in how often the key movement patterns were averagedly performed per fixture by correctly predicted players of the two positional groups (Table 3).

Having re-trained classification models based on the reduced dataset, the performances of the six classification models applied were slightly lower than the performances of the same classification models when applied to the full dataset but with the exception of the Naïve Bayes classification model (Table 4). The Random Forest model achieved the highest accuracy of 74.75% and f1-score, recall and precision scores of 0.75 respectively. Also, the Multi-Layered Perceptron model of 74.43% accuracy, 0.75 precision, f1-score



**Table 3** Effect sizes and average (in %) of performed key movement patterns classifying between RFL Forwards and Backs

| Movement Pattern | Decoded Pattern   | Cohen's Effect | Remark     | Percentage (%) average of performed movement pattern across fixtures |       |
|------------------|---|----------------|------------|--|-------|
|                  |   |                |            | Forwards   | Backs |
| mo               | Jog Deceleration Straight, Jog Deceleration Large-Change  | 0.281          | moderate   | 10.1   | 8.8   |
| jj               | Walk Acceleration Acute-Change, Walk Acceleration Acute-Change  | -0.454         | moderate   | 20   | 24.6  |
| HGH              | Run Acceleration Acute-Change, Run Acceleration Straight, Run Acceleration Acute-Change   | 1.048          | large      | 11.1   | 6     |
| iiie             | Walk Acceleration Straight, Walk Acceleration Straight, Walk Acceleration Straight, Walk Neutral Straight   | -1.242         | very large | 6  | 9     |
| mno              | Jog Deceleration Straight, Jog Deceleration Acute-Change, Jog Deceleration Large-Change   | 0.171          | small      | 8.3  | 7.8   |
| SSS              | Sprint Acceleration Straight, Sprint Acceleration Straight, Sprint Acceleration Straight  | 0.499          | moderate   | 13.5   | 10.2  |
| jjj              | Walk Acceleration Acute-Change, Walk Acceleration Acute-Change, Walk Acceleration Straight  | -0.456         | moderate   | 11.4   | 14.4  |
| ijjj             | Walk Acceleration Straight, Walk Acceleration Acute-Change, Walk Acceleration Acute-Change, Walk Acceleration Acute-Change  | -0.737         | large      | 5.5  | 7.9   |
| jiii             | Walk Acceleration Acute-Change, Walk Acceleration Straight, Walk Acceleration Straight, Walk Acceleration Straight, Walk Acceleration Straight                              | -1.513         | very large | 5.5  | 10    |
| ij               | Walk Acceleration Straight, Walk Acceleration Acute-Change,   | -0.412         | moderate   | 23.3   | 28.2  |
| HHG              | Run Acceleration Acute-Change, Run Acceleration Acute-Change, Run Acceleration Straight   | 0.634          | large      | 12.9   | 8.9   |
| iiii             | Walk Acceleration Straight, Walk Acceleration Straight, Walk Acceleration Straight, Walk Acceleration Straight, Walk Acceleration Straight                                  | -0.281         | moderate   | 11   | 13.2  |
| GH               | Run Acceleration Straight, Run Acceleration Acute-Change  | 0.496          | moderate   | 17.5   | 13.5  |
| HH               | Run Acceleration Acute-Change, Run Acceleration Acute-Change,   | 0.622          | large      | 13.6   | 9.8   |
| jjj              | Walk Acceleration Acute-Change, Walk Acceleration Acute-Change, Walk Acceleration Acute-Change  | -0.292         | moderate   | 9.7  | 11.3  |
