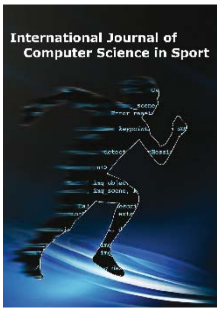
International Journal of Computer Science in Sport

Volume 19, Issue 1, 2020

Journal homepage: http://iacss.org/index.php?id=30

sciendo

DOI: 10.2478/ijcss-2020-0002



Performance of machine learning models in application to beach volleyball data.

Sebastian Wenninger, Daniel Link and Martin Lames

Technical University of Munich

Abstract

Driven by the increased availability of position and performance data, automated analyses are becoming the daily routine in many top-level sports. Methods from the domains of data mining and machine learning are more frequently used to generate new insights from massive amounts of data. This study evaluates the performance of four current models (multi-layer perceptron, convolutional network, recurrent network, gradient boosted tree) in classifying tactical behaviors on a beach volleyball dataset consisting of 1,356 top-level games. A three-way between-subjects analysis of variance was conducted to determine the effects of model, input features and target behavior on classification accuracy. Results show significant differences in classification accuracy between models as well as significant interaction effects between factors. Our models achieve classification performance similar to previous work in other sports. Nonetheless, they are not yet at the level to warrant practical application in day to day performance analysis in beach volleyball.

KEYWORDS: MACHINE LEARNING, SPORTS ANALYTICS, NEURAL NETWORKS, BEACH VOLLEYBALL

Introduction

Due to the large amount of available data and the resulting possibility to convert that data into useful knowledge, methods from artificial intelligence (AI) and data mining have become increasingly useful tools in many industry scenarios (Russel, & Norvig, 2010). This development also took place in the sports domain, where automated data analysis is a fast developing trend to assist the decision-making of practitioners (Link, 2018; Claudino et al, 2019).

However, predicting performance in beach volleyball is a subject that has not received much attention compared to other major sports. This may be due to the fact that teams in beach volleyball mostly operate for themselves, and data is not readily available. Unlike sports as football, hockey or basketball, where access to positional data from electronic position detection is common, beach volleyball only recently (at the world championships 2017 in Vienna) started to include sensor data to enhance media coverage. Additionally, the betting market, one of the biggest drivers for advances in performance prediction (Bunker, & Thabtah, 2019), mostly concentrates on other, financially stronger sports in order to optimize earnings.

The implementation and evaluation of different machine learning models on currently available performance data could serve as a first step to establish successful models from other application areas as viable methods of analysis. As Claudino et. al. (2019, p. 10) phrase it: Further evaluation research based on prospective methods is warranted to establish the predictive performance of specific AI techniques and methods. Additionally, the results of this study could inform future technological developments in the field of beach volleyball on the type of data and algorithms necessary for certain analyses.

Related Work

One of the earliest studies to consider AI in the analysis of sports performance was done by Lapham and Bartlett in 1995. They showed that the involvement of computers, with the help of methods from artificial intelligence, could be a rewarding future direction for the discipline. Since then, many different AI techniques were applied in a wide variety of sports, either to assess risk of injury (López-Valenciano et. Al., 2018), or to predict performance (Peterson, 2018).

In soccer, Perl & Memmert (2011) use DyCoN, a self-organizing neural network, to learn formation clusters from spatio-temporal data. These formation types are then fed to an analysis software to allow analysis of formation frequencies and interactions. Link and Hoernig (2017) employ a Bayesian network to classify ball control during individual ball possessions of soccer players with an accuracy of 96.7 percent. Bialkowski et al. (2015) learned the roles of players based on spatio-temporal data that describe their arrangement on the pitch. Using a minimum entropy model to partition the data into player roles, they showed that distinct formation classes can be discovered automatically. Spatio-temporal data was also used by Dick and Brefeld (2019) in combination with a deep convolutional network to automatically rate dangerous situations in soccer based on player positioning on the pitch.

The impact on spatio-temporal data on the possibilities of performance analysis can also be observed in basketball. Here it was used to compare self-organizing maps, a class of unsupervised neural networks, with a dynamical controlled network to classify three different pre-selected plays (Kempe, Grunz, & Memmert, 2015). Offensive play classification was also studied using deep recurrent neural networks by Wang and Zemel (2016), who state that the dynamic nature of team sports, especially the complexity of player interactions, makes this one of the hardest problems in sport analytics. Bianchi, Facchinetti and Zuccolotto (2017) combine self-organizing maps with a fuzzy clustering algorithm to identify five new positional player roles in basketball based on seasonal player statistics data. On the other hand, Leicht, Gómez and Wood (2017) try to predict the match outcome in the Olympic basketball tournament using both a linear regression and a conditional inference classification tree. They conclude that even though the classification tree shows a slightly lower classification accuracy of 81.4% correctly classified matches, its ability to resolve non-linear phenomena offers greater practical utility.

The working group around Schrapf and Tilp (Hassan, Schrapf, Ramadan, & Tilp, 2017a; Hassan, Schrapf, & Tilp, 2017b; Schrapf, Alsaied, & Tilp, 2017) employed various types of neural networks (self-organizing maps, radial basis function network, dynamical controlled network) to analyze tactical interaction patterns, tactical training outcome and to predict shot positions on the field in team handball.

