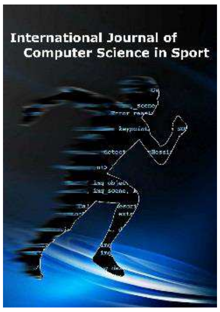
International Journal of Computer Science in Sport

Volume 23, Issue 2, 2024

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

sciendo

DOI: 10.2478/ijcss-2024-0014



Development of Anthro-Fitness Model for Evaluating Firefighter Recruits' Performance Readiness Using Machine Learning

*Borhanudin Mohd Yusof @ Mohamed[1], Rabiu Muazu Musa[*1], Mohamad Nizam Nazarudin[*2], Anwar P. P. Abdul Majeed[3], Naresh Bhaskar Raj[4], Mohd Azraai Mohd Razmaan[5]*

1 Centre for Fundamental and Continuing Education, Entrepreneurship and Sport Science Research Interest Group, Universiti Malaysia Terengganu, 21030 Kuala Nerus, Terengganu Malaysia

2Center for the Study of Education and Community Wellbeing, Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia

3School of Engineering and Technology (SET), Sunway University, No 5, Jalan Universiti, Bandar Sunway 47500 Selangor Darul Ehsan, Malaysia

4 Faculty of Health Science Universiti Sultan Zainal Abidin, 21300 Kuala Nerus, Terengganu, Malaysia.

5Innovative Manufacturing, Mechatronics and Sports Laboratory, Faculty of Manufacturing and Mechatronic Engineering Technology, Universiti Malaysia Pahang, 26600 Pekan, Pahang, Malaysia

**CORRESPONDING AUTHORS*

Abstract

The role of firefighters has evolved from traditional tasks like rescuing cats from trees and extinguishing house fires to more complex land, sea, and air rescues. The increasing demands for public safety necessitate rigorous training and high fitness levels for firefighters to manage their daily tasks effectively. In this study, final assessments of fitness and anthropometric parameters were gathered from 746 Malaysian firefighter recruits. A k-means clustering algorithm was utilized to group the performance levels of the firefighters whilst a quadratic discriminant analysis model was employed to predict the grouping of firefighters based on these parameters. Feature importance analysis was used to identify the most significant parameters contributing to model performance. Concurrently, the Mann-Whitney test was used to determine the essential anthro-fitness parameters differentiating between the groups of firefighters. The k-means clustering identified two performance groups: excellent and average anthro-fitness readiness (EFR and AFR) groups. The model demonstrated a mean performance accuracy of 91% for training and 87% for independent tests. Feature importance analysis revealed that inclined pull-ups, standing broad jump, shuttle run, 2.4 km run, age, and sit-ups were the most significant parameters. The Mann-Whitney test showed that the EFR group outperformed the AFR group in all anthro-fitness parameters except for height, weight, and age, which showed no significant difference. This study highlights the critical role of specific fitness and anthropometric parameters in distinguishing high-performing firefighters. By identifying the most significant contributors to overall fitness, fire departments can better prepare their personnel to meet the increasing public safety demands. The high accuracy of the predictive model also suggests its potential application in ongoing firefighter assessments and training optimization.

KEYWORDS: FIREFIGHTER FITNESS, PERFORMANCE ASSESSMENT, TRAINING OPTIMISATION, PUBLIC SAFETY DEMANDS, QUADRATIC DISCRIMINANT ANALYSIS.

Introduction

The role of firefighters has undergone significant evolution in recent years, expanding beyond traditional tasks to encompass more complex rescue operations across various terrains. This shift reflects the increasing demands placed on emergency services personnel in modern urban and rural environments[1]. As noted earlier by preceding researchers, the diversification of firefighter responsibilities now includes hazardous material management, advanced medical support, and complex rescue operations, necessitating a broader skill set and enhanced physical capabilities [2]. This expansion of duties underscores the critical importance of maintaining high levels of fitness and adaptability among firefighting personnel to effectively respond to a wide range of emergency scenarios.

The assessment of firefighter fitness and anthropometric parameters has become a crucial aspect of ensuring operational readiness and effectiveness. It was emphasised that a comprehensive evaluation of physical fitness and body composition provides valuable insights into a firefighter's capacity to perform under stressful conditions and sustain prolonged physical exertion [3]. Moreover, identifying the most significant parameters contributing to firefighter performance is essential for optimizing training programs and recruitment strategies. For instance, a recent study reported that specific physical attributes, including upper body strength, cardiovascular endurance, and agility, are strongly correlated with superior performance in firefighting tasks [4]. The authors opined that by focusing on these key attributes, fire departments can tailor their training regimens to enhance the overall fitness and operational effectiveness of their personnel.

Research on firefighter physical fitness has been conducted in various countries, each contributing to our understanding of the relationship between fitness parameters and firefighting performance. For instance, Nazari et al. (2018) examined the correlation between body composition and physical fitness among Canadian firefighters, highlighting the importance of maintaining a healthy body mass index for good performance [5]. A different study on Malaysian firefighters reported poor scores in hand grip strength and VO2 max, highlighting the need for further investigation into the factors contributing to this in Malaysian firefighters [6]. Xu (2020) and Michaelides (2008) both identified significant relationships between physical health parameters and performance on simulated firefighting tasks, with upper-body muscular strength and endurance, and low body composition being particularly important for Chinese firefighters [7]. Williford (1999) further emphasized the importance of physical fitness, particularly the 1.5-mile run, fat-free weight, and pull-ups, in relation to fire suppression tasks [8].

In the United States, Dennison et al. (2012) investigated the impact of a comprehensive fitness program on injury rates and work capacity among urban firefighters, demonstrating significant improvements in both areas [9]. Similarly, in a comprehensive review, Orr et al. (2019) investigated the risk of musculoskeletal injuries among firefighters from different countries due to the physical demands of their jobs, including carrying heavy loads and working in extreme conditions. This review identified lower extremities and back as common injury sites, with sprains and strains being prevalent. Key injury mechanisms include slips, trips, falls, and muscle stress from activities like bending and lifting [10].

Although, a range of studies have explored the physical fitness of firefighters, with a focus on specific parameters and their relationship to job performance. However, there is a notable gap in the literature regarding the specific physical fitness needs of Malaysian firefighters, particularly in the context of their evolving role and the unique demands they face. Given Malaysia's unique environmental challenges, including tropical climate conditions and diverse topography, the physical demands placed on its firefighters may differ from those in other countries. For instance, it has been reported that tropical climates and diverse topography require firefighters to have exceptional cardiovascular endurance, agility, and strength[11]. Consequently, training programs should prioritize cardiovascular endurance to cope with prolonged exertion in high heat and humidity, alongside agility for quick manoeuvres in hazardous environments. Strength training, especially in upper body and core muscles, becomes critical for lifting heavy equipment and handling physically taxing rescue operations. This phenomenon warrants further research to address this gap and develop tailored training programs to improve the physical fitness of Malaysian firefighters offering insights that are directly applicable to the local context and potentially informing region-specific training and assessment protocols.