| jje              | Walk Acceleration Acute-Change, Walk Acceleration Acute-Change, Walk Neutral Straight   | 0.466          | moderate   | 7.6  | 5.8   |
| iiij             | Walk Acceleration Straight, Walk Acceleration Straight, Walk Acceleration Straight, Walk Acceleration Straight, Walk Acceleration Acute-Change                              | -0.777         | large      | 6.8  | 8.5   |
| uuvuv            | Jog Acceleration Straight, Jog Acceleration Straight, Jog Acceleration Acute-Change, Jog Acceleration Straight, Jog Acceleration Straight, Jog Acceleration Acute-Change    | 0.288          | moderate   | 9.5  | 8.5   |
| GHH              | Run Acceleration Straight, Run Acceleration Acute-Change, Run Acceleration Acute-Change   | 0.694          | large      | 12.1   | 8.9   |
| uuvuvuv          | Jog Acceleration Straight, Jog Acceleration Straight, Jog Acceleration Straight, Jog Acceleration Acute-Change, Jog Acceleration Straight, Jog Acceleration Acute-Change    | 0.286          | moderate   | 7.5  | 7     |
| ff               | Walk Neutral Acute-Change, Walk Neutral Acute-Change  | -0.305         | moderate   | 20.9   | 24.3  |
| ijj              | Walk Acceleration Straight, Walk Acceleration Acute-Change, Walk Acceleration Acute-Change  | -0.413         | moderate   | 11.3   | 14    |
| ijji             | Walk Acceleration Straight, Walk Acceleration Acute-Change, Walk Acceleration Acute-Change, Walk Acceleration Straight  | -1.007         | large      | 6.5  | 9.1   |
| iiij             | Walk Acceleration Straight, Walk Acceleration Straight, Walk Acceleration Acute-Change  | -0.437         | moderate   | 13.7   | 17.1  |
| jjji             | Walk Acceleration Acute-Change, Walk Acceleration Acute-Change, Walk Acceleration Acute-Change, Walk Acceleration Straight  | -0.783         | large      | 6.5  | 8     |
| om               | Jog Deceleration Large-Change, Jog Deceleration Straight  | 0.584          | moderate   | 10.6   | 7.7   |
| je               | Walk Acceleration Acute-Change, Walk Neutral Straight   | -0.152         | small      | 11.9   | 12.8  |
| qv               | Jog Neutral Straight, Jog Acceleration Acute-Change   | 0.35           | moderate   | 11.1   | 9.2   |
| if               | Walk Acceleration Straight, Walk Neutral Acute-Change   | -0.106         | small      | 11.3   | 11.9  |
| uuvuvv           | Jog Acceleration Straight, Jog Acceleration Straight, Jog Acceleration Acute-Change Jog Acceleration Straight, Jog Acceleration Acute-Change, Jog Acceleration Acute-Change | 0.471          | moderate   | 11.1   | 7.7   |
| quv              | Jog Neutral Straight, Jog Acceleration Straight, Jog Acceleration Acute-Change  | 0.305          | moderate   | 8.6  | 7.8   |
| jjii             | Walk Acceleration Acute-Change, Walk Acceleration Acute-Change, Walk Acceleration Straight, Walk Acceleration Straight  | -0.748         | large      | 7.5  | 9.5   |
| iiij             | Walk Acceleration Straight, Walk Acceleration Straight, Walk Acceleration Acute-Change, Walk Acceleration Acute-Change  | -0.921         | large      | 7  | 9.5   |
| SS               | Sprint Acceleration Straight, Sprint Acceleration Straight  | 0.453          | moderate   | 13.7   | 10.7  |
| jiii             | Walk Acceleration Acute-Change, Walk Acceleration Straight, Walk Acceleration Straight, Walk Acceleration Straight  | -1             | large      | 8.5  | 13.2  |
| SSSSS            | Sprint Acceleration Straight, Sprint Acceleration Straight, Sprint Acceleration Straight, Sprint Acceleration Straight, Sprint Acceleration Straight                        | 0.275          | moderate   | 10.7   | 9.4   |

