Science & Sports (2019) 34, e241—e249

Disponible en ligne sur
ScienceDirect
www.sciencedirect.com

Elsevier Masson France
EM|consulte
www.em-consulte.com

ORIGINAL ARTICLE

The application of Artificial Neural Network and *k*-Nearest Neighbour classification models in the scouting of high-performance archers from a selected fitness and motor skill performance parameters

L'identification des méthodes de classification de réseaux de neurones artificiels et des k plus proches voisins dans le dépistage des archers de haute performance à partir d'une sélection de paramètres de performance physique et motrice

R. Muazu Musa[a,b,*], A.P.P. Abdul Majeed[b], Z. Taha[b], M.R. Abdullah[c], A.B. Husin Musawi Maliki[c], N. Azura Kosni[c]

[a] *Centre for Foundation and Liberal Education, Universiti Malaysia Terengganu, 21030 Kuala Nerus, Terengganu Malaysia*
[b] *Innovative Manufacturing, Mechatronics and Sports Laboratory, Faculty of Manufacturing Engineering, Universiti Malaysia Pahang, 26600 Pekan, Pahang, Malaysia*
[c] *Faculty of Applied Social Sciences, Universiti Sultan Zainal Abidin, 21300 Kuala Terengganu, Terengganu, Malaysia*

Received 10 April 2017; accepted 28 February 2019
Available online 15 June 2019

KEYWORDS
Archery;
Artificial intelligence;
Fitness parameters;
Motor skill;
Talent scouting

Summary
Objective. — The utilization of artificial intelligence has been demonstrated in the literature to be effective for classification and prediction. Nevertheless, the application of *k*-Nearest Neighbour (*k*-NN) and Artificial Neural Network (ANN) specifically the conventional feed forward Multilayer Perceptron (MLP) model for forecasting and scouting of high-performance archers have not been fully utilized. The current investigation predicted high and low-performance archers from a set of selected fitness and motor skill parameters trained on two distinct machine learning algorithms viz. ANN and *k*-NN.

* Corresponding author. Innovative Manufacturing, Mechatronics and Sports Laboratory, Faculty of Manufacturing Engineering, Universiti Malaysia Pahang, 26600 Pekan, Pahang, Malaysia.
E-mail address: rabiumuazu86@gmail.com (R. Muazu Musa).

https://doi.org/10.1016/j.scispo.2019.02.006

Methods. — A sample of 50 youth archers with the average age and standard deviation of (17.0 ± 0.56) recruited from varying youth archery schemes completed a one end archery score test. Standardize physical fitness and motor skill parameters measurements constituting of the hand grip, vertical jump, standing broad jump, static balance, upper muscle strength and the core muscle were carried out. The Hierarchical Agglomerative Cluster Analysis (HACA) with Mahalanobis' distance was employed to group the archers with regard to the performance parameters assessed. The t statistic and Cohen's d effect size analysis were carried out on the group defined by the HACA to view through the performance differences of the archers. The ANN (single hidden layer with ten neurons) and k-NN (fine Euclidean-based) models were trained based on the measured performance variables. The tenfold cross-validation technique was utilized in the study.

Results. — The HACA grouped the archers into two distinct clusters namely; high-performance archers (HPA) and low-performance archers (LPA). It was observed from the t-statistic as well as the effect-size analysis that the performance of the HPA archers differed from the LPA in standing broad jump, hand grip, upper muscle strength as well as the archery shooting score $P < 0.05$ with a large to moderate effect-size $d = 0.8–0.6$. It was established that the ANN model outperformed the k-NN in the present study. The ANN demonstrated reasonably excellent classification on the evaluated indicators with a classification accuracy of 92% and a stronger Matthews correlation coefficient, i.e. 0.816 amongst other performance metrics in comparison to the k-NN model in classifying the HPA and the LPA.

Conclusion. — These findings are invaluable to coaches and sports officials, particularly in the identification of high-performance archers from a consolidation of the selected few evaluated fitness and motor skill performance parameters. As a consequence, this approach, in turn, would save resources, time and energy during a talent search program.

MOTS CLÉS
Tir à l'arc ;
Intelligence artificielle ;
Indicateurs des capacités physiques ;
Capacité motrice

Résumé

Objectif. — L'utilisation de l'intelligence artificielle a été démontrée dans la littérature comme pouvant être efficace pour la classification et la prédiction. Néanmoins, l'application de la méthode des k plus proches voisins (k-NN) et le réseau de neurones artificiels (ANN), en particulier le modèle classique de Perceptron multicouche (MLP) pour la prédiction et le classement en tir à l'arc est encore à ses débuts. L'étude a permis de prédire les performances potentielles d'archers à partir d'ensemble de variables de compétences physiques et de capacités motrices, au moyen de deux algorithmes d'apprentissage automatique distincts, ANN et k-NN.

Méthodes. — Cinquante jeunes archers d'un âge moyen (et d'écart-type) de $17{,}0 \pm 0{,}56$ ans, recrutés à partir de programmes de jeunes ont réalisé un jet de tirs à l'arc tir au cours duquel le score a été mesuré. Des mesures de capacités physiques ont été réalisées à partir de tests standard, force de préhension, saut vertical, saut en longueur, équilibre statique, force musculaire des membres supérieurs. L'analyse de regroupement hiérarchique en sous-groupes (HACA) en utilisant la distance de Mahalanobis a été utilisée pour regrouper les archers en fonction des performances. L'analyse de la taille de l'effet en utilisant le t de Cohen d sur le groupe défini par la HACA a permis d'évaluer les différences de performance des archers.