Of course, artificial intelligence methods also found applications in volleyball. Tümer and Koçer (2017) were able to successfully predict league standings of teams between seasons with 98 percent accuracy using a multi-layer perceptron and match results and match location (home/away) as input features to their model. In order to improve the training process, dynamical programming was used in conjunction with a k-nearest neighbour classifier to detect jumps in training data and choose appropriate intensity intervals in the training process (Vales-Alonso, Chaves-Dieguez, Lopez-Matencio, Alcaraz, Parrado-Garcia, & Gonzalez-Castano, 2015). To a similar purpose, Wang, Zhao, Chan and Li (2018) use inertial sensors placed on the wrists of volleyball players to assess their skill in the spiking technique. Their support vector machine was able to differentiate between elite, sub-elite and amateur players with an average accuracy of 94 percent. In another approach, Van Haaren, Ben Shitrit, Davis and Fua (2016) utilize a relational learning technique called inductive logic programming to automatically detect play patterns in high level volleyball data. They present the top ranked, as well as the most distinguishing offensive patterns per team and compare playing patterns between men and women. Results show that attacks from outside positions are among the most successful offensive patterns. Furthermore, women's teams show more attack actions in the same number of rallies, hinting at a faster pace in men's volleyball, which makes it harder to gain control of the ball after an attack from the opponent.

However, even though beach volleyball is at least as popular as indoor volleyball, with 425,000 cumulative visitors at the London Olympics 2012 ("Net gains - the evolution of beach volleyball", 2016), and arguably easier to analyze due to the lesser number of players on the court, AI methods have not attracted much interest. The only study we found used a deep convolutional neural network based on wearable sensors to automatically monitor player loads in beach volleyball (Kautz, Groh, Hannink, Jensen, Strubberg, & Eskofier, 2017).

The aim of this paper is to evaluate the application of AI methods on both positional as well as event data in beach volleyball. As far as the authors are aware, there exists no previous work to predict technical or tactical behaviors in beach volleyball. We will focus on three different neural network architectures together with a gradient boosted classification tree, some of the most used AI techniques employed to predict sporting performance (Claudino et. al., 2019), in order to cover a wider range of methods.

Methods

Data acquisition

The dataset consists of 569 men's and 787 women's top-level games collected at FIVB world tour tournaments and championships in the years from 2013 to 2018. All data was annotated by professional beach volleyball analysts by using custom-made observation software (Link, 2014). Cohen's κ statistics showed substantial to perfect agreement between two observers based on a subset of 121 sequences (κ = 0.93 up to 1.0). For each rally in a game, the analysts collected more than 25 performance indicators (PIs) in addition to the discrete X and Y coordinates each for the serve, reception, set, approach and attack actions. These positions were tracked manually by the scouts through clicks on a calibrated projection of the beach volleyball field on the scouting video in the observation software. In total, our database contains 84,415 so-called standard sideouts. Standard sideouts in beach volleyball are classified as the rallies, in which the receiving team has the chance to attack after a structured build-up (total 3 contacts of the ball). These sideouts are known to be one of the most influencing factors to the result of a game (Zetou, Moustakidis, Tsigilis, & Komninakidou, 2007).

Data Preprocessing

Since not all games were analyzed in full detail by the analysts, we had to eliminate those rallies that did not contain valid X/Y coordinates. All X/Y coordinates were flipped so that the attacking position always appeared on the bottom half of the court and afterwards normalized to an area of 1 meter around the beach-volleyball court (see Fig.1) to account for positions outside the actual playing area (e.g. services). Additionally, we only used the PIs relevant to a single standard sideout from the collected data, leaving out general information such as tournament, match score and player information. We also did not include the defense position in the analysis, because for certain attacks, e.g. attack errors that were hit to the outside, this position also encoded the point where the ball hit the ground. Including this position would trivialize the prediction of success and attack direction. In total, this resulted in 14 PI and 10 positional (5 X/Y-coordinate pairs) input features for our models.

