

The identification of high potential archers based on relative psychological coping skills variables: A Support Vector Machine approach

Zahari Taha¹, Rabi Muazu Musa^{1,2}, A P P Abdul Majeed¹, Mohamad Razali Abdullah², Muhammad Aizzat Zakaria¹, Muhammad Muaz Alim¹, Jessnor Arif Mat Jizat¹ and Mohamad Fauzi Ibrahim³

¹ Innovative Manufacturing, Mechatronics and Sports Laboratory, Faculty of Manufacturing Engineering, Universiti Malaysia Pahang, 26600 Pekan, Pahang, Malaysia

² Faculty of Applied Social Sciences, Universiti Sultan Zainal Abidin, 21300 Kuala Terengganu, Terengganu, Malaysia

³ Research and Education Division, National Sports Institute, Kompleks Sukan Negara Bukit Jalil, 57700 Kuala Lumpur, Malaysia

anwarmajeed1983@gmail.com

Abstract. Support Vector Machine (SVM) has been revealed to be a powerful learning algorithm for classification and prediction. However, the use of SVM for prediction and classification in sport is at its inception. The present study classified and predicted high and low potential archers from a collection of psychological coping skills variables trained on different SVMs. 50 youth archers with the average age and standard deviation of (17.0 ± 0.056) gathered from various archery programmes completed a one end shooting score test. Psychological coping skills inventory which evaluates the archers level of related coping skills were filled out by the archers prior to their shooting tests. k-means cluster analysis was applied to cluster the archers based on their scores on variables assessed. SVM models, i.e. linear and fine radial basis function (RBF) kernel functions, were trained on the psychological variables. The k-means clustered the archers into high psychologically prepared archers (HPPA) and low psychologically prepared archers (LPPA), respectively. It was demonstrated that the linear SVM exhibited good accuracy and precision throughout the exercise with an accuracy of 92% and considerably fewer error rate for the prediction of the HPPA and the LPPA as compared to the fine RBF SVM. The findings of this investigation can be valuable to coaches and sports managers to recognise high potential athletes from the selected psychological coping skills variables examined which would consequently save time and energy during talent identification and development programme.

1. Introduction

Various researchers have endeavoured to provide insights on the impacts of psychological skills to the advancement of athlete's performance in different sports. Abdullah et al. studied the roles of psychological factors on the performance of elite soccer players and established that although, psychological factors alone could not forecast the performance of elite soccer players, but play an essential part in improving their performance [1]. Likewise, Lazarus [2] and Salim et al. [3] inferred



that psychological elements such as stress and worry could be a significant constituent in hindering athletes performance. They further emphasised that athletes who are unable to cope with stress and adversity effectively might likely suffer from performance declining and overall psychological well-being issues. Furthermore, Pasaei et al. [4] stated that in the contemporary archery competitions there are presents of various noises, tensions, and anxiety, and as such selective attention through the application of the psychological coping skills during the competition become paramount for better performance. This implies that the athletes with higher psychological coping skills are likely to outperform their opponents with lesser coping skills.

The employment of machine learning or artificial intelligence has gained popularity in prediction and classification in sports and exercise due to its superiority over conventional means. Bayesian network (BN) analysis was utilised to investigate a semi-professional team's performance and collective efficacy based on a number of psychological features [5]. In a different study, Fuster-Parra et al. [6] also employed BN in correlating negative psychological features with sportive behaviours of young team players in competitive sports.

It is evident from the literature that the employment of machine learning could be a useful tool for prediction as well as classification. Support vector machine (SVM) is a rather popular supervised learning algorithm amongst machine learning methods that have been employed by researchers in a variety of fields for both classification and regression analysis [7–10]. The underlying principle behind SVM is inherently statistical learning theory that was initially proposed by Cortes and Vapnik [11]. The present study aims at correlating the selected psychological coping skill strategies viz. coping with adversity, coachability, peaking under pressure, concentration, freedom from worry, confidence and achievement motivation, goal setting and mental preparation

2. Material and methods

2.1. Participants

A complete number of 50 archers were selected to take part in in this study. The participants consisted of 37 male and 13 female youth archers between the age's range of 13-20 with a mean and standard deviation of (17.0 ± 0.56) drawn from different archery programmes in Malaysia. The archers were under a development program for preparing both at university and the state level and consequently targeted to be preferred to state and national archers respectively. Written consent was obtained, and all the archers signed consent forms. All the procedures, protocol, and equipment for this study were authorised by the Research Ethics Board of the Terengganu Sports Institute (ISNT) with a memo number 04-04/T-01/Jid 2.