Résultats. — La HACA a permis de regrouper les archers en deux groupes distincts, à savoir : les archers à haut potentiel (HPA) et les archers à plus faible potentiel (LPA). L'analyse statistique ainsi que l'analyse de la taille des effets ont montré que les performances du saut en position debout, de force de préhension de la main, de la force musculaire des membres supérieurs, ainsi que du score de tir à l'arc des archers classés HPA étaient meilleures que celles mesurées chez les archers LPA ($p < 0{,}05$; taille d'effet grande à modérée $d = 0{,}8–0{,}6$). Comparativement à la méthode des k-NN, le ANN propose une meilleure classification du potentiel des archers (HPA et LPA) à partir les indicateurs évalués, avec une précision de classification de 92 %, et un fort coefficient de corrélation (Matthews, 0,816).

Conclusions. — Ces résultats sont de grande importance pour les entraîneurs et les officiels sportifs, en particulier pour identifier les archers de haut niveau, à partir de paramètres physiques et fonctionnels d'évaluation de la condition physique et de la motricité. En conséquence, cette approche économiserait des ressources, du temps et de l'énergie lors d'un programme de recherche de talents.

1. Introduction

Archery performance is assessed with regards to the frequency of the arrows scores shot by an archer that hit a predetermined target. Successful performance in the sport is believed to be influenced by various components including environmental conditions, fatigue, and the archer's skill that intermingle to regulate the behaviour of the bow and the arrow towards the target [1]. The placement of the arrows shot on the target board is understood to indicate a valuable clue to evaluate the achievement of the archer. This vital clue sheds light on the response of the archer to such complex elements influencing the outcome of the sport [2]. Research have long documented the relationship as well as the contribution of a number of fitness parameters and motor skill such as core body strength, upper body strength, handgrip, leg power and static balance towards the accomplishment of higher archery scores [1,2]. It was demonstrated that the aforesaid performance variables are vital since the nature of the sport involve several actions as well as a synergy of both the gross and fine activations of the human muscle, which regulate the body of the archer in the process of aiming and release of the arrow [3,4]. Furthermore, other researchers have recently reiterated the contributions of physical fitness and specific motor skill motor related performance parameters to the achievement of high archery shooting scores [5]. Moreover, Musa et al. [6] reported that the interactions of both fitness and motor skills could be a necessary prerequisite for a favourable performance in archery sport. They opined that these sets of vital performance related variables could hold up the archer's resistance in regulating the recommended posture and thus fit the archer's body to the possessions of the necessary skills akin to the sport.

Different machine learning models have been utilised over the years to predict and classify the performance of different sporting activities, activity types as well as physiological properties, primarily due to its advantage against classical methods [5]. Hidden Markov Models (HMM) have been employed in classifying different physical activities through the utilisation of on-body accelerometers [6]. The accelerometers were placed on the hip, wrist, ankle, thigh and ankle. It was shown that the developed model, i.e., the cHMM-based sequential classifier was able to demonstrate a classification accuracy of 99.1%.

Artificial Neural Networks (ANN) has also been used to predict different physical activity (PA) type as well as the energy expenditure (EE) associated with it amongst youth [7]. Oxygen uptake as well accelerometer data was taken from 100 participants between the age of 5 to 15 years old with equal representation from both genders. Two different ANN models were developed for both the PA and EE, respectively. Both models consist of two hidden neurons, whereby the activation function for the hidden and output layer for the PA model is the logistic function, whilst for the EE model the activation function of the hidden and output layer is the logistic and linear function. It was shown from the investigation that the classification accuracy attained via the ANN method is 88.4% whilst the root mean square error (RMSE) accounted for the metabolic equivalents (METs) is 0.9. Moreover, it was demonstrated that the ANN model is able to outperform other regression models.

The classification of PA as well as the prediction of EE by means of random forest classifier as well as random forest regression trees was investigated by Ellis et al. [8]. The data were obtained via heart rate (HR) data and accelerometer data. The accelerometer data was extracted from the right hip, left hip apart from the non-dominant wrist from forty adults. A sensitivity test was carried out on the inclusion of either the hip, wrist or the combination of both towards the classification accuracy of the PA. The leave-one-out cross-validation technique was utilised in the study. It was shown that the inclusion of HR information improved the MET estimation, nonetheless it did not significantly improve the accuracy of the PA classification.

The identification of different activity type of preschool children via different machine learning models was investigated by Hagenbuchner et al. [9]. The typical ANN model, i.e., feedforward multilayer perceptron (MLP), as well as the deep learning ensemble network (DLEN), were investigated of its efficacy to classify sedentary, moderate to vigorous, light, walking and running activities. The data were extracted from accelerometers placed on the right hip of eleven children between the age of three and six years old. A sensitivity test was carried out by evaluating the effect of window size for the accelerometer data namely 10, 30 and 60 seconds towards the classification accuracy. The leave-one-out cross-validation technique was also utilised in the investigation, primarily owing to the limited number of participants. It was demonstrated that the DLEN provided a better overall classification accuracy of 82.6% whilst the MLP ANN, 69.7% in the event that the 1960s window is used.

Staudenmayer et al. [10] investigated the efficacy of ANN in estimating MET as well as classifying activity type via data collected through Actigraph accelerometers. The physical activity types evaluated includes sedentary, light, moderate, and vigorous intensities. The leave-one-out cross-validation technique was utilised to assess the performance of the ANN models developed. It was demonstrated that the MET ANN predictive model is able to estimate well the MET with a root-mean-squared error of 1.22, whilst the activity type classification model is able to provide a classification accuracy of 88.8%, in turn suggesting the promising capability of ANN to predict and classify.