**Table 4** Performance of all classification models for Forwards and Backs using key movement patterns

| Models                   | Accuracy (%) | Precision   | Recall      | F1-Score    | No of Features | Dataset        |
|--------------------------|--------------|-------------|-------------|-------------|----------------|----------------|
| Decision Tree            | 64.73        | 0.65        | 0.65        | 0.65        | 36             | Reduced        |
| Naïve Bayes              | 68.41        | 0.7         | 0.68        | 0.67        | 36             | Reduced        |
| Random Forest            | <b>74.75</b> | <b>0.75</b> | <b>0.75</b> | <b>0.75</b> | <b>36</b>      | <b>Reduced</b> |
| Logistic Regression      | 74.02        | 0.74        | 0.74        | 0.74        | 36             | Reduced        |
| Multi-Layered Perceptron | 74.43        | 0.75        | 0.74        | 0.74        | 36             | Reduced        |
| 5-Nearest Neighbour      | 71.62        | 0.72        | 0.72        | 0.72        | 36             | Reduced        |

and recall scores of 0.75 respectively is not significantly lower than the Random Forest classification model.

The comparative analysis of the six different classification models’ performances reveals there is no drastic difference in the models’ performance (in terms of accuracy) on the data with 18,173 unique movement patterns or the data with 36 key movement patterns.

Also, a cross-tabulation of the top-36 movement patterns used by the Logistic regression of 77.58% accuracy (Table 2) with the key movement patterns used by the Random forest of 74.75% accuracy (Table 4) revealed seven overlapped patterns vis-a-vis: “GH”, “HH”, “HHG”, “uuvuv”, “GHH”, “mo”, “qv” decoded in Table 3.

## 5 Discussion

Broad positional groups exist in rugby league football [4] and an existing clustering analysis identified distinct playing positions within the same cluster (i.e., positional groups) [3] but did not reveal the movement activities that brought about the similarity because traditional indicators are limited. This study solves this problem via the classification of players into positional groups (i.e., forwards and backs) using data characterised by movement patterns, identifies the key movement patterns necessary for such classification, and quantifies the differences in the movement activities performed by players of both positional groups. The LCCspm algorithm was used to extract frequent and closed contiguous movement patterns and a total of 18,173 unique movement patterns were derived for both elite rugby league forwards and backs. The use of all unique movement patterns as independent variables produced the more accurate classification

models among five of the six classification models in comparison with using only the key movement patterns. The highest classification accuracy of 77.58%, recall of 0.77 and f1-score and precision scores of 0.78 respectively was achieved through a Logistic regression model on the dataset with 18,173 unique movement patterns (Table 2). Also, 36 of 18,173 unique movement patterns are identified as key independent variables (Table 3) and produced its highest classification accuracy of 74.75% with f1-scores, recall and precision scores of 0.75 respectively through the Random Forest classification model (Table 4). Eleven (11) and two (2) of the 36 key movement patterns had large and very large Cohen’s effect sizes between elite rugby league forwards and backs (Table 3).

To reveal the players’ movement activities within the same positional group but distinct playing positions, this study is based on movement patterns generated from GPS data through the application of the SMP framework [24] for generating discrete movement sequences and the application of LCCspm pattern mining algorithm [11] to extract frequent and closed contiguous movement patterns. Rugby league backs comprised more distinct playing positions (n=5) than forwards (n=4) and their movement patterns are often characterised by more backwards running (e.g., “LLyzBBmmnnnpnmmmmmvu” and “zABoyzzDonnpppnmabbn”) than forwards which corroborate with study [21] that reported rugby league backs have more space during fixtures and have to face opposition when retreating to the defensive line. Elite rugby league backs performed more (unique) movement patterns than forwards, which may be due to the greater number of backs being fielded during matches (n=7) than forwards (n=6) as well as backs comprising more distinct playing positions,

despite forwards having more observations in the data. Rugby league forwards movement patterns are often characterised by more accelerated running and sprinting (e.g., “GGGGGGGGGC-CCCCCCCC” and “SSSTSSTS”). These movement patterns can help with training prescription wherein elite rugby league players are conditioned based on their exact movements during competitive matches.

Finding the best among distinct models to classify elite rugby league players into forwards and backs reveals the highest classification accuracy of 77.58% via Logistic Regression using all 18,173 unique movement patterns. Also, this study’s classification analyses involved the use of the relative frequency of performed movement patterns as independent variable values. This study used alternative movement patterns’ values, in the form of relative frequency, in comparison to [13] that used values that only indicated if a player performed a movement pattern within a fixture or not. This demonstrates that the quantification of how frequently movement patterns are performed by players can also be used for classifying rugby league elite players into positional groups at an excellent classification accuracy.