2.2. Psychological coping skills inventory assessment

The Athletic Coping Skills Inventory (ACSI), an instrument for assessing athlete's psychological skills, produced by Smith et al. [12] was selected. This instrument is considered apt for evaluating the psychological skills of the archers in this study because of its connection to the nature of archery game as an individual sport rather than a team sport. The instrument was distributed to the archers prior to the shooting test, and their responses were compiled and analysed. A simulated shooting competition area was set up, and all the archers' shoot six arrows (one end) over a distance of 50 meters. All the archers were given trials of four arrows shot before recording the final six arrows scores.

2.3. Data analysis

2.3.1. Clustering: k-means cluster analysis

In the current study, the *k*-means clustering algorithm, which is essentially a type of unsupervised learning, is used to separate the classes of the related performance variables assessed via Orange 2.7. It has been reported in the literature that the *k*-means clustering method is more reliable than the hierarchical agglomerative clustering in catering large datasets [13]. This is primarily due to the fact that it operates on actual observations rather than the dissimilarity measures employed in hierarchical clustering. The number of clusters, *k* is selected to be two as the data is consists of HPPA and LPPA.

2.3.2. Classification: Support vector machines (SVMs)

As for classification problems, the SVM is utilised to acquire the optimal hyperplane that correctly classifies or divides the data into two distinct class which in this study, these categories are represented by HPPA and LPPA. Nonetheless, it is worth to note that SVM may also be extended to cater multiclass problems [14]. The acquisition of the aforementioned optimal hyperplane is through the identification of the maximal distant from the classes that in turn minimises the risk of misclassification of both the training and validation data set. For a more in-depth mathematical treatment of the subject matter, the readers are referred to [11]. In this study, two different kernel

functions are investigated namely, linear and fine RBF. The scale of the fine RBF is defined by $\frac{\sqrt{P}}{4}$,

where P is the number of predictors, i.e. seven. The solver used for the training is the sequential minimal optimisation algorithm.

2.3.3. Model training and testing

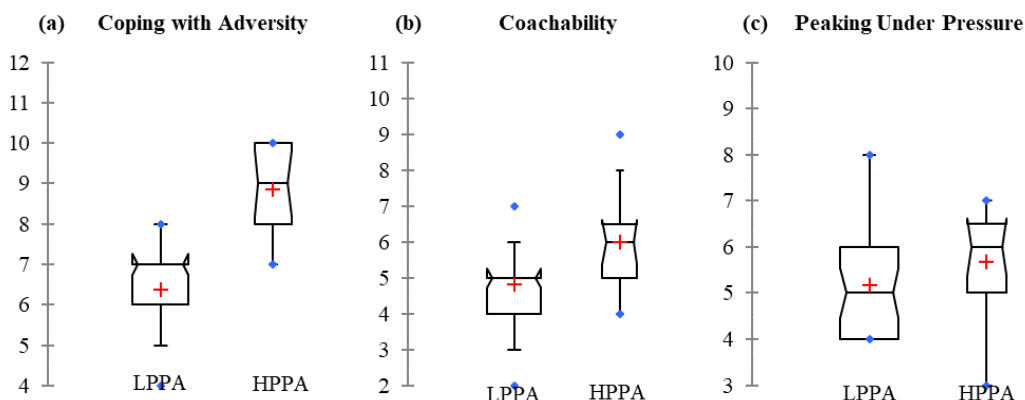
A fivefold cross-validation method was utilised in this study. This form of validation technique is desirable as it mitigates the notion of overfitting through partitioning the dataset into a number of folds and estimating the accuracy of each fold. The data (50 observations) is randomly split into five subsets, and for each iteration, one of the five subsets are used as the testing data, whilst the remaining four will be used as the training data. Then, the average performance over all the folds is then computed. The SVM analysis and assessment were conducted by means of MATLAB 2016a (Mathworks Inc., Natick, USA).

