Author(s): Ashfaq, Atiqa; Cronin, Neil; Müller, Philipp

Title: Recent advances in machine learning for maximal oxygen uptake (VO2 max) prediction : A review

Year: 2022

Version: Published version

Please cite the original version:

Ashfaq, A., Cronin, N., & Müller, P. (2022). Recent advances in machine learning for maximal oxygen uptake (VO2 max) prediction : A review. Informatics in Medicine Unlocked, 28, Article 100863. https://doi.org/10.1016/j.imu.2022.100863

Informatics in Medicine Unlocked 28 (2022) 100863

Contents lists available at ScienceDirect

Informatics in Medicine Unlocked

journal homepage: www.elsevier.com/locate/imu

Recent advances in machine learning for maximal oxygen uptake (VO_2 max) prediction: A review

Atiqa Ashfaq [a,*], Neil Cronin [b], Philipp Müller [c]

[a] Computing Sciences Unit, Faculty of Information Technology and Communication Sciences, Tampere University, 33720, Tampere, Finland
[b] Neuromuscular Research Centre, Faculty of Sport and Health Sciences, University of Jyväskylä, 40014, Jyväskylä, Finland
[c] Computing Sciences Unit, Faculty of Information Technology and Communication Sciences, Tampere University, 33720, Tampere, Finland

ARTICLE INFO

Keywords:
Maximal oxygen uptake (VO_2 max)
Machine learning
Graded exercise tests
Artificial neural network
Prediction models
Error metrics

ABSTRACT

Maximal oxygen uptake (VO_2 max) is the maximum amount of oxygen attainable by a person during exercise. VO_2 max is used in different domains including sports and medical sciences and is usually measured during an incremental treadmill or cycle ergometer test. The drawback of directly measuring VO_2 max using the maximal test is that it is expensive and requires a fixed and controlled protocol. During the last decade, various machine learning models have been developed for VO_2 max prediction and numerous studies have attempted to predict VO_2 max using data from submaximal and non-exercise tests. This article gives an overview of the machine learning models developed over the past five years (2016–2021) for the prediction of VO_2 max. Multiple linear regression, support vector machine, artificial neural network and multilayer perceptron are some of the techniques that have been used to build predictive models using different combinations of predictor variables. Model performance is generally assessed using correlation coefficient (R-value), standard error of estimate (SEE) and root mean squared error (RMSE), computed between ground truth and predicted values. The findings of this review indicate that models using ANN typically outperform other machine learning techniques. Moreover, the predictor variables used to build the model have a large influence on the model's predictive performance.

1. Introduction

Measurement of maximal oxygen uptake (VO_2 max) using oxygen uptake equipment is the gold standard for assessing a person's aerobic fitness level. VO_2 max is defined as the highest amount of oxygen attained during exhaustive exercise and is measured in millilitres of oxygen used per kilogram of body mass per minute ($ml\,kg^{-1}min^{-1}$) [1]. By measuring VO_2 max, an indication of the upper bound of an athlete's performance capacity can be obtained. VO_2 max testing is the most precise way to assess cardiovascular health and aerobic fitness since it measures both muscular and aerobic endurance [2]. Generally, oxygen uptake measurements are collected during maximum effort activity on a treadmill or a cycle ergometer [3]. These maximal tests directly measure VO_2 max with high accuracy but are expensive, requiring fixed, controlled protocol and trained personnel. These tests also pose several health risks, especially for people with existing conditions such as heart disease. Therefore, alternative approaches for predicting a person's VO_2 max have been proposed [3–7].

One approach is a submaximal test, which is an indirect way of estimating VO_2 max that is less expensive, more convenient, and faster, whilst also minimizing severe health risks. The subjects are allowed to do physical activity at a self-chosen pace. However, the predictive accuracy of submaximal tests is lower than that of the maximal test. Another approach is non-exercise tests, which use self-reported data from questionnaires to estimate VO_2 max. These tests are feasible, can be used to study large populations, and are independent of any laboratory equipment but require honest self-reported data from the subjects [3]. Hybrid models can also be created by using a combination of maximal, submaximal and non-exercise tests to achieve higher accuracy in terms of R and SEE values.

Predictive approaches usually employ machine learning algorithms and use predictor variables such as speed and heart rate to predict VO_2 max. For example, an algorithm can be used to return an estimate of the maximal oxygen uptake from long-term data collected with wearables and previously defined rules [8]. By combining machine learning models with relevant predictive data, it is possible to estimate maximum oxygen uptake indirectly, requiring less time and effort, and reducing health risks associated with exercise testing [9]. Algorithms that have

* Corresponding author.
E-mail addresses: atiqa_ashfaq3@yahoo.com, atiqa.ashfaq@tuni.fi (A. Ashfaq).

https://doi.org/10.1016/j.imu.2022.100863
Received 22 October 2021; Received in revised form 12 January 2022; Accepted 19 January 2022
Available online 21 January 2022

been used to predict VO_2 max based on different predictor variables include but are not limited to multiple linear regression (MLR), support vector machine (SVM), multilayer perceptron (MLP) and artificial neural networks (ANN) [6].