The total number of unique movement patterns used as independent variables for classification modelling is extremely high and not feasible for practical application. This study resolved this high dimensionality data problem by identifying the key movement patterns from the 18,173 extracted movement patterns. Thirty-six (36) of 18,173 movement patterns were identified by the Correlation-based feature subset filter method as the key movement patterns (Table 3). This total number of the identified key movements is about 0.2% of the total number of extracted movement patterns. Upon re-training and re-evaluating the models on data characterised by only the key movement patterns, their performances were not drastically lower than those of the full dataset despite the drastic reduction in the number of independent variables. For example, the Decision Tree classification model had an accuracy of 65.53% on the full dataset but had an accuracy of 64.73% accuracy on the reduced dataset. On the other hand, the Naive Bayes model increased its accuracy from 52.33% to 68.41% which highlights the problem of multi-collinearity in the full

dataset. Hence, the 36 key movement patterns are encouraged to be used in practice.

Significant differences in forwards and backs of elite rugby league players were observed in the key movement patterns based on Cohen’s effect size and percentage average of performed movement patterns. Two (2) of the 36 key movement patterns had very large effect sizes between forwards and backs. Twenty (20) and eleven (11) of the 36 key movement patterns had moderate and large effect sizes respectively (Table 3). The effect sizes indicate that there are significant differences in how frequently key movement patterns are performed between correctly predicted elite rugby league forwards and backs. Examples from Table 3 include the movement pattern “HGH” (denoted as Run Acceleration Acute-Change, Run Acceleration Straight, Run Acceleration Acute-Change) had a large Cohen effect size and was averagely performed by backs at a relative frequency of 11.1% per fixture while it was averagely performed by forwards at a relative frequency of 6% per fixture. The movement pattern “uuvuvv” (denoted as Jog Acceleration Straight, Jog Acceleration Straight, Jog Acceleration Acute-Change Jog Acceleration Straight, Jog Acceleration Acute-Change, Jog Acceleration Acute-Change) had a moderate Cohen effect size and was averagely performed by forwards at 11.1% relative frequency per fixture while being performed by backs at 7.7% relative frequency per fixture. The movement pattern “jiiii” (denoted as Walk Acceleration Acute-Change, Walk Acceleration Straight, Walk Acceleration Straight, Walk Acceleration Straight) had a very large Cohen’s effect size and was averagely performed at 5.5% relative frequency by forwards while it was averagely performed by backs at 10% relative frequency. Another interesting key movement pattern is “SS” (denoted as Sprint Acceleration Straight, Sprint Acceleration Straight), performed by forwards per fixture at 13.7% relative frequency and backs at 10.7% relative frequency with a moderate Cohen effect size. These key movement patterns, especially those involved in accelerated running and sprinting, differ from the study [18] that reported rugby league backs undertake more running than forwards during fixtures.

The overall performance of the classification models for classifying rugby league players into

forwards and backs is higher than similar performances of classification models using traditional performance indicators in sports analytics. For example, in [28], an accuracy of 70.1% for classifying elite junior Australian football players into midfield, defence, forwards or rucks positional groups using technical performance indicators was obtained and published. The classification results of this study based on movement patterns (i.e., 77.58% using all unique movement patterns and 74.75% using key movement patterns) provide useful predictive models for new rugby league players in practice. Also, the identified key movement patterns can be used by rugby league coaches to improve the training specificity of players between the groups as well as help with talent identification and recruitment. Additionally, in [3] clusters containing players of distinct players positions were identified based on traditional indicator that does not reveal the match-based locomotive activities that led to such similarity groupings. This study revealed the match-based movement activities that led to such grouping and presented a model for classifying rugby league players into positional groups.

## 6 Conclusions and Future Works

This study fulfilled its aim by extracting frequent closed contiguous movement patterns from elite rugby league players' sets of movement sequences generated from GPS data to classify players into broad positional groups (i.e., forwards and backs) and identify key movement patterns for practical use. The classification analyses measure the extent to which exact player movement activities can classify elite rugby league players into forwards and backs as well as help identify the key movement patterns necessary for such classification. Logistic Regression fitted on the dataset with 18,173 unique movement patterns as independent variables with a relative frequency of performed patterns as values were the most accurate classification model with 77.58% accuracy. Thirty-six (36) key movement patterns were identified which also produced another high-performing classification model of Random Forest of 74.75% accuracy. Differences in match-based movement activities of elite rugby league forwards and backs as well as

the similarities in the movement patterns of distinct playing positions within each group were identified. Based on the findings of this study, it is now established that movement patterns can be used to classify elite rugby league players into forward and back rugby league positional groups, thereby enabling coaches and trainers to develop position-specific training programs while helping players' scouts to identify and recruit talent. More so, the Random Forest classification model with the key movement patterns can be used in practice as a predictive model for predicting the positional group (i.e., forward or back) a rugby player belongs to. In the future, the (parameter) optimisation of the classification model for positional groups will be considered. Also, the classification of players into nine rugby league playing positions, the identification of signature movement pattern(s) specific to each playing position, and the movement patterns' contribution to each playing position will be considered.