Additionally, the use of advanced statistical techniques, such as k-means clustering and quadratic discriminant analysis, provides a sophisticated approach to categorizing and predicting firefighter performance levels. These methods offer a holistic understanding of the factors that contribute to identifying capable firefighters during the recruitment evaluation process. Moreover, developing accurate predictive models based on these parameters can aid in ongoing assessment and career development for firefighters, ensuring they maintain the high standards required for public safety. This study addresses the gap by exploring the association between anthropometric and fitness parameters among Malaysian firefighter recruits. To guide the conduct of the current study, the following objectives (OBJ) and hypotheses (H) are formulated:

Objectives

OBJ1: To cluster Malaysian firefighter recruits into performance groups using a k-means clustering algorithm based on their fitness and anthropometric parameters.

OBJ2: To predict the performance groups of the firefighters based on fitness and anthropometric parameters using a Quadratic Discriminant Analysis (QDA) model.

OBJ3: To determine the most significant fitness and anthropometric parameters differentiating the firefighters' performance groups using feature importance analysis and the Mann-Whitney U test.

Hypotheses

H1: The k-means clustering algorithm will effectively classify Malaysian firefighter recruits into distinct performance groups based on their fitness and anthropometric parameters.

H2: The QDA model will provide a reasonable prediction of the performance levels of firefighters, demonstrating that fitness and anthropometric parameters can be a basis for detecting an individual's suitability at the entrance level in firefighter units.

H3: Feature importance analysis and the Mann-Whitney U test will identify specific fitness and anthropometric parameters as key attributes that differentiate the performance groups of the firefighters.

Methods

Participants

The study involved 746 recruits (697 males and 49 females) with the following characteristics: Male (BMI 23.09 ± 2.25; height 1.70 ± 0.06 m; weight 64.34 ± 6.86 kg; age 26.51 ± 3.90 yrs); Female (BMI 22.88 ± height 1.67 ± 0.08 m; weight 63.30 ± 6.59 kg; age 24.55 ± 3.52 yrs) mean and standard deviation respectively. The participants of the current study were enrolled from one of the firefighter academies in Terengganu Malaysia. It is worth highlighting that the participants were selected from all the states in Malaysia. Before the commencement of the study, the firefighters and the trainers were informed about the purpose of the study and verbal informed consent was obtained. Approval to conduct the study was obtained from the departmental research ethics committee (MC/UMT/ PSH/2023/01).

Anthro-fitness assessments

The anthro-fitness assessments in the current study involve a variety of measurements to evaluate the body attributes and physical fitness of the participants, including height, weight, body mass index (BMI), 2.4 km run, shuttle run, inclined pull-up, standing broad jump, and sit-ups. The height of the participants was measured while they were standing barefoot with their backs against a stadiometer, from the floor to the top of the head, in meters. The participants were weighed in kilograms using a calibrated scale, without shoes and in minimal clothing. BMI was calculated using the formula: weight (kg) divided by height (m) squared (weight/height^2). For the 2.4 km run, the participants ran a distance of 2.4 kilometres (1.5 miles) on a flat surface (track), with time recorded in minutes and seconds. This test assesses cardiovascular endurance and aerobic capacity. The shuttle run involved participants running back and forth between two lines set 20 meters apart, following the pace set by audio signals. The test continued until the participants could no longer keep up with the increasing pace, with the time and number of laps completed recorded. This measures aerobic fitness and agility. For the inclined pull-up test, participants performed as many pull-ups as possible on an inclined bar, with the number of correctly executed repetitions counted. This test evaluates upper body and back strength. The standing broad jump was assessed by having participants jump horizontally as far as possible from a standing position. The distance from the starting line to the back of the heel closest to the starting line was measured in centimetres, assessing lower body explosive power. For the sit-up test, participants performed as many sit-ups as possible within one minute, with the number of correctly executed repetitions counted. This measures abdominal muscular endurance. Finally, the participants' age was recorded in years based on their date of birth.

Data Treatment and Analysis

The data on the anthro-fitness parameters were gathered and pre-processed. The data were also normalized using min-max scaling to avoid bias from different units of measurement [12]. This normalization technique ensures that all variables contribute equally to the analysis, preventing any single parameter with larger numerical values from dominating the model. Moreover, The Shapiro-Wilks test was used to analyse the performance distribution of firefighters, including both males and females[13]. No violation of normality distribution between genders was observed, as both are expected to meet a certain fitness standard upon recruitment. Therefore, gender was not isolated for analysis in this study.

Fitness Score Grading

The firefighter department set performance threshold scores for each fitness parameter, with scores ranging from 0 (lowest) to 5 (highest). For example, completing 37 or more sit-ups and chin-ups earns 5 points, while 5 fewer reps result in 4 points. A standing broad jump of 240 cm or more also scores 5 points, with each 10 cm less yielding 4 points. Similarly, a shuttle run completed in 10 seconds or a 2.4 km run in 11 minutes both score 5 points, with a 0.5-second difference reducing the score to 4 points. The total points accumulated across the five fitness parameters were used for further analysis, i.e., clustering.

Clustering Analysis

K-means clustering is a form of cluster analysis that divides a set of data into k-predefined, non-overlapping sub-groups called clusters, with each data point assigned to only one cluster [14–16]. The method aims to maximize the similarity of data points within each cluster (intra-cluster) while minimizing the similarity between data points in different clusters (inter-cluster). In this study, k-means cluster analysis was used to group recruits based on their scores in the performance of physical fitness and anthropometric attributes. Euclidean distance was employed as the distance metric for the formation of optimum clusters.

Development of the Quadratic Discriminant Model

In this study, the Quadratic Discriminant Analysis (QDA) model was employed to predict the performance levels of firefighters based on their fitness and anthropometric parameters due to its ability to handle non-linear relationships and complex decision boundaries which allows for different covariance structures within each performance group, offering greater flexibility in modelling the unique variance and covariance within the dataset[17]. Firefighters' performance, particularly when classified into "excellent" and "average" groups, is likely influenced by numerous interacting factors, such as strength, endurance, body composition, and cardiovascular capacity. These variables may not exhibit linear separability across performance levels due to the intrinsic complexity of how physical and anthropometric traits combine to influence outcomes. The QDA model, by accounting for these non-linear relationships, is perceived to be well-suited to capture the variations within these multidimensional fitness and body metrics.