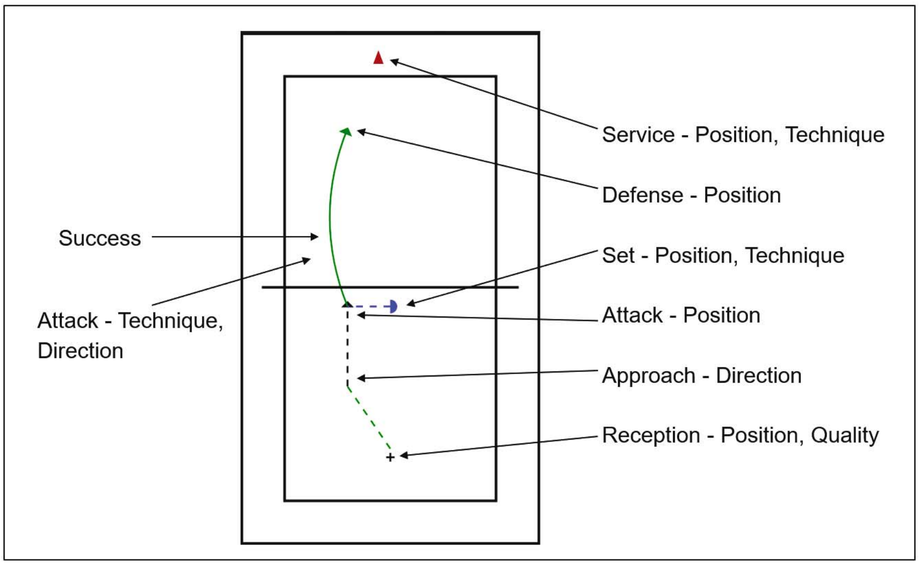


Figure 1. Positions and event features captured in a typical standard sideout. The figure sketches the 8x8m area of the beach volleyball court as well as the 1 meter wide area around it that accounts for positions outside the playing field. The black and green dotted lines depict player movement from reception to attack, while the blue dotted line represents the ball movement after the set.

Models

Except for the gradient boosted tree, all models were implemented using the PyTorch (Paszke et. al., 2017) and fast.ai (Howard et. al., 2018) python libraries, running on Google Compute Engine (GCE) virtual machines. The boosted tree was implemented leveraging the XGBOOST

(Chen, & Guestrin, 2016) python library and also ran on GCE.

Input features

We presented the models two different sets of input features: One with only the 5 positional X/Y tuples and one with the positional data combined with the performance indicators. In order to combine positional data and PIs for the neural networks, PIs were passed through automatically sized embedding layers (Kocmi, & Bojar, 2017) before concatenating the resulting output vector with the positional data. For the XGBOOST model, PIs were encoded using a One-Hot encoding scheme before passing them to the model along with the positional features.

Target variables

In addition to the different input features, we also evaluated the performance of our predictions for different target variables. We tried to predict the SUCCESS of a rally, the attack DIRECTION and attack TECHNIQUE based on the events and/or positions that occurred in the rally before. The SUCCESS was encoded as point/no point in the sideout, while the TECHNIQUE could take the values smash (strong attack) or shot (more precise, less strong attack). Finally, six values (diagonal, short diagonal, dink, cut, middle, line) were captured for the DIRECTION variable.

Training

Each model was trained six times, one time per combination of input features and target variable. We randomly split our data in training and test sets using a split size of 20 percent for the test set. The learning rate for the neural networks was set to 5e-3, as determined by a learning rate finder (Smith, 2017). Training then consisted of 10 epochs using the 1-cycle policy (Smith, 2018) and the Adam optimizer (Kingma, & Ba, 2014), that were iterated 50 times. The model with the best accuracy on the test set per 10-epoch-iteration was saved for model evaluation, resulting in 50 accuracy values per model-input-target combination. The loss function used in all neural network training cases was the cross-entropy loss, which combines a logarithmic softmax-layer with the negative log likelihood loss.

XGBOOST

The gradient boosted tree model was set up to fit 50 trees with a maximum depth of 5 nodes. The objective function was set to logistic regression for binary classification for the SUCCESS and TECHNIQUE target variables, and to softmax for multiclass classification for the DIRECTION target.

MLP

In the multi-layer perceptron (Rosenblatt, 1961), the 10 positional input features are first passed through a 1-dimensional batch-normalization layer (Ioffe, & Szegedy, 2015), with momentum set to 0.1. We then use three blocks of linear layers followed by rectified linear unit (ReLU) activation functions (Maas, Hannun, & Ng, 2013), and batch normalization layers with sizes of 128, 64 and 32 neurons each. The final linear layer then maps to the class size of the respective target variable. In the combined input feature case, the embedding output is concatenated to the output of the first batch-normalization layer before passing it to the linear layers.

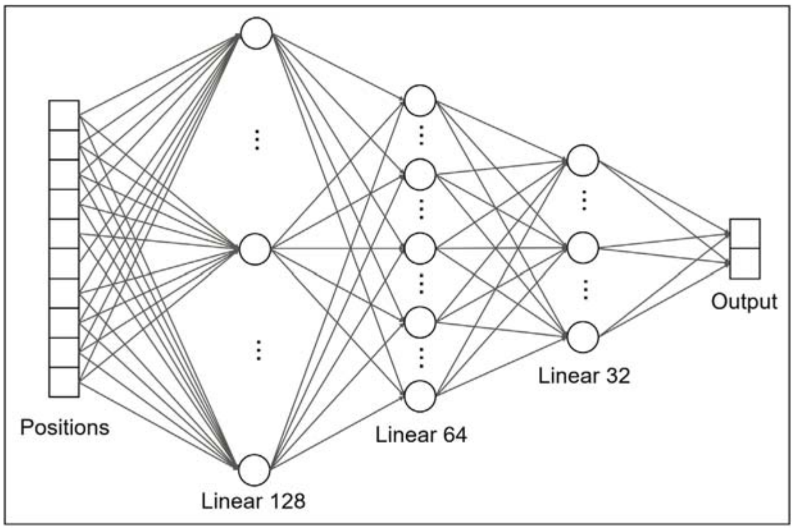


Figure 2. MLP architecture for positions only input, batch normalization and activation layers are omitted

CNN

Our convolutional network consisted of five 1-dimensional convolutional blocks with different kernel sizes, strides and dilations (Fig. 3) in order to capture different relative patterns between positions in the input. The concatenated output from those five blocks was then fed to three linear layers with batch-normalization and ReLU activations for classification. Again, we combined this architecture with embedding layers for the performance indicators, whose output was concatenated with the output of the convolutional blocks before it was passed to the linear layers.