**Supplementary information.** An example of converting GPS data into movement sequence is given in Appendix A.

**Acknowledgments.** The authors would like to acknowledge The Rugby Football League (RFL) for the access to GPS data.

## Ethics approval

The study got the ethics approval of the Institutions Ethics Committee and written informed consent from the organisation representing all participants.

## Conflict of interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Appendix A GPS data processing

**Table A1** Example of Conversion of GPS data into movement sequence

| Longitude | Latitude | Velocity | Acceleration | Seconds | D.Vel | D.Accel      | D.Turning Angles | Movement Unit (MU)           | Encoded_MU |
|-----------|----------|----------|--------------|---------|-------|--------------|------------------|------------------------------|------------|
| 0         | 0        | 0        | 0            | 1       | Walk  | Neutral      | Straight         | WalkNeutralStraight          | e          |
| 0         | 0        | 0        | 0            | 1.1     | Walk  | Neutral      | Straight         | WalkNeutralStraight          | e          |
| 0         | 0        | 0        | 0            | 1.2     | Walk  | Neutral      | Straight         | WalkNeutralStraight          | e          |
| 0         | 0        | 0        | 0            | 1.3     | Walk  | Neutral      | Straight         | WalkNeutralStraight          | e          |
| 0         | 0        | 0        | 0            | 1.4     | Walk  | Neutral      | Straight         | WalkNeutralStraight          | e          |
| 0         | 0        | 0        | 0            | 1.5     | Walk  | Neutral      | Straight         | WalkNeutralStraight          | e          |
| -2.654    | 53.548   | 1.063    | 1.344        | 1957.9  | Walk  | Acceleration | Acute-Change     | WalkAccelerationAcute-Change | j          |
| -2.654    | 53.548   | 1.184    | 1.313        | 1958    | Walk  | Acceleration | Straight         | WalkAccelerationStraight     | i          |
| -2.654    | 53.548   | 1.442    | 1.630        | 1958.1  | Walk  | Acceleration | Acute-Change     | WalkAccelerationAcute-Change | j          |
| -2.654    | 53.548   | 1.768    | 2.035        | 1958.2  | Jog   | Acceleration | Acute-Change     | JogAccelerationAcute-Change  | v          |
| -2.654    | 53.548   | 1.998    | 2.101        | 1958.3  | Jog   | Acceleration | Acute-Change     | JogAccelerationAcute-Change  | v          |
| -2.654    | 53.548   | 2.317    | 2.373        | 1958.4  | Jog   | Acceleration | Straight         | JogAccelerationStraight      | u          |
| -2.654    | 53.548   | 2.532    | 2.318        | 1958.5  | Jog   | Acceleration | Straight         | JogAccelerationStraight      | u          |
| -2.654    | 53.548   | 2.962    | 2.813        | 1958.6  | Jog   | Acceleration | Straight         | JogAccelerationStraight      | u          |
| -2.654    | 53.548   | 3.145    | 2.567        | 1958.7  | Jog   | Acceleration | Straight         | JogAccelerationStraight      | u          |
| -2.654    | 53.548   | 3.353    | 2.446        | 1958.8  | Jog   | Acceleration | Straight         | JogAccelerationStraight      | u          |
| -2.654    | 53.548   | 3.961    | 3.354        | 1958.9  | Run   | Acceleration | Straight         | RunAccelerationStraight      | G          |
| -2.654    | 53.548   | 4.361    | 3.516        | 1959    | Run   | Acceleration | Straight         | RunAccelerationStraight      | G          |
| -2.654    | 53.548   | 4.700    | 3.484        | 1959.1  | Run   | Acceleration | Straight         | RunAccelerationStraight      | G          |
| -2.654    | 53.548   | 4.836    | 2.953        | 1959.2  | Run   | Acceleration | Acute-Change     | RunAccelerationAcute-Change  | H          |