Given the complexity of interactions among these parameters, the quadratic decision boundaries that QDA introduces allow for more precise classification compared to models with linear assumptions. This makes QDA an ideal choice for ensuring higher accuracy in distinguishing performance levels in firefighters, thus enhancing the robustness of the predictive model in the study.

The QDA model was developed using a dataset comprising 746 observations per parameter of the participants. A stratified 5-fold cross-validation technique was employed, wherein the data was divided using a holdout approach. Specifically, 70% of the data (522 observations) was allocated for training and validation i.e., 365 for training and 157 observations were held out for testing. The remaining 30% (224 observations) were not exposed to the model during training and testing and hence, reserved for independent testing as a fresh dataset. The stratified 5-fold cross-validation technique was chosen over the typical standard k-fold cross-validation to ensure that the training and validation sets were representative of the entire dataset, particularly with respect to class distribution [18]. Moreover, by reserving a separate independent test set, we aimed to provide a more robust evaluation of the model's generalization capability on completely unseen data. The PyCaret libraries were utilized for the development of the QDA model via the Spyder Integrated Development Environment (IDE). Other statistical analyses were conducted via Jamovi version 2.4.14 for Windows.

Model Evaluation Metrics

To evaluate the efficacy of the developed model, various performance metrics were utilized, including classification accuracy (ACC), area under the curve (AUC), recall, precision (PREC), F1 score, Kappa, and Matthew's correlation coefficient (MCC). These metrics provided a comprehensive assessment of the model's performance across different dimensions: The ACC represents the proportion of correct predictions (both true positives and true negatives) among the total number of cases, offering a general measure of the model's overall effectiveness. The AUC measures the area under the Receiver Operating Characteristic (ROC) curve, which plots the true positive rate against the false positive rate at various threshold settings, indicating the model's ability to discriminate between positive and negative classes. Recall refers to the proportion of true positive predictions among all actual positive cases (sensitivity), reflecting the model's ability to identify all relevant positive cases while serving as the proportion of true positive predictions among all positive predictions made by the model, indicating the accuracy of the positive predictions. The F1 score reflects the harmonic mean of precision and recall, providing a single metric that balances both, especially useful for imbalanced class distributions. Kappa is the statistical measure of inter-rater agreement for categorical items, adjusted for agreement occurring by chance, evaluating the degree of agreement between predicted and actual classifications while the MCC measures the correlation coefficient between observed and predicted binary classifications, ranging from -1 to +1, offering a balanced measure that considers true and false positives and negatives, even in the case of class imbalance. These metrics collectively offer a comprehensive evaluation of the model's performance, highlighting its strengths and weaknesses.

Feature Importance Investigation

Symmetrical uncertainty attribute evaluation was also used to investigate the anthro-fitness variables that have a higher influence on model performance. This technique was chosen because it effectively measures the relevance of each attribute by accounting for the mutual information between attributes and the target variable, providing a balanced assessment of both linear and non-linear relationships within the dataset [19,20]. By identifying the most influential variables, this method helps highlight the critical parameters that significantly impact performance, thereby enhancing model interpretability and predictive power.

Results

Figure 1 displays the differences between the two groups with respect to the fitness scores. Two clusters are identified based on the k-means analysis namely, excellent anthro-fitness readiness (EFR) group and the average anthro-fitness readiness (AFR) group. It could be observed the EFR is associated with a higher fitness score in comparison with the AFR. The EFR has an average score of 21.15 while the AFR has 13.89 with a mean difference of 7.26. This reflects that the cluster analysis is effective in separating the groups of firefighters based on their fitness scores.

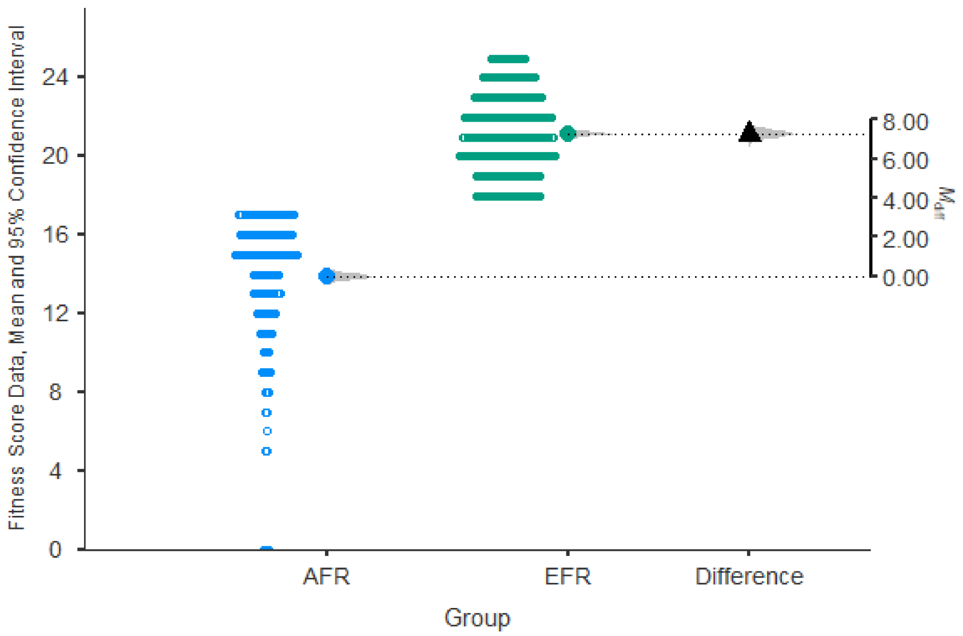


Figure 1: Grouping of the firefighters based on their fitness score. (AFR = Average anthro-fitness readiness; EFR = Excellent anthro-fitness readiness; Mdiff = Mean difference)

Table 1 summarizes the performance of the QDA model in predicting the performance level of the firefighters. The model achieved a mean accuracy score of 91%, indicating a strong ability to predict excellent or average performance groups. The Area Under the Curve (AUC) was 0.94, which signifies excellent modelling performance in predicting the groups. The F1 score, which is a weighted average of precision and recall, further supports this conclusion.