Pavey et al. [11] have also utilised RF classifier to classify a variety of PA namely sedentary, stationary+, walking and running via acceleration data attained via accelerometer placed on the wrist. It is worth noting that a variety of time and frequency domain features were extracted from the acceleration data. The performance of the model was evaluated via the leave-one-out-cross-validation technique. It was demonstrated from the study that the RF algorithm yields an overall classification accuracy of 92.7%, suggesting the efficacy of the classifier in distinguishing the assessed PAs.

Trost et al. [12] investigated the efficacy of a regularised logistic regression (LR) model to classify different activity classes through acceleration signals obtained from the wrist and hip from 52 children between the age of ten to sixteen years old. The activities intended to be classified are lying down, standing, sitting, running, walking, dancing as well as basketball. The time-domain features extracted from the acceleration data include mean, standard deviation, the coefficient of variation, percentiles (10th, 25th, 50th, 75th, 90th), lag-one autocorrelation, skewness, kurtosis, signal power, log energy, peak intensity, zero crossings, as well as cross-axis correlation. Efficacy of the developed LR model is appraised via a three-fold cross-validation method. It was demonstrated that the overall classification accuracy attained from the hip and wrist data were 91% and 88.4%, respectively, suggesting that the LR model is able to well discriminate the evaluated PAs.

It is apparent from the literature that the use of the aforesaid intelligent techniques amongst others is beneficial for the investigation of both classifications as well as prediction. The multilayer perceptron (MLP) is a feed-forward ANN model that consists of three layers, i.e. input, hidden and output layers. It is a common yet powerful black-box supervised learning algorithm that is able to cater complex nonlinear mapping both for classification as well as a prediction as demonstrated in the aforementioned literature. More often than not, the number of neurons in the hidden layer is varied to obtain desirable results, however, it is noteworthy to mention that care must be taken whilst varying the number of neurons to avoid overfitting [13].

The *k*-Nearest Neighbour (*k*-NN) is one of the simplest types of supervised machine learning algorithms that could be employed in both regression and classification problems [14]. This elementary machine learning model is also recognised as 'lazy learning' or 'instance-based learning' as it does not require learning, in other words, the computation of the algorithm occurs during runtime as it memorises the training dataset [15]. It is worth noting that the *k*-NN model has been applied successfully for a number of classification problems in diverse disciplines [16—20].

To the best of the authors' knowledge, the efficacy of ANN and *k*-NN to classify high-performance athletes in the sport of archery has not yet been fully utilized. The objective of the present investigation is to examine the relationship of the selected fitness and motor skill parameters (hand grip, vertical jump, standing broad jump, static balance, upper muscle strength and core muscle strength), towards the performance of archery sport and to utilise the parameters in the scouting of future performance archers with potential. The HACA grouping method is used to group the archers on the basis of their performances in the selected parameters as well as their shooting scores as a result of which two groups are identified, i.e. high-performance archers (HPA) and low-performance archers (LPA). In addition, the efficacy of the ANN and *k*-NN learning algorithms are assessed in classifying the HPA and LPA. A detailed methodology of this process is provided in the subsequent section.

2. Methods

Fifty archers were enlisted in this investigation. The archers comprised of 37 male and 13 female youth between the age's range of 13—20 with a mean, and standard deviation of (17.0 ± 0.56) gathered from various archery schemes in Malaysia. The archers were under an expansion program for development both at university and the state level, which in turn, consequently set to be selected for representation both at state and national archery level competitions. Normality test was conducted with regards to the performance parameters evaluated i.e. sit up, push up, standing stork test, handgrip test, standing broad jump test and vertical jump test as well as the archery shooting score by means of Shapiro-Wilk, and the scores of the archers were observed to be equally distributed. It is worth to mention that prior to any data collection in the present study, the coaches and the managers of all the programmes were fully informed about the purpose of the research. Signed permission was obtained, and all the archers signed consent forms. All the procedures, protocol, and equipment for this study were reviewed and allowed by the Research Ethics Board of the Terengganu Sports Institute (ISNT) with an endorsement number of 04-04/T-01/Jid 2.

A standardised fitness and motor skill related performance parameters constituting of sit up, push up, standing stork test, handgrip test, standing broad jump test and vertical jump test was carried out as per the recommendation of ideal physical fitness assessments [21]. The archers were instructed to perform a warm-up involving some stretches and a light jog lasting for about 5- to 10-min before the assessments sessions commence. The push-up and sit up tests were conducted utilising a mat spread on the floor. The archers performed the tests alternately under a time span of 1-minute once for each test. The standing stock test was administered employing a stopwatch. The archers removed their shoes and rested their hands on their hips, lifted up their heel to settle on the ball of the foot then placed the non-raised foot against the inner part of the supporting leg. The test is terminated when the archers could no longer maintain the precise position. The grip strength of both hands was evaluated applying a standard adjustable grip strength dynamometer (Takei Scientific Instruments Co., LTD). The archers were directed to remain in a straight posture with the shoulder in abduction and neutral rotation with an elbow in maximum extension. The standing broad jump was measured using a landing mat placed on the flat synthetic surface, and the take-off line explicitly indicated from the back of the mat. Take-Off with the two feet is permitted with the swaying of the arms and bending of the knees to give forward momentum. Three trials were allowed, and the best was used for the statistical analysis. The vertical jump was determined using Vertec (M-F Athletic Co., Cranston, Rhode Island). The researchers arranged the length of the colour-coded plastic vanes such that it corresponded to the archer's standing height. The archers were instructed to flex the ankles, knees, and hips and oscillated the arms in an upward motion whilst touching the highest possible vane utilising the fingers of their dominant hand. The performance of the best jump out of the three trials was subjected to further statistical analysis. The archery evaluation test score was undertaking before the aforesaid physical fitness motor skill evaluations. The archers shot a total of 10 arrows over 50 meters range using a typical archery competition shooting target board. The archers were permitted to take a shot of four arrows as trials whilst the final six arrows scores are utilised for statistical analysis [21,22].