| -2.654    | 53.548   | 4.802    | 2.130        | 1959.3  | Run   | Acceleration | Acute-Change     | RunAccelerationAcute-Change  | H          |
| -2.654    | 53.548   | 4.565    | 1.005        | 1959.4  | Run   | Acceleration | Acute-Change     | RunAccelerationAcute-Change  | H          |
| -2.654    | 53.548   | 4.161    | -0.256       | 1959.5  | Run   | Deceleration | Acute-Change     | RunDecelerationAcute-Change  | z          |
| -2.654    | 53.548   | 3.792    | -1.115       | 1959.6  | Jog   | Deceleration | Straight         | JogDecelerationStraight      | m          |
| -2.654    | 53.548   | 3.505    | -1.554       | 1959.7  | Jog   | Deceleration | Acute-Change     | JogDecelerationAcute-Change  | n          |
| -2.654    | 53.548   | 3.126    | -2.113       | 1959.8  | Jog   | Deceleration | Acute-Change     | JogDecelerationAcute-Change  | n          |
| -2.654    | 53.548   | 3.041    | -1.797       | 1959.9  | Jog   | Deceleration | Acute-Change     | JogDecelerationAcute-Change  | n          |
| -2.654    | 53.548   | 2.899    | -1.703       | 1960    | Jog   | Deceleration | Straight         | JogDecelerationStraight      | m          |
| -2.654    | 53.548   | 2.779    | -1.577       | 1960.1  | Jog   | Deceleration | Straight         | JogDecelerationStraight      | m          |
| -2.654    | 53.548   | 2.582    | -1.675       | 1960.2  | Jog   | Deceleration | Large-Change     | JogDecelerationLarge-Change  | o          |
| -2.654    | 53.548   | 2.536    | -1.371       | 1960.3  | Jog   | Deceleration | Acute-Change     | JogDecelerationAcute-Change  | n          |
| -2.654    | 53.548   | 1.978    | -2.424       | 1960.4  | Jog   | Deceleration | Large-Change     | JogDecelerationLarge-Change  | o          |
| -2.654    | 53.548   | 1.761    | -2.360       | 1960.5  | Jog   | Deceleration | Acute-Change     | JogDecelerationAcute-Change  | n          |
| -2.654    | 53.548   | 1.377    | -2.730       | 1960.6  | Walk  | Deceleration | Acute-Change     | WalkDecelerationAcute-Change | b          |
| -2.654    | 53.548   | 1.454    | -1.855       | 1960.7  | Walk  | Deceleration | Straight         | WalkDecelerationStraight     | a          |
| -2.654    | 53.548   | 1.562    | -1.121       | 1960.8  | Walk  | Deceleration | Acute-Change     | WalkDecelerationAcute-Change | b          |
| -2.654    | 53.548   | 0.000    | 0.000        | 1960.9  | Walk  | Neutral      | Acute-Change     | WalkNeutralAcute-Change      | f          |
| -2.654    | 53.548   | 1.482    | -1.041       | 1961    | Walk  | Deceleration | Acute-Change     | WalkDecelerationAcute-Change | b          |
| -2.654    | 53.548   | 1.766    | -0.068       | 1961.1  | Jog   | Neutral      | Acute-Change     | JogNeutralAcute-Change       | r          |
| -2.654    | 53.548   | 1.973    | 0.464        | 1961.2  | Jog   | Acceleration | Acute-Change     | JogAccelerationAcute-Change  | v          |
| -2.654    | 53.548   | 2.294    | 1.150        | 1961.3  | Jog   | Acceleration | Straight         | JogAccelerationStraight      | u          |
| -2.654    | 53.548   | 2.342    | 0.985        | 1961.4  | Jog   | Acceleration | Acute-Change     | JogAccelerationAcute-Change  | v          |
| -2.654    | 53.548   | 2.345    | 0.746        | 1961.5  | Jog   | Acceleration | Large-Change     | JogAccelerationLarge-Change  | w          |
| -2.654    | 53.548   | 2.427    | 0.762        | 1961.6  | Jog   | Acceleration | Acute-Change     | JogAccelerationAcute-Change  | v          |
| -2.654    | 53.548   | 2.088    | -0.276       | 1961.7  | Jog   | Deceleration | Straight         | JogDecelerationStraight      | m          |