The precision and recall scores were both 0.91. These scores demonstrate that the model correctly predicted over 91% of positive cases and accurately identified 91% of the actual positive classes. The Kappa scores and Matthew's Correlation Coefficient (MCC) were 0.79, indicating good reliability, reflecting strong predictive power. Overall, these findings suggest that the QDA model performed well in predicting EAF and the AAF group.

Table 1: Performance evaluation of the DA model for predicting fitness levels of the firefighters

	Accuracy	AUC	Recall	Prec.	F1	Kappa	MCC
Mean	0.913	0.942	0.912	0.913	0.911	0.788	0.792
Std	0.032	0.024	0.032	0.033	0.034	0.079	0.076

Note: AUC=Area under the curve; Prec.=Precession; MCC=Mathieu Correlation Coefficient; Std = Standard deviation.

Figure 2 illustrates the training and cross-validation scores of the model on the holdout dataset. The training score was 0.93, whereas the cross-validation score was 0.91. The slightest drop observed in scores after cross-validation can be attributed to the data-splitting process for training and testing. It is plausible that the training score before cross-validation might have included noise and errors, leading to an inflated accuracy rate [2,21]. However, during the cross-validation stage, the model was adjusted and refined, resulting in a prediction accuracy score of 91%. This outcome accentuates the efficiency of the formulated model, demonstrating its robust predictive capability while mitigating any overfitting that may arise in the initial training phase.

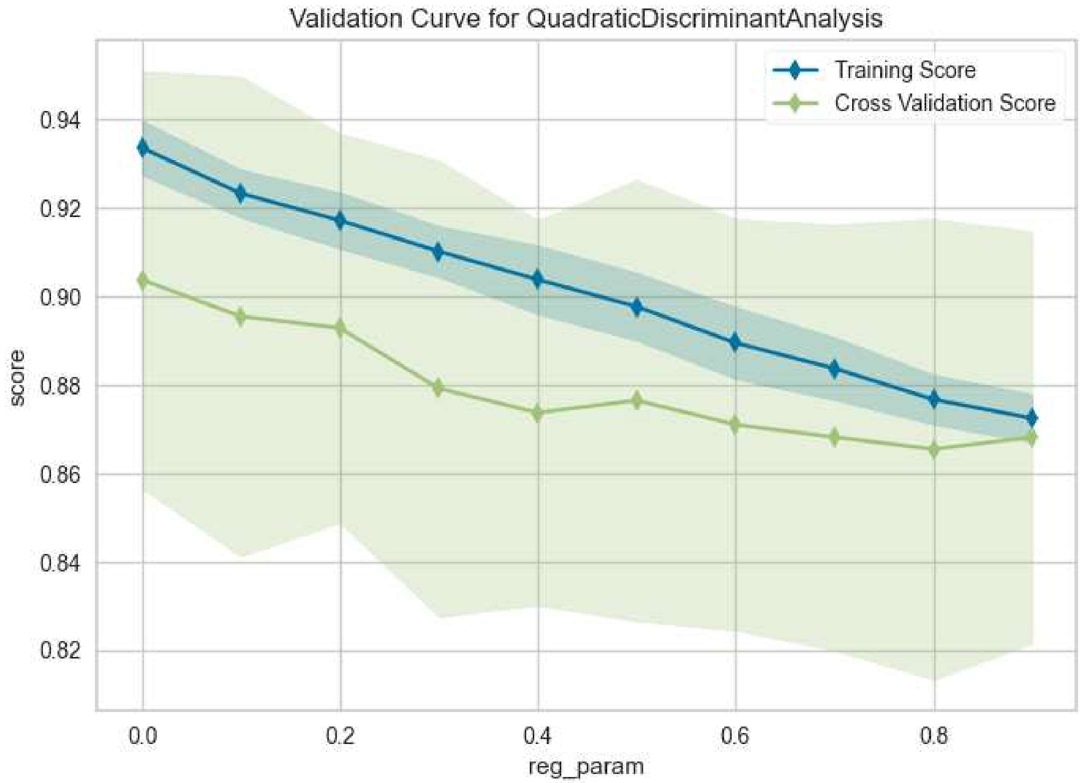


Figure 2: Performance of the Quadratic Discriminant Analysis model on validation data.

Figure 3 highlights the results of the QDA analysis for the classification of excellent (1) and average fitness (0) instances during cross-validation. Cross-validation is a critical process for validating the effectiveness of the formulated model on hold-out test data. In this stage, the data held out (157 observations) were used to test the previously formulated QDA model to examine the accuracy of the developed model. It is worth noting that the F1, precision, and recall scores remain robust during cross-validation for predicting correctly all 157 instances of the excellent fitness group (109 instances) and average fitness (48 instances). These scores indicate that the model maintained high performance and reliability even when tested on reserve data, reinforcing its predictive accuracy and generalizability.