The in the present investigation, the hierarchical agglomerative cluster analysis (HACA) is used to ascertain the

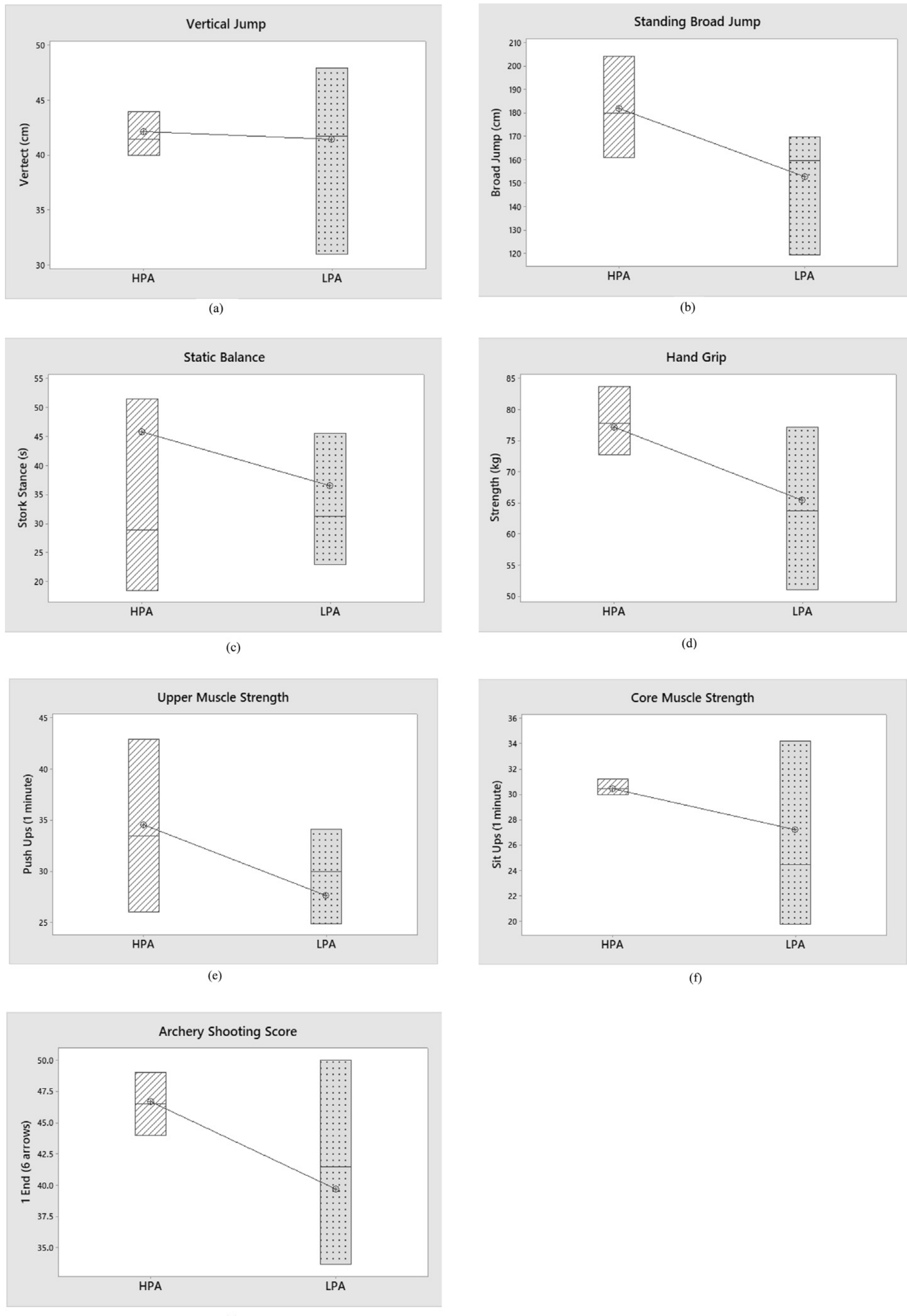

Figure 1 Appraisals of performance differences of the archers in relation to the variables evaluated: vertical jump (a); standing broad jump (b); static balance (c); hand grip (d); upper muscle strength (e); core muscle strength (f); Archery shooting score (g).

grouping of the archers with regard to all the performance parameters measured in the study. HACA is an exploratory and unsupervised technique in which a hierarchy of clusters are formed starting from one observation, and subsequently, identical observations are merged into a single cluster as the hierarchy are built from a dataset [23]. The merges and the splits of the data are ascertained in a greedy manner, i.e. based on the peculiar similarity of the observations, and the results are generally displayed in a dendrogram [24]. The number of clusters is demonstrated from the dendrogram as well as their proximity, which are observed to be two, i.e. HPA and LPA. The Mahalanobis' distance was used in this investigation, and the validation of the clustering was performed using the class centroids.