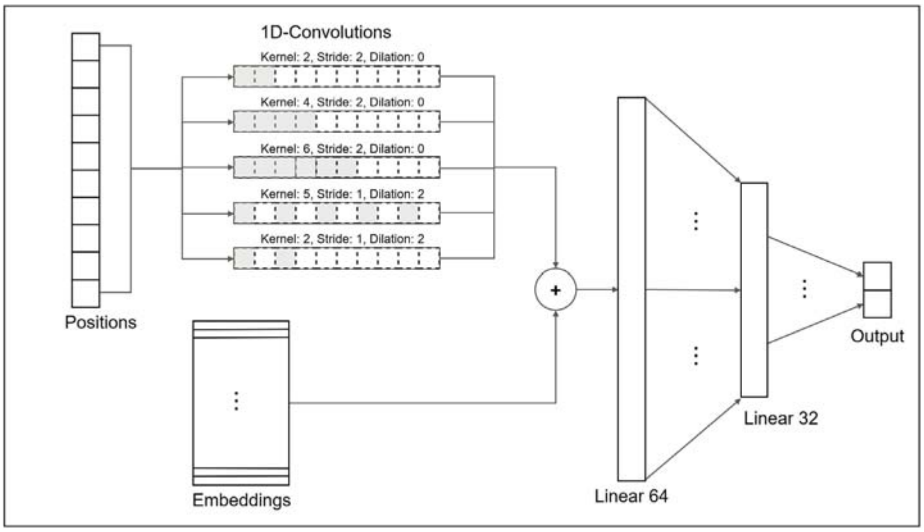


Figure 3. CNN architecture for the combined input case. BN-layers and activations are omitted for the linear layers.

RNN – GRU

Our choice of recurrent neural network was a so called Gated-Recurrent-Unit (GRU) (Cho, Van Merriënboer, Gulcehre, Bahdanau, Bougares, Schwenk, & Bengio, 2014), a kind of Long-Short-Term memory (LSTM) neural network that has proven to be easier to train on smaller datasets (Chung, Gulcehre, Cho, & Bengio, 2014). Our model consisted of a 5-layer GRU with a dropout rate of 20 percent, followed by two linear layers with batch-normalization and ReLU activations for classification. Since recurrent networks expect input data to be sequenced, the positional data was reshaped to a sequence of five X/Y-coordinate tuples for the playing positions, ordered by occurrence in a beach volleyball rally. To combine positional and PI data, embedding layers were used for the PIs. The output of the embeddings was then concatenated with the output of the GRU before passing it to the final linear classification layer.

Evaluation

For the evaluation we performed a three-way ANOVA with classification accuracy as dependent variable and model, target variable and input features as factors. All statistical tests were performed using IBM SPSS Statistics for Windows, Version 23. G*Power analysis revealed that with a total sample size of 1200 accuracy values, small effect sizes ($\eta 2$ = .018) could be detected with a power of 0.96.

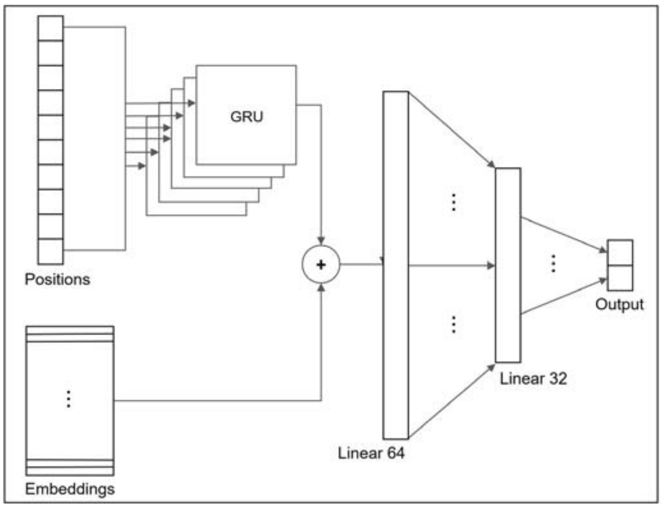


Figure 4. RNN architecture for the combined input case. Again, BN-layers and activations are omitted for the linear layers.

Results.