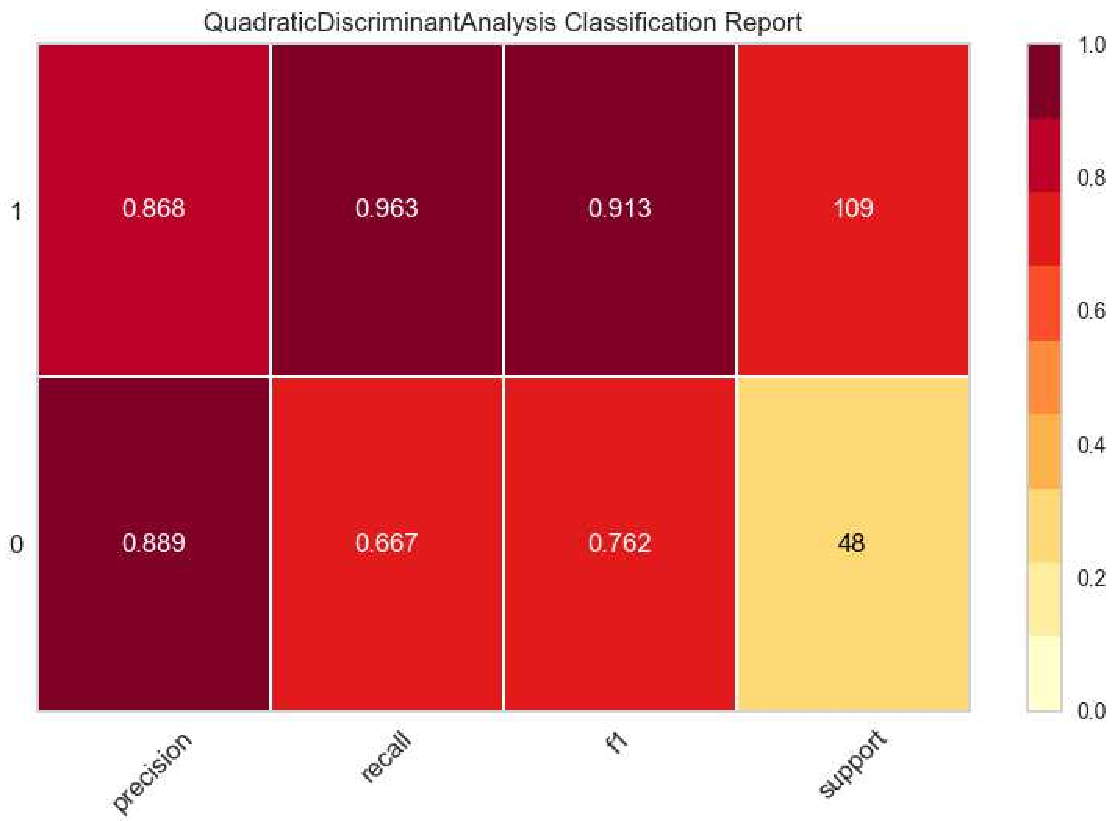


Figure 3: Classification report of the validation data in the developed model (1 represents EFR; 0 represents AFR)

Table 2: Independent testing for the Performance evaluation of the model against fresh/unseen data.

Observations	Actual	Predicted	Prediction score
1	Average Anthro Fitness	Average Anthro Fitness	0.52
2	Average Anthro Fitness	Average Anthro Fitness	0.54
3	Excellent Anthro Fitness	Excellent Anthro Fitness	0.95
4	Excellent Anthro Fitness	Excellent Anthro Fitness	0.97
5	Excellent Anthro Fitness	Excellent Anthro Fitness	0.98
6	Excellent Anthro Fitness	Excellent Anthro Fitness	1.00
7	Excellent Anthro Fitness	Excellent Anthro Fitness	1.00
8	Excellent Anthro Fitness	Excellent Anthro Fitness	0.85
9	Excellent Anthro Fitness	Excellent Anthro Fitness	0.95
10	Excellent Anthro Fitness	Excellent Anthro Fitness	0.96
11	Excellent Anthro Fitness	Excellent Anthro Fitness	0.91
12	Excellent Anthro Fitness	Excellent Anthro Fitness	1.00
13	Excellent Anthro Fitness	Excellent Anthro Fitness	1.00
14	Excellent Anthro Fitness	Excellent Anthro Fitness	0.98
15	Excellent Anthro Fitness	Excellent Anthro Fitness	0.99

Table 2 presents the performance of the QDA model on unseen data. In this stage of model integrity analysis, 30% (224 observations) of the original data, which were not exposed to the model during training and validation processes, were used. The model demonstrated 87% accuracy, correctly classifying all 195 observations out of 224. This excellent performance and consistency highlight the model's strong ability to accurately classify the performance levels of the firefighters, highlighting its effectiveness in practical applications. It is worth noting that Table 2 displays the first 15 observations. The results confirm the robustness of the QDA model and its reliability in predicting firefighter recruit performances based on anthro-fitness performance variables.

Figure 4 depicts the Symmetrical Uncertainty attribute evaluation applied to investigate the anthro-fitness variables that have a higher influence on the model performance. It is observed that shuttle run, inclined pull-ups, standing broad jump, 2.4 km run, sit-ups and age significantly contributed to the model's accuracy due to their scores greater than 0.01 threshold. These parameters are therefore considered vital for determining the fitness status of the firefighters. Hence, it can be inferred that trainers should prioritize these parameters for the identification and evaluation of firefighters to increase the likelihood of competency and fitness ability.

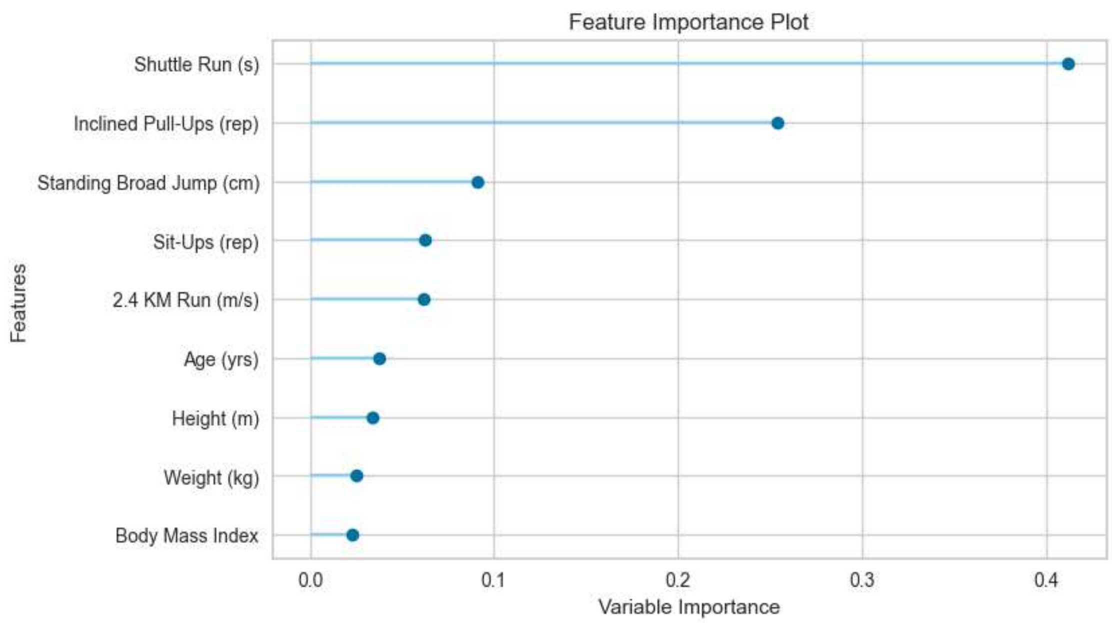


Figure 4: Variables contribution towards the Quadratic Discriminant Analysis model performance.

Table 3 tabulates the performance differences in the measured anthro-fitness variables between the two groups of firefighters. The grouping, number of participants in each, median, standard deviation, U statistics of the Mann-Whitney test as well as its associated p-values are shown. It could be observed from the table that the EFR group outperformed the AFR group in the 2.4km run, shuttle run, inclined pull-up, standing broad jump, sit-up and lower BMI status $p < 0.05$. Whereas height, weight, and age, showed no significant difference.