Additionally, the *t*-statistic, as well as the Cohen's *d* effect size analysis, were applied on the group defined by the HACA in order to ascertain the performance differences of the archers based on parameters measured. It is worth to mention that the effect size analysis and t-statistics are further carried out in the present study due to their relative importance in facilitating the interpretation of the differences in performances of the two groups as opposed to only reporting a statistical significance or a graphical presentation [25]. It should also be noted that in an effect-size analysis, a d of 0.20, 0.50 and 0.80 with a Pearson r of 0.10, 0.30, and 0.50 represents a mean difference of small, medium and large respectively [26].

The supervised learning for classification, on the other hand, is carried out by means of machine learning methods viz. the MLP based ANN and the k-NN. As for the MLP model, a single hidden layer model with a Rectified Linear Units (ReLU) activation function with the Adaptive Moment Estimation (Adam) optimisation algorithm was employed. The number of hidden neurons was selected to be 100. The concept of k-NN is fairly straightforward in which a sample is assigned to a predefined class per the majority of its k-nearest neighbour in the data space. The distance matrix is then employed to compute the distance of the individual samples from all the other samples prior to being sorted based on the distance. In this study, the number of neighbours, k selected for the k-NN model is 1, and as for the distance metrics, the Euclidean distance is utilised.

In the present investigation, the cross-validation technique employed is the k-fold method, where the number of fold selected is ten. This technique is employed as it has been reported to mitigate the notion of overfitting by subdividing the dataset into a number of folds and evaluating the accuracy of each fold [11,27,28]. Fifty observations are randomly divided into ten subgroups, and for each iteration, one of the ten subsets is put forth as the testing data, whilst the remaining nine will be used as the training data. Then, the average performance over all the folds is then computed. The performance of both the MLP ANN and the k-NN models were assessed and appraised via Orange 2.7.

The classification models employed in this study are assessed by means of classification accuracy (ACC), specificity (SPEC), precision (PREC), sensitivity (SENS), error rate (ERR), as well as the Matthew's correlation coefficient (MCC). The ACC is essentially the ratio between the number of correctly classified observations and the total number of observations. The SENS and the SPEC are the true positive rate or the positive class accuracy as well as the true negative rate or negative class accuracy, respectively. The PREC computes the number of correct positive predictions over the total number of positive predictions. The ERR, on the other hand, appraises all misclassifications over the number of total observations. Conversely, the MCC measures the quality of binary classification, and it has a range of -1 to 1 whereby 1 suggests an entirely accurate binary classifier.

The confusion matrix allows the observation of correctly classified and misclassified observations that ensues between the defined classes. The detailed method of acquiring the aforementioned assessment parameters is given in Appendix A.

3. Results

Fig. 1 displays the performance differences of the archers based on the seven related performance parameters measured. It can be seen from the box plots that the average performances of HPA are higher than LPA in relation to all the seven-physical fitness and motor skills tested in the study.

Table 1 Inferential statistics and the effect-size analysis of the assigned HPA and LPA archers.

Performance variables	Mean Diff	Sig	Cohen's d	Effect-Size r
Vertical jump	0.715	0.789	0.074	0.037
Standing broad jump	29.07	0.011*	0.801	0.376
Stork balance	9.314	0.291	0.288	0.142
Hand grip	11.723	0.015*	0.771	0.359
Push muscle strength	6.933	0.042*	0.613	0.296
Core muscle strength	3.221	0.130	0.425	0.208
Archery shooting score	7.018	0.038*	0.7001	0.330

* $P < 0.05$.

Table 1 demonstrates the inferential statistics as well as the effect size analysis of the archers under study. The analysis is essentially carried out to examine the statistical differences between the HPA and LPA archers with regards to the performance parameters examined. It could be seen from the table that the performance of the HPA archers differ from the LPA in standing broad jump, hand grip, upper muscle strength as well as the archery shooting score $P < 0.05$. Moreover, to further ascertain the differences between the group of the archers, the effect-size analysis of the Cohen's d and the Pearson r are computed. A large to moderate effect-size is observed in the performance of the aforementioned performance parameters ($d = 0.8$-0.6). Correspondingly, an r value ranging from 0.4 to 0.3 are found in the significant performance variables revealing the strength of the differences in performance between the HPA and the LPA. Thus, the analysis has shown that the differences in the performance of the parameters by the archers are non-trivial. In other words, the greater performance detected from the HPA could not have been by chance as such standing broad jump, hand grip, upper muscle strength as well as the archery shooting score distinguish the HPA from the LPA archers.

Table 2 tabulates the performance of the evaluated between the k-NN and the MLP ANN algorithms employed. It is evident that the MLP outperforms k-NN across all performance metrics. Furthermore, a relatively strong correlation via the MCC is observed between the classification and the selected performance variables for both models, suggesting the significance of the selected parameters. Fig. 2 depicts the performance of the confusion matrix of the classification models.

