COMPARISON OF MACHINE LEARNING CLASSIFIERS FOR DIFFERENTIATING LEVEL AND SPORT USING MOVEMENT DATA

Gwyneth B. Ross[1*], Allison L. Clouthier[1], Alistair Boyle[2], Steven L. Fischer[3], and Ryan B. Graham[1]
[1]School of Human Kinetics, Faculty of Health Sciences, University Ottawa, Ottawa, ON, Canada
email: gwyneth.ross@uottawa.ca*

Introduction

Movement screens are used to identify aberrant movement patterns believed to increase risk of injury and/or impede performance. Common criticisms of movement screens are that they are not able to predict injury risk, which is thought to be due to a lack of sensitivity within the scoring criteria[1] and the poor inter- and intra-rater reliability due to scores being based on visual appraisal[2]. Furthermore, the scoring criteria does not account for differences in skill level or sport played[2]. Therefore, our previous research has focused on the development of an objective movement screening tool that uses 3D motion capture data, principal component analysis (PCA), ensemble feature selection, and traditional machine learning to differentiate between elite and novice athletes[3]. However, machine learning algorithms designed for time-series data, such as recurrent neural networks (RNNS), may provide a better method for differentiating the movement patterns of athletes. Therefore, the purpose of this study was two-fold: 1) to determine if RNNs designed for time-series analyses can outperform the previously used traditional classifiers at classifying athlete skill level, and 2) if athletes can be differentiated based on sport played, and if so, to identify which machine learning algorithm(s) performs the best.

Methods

Kinematic data were collected from 542 athletes ranging in competition level from youth to professional and competing in 11 different sports. Athletes competing at the collegiate level or above were considered elite athletes, whereas athletes competing below collegiate were considered novice. For classifying sport, only elite athletes competing in football, baseball, basketball, and soccer were analyzed. The protocol consisted of a movement screening battery consisting of 7 tasks: bird-dog, drop-jump, hop-down, L-hop, lunge, step-down, and T-balance[3]. In Visual3D, a 3D model was developed, and positional data of all major joint centres were calculated and time-normalized to 500 frames.

For the traditional machine learning algorithms (linear discriminant analysis (LDA), binary logistic regression (BLR), support vector machine with a linear (SVM) and radial basis function kernel (RBF), decision tree (DT), naïve bayes (NB), and *k*-nearest neighbors (KNN)), a matrix for each task was constructed where the rows were each subject and the columns were the frames, axes, and joint centres and PCA was applied to each matrix. Using the principal component (PC) scores as features to classify skill level and sport, ensemble feature selection was used to rank the PC scores based on contribution to the model for each movement task and classifier. To classify the data, the top PC scores from the ensemble feature selection were used as inputs and either sport or level were used as the class. For the RNNs (reservoir computing (RC) and Long-short term memory (LSTM)), a 3D matrix was constructed (subjects x frame x joint centre/axis) for each task. For both RC and LSTM, the hyperparameters were tuned using grid search and a train-test split of 80-20. All classifiers were validated using 10 iterations of random subsampling with a train-test split of 80-20. Lastly, a naïve algorithm was used that predicted all athletes in the testing set as the majority group for that training set. One-way ANOVAs with Tukey post-hoc tests were run in SPSS for each class and movement task between the ten different

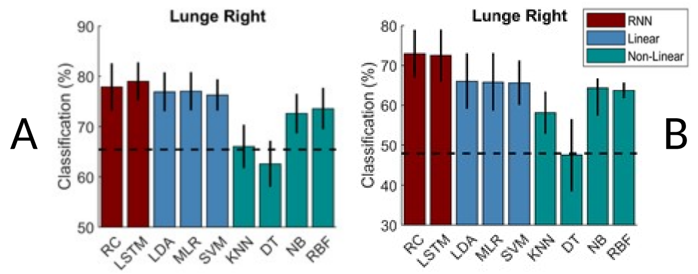

Figure 1. The mean classification rates and standard deviations when classifying for level (A) and sport (B) for the lunge right. The dotted black line denotes the naïve classification rates.

Results and Discussion

For skill level, looking at just the average classification rates for the lunge right, RC, LSTM, LDA, BLR, and SVM all had significantly greater classification rates than KNN, DT, and naïve ($p<0.001$); however, there were no significant differences between RC, LSTM, LDA, MLR, or SVM ($p > 0.05$; Figure 1A). For sport, looking at just the lunge right, RC and LSTM had a significantly greater classification rates than KNN, DT, and naïve ($p<0.001$); however, there were no significant differences between RC and LSTM ($p = 1.00$; Figure 1B). Similar trends were observed across all tasks.

The RNNs and the linear classifiers had significantly better classification rates than the naïve classifier, suggesting that the classifiers were classifying based on actual differences between classes and not noise or by chance. Due to having one of the highest classification rates and taking the least amount of time to train for all tasks, going forward, it is suggested to use RC for these types of analyses. In addition, it was possible to classify athletes based on sport, which suggests that athletes move differently based on the sport they play.

Significance

Currently, popular movement screens do not consider athlete-specific demographics such as competition level or sport. Based on athletes moving differently depending on their level of play and sport, there should be sport- and level- specific scoring criteria for movement competency assessments, which may increase their ability to predict injury risk. Future research should study the misclassified athletes and whether novice athletes misclassified as elite athletes have a higher likelihood of making it to an elite level of play or if elite athletes misclassified as novice athletes have a higher risk of injury

Acknowledgments

Thank you to the athletes who partook in this study.

References

1. Frost et al., 2015. *Phys. Ther. Sport* 16:324. 2. McCunn et al., 2016. *Sport Med.* 46:763. 3. Ross et al., 2018. *Med. Sci. Sport. Exerc.* 50:1457.