Table 3: Performance differences between the two groups of firefighters with respect to the measured anthro-fitness variables

Anthro-Fitness Variables	Group	N	Median	SD	U Statistics	p-value
2.4 KM Run (m/s)	AFR	226	10.54	1.23	36373	0.001*
	EFR	520	10.08	1.08		
Shuttle Run (s)	AFR	226	10.69	0.74	22728	0.001*
	EFR	520	10	0.59		
Chin-Up / Inclined Pull-Up (reps)	AFR	226	6	3.59	22344	0.001*
	EFR	520	10	6.20		
Standing Broad Jump (cm)	AFR	226	200	20.08	22339	0.001*
	EFR	520	220	21.70		
Sit-Up (reps)	AFR	226	50	5.86	41979	0.001*
	EFR	520	53	6.73		
Body Mass Index	AFR	226	23.62	2.41	49585	0.001*
	EFR	520	22.7	2.16		
Height (m)	AFR	226	1.7	0.06	55625	0.246
	EFR	520	1.7	0.06		
Weight (kg)	AFR	226	63.1	7.63	56720	0.451
	EFR	520	64	6.48		
Age (yrs.)	AFR	226	26	3.64	55542	0.233
	EFR	520	26	4.01		

Note: Mann-Whitney U test; $p < .01$; AFR = Average anthro-fitness readiness; EFR = Excellent anthro-fitness readiness; N = Number of observations; SD = Standard deviation

Discussion

The purpose of the current study was to examine the anthro-fitness variables that contribute to the selection and performance of firefighters and to develop a machine learning-based classification model for the identification of excellent and average-performance firefighters. The model demonstrated a mean performance accuracy of 91% for training and 87% for independent tests against fresh or unseen data. Feature analysis revealed that inclined pull-ups, standing broad jumps, shuttle runs, 2.4km runs, age, and sit-ups were the most significant parameters in the evaluation of firefighters' performance. The Mann-Whitney test showed that the EFR group outperformed the AFR group in all anthro-fitness parameters except for height, weight, and age, which showed no significant difference.

The effective grouping of the firefighters based on the anthro-fitness parameters demonstrated the capability of k-means clustering in segregating the performance levels of the firefighters as previously hypothesised (H1). The high accuracy achieved corroborates our second hypothesis (H2) and highlights QDA's effectiveness in handling multivariate data, making it suitable for fitness and performance assessments. Accurate classification and prediction could enable personalized training programs[22,23]. The identification of certain key fitness attributes, such as lower body strength and cardiovascular endurance, supports our third hypothesis (H3) emphasising the need for training that can focus on exercises like squats, deadlifts, and aerobic conditioning. This targeted approach ensures firefighters train the most relevant fitness components, enhancing overall performance and readiness.

The feature importance analysis identifies inclined pull-ups, standing broad jumps, shuttle runs, 2.4 km runs, age, and sit-ups as key physical attributes for firefighters. Each parameter highlights different aspects of fitness, strength, power, endurance, and agility that are non-trivial for firefighting tasks. This accentuates the need for a well-rounded fitness regimen. Inclined pull-ups indicate upper body strength, essential for lifting equipment, climbing, and rescues, with studies showing a strong correlation to job performance[24]. The standing broad jump measures lower body explosive power, vital for tasks like breaking down doors and quick evacuations, as documented in the previous investigation[25].

The shuttle run measures agility and aerobic fitness, crucial for quick direction changes and speed maintenance in firefighting. Agility and aerobic capacity are vital for navigating hazardous environments, as noted in the previous researchers[9]. The 2.4 km run assesses cardiovascular endurance, essential for prolonged physical exertion under stress. High aerobic fitness helps firefighters perform strenuous tasks without excessive fatigue, with research by Poplin et al. (2016) linking it to job performance and injury prevention in firefighters [26].

Sit-ups measure core strength and endurance, crucial for stability and support in firefighting tasks. A strong core prevents injuries and enhances performance, as supported by the preceding investigators[27]. Age influences physical performance, with younger firefighters excelling in strength and agility, while older ones bring experience and resilience. Although fitness declines with age, regular training can mitigate this effect. The inclusion of age highlights the balance between physical capabilities and the experience of older firefighters[28].

The Mann-Whitney U test showed significant differences between the excellent anthro-fitness readiness (EFR) and average anthro-fitness (AFR) groups in key parameters. The EFR group's superior performance in the 2.4 km and shuttle runs aligns with research emphasizing cardiovascular fitness in firefighting. Nazari et al. (2020) found higher aerobic capacity strongly linked to better performance in simulated tasks[5]. The 2.4 km run measures aerobic capacity, essential for prolonged exertion, with studies showing higher fitness correlates with better job performance and lower injury rates[26]. The shuttle run assesses aerobic capacity and agility, which are crucial for navigating hazardous environments. These differences highlight the importance of cardiovascular endurance in firefighter effectiveness and safety.

Inclined pull-ups and sit-ups measure muscular strength and endurance. The EFR group's superior performance in these tests aligns with the findings of the previous researchers, who found the upper body and core strength crucial for firefighting tasks[29]. These results underline the need for targeted strength training in firefighter fitness programs. Attributes like lifting heavy equipment, breaking barriers, and rapid movement through obstacles are vital. Gledhill and Jamnik (1992) also emphasize the necessity of significant muscle strength and power for safe and effective task performance[25].

The lower BMI of the EFR group aligns with Dawes et al. (2021), who found that firefighters with lower BMI perform better in job-specific tasks and have a lower injury risk[30]. This indicates better overall fitness and lower body fat, linked to improved performance and reduced health risks. Similarly, it was demonstrated that firefighters with lower BMI have better cardiovascular health and are less prone to injuries[31].

The lack of significant differences in height, weight, and age between the EFR and AFR groups suggests these measures do not decisively impact fitness performance. While height and weight are important for BMI, their direct effect on fitness attributes appears to be minimal. Age-related declines in performance can be mitigated by fitness and training, as documented in the previous research[28]. This indicates that excellent performance is not necessarily tied to these measures or age. Stevenson et al. (2022) also found that targeted fitness training improves firefighters' job performance across various ages and body types[32].