2.3.4. Model evaluation

The variations of the SVM employed in this study are evaluated by means of classification accuracy (ACC), sensitivity (SENS), specificity (SPEC), precision (PREC), error rate (ERR) as well as Matthew's correlation coefficient (MCC). The confusion matrix allows the observation of correctly classified and misclassified observations that transpires between the defined classes.

3. Results

3.1. Clustering



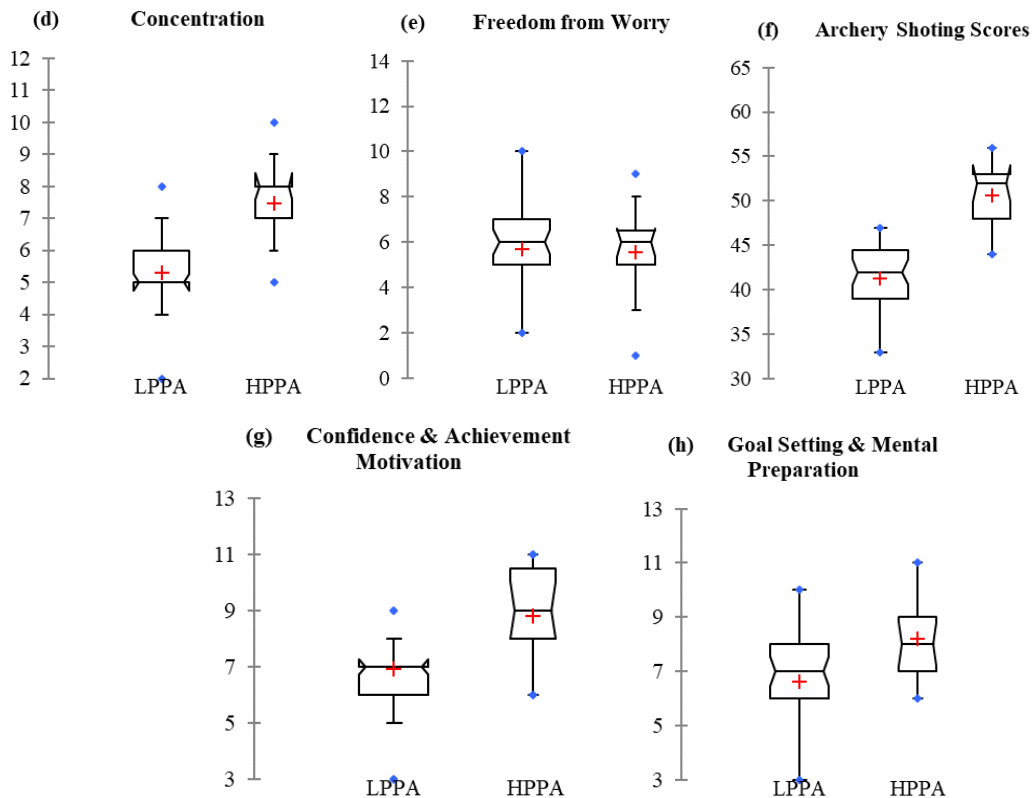


Figure 1. Comparisons of performance differences of the archers based on seven variables evaluated based on ACSI (a) Coping with Adversity; (b) Coachability; (c) Peaking Under Pressure; (d) Concentration; (e) Freedom from Worry; (f) Archery shooting score; (g) Confidence and Achievement Motivation; (h) Goal Setting and Mental Preparation

Figure 1 displays the performance differences of the archers based on the eight performance variables assessed that were clustered via *k*-means. It can be observed from the box plots that the mean performances of HPPA are greater than LPPA across all the seven psychological coping skills measured in the study with the exception of ‘freedom from worry’ (Figure 1(e)). This observation is rather natural as those who are more worried tend to have less performance. These variables are, therefore, considered as essential attributes that distinguish HPPA from LPPA.