## References

- [1] Saal, C., Helm, N., Prieske, O.: Anthropometry, Physical Fitness, Sport-Specific Performance and the Prediction of Performance Level in Young Canoe Sprint Athletes. In: International Conference on Security, Privacy, and Anonymity in Computation, Communication and Storage. pp. 74–77 (2022)
- [2] Whitehead, S., Till, K., Jones, B., Beggs, C., Dalton-Barron, N., Weaving, D.: The use of technical-tactical and physical performance indicators to classify between levels of match-play in elite rugby league. *Science and Medicine in Football* **5**(2), 121–127 (2020). <https://doi.org/10.1080/24733938.2020.1814492>, <http://dx.doi.org/10.1080/24733938.2020.1814492>
- [3] Dalton-Barron, N., Palczewska, A., Weaving, D., Rennie, G., Beggs, C., Roe, G. and Jones, B.: Clustering of match running and performance indicators to assess between-and within-playing position similarity in professional rugby league. *Journal of Sports Sciences* **40**(15), 1712-1721 (2022)
- [4] Austin, D.J. and Kelly, S.J.: Positional differences in professional rugby league match play through the use of global positioning systems. *The Journal of Strength & Conditioning*

- [5] Kempton, T., Sirotic, A.C., Coutts, A.J.: A Comparison of Physical and Technical Performance Profiles Between Successful and Less-Successful Professional Rugby League Teams. *International Journal of Sports Physiology and Performance* **12**(4), 520–526 (2017). <https://www.cochranelibrary.com/central/doi/10.1002/central/CN-01787161/full>
- [6] Gabbett, T.J., Hulin, B.T.: Activity and recovery cycles and skill involvements of successful and unsuccessful elite rugby league teams: A longitudinal analysis of evolutionary changes in National Rugby League match-play. *Journal of Sports Sciences* **36**(2), 180–190 (2017). <https://doi.org/10.1080/02640414.2017.1288918>, <http://dx.doi.org/10.1080/02640414.2017.1288918>
- [7] Carey, D.L., Ong, K., Whiteley, R., Crossley, K.M., Crow, J., Morris, M.E.: Predictive Modelling of Training Loads and Injury in Australian Football. *International Journal of Computer Science in Sport* **17**(1), 49–66 (2018). <https://doi.org/10.2478/ijcss-2018-0002>, <http://dx.doi.org/10.2478/ijcss-2018-0002>
- [8] Sweeting, A.J., Aughey, R.J., Cormack, S.J., Morgan, S.: Discovering frequently recurring movement sequences in team-sport athlete spatiotemporal data. *Journal of Sports Sciences* **35**(24), 2439–2445 (2017). <https://doi.org/10.1080/02640414.2016.1273536>, <https://doi.org/10.1080/02640414.2016.1273536>
- [9] Whitehead, S., Till, K., Weaving, D., Jones, B.: The use of microtechnology to quantify the peak match demands of the football codes: a systematic review. *Sports medicine* **48**(11), 2549–2575 (2018)
- [10] Adeyemo, V.E., Palczewska, A., Jones, B.: LCCspm: L-Length Closed Contiguous Sequential Patterns Mining Algorithm to Find Frequent Athlete Movement Patterns from GPS. In: *Proceedings - 20th IEEE International Conference on Machine Learning and Applications, ICMLA 2021*. pp. 455–460. IEEE (2021). <https://doi.org/10.1109/ICMLA52953.2021.00077>
- [11] Adeyemo, V.E., Balogun, A.O., Mojeed, H.A., Akande, N.O., Adewole, K.S.: Ensemble-Based Logistic Model Trees for Website Phishing Detection. *Communications in Computer and Information Science* **1347**(February), 627–641 (2021). <https://doi.org/10.1007/978-981-33-6835-4-41>
- [12] Adeyemo, V.E.: Using Movement Patterns to Classify Elite RFL Forwards and Backs into Positions. <https://github.com/arhvel/MovementPatterns4ForwardsBacksClassification> (2023), [Online; accessed 23-August-2023]
- [13] Adeyemo, V.E., Palczewska, A., Jones, B., Weaving, D.: Identification of pattern mining algorithm for rugby league players positional groups separation based on movement patterns. arXiv preprint arXiv:2302.14058 -2023
- [14] Adeyemo, V.E., Palczewska, A., Jones, B., Weaving, D., Whitehead, S.: Optimising classification in sport: a replication study using physical and technical-tactical performance indicators to classify competitive levels in rugby league match-play. *Science and Medicine in Football* pp. 1–8 (2022). <https://doi.org/10.1080/24733938.2022.2146177>, <http://dx.doi.org/10.1080/24733938.2022.2146177>
- [15] Alsariera, Y.A., Elijah, A.V., Balogun, A.O.: Phishing Website Detection: Forest by Penalizing Attributes Algorithm and Its Enhanced Variations. *Arabian Journal for Science and Engineering* **45**(12), 10459–10470 (2020). <https://doi.org/10.1007/s13369-020-04802-1>
- [16] Collins, N., White, R., Palczewska, A., Weaving, D., Dalton-Barron, N., Jones, B.: Moving beyond velocity derivatives; using global positioning system data to extract sequential movement patterns at different levels of rugby league match-play. *European Journal of Sport Science* pp. 1–9 (2022). <https://doi.org/10.1080/17461391.2022.2027527>