The results of this study are essentially relevant for Malaysia, given its unique environmental conditions, including the tropical climate and diverse topography, which pose specific physical demands on firefighters. Compared to other countries, Malaysian firefighters face extreme heat, high humidity, and variable terrain, requiring high levels of cardiovascular endurance, agility, and strength. Key fitness parameters such as the shuttle run (for agility and aerobic capacity), 2.4 km run (for cardiovascular endurance), and standing broad jump (for lower body strength) are crucial for operating in such environments[33]. These metrics emphasize the need for a well-rounded fitness regimen tailored to the challenging conditions Malaysian firefighters encounter, compared to other nations where climate and operational demands may differ. Likewise, the findings could also be applicable to regions with similar environmental and operational challenges, such as tropical climates, diverse topographies, and high urbanization rates. For instance, fire departments in Southeast Asia, parts of South America, or Central Africa, which share Malaysia's climatic and geographical constraints, could find the results directly relevant to tailoring their recruitment and training protocol.

Moreover, the lower BMI of the excellent anthro-fitness group highlights the importance of maintaining a lean physique, essential in a hot and humid climate to reduce fatigue and prevent injuries. Unlike colder climates where different physical stressors may dominate, Malaysia's tropical environment requires firefighters to sustain prolonged exertion under intense heat, reinforcing the significance of aerobic fitness and efficient body composition for ideal performance[6]. Thus, the model and results align closely with the operational and environmental demands specific to Malaysia, while offering insights that could inform region-specific firefighter training and assessment protocols.

These findings pave the way for future research to explore the use of QDA and other advanced statistical models in different contexts. Further studies could investigate the integration of additional variables, such as psychological factors or environmental conditions, to enhance model accuracy and applicability. Moreover, longitudinal studies could examine how changes in fitness parameters over time impact performance, providing deeper insights into the long-term effectiveness of tailored training programs. It is also important to highlight that the issue of collinearity was not explicitly assessed in this study, as the primary multivariate approaches used (k-means clustering and quadratic discriminant analysis) do not inherently require collinearity diagnostics. However, addressing collinearity could enhance the interpretability and robustness of the feature selection process. This issue could be addressed in the future study.

Practical Implication and Future Direction

Given the high accuracy of the QDA model, we developed a web-based prototype for automatically evaluating and predicting firefighters' anthro-fitness performance, as shown in Figure 5 where an individual recruit was predicted to be in the excellent category with a 99% prediction score based on the inputs in his anthro-fitness performance. We believe this predictive model can enable a more personalized approach to training and enhancing firefighters' effectiveness thereby serving as a basis for detecting individual readiness and suitability at the entrance level in firefighter units. Additionally, the anthro-fitness predictive application has broader implications for fitness assessments beyond firefighting. The methodology can be applied to other professions and sports where physical fitness is critical. Drawing parallels between professional athletes and firefighters could enrich the discussion, as both require exceptional physical fitness for demanding tasks. Just as sports talent identification protocols assess strength, agility, endurance, and explosive power, firefighter fitness assessments ensure readiness for peak performance. For instance, in East Germany's former DDR system, basic fitness tests like sprinting, jumping, and strength evaluations were used to screen children for athletic potential. Similarly, this study's fitness parameters i.e., inclined pull-ups, shuttle runs, standing broad jumps, and 2.4 km runs evaluate essential traits for firefighting, such as upper body strength, cardiovascular endurance, agility, and explosive power. Both protocols aim to optimize physical capability for specific performance demands. Highlighting these parallels underscores how systematic fitness assessments can enhance performance outcomes and tailor training programs, akin to athlete development pipelines. This approach strengthens the study's broader applicability and theoretical foundation.

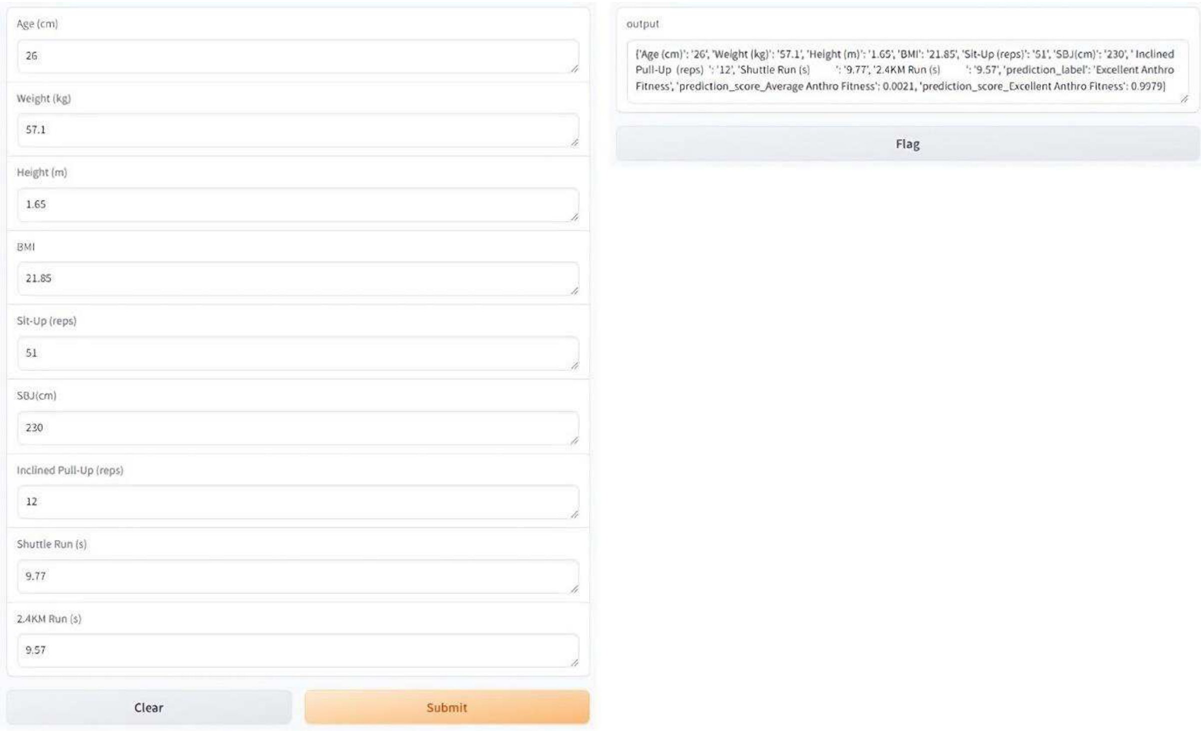


Figure 5: Prototype of a Web-Based Anthro-Fitness Predictive Application.