A three-way ANOVA was conducted to determine the effects of model, target variable and input features on the classification accuracy. There was a statistically significant three-way interaction between model, target variable and input features ($F(6, 1176) = 92.37$, $p < .001$, $\eta 2$=.001). Statistical significance was accepted at the $p < .025$ level for simple two-way interactions and at $p < .016$ for simple simple main effects. There was a statistically significant simple two-way interaction between target and model for positional input ($F(6, 1176) = 122.5$, $p < .001$), as well as for the combined input ($F(6, 1176) = 148.75$, $p < .001$). There were statistically significant simple simple main effects of the model for the target variables Direction ($F(3, 1176) = 72.004$, $p < .001$), Success ($F(3, 1176) = 53.282$, $p <.001$) and Technique ($F(3, 1176) = 323.47$, $p < .001$), when training with the combined input features. The positional input also showed statistically significant ($p < .001$) simple simple main effects of model for all target variables (Direction: $F(3, 1176) = 93.41$, Success: $F(3, 1176) = 63.671$, Technique: $F(3, 1176) = 221.202$). All simple pairwise comparisons were run with a Bonferroni adjustment applied. Only two comparisons did not show significant differences in classification accuracy between models. The mean difference between the gradient boosted tree and the gated recurrent unit targeting the success variable, .001 percentage points (95% CI [-.001, .002]), was not statistically significant, $p = .783$. Additionally, the means between the convolutional neural net and the gated recurrent unit, .001 percentage points (95% CI [.000, .003]) showed no significant differences, $p = .119$.

Table 1: : 3-factor ANOVA-Results, sum of squares, F-value, significance p and effect size η2 for each factor, three twofold interactions and the threefold interaction (*). Cell means are depicted in figure 5.

	SS	***F***	***p***	η^2
Factor Model	.0062	282.02	< .001	.001
Factor Features	.7254	99727.01	< .001	.115
Factor Target	4.6522	319803.89	< .001	.735
Model*Features	.0001	2.52	.056	< .001
Model*Target	.0078	178.89	< .001	.001
Features*Target	.9267	63702.91	< .001	.146
Model*Features*Target	.004	92.37	< .001	.001
Error	.0086			.001

Figure 5 shows the profile plot for the three-way interaction effect, together with the mean classification accuracies per combination of model, input data and target variable. Classification accuracies for the prediction of the attack direction showed values at roughly 37% for the positions-only input, and 51% for the combined input features. In contrast, the prediction of the attack technique showed relatively similar accuracies for both the combined (55%) and the positions-only input (54%). Finally, the difference in classification precision between the different input feautres for the predicition of the success showed even smaller discrepancies with roughly 59% correctly recognized rallies for both cases.

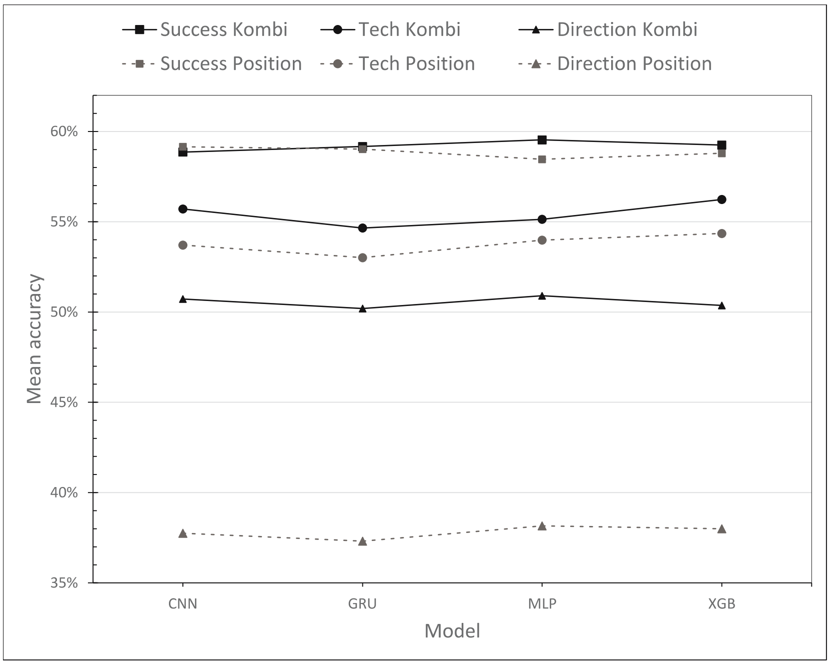


Figure 5. Profile plot for the three-way interaction effect. The values on the y-axis depict the mean classification accuracy, while the x-axis shows the employed models.

Discussion

Results revealed consistent significant differences between models in terms of classification accuracy. Our classification results achieve better performance than strict guessing in all cases (see figure 5), with prediction accuracies ranging from around 37% for the forecasting of the attack direction (guessing: 17%) to almost 60% for the prediction of success and attack technique (guessing: 50%). The absolute differences in accuracy between models are in the range of .001 percentage points, however. The significant differences are most likely caused by the small within-group variances with coefficients of variation between .02% and 1.2%, which are the result of our evaluation method.

However, these small variations also give reason to believe that there exists an upper bound to accuracy. Similar studies have found these upper bounds in other sports (Weissbock & Inkpen, 2014) and gave difference in skill, as well as random chance as reasons for their existence. Given the available data and the innate complexity of the decision-making process, we believe that these upper bounds also exist in beach volleyball.