This article aims to provide the reader with a detailed review of recent advances in machine learning for VO_2 max prediction. Therefore, studies published between 2016 and 2021 are discussed. Comprehensive overviews of studies published before 2016 can be found in Abut et al., 2016 [3] and Haneen Alzamer et al., 2021 [9]. Haneen Alzamer et al., 2021 also provided short reviews on five references published since 2016. In contrast, our article discusses 18 studies conducted between 2016 and 2021 in more detail, focusing on machine learning techniques used for VO_2 max prediction. In addition, a dedicated section describes the various machine learning techniques used in the studies, their advantages and disadvantages, as well as error metrics used to evaluate the performance of VO_2 max estimation techniques.

The rest of the paper is arranged as follows. Section 2 explains the methods used to collect the information presented in this paper. Section 3 provides a detailed review of machine learning methods for VO_2 max prediction published between 2016 and 2021. Section 4 states the results. A comparison and discussion of machine learning methods used for VO_2 max prediction is given in Section 5. Finally, Section 6 provides conclusions, limitations, and future recommendations.

2. Methods

2.1. Search strategy

The PRISMA statement was used as the basis for conducting and reporting this comprehensive review [10]. Google Scholar, Elsevier, Sensors, Scientific Reports, ScienceDirect, SPROC, IEEE Xplore, and PubMed were used to search for information about VO_2 max prediction using machine learning methods. The key search terms included VO_2 max, maximal oxygen uptake, fitness level, machine learning, prediction models, gait analysis, running analysis, maximal exercise test, submaximal exercise test, graded exercise test and estimation of VO_2 max. The most recent database search took place in December 2021. Citations within the studied articles and research led to the discovery of further references. Duplicate citations were removed. The findings were then examined to see whether they might be included. In total, 18 relevant articles were included in the literature review. Results from 16 articles between 2016 and 2021 were analyzed.

2.2. Machine learning methods

Machine learning is an umbrella term for techniques in which past information is used for future decision making. Various industry fields such as healthcare, banking and life sciences use machine learning for problem-solving. Input data is observed and analyzed to understand the type, issues and relationship among data elements. The data can be preprocessed, for example, to remove noise, highlight certain features, make features comparable, etc. Different models can be trained on the input data and their performance can be evaluated and compared. Fig. 1 illustrates the process of machine learning.

This section provides a short description of various machine learning techniques that have been used for VO_2 max prediction in recent studies conducted between 2016 and 2021.

2.2.1. Support vector machines

As one of the most prominent methods for supervised learning, support vector machine (SVM) is used for both classification and regression tasks. In machine learning, SVM is mostly used for classification problems [11]. For n-dimensional data, SVM aims to find the n-dimensional optimal decision boundary, known as a hyperplane, to divide and classify data points belonging to two different classes [12]. The algorithm selects the support vectors/maximum points that help to find the hyperplane. SVM performs well in high dimensional spaces when there is a clear distinction between classes and requires only a small amount of memory. However, for bigger datasets, SVM might not perform well due to noise and overlap between classes. If data contains points from more than two classes, then multiple SVM models are needed to distinguish between all available classes. This type of SVM is called Multiclass SVM.

2.2.2. Decision tree

Decision trees, a supervised machine learning technique, consist of nodes and leaves. Data are split according to a certain condition. Decision trees are often used for classification because they are easy to implement and interpret, but they can also be applied to regression problems and can deal with continuous and categorical variables. However, decision trees have a tendency to overfit, meaning that the decision tree model learns the specific features of the dataset that it was trained on but fails to accurately classify unseen test data, which often results in low prediction accuracy [13]. Tree boost (TB) is a tree-structured regression method. The parameters of the method include the maximum number of trees used, the depth of the trees and the minimum size node to split.

2.2.3. Random forest

Random forest is a supervised machine learning technique that can be applied to both classification and regression problems. Random forest is a classifier comprising of several decision trees on a given dataset. It forecasts or classifies the outcome by combining predictions from all trees to enhance prediction accuracy [13]. Because the random forest does not depend on a single decision tree, the risk of overfitting is significantly reduced compared to the decision tree. Random forest can offer higher accuracy but requires more computational power.