4. Discussion

It has been established from the present investigation that high-performance archers could be determined in relation to their performance on the set of performance parameters evaluated. The shooting scores allow us to group the archers with respect to the measured fitness and motor skill performance parameters particularly; hand grip, vertical jump, standing broad jump, static balance, upper muscle strength, and the core muscle strength using HACA clustering technique. The HACA demonstrated two groups of archers, i.e. HPA and LPA as shown in Fig. 1. Moreover, the box plot portrayed in Fig. 1 as well as Table 1 implies that the mean performances of HPA are higher than LPA in reference to

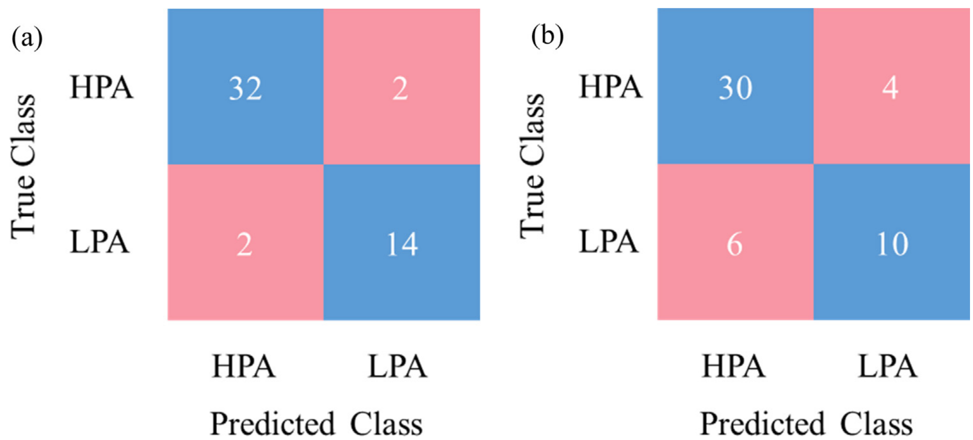

Figure 2 Figure depicts the performance of the confusion matrix of the classification models. Confusion matrix. MLP (a); *k*-NN (b).

Table 2 The performance evaluation of the different machine learning algorithms.

Algorithm	ACC (%)	SENS (%)	SPEC (%)	PREC (%)	ERR (%)	MCC
ANN	92	94.12	87.50	94.12	8	0.816
k-NN	80	83.33	71.43	88.24	20	0.527

hitherto performance related parameters assessed specifically standing broad jump, hand grip as well as upper muscle strength.

It was demonstrated from the present investigation that both the *k*-NN and MLP ANN models evaluated are able to classify the performance of the archers reasonably well as it was shown that both models are able to attain a classification accuracy above 80%. Nonetheless, it is apparent that the ANN model is more superior than the *k*-NN model as it was able to classify accurately the LPA and HPA to up to 92%. This is evident as it misclassifies only two archers from both LPA and HPA classes as shown in Fig. 2. In a nutshell, the findings from the current investigation have demonstrated that the performance parameters assessed do have a relationship with archery performance as the usage of the parameters provide a reasonably accurate model for the identification of the HPA and the LPA. The significance of the chosen performance related variables has been pointed out as essential variables for a successful performance of archery sport in various literatures especially standing broad jump, hand grip and the upper muscle strength.

The ability of an individual to execute a higher standing broad jump is to a large extent determines the leg as well as the power of the lower body [29]. Tinazci [30] inferred that one of the essential components of attaining better stability in archery is having a proper stance, which is achieved, by a greater level of lower body strength. The term stance mainly refers to the standing posture of the archers. The attainment of a proper stance entails the strength in the legs, and the right stance can help to maintain the stability while standing for an extended duration. Therefore, during the arrow release, an archer is required to maintain a static balanced in order to achieve a consistent and precise shot. The previous researchers have also recognized that in the sport of archery, the archers are needed to stand erectly for a long time as a result of which the lower body strength supported by legs helps them to sustain the necessary body balance needed during the overall arrow shooting period [31]. Thus, leg power enables a good stance that could assist the archer to sustain the steadiness for a longer time, that in turn, benefits them to aim to the target board consistently. When the body is steady, the archer's shots are bound to be more consistent and less frustrating.

Kolayis, Çilli, Ertan, and Knicker [32] have demonstrated that successful performance in archery is connected with a tremendous upper muscle strength that provides the archers with the ability to sustain the more extended hour of lifting the bow, pulling string as well as shooting. The fusion of the upper muscles to react to the request placed upon them during the execution of the aforesaid activity enables the archer to shoots the arrows consistently without undue fatigue. Moreover, previous researchers described that drawing a bowstring induces stress on the muscles of the arms, chest, shoulders and the back [33]. In tandem with weightlifting, this stress is generally sustained for several seconds prior to launching the arrow toward the target. Hence, the repetitive recurrence of lifting the bow, drawing as well as shooting the arrow results in the development of muscle stress in most of the major muscle groups of the upper body. The level of the stress-induced depends on both the draw strength of the bow and the amount of time that the archer spends. Hence, the ability of the archer to maintain steady shoot at the target is dependent mainly on the strength of the upper muscles to cope with the stress placed upon them.

Handgrip ability has been demonstrated to predict success in many sporting activities [34,35]. Likewise, in the sport of archery handgrip strength is described as one of the subcomponents of achieving higher accuracy and consistency of a shoot [36]. The gripping ability provides an advantage for limiting the sway of the bow, which in turn increases the aiming precision as well as the consistency of an arrow towards the set target. The effect is that, when pulling and holding the bow, the archer requires a high stability of the bow. Furthermore, evidence has demonstrated that in a precision sport like archery, the steadiness, as well as the precision of the arrow shot, is usually affected by some natural condition, such as the wind blow [37]. When there is a higher wind blows it results in shifting the arrow away from its target and thus the more significant variance of the arrow's bearing.