3.2. Classification

Table 1. Model evaluation

Kernel Functions	ACC (%)	SENS (%)	SPEC (%)	PREC (%)	ERR (%)	MCC
Linear	92	100	73.33	89.74	8	0.8112
Fine RBF	70	100	0	70	30	-

It could be observed from the tabulated results that the linear kernel function based SVM model is able to produce exceptional classification through the evaluation of all assessment parameters as compared to the fine RBF. It is apparent that the linear-based model could provide a reasonably accurate classification with an error rate of 8 %. Conversely, the fine RBF-SVM model is rather unsuitable for predicting the correct classification of HPPA and LPPA, as it yields the highest error rate at 30% on top of a poor negative class accuracy demonstrated via the SPEC evaluation that in turn produces an uncorrelated model. Through the present study, it is evident that the linear-based kernel functions

SVM model may be used for the purpose of talent identification for archers based on the predefined psychological coping skills. The confusion matrix s of the evaluated SVMs are depicted in Figure 2.

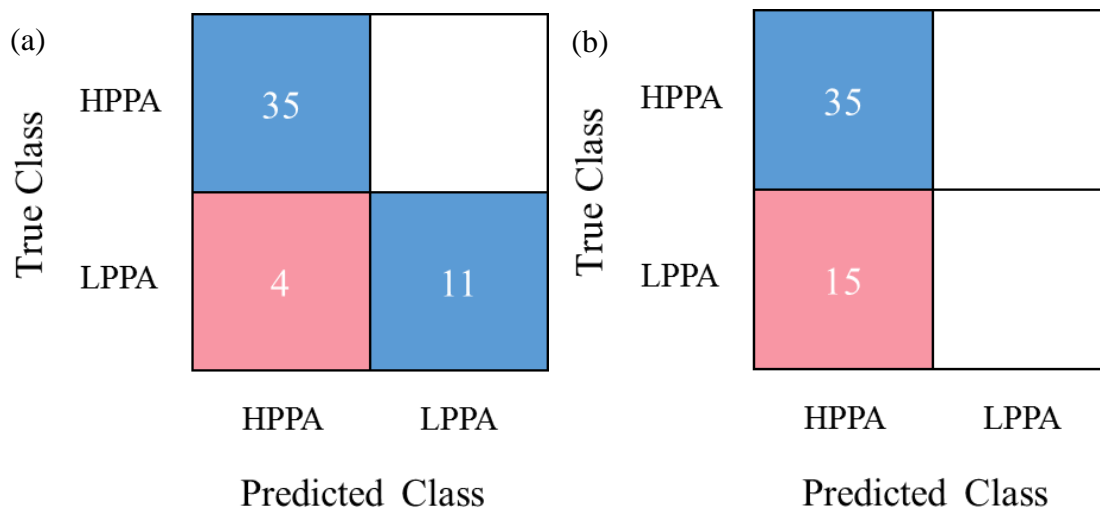


Figure 2. Confusion matrix. (a) Linear SVM; (b) Fine RBF SVM;

4. Discussion

The finding of the present study has indicated that the selected psychological coping skills variables established are able to predict or classify well the performance of the archers, i.e. HPPA and LPPA through the shooting score of the respective archers. The shooting scores enable us to cluster the athletes with respect to measured psychological coping skills via k -means as illustrated in the box plot shown in Figure 1. It was demonstrated that the linear-based kernel function SVM is able to classify the LPPA and HPPA reasonably well.

The finding underscored the necessity for psychological skills such as reliance and self-motivational strategies in the sport. Archery is closed skill as well as an individual competitive sport that calls for the archers to compete against one another. For an archer to obtain a higher score, he or she must be capable of maintaining confidence and be self-motivated to compete. An evolving body of evidence for sports indicates that one's awareness of competence or self-confidence is the crucial factor leading to the success of any chosen sport [15]. Moreover, it has been described that confidence and self-motivation are the symbols of a winner. They are the secret ingredients that all outstanding athletes seem to acquire, irrespective of what level they contest since constructive mental attitude keeps an athlete working hard regardless of how often he/she may fail or how many obstacles get tossed in his/her direction [16]. Confidence and self-motivation can provide a normal athlete or team the determination and focus on defeating a stronger opponent. These psychological skills can inspire an athlete to undertake and achieve the unthinkable. Similarly, lacking confidence and self-motivation could make athletes continually perform way below their potential. Low confidence and self-motivation can wipe out an athlete's indulgence of the sport and turn him/her into a dropout and a consistent loser [1]. This study is somewhat in its preliminary nature, and owing to its limited sample size, the results could not be generalised to other level of archery participation as well as other demography.