Effect sizes reveal, that most of the variation in classification accuracy is explained by the target of the prediction. This can also be observed in figure 5 by comparing the models in the interaction plot, where the differences between the models for the different target variables are more pronounced compared to the differences between models for varying input features. The strongest increase in classification accuracy for this case can be observed when using both position and PI data for the prediction of the attack direction compared to the prediction using only positional data. This is caused by the inclusion of the attack technique as an input feature in our opinion, which would also explain the high effect sizes of input features and the interaction between target variable and input features. Different attack techniques are used as a solution to certain block and defense constellations, which in turn imply specific shot directions due to standard defensive formations. One such example would be a line block in combination with a diagonal defense, which encourages a shot to the line as offensive solution. Interestingly, all effects involving the different models show eta-squared values lower or equal than .001, suggesting that it is not the choice of model, but the appropriate choice of input features and target variable in the sports context that affect classification accuracy the most. This also means that future studies should try to include the positions of the opponent team as model inputs, since top-level beach volleyball players can perceive the opponent's defensive formation during the approach and include this information in the decision-making process for the own attack. We did not have access to this type of data and it will most likely be only available if it becomes mandatory for all athletes to wear sensors providing positional data during tournaments in the future, however.

Considering the absolute classification performance, our models are not yet at the level to warrant practical application in day to day performance analysis in beach volleyball. However, even though there are no comparable studies in beach volleyball, work in other sports has shown that performance prediction in team sports is one of the hardest problems due to the dynamical interaction process of the playing parties. As examples, Wang and Zemel (2016) reach 77.9% top-3 accuracy in classifying NBA offensive plays, while Weissbock, Viktor and Inkpen (2014) were able to correctly classify success (Win/Loss) in Ice Hockey with an accuracy of 60.25 %. Parmar, James, Hughes, Jones and Hearne (2017) achieve a prediction accuracy of 85.5% in predicting team wins in rugby from PIs, even though they use data aggregated over whole games. The performance of our models is in a similar accuracy range, considering we are not able to use continuous position data, as Wang and Zemel for example. In contrast to previous work, we are also not trying to predict match outcomes, but actions and success in single rallies, using only the prior sequence of events from that rally as input. Under these circumstances, the results are encouraging.

Even though our model parameters were determined after a thorough experimentation phase, we are confident that more specialized models, with even more sophisticated feature engineering can achieve higher accuracy values. Further studies could provide deeper comparisons between different network configurations per model (e.g. more linear layers or different hyper-parameter configurations) in order to increase classification performance. Especially the inclusion of continuous positional data, as provided by electronic tracking systems, may boost prediction performance considerably, since it inherently contains temporal information. Since we grouped the rallies for both men and women together, our models are not able to differentiate gender-specific behaviors. It could also make sense to train models for specific players or teams in order to gain more specific insights into individual tactics. This is especially important for practical performance analysis (Lames & McGarry, 2007), since it directly influences the possibilities to use our models e.g. for strategy generation against specific opponents as well as retrospective analysis of the own team for coaches and scouts. At the same time, this would greatly reduce the available training data and could lead to the inability to successfully train certain “deep” network models, however.

We did not compare models using only PI data for several reasons. Some models (e.g. RNNs and CNNs) expect data to be spatially or temporally structured in order to generate best results, which is not the case for categorical performance data. Moreover, the long-term goal using AI in performance analysis would be to rely only on positional data to automatically predict performance measures in a beach volleyball rally, without the help of (manually) collected PI data. Lastly, even though the absolute classification performance between models didn’t differ much, the gradient boosted tree model still has a particular advantage, since it allows easy interpretation of its decisions and can even assign importance to different input features to rank them for the experts. Yet, considering the recent trend in interpretability research for neural networks (Melis, & Jaakkola, 2018), it can be expected that at least some of these features will be available for neural networks in the near future too.

Future applications could also try to predict set/match outcome, for example employing recurrent neural networks to incorporate the course of the set/match. Another interesting approach would be to leverage the current advances in image recognition to classify successful rallies. Given the availability of continuous position data, this could be done by encoding player, and potentially ball positions together with additional information (e.g. velocity) in an image and using transfer learning to apply already trained models to the image classification task. This approach is already in use, for example to identify fraudulent website access via analysis of mouse movements (Esman, 2017). Another possible extension could try to group multiple models together using an ensemble method to increase classification accuracy for different target variables.

Regardless of which method will be used in the future, the prediction of tactical behavior could influence many applications and provide valuable insights in successfull rally patterns in order to support both the scouting as well as the training process in beach volleyball.

Scouts could be able to rely on models to automatically detect and update behaviors in real-time, minimizing the required inputs to capture all tactical details per rally during live-scouting. Additionally, sufficiently accurate models could be able to fill out manually scouted rallies, again reducing the workload for the scouts, who currently have to analyze beach volleyball games in a two-stage process that takes roughly three hours per game. In turn, faster analysis of games could lead to a competitive advantage in tournaments, where athletes often have less then a day to prepare for their next match.

Coaches could use models trained on single player’s or opponent team’s data in conjunction with a simulative approach to find tactical behaviours that provide the highest chance at success against certain opponents. In the same way, they could simulate variances of the behavior of their own teams to find strategies for improvement in training.