2.2.4. Multiple linear regression

Multiple linear regression (MLR) is a regression algorithm that describes the linear connection between a dependent variable and several independent variables. Each feature variable must represent the linear relation between the dependent and independent variables. MLR aims to fit a regression line in a multidimensional space. MLR models can be used to identify how predictor variables relate to a criteria value and can also be used to identify outliers [14].

2.2.5. Neural networks

Artificial neural network (ANN), inspired by biology and patterned loosely after the brain, consists of nodes (or 'neurons') that are arranged in several inter-connected layers to form a network [15]. In the learning process of ANN, the nodes are associated with weights, and these weights are adjusted during training in response to the training data. ANNs can operate with limited or no 'knowledge', and exhibit fault tolerance by responding to a failure if one or more components are lost.

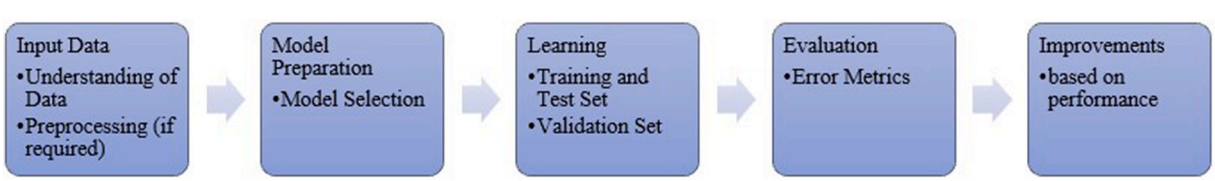

Fig. 1. Steps involved in a Machine Learning Pipeline.

A single-layer feed-forward neural network consists of only two layers, namely the input layer and the output layer. The labels in a classification task are generated in the output layer. A multilayer feedforward network involves multiple hidden layers between the input and output layers, whereas recurrent neural networks (RNN) include a feedback loop from the output layer to the input layer neurons [16]. Long short-term memory network (LSTM) is a type of RNN, which can solve long-term and short-term problems, i.e., where time has an important impact on the variable being predicted. LSTM contains memory cells, which replace the hidden layers, providing the ability to identify the cells activated or suppressed depending on the state [17]. A deep neural network is formed when the number of hidden layers in a multi-layer network exceeds three.

2.2.6. Multilayer perception

The multilayer perceptron (MLP) is a feed-forward ANN in which there are at least three layers. All nodes of the MLP, except for those in the input layer, are neurons employing non-linear activation functions to process information, which enables MLP to map more complex relationships between input data and predicted outputs. Each neuron in an MLP is trained through backpropagation, an iterative process through which the network 'learns' patterns in data [18]. Although it requires a long computation time for training, it can provide quick and accurate predictions once trained and is efficient with larger data sets. MLP and other neural network-based classification methods are capable of 'learning' from data, and can be used to solve different real-world and complex problems [19].

2.2.7. Feature selection

Feature selection is used in data mining and statistics. Features are selected either automatically or manually based on their contribution to the prediction variable or output. Feature selection and machine learning can be combined to identify discriminative features, thereby helping to eliminate irrelevant features whilst producing robust and simple models [11,19]. This basic approach is then used to predict the class of a new data point accurately [20].

2.2.8. Relief-F

The Relief-F algorithm estimates the relevance of all features, whether they are discrete or continuous. However, Relief-F is used as a preprocessor, which excludes irrelevant features before the learning process begins.

2.2.9. Radial basis function

Radial basis functions (RBF) are used as function approximators. They consist of three layers: input layer, hidden layer and output layer. Each input in the domain of RBF is assigned an absolute number. RBF produces a value that can never be negative since it is a real-valued function that depends on the distance between the input and a fixed point [21]. RBF approximation is similar to the way neural network's function, and has been shown to be efficient for solving complex problems but can result in high classification costs [16].

2.2.10. Cluster analysis

Clustering Analysis is a popular statistical data analysis and machine learning approach that separates unlabeled data points into clusters based on their similarity to one another. Different measures of the similarity between two data points can be used. The algorithm uses unsupervised learning and works with unlabeled data [22]

2.3. Error measures

In the current context, the correlation coefficient (R-value) measures the similarity between the actual value of the variable (i.e., the ground truth) and its predicted value and is defined as

$$R = \sqrt{1 - \frac{\sum_{n-1}^{N}\left(\widehat{x_n} - x_n\right)^2}{\sum_{n-1}^{N}\left(\overline{x_n} - x_n\right)^2}}$$