5. Conclusion

The current study has successfully demonstrated that some psychological coping skills variables, i.e. coping with adversity, coachability, peaking under pressure, concentration, freedom from worry, confidence and achievement motivation, goal setting and mental preparation does influence the determination the performance class of the archers. The study also has revealed the use of such machine learning methods is nontrivial as it allows coaches to correctly identify high potential athletes in the sport of archery by considering certain psychological coping skills elements. The current

findings could be useful to coaches and sports managers in discovering talents which would eventually save cost and energy during talent identification and development programmes. Future study should consider other related performance variables associated with the sport as well as other non-conventional classification methods.

Acknowledgement

This work is funded by the National Sports Institute of Malaysia (ISNRG: 8/2014-12/2014).

References

- [1] Abdullah M R, Musa R M, Maliki A B H M B, Kosni N A and Suppiah P K 2016 Role of psychological factors on the performance of elite soccer players *J. Phys. Educ. Sport* **16** 170
- [2] Lazarus R S 2000 How emotions influence performance in competitive sports *Sport Psychol.* **14** 229–52
- [3] Smith R E and Christensen D S 1995 Psychological skills as predictors of performance and survival in professional baseball *J. Sport Exerc. Psychol.* **17** 399–415
- [4] Pasaei N, Abdoli B, Kazem S, Mousavi V and Aslankhani M 2013 Does Predict Personality Introversion-Extraversion Dimension the Archer's Performance? *Middle-East J. Sci. Res.* **13** 760–3
- [5] Fuster-Parra P, García-Mas A, Ponseti F J and Leo F M 2015 Team performance and collective efficacy in the dynamic psychology of competitive team: A Bayesian network analysis *Hum. Mov. Sci.* **40** 98–118
- [6] Fuster-Parra P, García-Mas A, Ponseti F J, Palou P and Cruz J 2014 A Bayesian network to discover relationships between negative features in sport: a case study of teen players *Qual. Quant.* **48** 1473–91
- [7] Azamathulla H M and Wu F-C 2011 Support vector machine approach for longitudinal dispersion coefficients in natural streams *Appl. Soft Comput. J.* **11**
- [8] De Yong D, Bhowmik S and Magnago F 2015 An effective Power Quality classifier using Wavelet Transform and Support Vector Machines *Expert Syst. Appl.* **42** 6075–81
- [9] Guyon I, Weston J, Barnhill S and Vapnik V 2002 Gene Selection for Cancer Classification using Support Vector Machines *Mach. Learn.* **46** 389–422
- [10] Khedher L, Ramírez J, Górriz J M, Brahim A and Segovia F 2015 Early diagnosis of Alzheimer's disease based on partial least squares, principal component analysis and support vector machine using segmented MRI images *Neurocomputing* **151** 139–50
- [11] Cortes C and Vapnik V 1995 Support-vector networks *Mach. Learn.* **20** 273–97
- [12] Smith R E, Schutz R W, Smoll F L and Ptacek J T 1995 Development and validation of a multidimensional measure of sport-specific psychological skills: The Athletic Coping Skills Inventory-28 *J. Sport Exerc. Psychol.* **17** 379–98
- [13] Collins B M, Lydersen J M, Fry D L, Wilkin K, Moody T and Stephens S L 2016 Variability in vegetation and surface fuels across mixed-conifer-dominated landscapes with over 40 years of natural fire *For. Ecol. Manage.* **381** 74–83
- [14] Brereton R G and Lloyd G R 2010 Support Vector Machines for classification and regression *Analyst* **135**
- [15] Bird A M and Brame J M 1978 Self versus team attributions: A test of the "I'm OK, but the team's so-so" phenomenon *Res. Quarterly. Am. Alliance Heal. Phys. Educ. Recreat.* **49** 260–8
- [16] Hardy L, Jones J G and Gould D 1996 *Understanding psychological preparation for sport: Theory and practice of elite performers.* (John Wiley & Sons Inc)