In addition to the betting market, who could employ advanced models for win/loss-prediction to offer improved betting odds, the media could use such models to enhance their broadcasts by showing live analysis (e.g. overlays for the most probable attack technique/direction used) and statistics of beach volleyball rallies, which would increase viewer engagement and help the sport to grow in popularity.

Conclusion

In conclusion, since all models show relatively similar performance, our results suggest that for the moment the available data is the limiting factor for a implementation of rally classification in beach volleyball practice. Given that our dataset may be one of the biggest collected records of beach volleyball performance data to date, the impact of machine learning techniques on todays practical applications seems small. However, if we look at the evolution of data in other sports together with the recent development in beach volleyball, the use of continuous positional data for prediction may be possible in the near future. For this reason, and because we think the search for efficient solutions in performance analysis is one of the main tasks of computer science in sports, we encourage further studies in the AI domain to evaluate the performance of methods in different sports, using different data representations.

Acknowledgements

This study was supported by a grant from the German Federal Institute of Sports Science (BISp ZMVI4-070504/19-20). We would also like to thank the scouts of the Olympic Training Center (OSP) Stuttgart, as well as the coaches of the national teams for their support in collecting the data.

References

Arndt, C., & Brefeld, U. (2016), Predicting the future performance of soccer players. *Statistical Analysis and Data Mining: The ASA Data Science Journal. 9*. 373-382.

Bialkowski, A., Lucey, P., Carr, P., Yue, Y., Sridharan, S., & Matthews, I. (2015). Large-Scale Analysis of Soccer Matches Using Spatiotemporal Tracking Data. *In Proceedings - IEEE International Conference on Data Mining*, ICDM.

Bianchi, F., Facchinetti, T., & Zuccolotto, P. (2017). Role revolution: Towards a new meaning of positions in basketball. *Electronic Journal of Applied Statistical Analysis, 10*(3), 712–734.

Bunker, R. P., & Thabtah, F. (2019). A machine learning framework for sport result prediction. *Applied Computing and Informatics, 15*(1), 27–33.

Chen, T., & Guestrin, C. (2016). {XGBoost}: A Scalable Tree Boosting System. In Proceedings of the *22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785–794). New York, NY, USA: ACM.

Cho, K., Van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). Learning phrase representations using RNN encoder-decoder for statistical machine translation. *arXiv preprint* arXiv:1406.1078.

Chung, J., Gulcehre, C., Cho, K., & Bengio, Y. (2014). Empirical evaluation of gated recurrent neural networks on sequence modeling. *arXiv preprint* arXiv:1412.3555.

Claudino, J. G., Capanema, D. de O., de Souza, T. V., Serrão, J. C., Machado Pereira, A. C., & Nassis, G. P. (2019). Current Approaches to the Use of Artificial Intelligence for Injury Risk Assessment and Performance Prediction in Team Sports: A Systematic Review. *Sports Medicine - Open, 5*(1).

Dick, U., & Brefeld, U. (2019). Learning to Rate Player Positioning in Soccer. *Big Data, 7*(1), 71–82.

Esman, G. (2017, April 18). Splunk and Tensorflow for Security: Catching the Fraudster with Behavior Biometrics. Retrieved from https://www.splunk.com/blog/2017/04/18/deep-learning-with-splunk-and-tensorflow-for-security-catching-the-fraudster-in-neural-networks-with-behavioral-biometrics.html

Hassan, A., Schrapf, N., Ramadan, W., & Tilp, M. (2017a). Evaluation of tactical training in team handball by means of artificial neural networks. *Journal of Sports Sciences*, *35*(7).

Hassan, A., Schrapf, N., & Tilp, M. (2017b). The prediction of action positions in team handball by non-linear hybrid neural networks. *International Journal of Performance Analysis in Sport, 17*(3), 293–302.

Howard, J., et. al. (2018). fast.ai. GitHub. Retrieved from https://github.com/fastai/fastai

Ioffe, S., & Szegedy, C. (2015). Batch normalization: Accelerating deep network training by reducing internal covariate shift. *arXiv preprint* arXiv:1502.03167.

Kautz, T., Groh, B. H., Hannink, J., Jensen, U., Strubberg, H., & Eskofier, B. M. (2017). Activity recognition in beach volleyball using a Deep Convolutional Neural Network. *Data Mining and Knowledge Discovery, 31*(6), 1678–1705.

Kempe, M., Grunz, A., & Memmert, D. (2015). Detecting tactical patterns in basketball: Comparison of merge self-organising maps and dynamic controlled neural networks. *European Journal of Sport Science, 15*(4), 249–255.

Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint* arXiv:1412.6980.

Kocmi, T., & Bojar, O. (2017). An exploration of word embedding initialization in deep-learning tasks. *arXiv preprint* arXiv:1711.09160.

Lames, M., & McGarry, T. (2007). On the search for reliable performance indicators in game sports. *International Journal of Performance Analysis in Sport, 7*(1), 62-79.