Although archery performance could be affected by a myriad of factors since the nature of the sport involves several encompassing performance elements. Nonetheless, it can be deduced from the findings of the present study that there is a relationship between the upper muscular strength, leg power as well as the hand gripping ability towards the attainment of a successful archery performance. Therefore, to accomplish an effective archery technique of lifting as well as pulling of the bowstring, there is a necessity for higher arm strength, hand grip and the relating efficiency to sustain balance.

5. Conclusions

The findings from the current investigation have demonstrated certain performance-related parameters especially hand gripping ability, leg power, as well as the upper muscle strength are capable of ascertaining the performance quality of the archers. The study likewise discovered that the application of artificial intelligence-based models, specifically, the *k*-NN and MLP ANN can reasonably predict the class of the archers in light of the selected performance-related variables investigated at a reasonable level. The accompanying classifier of ANN has displayed an excellent level of precision and accuracy with a fewer misclassification rate throughout the training and validation processes. It could, therefore, be inferred that the application of machine learning techniques is vital particularly in the identification of high-performance archers from a consolidation of the selected few evaluated fitness and motor skill performance parameters. As a consequence, this approach, in turn, would save resources, time and energy during a talent search programme. Future studies should be directed towards the provision of insightful information regarding other related performance parameters that impact performance in the sport by means of non-conventional classification techniques.

Disclosure of interest

The authors declare that they have no competing interest.

Acknowledgement

The researchers wish to thank the National Sports Institute of Malaysia for providing the grant for the current study (ISNRG: 8/2014-12/2014). The researchers have no conflict of interest to declare.

Appendix A. Appendix A

Where TP, TN, FP and FN are true positives (the number of positive samples correctly predicted), true negatives (the number of negative samples predicted correctly), false positives (negative samples predicted as positive) and false negatives (number of positive samples predicted as negative), respectively. The formulae of the classifier performance assessment parameters are as follows:

Table 2 Table 1 Confusion Matrix.

	Predicted class	
Actual Class	TP	FN
	FP	TN

$$ACC = \frac{TP + TN}{TP + TN + FP + FN}$$

$$SENS \frac{TP}{TP + FN}$$

$$SPEC = \frac{TN}{TN + FP}$$

$$PREC = \frac{TP}{TP + FP}$$

$$ERR = \frac{FP + FN}{TP + TN + FP + FN}$$

References

[1] Kim H-B, Kim S-H, So W-Y. The relative importance of performance factors in Korean Archery. J Strength Cond Res 2015;29:1211–9, http://dx.doi.org/10.1519/JSC.0000000000000687.

[2] Humaid H. Influence of arm muscle strength, draw length and archery technique on archery achievement. Asian Soc Sci 2014;10:28.

[3] Taha Z, Musa RM, Abdullah MR, Razman MAM, Lee CM, Adnan FA, et al. The application of inertial measurement units and wearable sensors to measure selected physiological indicators in archery. Asian J Pharm Res Heal Care 2017;9:85–92.

[4] Wang Y, Ma Y. Physique characteristics of ethnic archers based on R-cluster matrix; 2015.

[5] Muazu Musa R, Taha Z, Abdul Majeed PP, Abdullah AMR. Machine learning in sports. Singapore: Springer Singapore; 2019, http://dx.doi.org/10.1007/978-981-13-2592-2.

[6] Mannini A, Sabatini AM, Mannini A, Sabatini AM. Machine learning methods for classifying human physical activity from on-body accelerometers. Sensors 2010;10:1154–75, http://dx.doi.org/10.3390/s100201154.

[7] Trost SG, Wong W-K, Pfeiffer KA, Zheng Y. Artificial neural networks to predict activity type and energy expenditure in youth. Med Sci Sports Exerc 2012;44:1801–9, http://dx.doi.org/10.1249/MSS.0b013e318258ac11.

[8] Ellis K, Kerr J, Godbole S, Lanckriet G, Wing D, Marshall S. A random forest classifier for the prediction of energy expenditure and type of physical activity from wrist and hip accelerometers. Physiol Meas 2014;35:2191–203, http://dx.doi.org/10.1088/0967-3334/35/11/2191.

[9] Hagenbuchner M, Cliff DP, Trost SG, Van Tuc N, Peoples GE. Prediction of activity type in preschool children using

machine learning techniques. J Sci Med Sport 2015;18:426—31, http://dx.doi.org/10.1016/j.jsams.2014.06.003.

[10] Staudenmayer J, Pober D, Crouter S, Bassett D, Freedson P. An artificial neural network to estimate physical activity energy expenditure and identify physical activity type from an accelerometer. J Appl Physiol 2009;4:107.

[11] Pavey TG, Gilson ND, Gomersall SR, Clark B, Trost SG. Field evaluation of a random forest activity classifier for wrist-worn accelerometer data. J Sci Med Sport 2017;20:75—80, http://dx.doi.org/10.1016/j.jsams.2016.06.003.

[12] Trost SG, Zheng Y, Wong W-K. Machine learning for activity recognition: hip versus wrist data. Physiol Meas 2014;35:2183—9, http://dx.doi.org/10.1088/0967-3334/35/11/2183.

[13] Yusri IM, Abdul Majeed APP, Mamat R, Ghazali MF, Awad OI, Azmi WH. A review on the application of response surface method and artificial neural network in engine performance and exhaust emissions characteristics in alternative fuel. Renew Sustain Energy Rev 2018;7:90, http://dx.doi.org/10.1016/j.rser.2018.03.095.