Conclusion

This study comprehensively evaluated firefighters' physical fitness and anthropometric parameters using advanced statistical methods consisting of k-means clustering and quadratic discriminant analysis. The findings revealed that specific fitness attributes, such as inclined pull-ups, standing broad jumps, shuttle runs, 2.4 km runs, age, and sit-ups, are essential determinants of distinguishing firefighters' groups of performance. The Mann-Whitney U test further demonstrated that the excellent anthro-fitness readiness group (EFR) significantly outperformed the average anthro-fitness readiness group (AFR) in these key fitness parameters, highlighting the importance of targeted training programs. Importantly, the study also found no significant differences in height, weight, and age between the groups, suggesting that while these anthropometric measures are relevant for general health assessments, they do not directly impact the specific physical capabilities required for firefighting tasks. These insights accentuate the necessity for a well-rounded fitness regimen focusing on cardiovascular endurance, agility, muscular strength, power, and core endurance to enhance overall firefighting effectiveness and safety. Hitherto, it is important to note that the basic fitness characteristics assessed in this study can help determine individual suitability for entry into firefighter units. However, these parameters alone cannot predict which recruits will excel professionally, given the diverse and demanding nature of firefighting tasks.

References

Bok Bok-choi, Ha Seong-bak, Ha Seong-gong. Field-Driven Physical Fitness Assessment for Firefighters: Overcoming Challenges and Improving Standards. J Korean Soc Hazard Mitig 2023; 23: 111–122.

Carrick RT, Park JG, McGinnes HL, et al. Clinical predictive models of sudden cardiac arrest: a survey of the current science and analysis of model performances. J Am Heart Assoc 2020; 9: e017625.

Johnson BVB. An Exploratory Analysis of Firefighters' Nutrient Intake Related to Obesity, Musculoskeletal Injury, Sleep, and Physical Fitness.

Gonzalez DE, Lanham SN, Martin SE, et al. Firefighter Health: A Narrative Review of Occupational Threats and Countermeasures. In: Healthcare. MDPI, 2024, p. 440.

Nazari G, MacDermid JC, Sinden KE, et al. The relationship between physical fitness and simulated firefighting task performance. Rehabil Res Pract 2018; 2018: 3234176.

Atikah CW, Nihayah M, Leonard JH, et al. A cross-sectional evaluation on physical fitness of Malaysian firefighters. Sains Malaysiana 2015; 44: 1461–1466.

Xu D, Song Y, Meng Y, et al. Relationship between firefighter physical fitness and special ability performance: predictive research based on machine learning algorithms. Int J Environ Res Public Health 2020; 17: 7689.

Williford HN, Duey WJ, Olson MS, et al. Relationship between fire fighting suppression tasks and physical fitness. Ergonomics 1999; 42: 1179–1186.

Dennison KJ, Mullineaux DR, Yates JW, et al. The effect of fatigue and training status on firefighter performance. J Strength Cond Res 2012; 26: 1101–1109.

Orr R, Simas V, Canetti E, et al. A profile of injuries sustained by firefighters: A critical review. Int J Environ Res Public Health 2019; 16: 3931.

Chizewski A, Box A, Kesler R, et al. Fitness fights fires: exploring the relationship between physical fitness and firefighter ability. Int J Environ Res Public Health 2021; 18: 11733.

Ab Rasid AM, Muazu Musa R, Abdul Majeed APP, et al. Physical fitness and motor ability parameters as predictors for skateboarding performance: A logistic regression modelling analysis. PLoS One 2024; 19: e0296467.

Abdullah MR, Musa RM, Maliki ABHM, et al. Development of tablet application based notational analysis system and the establishment of its reliability in soccer. J Phys Educ Sport 2016; 16: 951.

Azahari H, Juahir H, Abdullah MR, et al. A multivariate analysis of cardiopulmonary parameters in archery performance. Hum Mov 2019; 19: 35–41.

Razali MR, Alias N, Maliki A, et al. 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: 100–105.

Taha Z, Haque M, Musa RM, et al. Intelligent prediction of suitable physical characteristics toward archery performance using multivariate techniques. J Glob Pharma Technol 2009; 9: 44–52.

Hastie T, Tibshirani R, Friedman JH, et al. The elements of statistical learning: data mining, inference, and prediction. Springer, 2009.

Muazu Musa R, PP Abdul Majeed A, Taha Z, et al. A machine learning approach of predicting high potential archers by means of physical fitness indicators. PLoS One 2019; 14: e0209638.

Witten IH, Frank E, Hall MA, et al. Practical machine learning tools and techniques. In: Data mining. Elsevier Amsterdam, The Netherlands, 2005, pp. 403–413.

Hall MA. Correlation-based feature selection for machine learning.

Tabe-Bordbar S, Emad A, Zhao SD, et al. A closer look at cross-validation for assessing the accuracy of gene regulatory networks and models. Sci Rep 2018; 8: 1–11.
McLachlan GJ. Discriminant analysis and statistical pattern recognition. John Wiley & Sons, 2005.
Friedman JH. Regularized discriminant analysis. J Am Stat Assoc 1989; 84: 165–175.
Williams-Bell FM, Villar R, Sharratt MT, et al. Physiological demands of the firefighter Candidate Physical Ability Test. Med Sci Sports Exerc 2009; 41: 653–662.
Gledhill N, Jamnik VK. Characterization of the physical demands of firefighting. Can J Sport Sci J Can des Sci du Sport 1992; 17: 207–213.
Poplin GS, Roe DJ, Peate W, et al. The association of aerobic fitness with injuries in the fire service. Am J Epidemiol 2014; 179: 149–155.
Rhea MR, Alvar BA, Gray R. Physical fitness and job performance of firefighters. J Strength Cond Res 2004; 18: 348–352.
Henderson ND, Berry MW, Matic T. Field measures of strength and fitness predict firefighter performance on physically demanding tasks. Pers Psychol 2007; 60: 431–473.
Siddall AG, Stevenson RDM, Turner PFJ, et al. Development of role-related minimum cardiorespiratory fitness standards for firefighters and commanders. Ergonomics 2016; 59: 1335–1343.
Dawes JJ, Lindsay K, Bero J, et al. Physical fitness characteristics of high vs. low performers on an occupationally specific physical agility test for patrol officers. J Strength Cond Res 2017; 31: 2808–2815.
Soteriades ES, Hauser R, Kawachi I, et al. Obesity and cardiovascular disease risk factors in firefighters: a prospective cohort study. Obes Res 2005; 13: 1756–1763.
Stevenson RDM, Siddall AG, Turner PFJ, et al. Physical employment standards for UK firefighters: Minimum muscular strength and endurance requirements. J Occup Environ Med 2017; 59: 74–79.
Ras J, Smith DL, Soteriades ES, et al. Association between physical fitness and cardiovascular health in firefighters. Int J Environ Res Public Health 2023; 20: 5930.