Lapham A.C. , & Bartlett, R.M. (1995) The use of artificial intelligence in the analysis of sports performance: A review of applications in human gait analysis and future directions for sports biomechanics, *Journal of Sports Sciences, 13*(3), 229-237.

Leicht, A. S., Gómez, M. A., & Woods, C. T. (2017). Explaining match outcome during the men’s basketball tournament at the olympic games. *Journal of Sports Science and Medicine, 16*(4), 468–473.

Link, D. (2014). A toolset for beach volleyball game analysis based on object tracking. *Int. J. Comp. Sci. Sport 13*, 24–35

Link, D., & Hoernig, M. (2017). Individual ball possession in soccer. *PLoS ONE. 12*(7): e0179953. https://doi.org/10.1371/journal.pone.0179953

Link, D. (2018). Data Analytics in Professional Soccer. Springer Fachmedien Wiesbaden.

López-Valenciano A., Ayala F., Puerta J.M., De Ste Croix M., Vera-Garcia F.J., Hernández-Sánchez S., Ruiz-Pérez I., & Myer G.D. (2018). A preventive model for muscle injuries: a novel approach based on learning algorithms. *Med Sci Sports Exerc. 50*(5). 915–927

Maas, A. L., Hannun, A. Y., & Ng, A. Y. (2013). Rectifier nonlinearities improve neural network acoustic models. *In Proc. Icml. 30*(1), 3-9.

Melis, D. A., & Jaakkola, T. (2018). Towards robust interpretability with self-explaining neural networks. In *Advances in Neural Information Processing Systems* (pp. 7775-7784).

Net gains - the evolution of beach volleyball. (2016, July 21). Retrieved from https://www.olympic.org/news/net-gains-the-evolution-of-beach-volleyball

Parmar, N., James, N., Hughes, M., Jones, H., & Hearne, G. (2017). Team performance indicators that predict match outcome and points difference in professional rugby league. *International Journal of Performance Analysis in Sport, 17*(6), 1044–1056.

Paszke, A., Gross, S., Chintala, S., Chanan, G., Yang, E., DeVito, Z., ... & Lerer, A. (2017). Automatic differentiation in PyTorch. In Long Beach, California, USA: Autodiff Workshop.

Perl, J., & Memmert, D. (2011). Net-Based Game Analysis by Means of the Software Tool SOCCER. *International Journal of Computer Science in Sport, 10*(2), 77–84.

Peterson, K. D. (2018) Recurrent Neural Network to Forecast Sprint Performance, *Applied Artificial Intelligence, 32*(7-8), 692-706

Rosenblatt, F. (1961). Principles of neurodynamics. perceptrons and the theory of brain mechanisms (No. VG-1196-G-8). Cornell Aeronautical Lab Inc Buffalo NY.

Russell, S., & Norvig, P. (2010). Artificial Intelligence: A Modern Approach - Third Edition. Pearson.

Schrapf, N., Alsaied, S., & Tilp, M. (2017). Tactical interaction of offensive and defensive teams in team handball analysed by artificial neural networks. *Mathematical and Computer Modelling of Dynamical Systems, 23*(4), 363–371.

Smith, L. N. (2017). Cyclical learning rates for training neural networks. In *2017 IEEE Winter Conference on Applications of Computer Vision* (WACV) (pp. 464-472). IEEE.

Smith, L. N. (2018). A disciplined approach to neural network hyper-parameters: Part 1-- learning rate, batch size, momentum, and weight decay. *arXiv preprint* arXiv:1803.09820.

Tümer, A. E., & Koçer, S. (2017). Prediction of team league's rankings in volleyball by artificial neural network method. *International Journal of Performance Analysis in Sport, 17*(3), 202–211.

Vales-Alonso, J., Chaves-Dieguez, D., Lopez-Matencio, P., Alcaraz, J. J., Parrado-Garcia, F. J., & Gonzalez-Castano, F. J. (2015). SAETA: A Smart Coaching Assistant for Professional Volleyball Training. *IEEE Transactions on Systems, Man, and Cybernetics: Systems, 45*(8), 1138–1150.

Van Haaren, J., Ben Shitrit, H., Davis, J., & Fua, P. (2016). Analyzing volleyball match data from the 2014 world championships using machine learning techniques. In *Proceedings of the 22Nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '16, (pp. 627-634)., New York, NY, USA. ACM.

Wang, K., & Zemel, R. (2016). Classifying NBA Offensive Plays Using Neural Networks. *MIT Sloan Sports Analytics Conference*, 1–9.

Wang, Y., Zhao, Y., Chan, R. H. M., & Li, W. J. (2018). Volleyball Skill Assessment Using a Single Wearable Micro Inertial Measurement Unit at Wrist. *IEEE Access, 6*, 13758–13765.

Weissbock, J., & Inkpen, D. (2014). Combining textual pre-game reports and statistical data for predicting success in the national hockey league. In *Canadian Conference on Artificial Intelligence* (pp. 251-262). Springer, Cham.

Zetou, E., Moustakidis, A., Tsigilis, N., & Komninakidou, A. (2007). Does Effectiveness of Skill in Complex I Predict Win in Men's Olympic Volleyball Games? *Journal of Quantitative Analysis in Sports, 3*(4).