[14] Han J, Pei J, Kamber M. Data mining: concepts and techniques. Elsevier; 2011.

[15] Cunningham P, Delany SJ. k-Nearest neighbour classifiers. Mult Classif Syst 2007;34:1—17.

[16] Adeniyi DA, Wei Z, Yongquan Y. Automated web usage data mining and recommendation system using K-Nearest Neighbor (KNN) classification method. Appl Comput Informatics 2016;12:90—108, http://dx.doi.org/10.1016/j.aci.2014.10.001.

[17] Bidder OR, Campbell HA, Gómez-Laich A, Urgé P, Walker J, Cai Y, et al. Love Thy Neighbour: automatic animal behavioural classification of acceleration data using the K-Nearest Neighbour Algorithm. PLoS One 2014;9:e88609, http://dx.doi.org/10.1371/journal.pone.0088609.

[18] Holmström H, Nilsson M, Ståhl G. Simultaneous estimations of forest parameters using aerial photograph interpreted data and the *k* Nearest Neighbour Method. Scand J For Res 2001;16:67—78, http://dx.doi.org/10.1080/028275801300004424.

[19] Horton P, Nakai K. Better prediction of protein cellular localization sites with the k nearest neighbors classifier. Proc Int Conf Intell Syst Mol Biol 1997;5:147—52.

[20] Mehrotra R, Sharma A. Conditional resampling of hydrologic time series using multiple predictor variables: a K-nearest neighbour approach. Adv Water Resour 2006;29:987—99, http://dx.doi.org/10.1016/j.advwatres.2005.08.007.

[21] Taha Z, Musa RM, Abdul Majeed PP, Alim A, Abdullah MMMR. The identification of high potential archers based on fitness and motor ability variables: a Support Vector Machine approach. Hum Mov Sci 2018;57:184—93, http://dx.doi.org/10.1016/j.humov.2017.12.008.

[22] Spratford W, Campbell R. Postural stability, clicker reaction time and bow draw force predict performance in elite recurve archery. Eur J Sport Sci 2017;17:539—45, http://dx.doi.org/10.1080/17461391.2017.1285963.

[23] Lavra N. Data Mining and Knowledge Discovery Handbook. Springer; 2010, http://dx.doi.org/10.1007/978-0-387-09823-4.

[24] Abdullah MR, Maliki ABHM, Musa RM, Kosni NA, Juahir H, Haque M. Multi-hierarchical pattern recognition of athlete's relative performance as a criterion for predicting potential athletes. J Young Pharm 2016;8:463.

[25] Lininger M, Riemann BL. Statistical primer for athletic trainers: using confidence intervals and effect sizes to evaluate clinical meaningfulness. J Athl Train 2016;51:1045—8, http://dx.doi.org/10.4085/1062-6050-51.12.14.

[26] Cohen J. A power primer. Psychol Bull 1992;112:155—9.

[27] Muazu Musa R, Taha Z, Abdul Majeed PP, Abdullah AMR. Bio-physiological indicators in evaluating archery performance; 2019. p. 13—20, http://dx.doi.org/10.1007/978-981-13-2592-2_2.

[28] Muazu Musa R, Taha Z, Abdul Majeed PP, Abdullah AMR. Psychological variables in ascertaining potential archers; 2019. p. 21—7, http://dx.doi.org/10.1007/978-981-13-2592-2_3.

[29] Razali MR, Alias N, Maliki A, Musa RM, Kosni LA, Juahir H. Unsupervised pattern recognition of physical fitness related performance parameters among Terengganu Youth Female Field Hockey Players. Int J Adv Sci Eng Inf Technol 2017;7:100—5.

[30] Tinazci C. Shooting dynamics in archery: a multidimensional analysis from drawing to releasing in male archers. Procedia Eng 2011;13:290—6.

[31] Squadrone R, Rodano R, Gallozzi C. Fatigue effects on shooting archery performance. ISBS - Conf Proc Arch 1994;1:274—7.

[32] Kolayis İE, Çilli M, Ertan H, Knicker JA. Assessment of target performance in archery. Procedia Soc Behav Sci 2014;152:451—6.

[33] Suppiah PK, Muazu Musa R, Wong T, Kiet K, Abdullah MR. Sensitivity prediction analysis of the contribution of physical fitness variables on Terengganu Malaysian Youth Archers' Shooting Scores. Int J Pharm Sci Rev Res 2017;43:133—9, http://dx.doi.org/10.1080/02640414.2015.1035738.

[34] Limonta E, Cè E, Gobbo M, Veicsteinas A, Orizio C, Esposito F. Motor unit activation strategy during a sustained isometric contraction of finger flexor muscles in elite climbers. J Sports Sci 2016;34:133—42, http://dx.doi.org/10.1080/02640414.2015.1035738.

[35] MacKenzie SJ, Boucher DE. The influence of golf shaft stiffness on grip and clubhead kinematics. J Sports Sci 2017;35:105—11, http://dx.doi.org/10.1080/02640414.2016.1157262.

[36] Iermakov SS, Podrigalo LV, Jagiełło W. Hand-grip strength as an indicator for predicting the success in martial arts athletes. Arch Budo 2016;12:179—86.

[37] Painter KB, Haff GG, Ramsey MW, McBride J, Triplett T, Sands WA, et al. Strength gains: block versus daily undulating periodization weight training among track and field athletes. Int J Sports Physiol Perform 2012;7:161—9, http://dx.doi.org/10.1123/ijspp.7.2.161.