- <https://doi.org/10.1080/17461391.2022.2027527>
- [17] Gabbett, T.J.: Influence of playing standard on the physical demands of professional rugby league. *Journal of Sports Sciences* **31**(10), 1125–1138 (2013). <https://doi.org/10.1080/02640414.2013.773401>, <http://dx.doi.org/10.1080/02640414.2013.773401>
- [18] Geeson-Brown, T., Jones, B., Till, K., Chantler, S., Deighton, K.: Body composition differences by age and playing standard in male rugby union and rugby league: A systematic review and meta-analysis. *Journal of sports sciences* **38**(19), 2161–2176 (2020)
- [19] Morgulev, E., Azar, O.H., Lidor, R.: Sports analytics and the big-data era. *International Journal of Data Science and Analytics* (2018). <https://doi.org/10.1007/s41060-017-0093-7>, <https://doi.org/10.1007/s41060-017-0093-7>
- [20] Pasquier, N., Bastide, Y., Taouil, R., Lakhal, L.: Discovering frequent closed itemsets for association rules. In: *International Conference on Database Theory*. pp. 398–416. Springer (1999)
- [21] Sirotic, A.C., Knowles, H., Catterick, C., Coutts, A.J.: Positional match demands of professional rugby league competition. *The Journal of Strength & Conditioning Research* **25**(11), 3076–3087 (2011)
- [22] Thornton, H.R., Delaney, J.A., Duthie, G.M., Dascombe, B.J.: Importance of various training-load measures in injury incidence of professional rugby league athletes. *International journal of sports physiology and performance* **12**(6), 819–824 (2017)
- [23] Wedding, C., Woods, C., Sinclair, W., Gomez, M., Leicht, A.: Examining the evolution and classification of player position using performance indicators in the National Rugby League during the 2015–2019 seasons. *Journal of Science and Medicine in Sport* **23**(9), 891–896 (2020)
- [24] White, R., Palczewska, A., Weaving, D., Collins, N., Jones, B.: Sequential movement pattern-mining (SMP) in field-based team-sport: A framework for quantifying spatiotemporal data and improve training specificity? *Journal of Sports Sciences* pp. 1–10 (2021)
- [25] Williamson, B.D., Gilbert, P.B., Carone, M., Simon, N.: Nonparametric variable importance assessment using machine learning techniques. *Biometrics* **77**(1), 9–22 (2021)
- [26] Witten, I.H., Frank, E.: *Data mining: practical machine learning tools and techniques with java implementations*. *Acm Sigmod Record* **31**(1), 76–77 (2002)
- [27] Woodhouse, L.N., Tallent, J., Patterson, S.D., Waldron, M.: International female rugby union players’ anthropometric and physical performance characteristics: A five-year longitudinal analysis by individual positional groups. *Journal of Sports Sciences* **40**(4), 370–378. (2022)
- [28] Woods, C.T., Veale, J., Fransen, J., Robertson, S., Collier, N.F.: Classification of playing position in elite junior Australian football using technical skill indicators. *Journal of sports sciences* **3**(1), 97–103 (2018)