The R-value is always between −1 (perfect negative correlation) and 1 (perfect positive correlation) and is a measure of the strength of the relationship between two variables. Coefficient of determination or R-square (R^2) is also commonly used. Values always lie between 0 and 1 since it is a square of the R-value. A higher R-value or (R^2) value indicates more accurate model performance. Two other metrics are commonly used to assess the predictive error of machine learning models:

Root mean square error (RMSE)

$$R = \sqrt{\frac{\sum_{n-1}^{N}\left(\widehat{x_n} - x_n\right)^2}{N}}$$

and standard error of estimate (SEE)

$$R = \sqrt{\frac{\sum_{n-1}^{N}\left(\widehat{x_n} - x_n\right)^2}{N-2}}$$

These metrics help to assess the absolute fit of a regression model. RMSE is easier to calculate and is commonly used for regression tasks. RMSE values are only comparable with other RMSE values to determine the accuracy of different models. The SEE, on the other hand, provides the standard deviation or variation of the predicted value. Smaller values of SEE and RMSE indicate higher model accuracy.

3. Literature review

In Abut et al., 2016 [3], the performance of hybrid models including SVM with Relief-F was compared to MLP and TB-based models. Hybrid models used a combination of data acquired from maximal and submaximal tests as well as questionnaires. 100 healthy participants (50 females, 50 males) aged 18 to 65 were included. The submaximal test included an exercise test on a treadmill consisting of three stages (walking, jogging and running) at a self-chosen speed to determine exercise heart rate, treadmill submaximal speed (SM-ES) and exercise stage (SM-Stage). The maximal test involved an increase in speed until the participants attained the highest level of exertion. This was used to determine maximum heart rate (MX-HR), rating of perceived exertion (MX-RPE), treadmill grade and maximal respiratory exchange ratio. Questionnaires (perceived functional ability (Q-PFA) and physical activity rating (Q-PAR)) were used to assess the level of physical activity of the participants. Gender, age, MX-HR, SM-ES and Q-PFA questionnaires were the most suitable variables for predicting VO_2 max. 10-fold cross-validation was used to assess the generalization error of the prediction models, and model performance was quantified with R and RMSE values. The suggested hybrid prediction SVM based model's usefulness was tested using an ANN with different numbers of perceptron layers. The study showed that the number of hidden layer neurons has a major impact on performance; 3–11 hidden neurons were used, and a linear function was applied to the hidden and output layers to optimize performance and minimize error. The performance of an MLP-based ANN led to a reduced RMSE and improved performance compared to the TB model with 100–450 trees. The SVM-based model outperformed the MLP and TB models with an R-value of 0.94 and RMSE value of 2.92, yielding 5.03% and 10.22% lower RMSE on average than MLP and TB respectively. The authors of Abut et al., 2016 recommended testing whether other regression methods such as Decision Tree Forest (DTF) could increase the prediction accuracy of the proposed hybrid models. In addition, different feature selection techniques should be integrated with these regression methods and compared to the Relief-F algorithm.

Dincer et al., 2016 [23] predicted VO_2 max in college-aged students

using MLR with hybrid data from 26 university students. 24 models were created using predictor variables from exercise data and questionnaire variables including age, gender, height, weight, body mass index (BMI), maximum heart rate (HRmax), test duration (TT), perceived functional ability score (PFA-1 and PFA-2), and physical activity rating (Q-PAR). R-value and SEE were used to evaluate model performance. The prediction equation:

$$VO_2max = -(7.42 * gender) + (4.26 * age) - (1.44 * BMI) + (4.31 * HRmax) + (3.64 * TT) - (0.16 * PFA1) + (0.75 * PFA2) + (0.61 * PAR) - 895.26$$

resulted in the lowest SEE of 4.22 and the highest R-value of 0.79.

SVM, MLP, and Single Decision Tree (SDT) were used in Kaya et al., 2016 [24] on a dataset obtained from 48 athletes (38 males and 10 females, age 19.4 ± 4.0 yrs, height 172.5 ± 8.2 cm, body mass 64.3 ± 10.4 kg) during maximal exercise tests conducted on a treadmill. Age, height, weight, BMI, TT, and HRmax were used as predictor variables to predict VO_2 max. SEE and R values were calculated to evaluate the performance of the 12 models created using different combinations of predictor variables. The results showed reasonable error rates and SVM outperformed other techniques with 5.19% lower SEE (on average) than the MLP model, and 19.17% lower SEE than the SDT model. SVM based prediction model yielded the highest R-value of 0.72 and the lowest SEE value of 8.03.

Ozciloglu et al., 2016 [25] used SVM, MLP, and MLR to create 14 submaximal VO_2 max prediction models for 65 university students (37 males and 28 females, aged 18–37). The prediction of VO_2 max was based on the use of two categories. First: gender, age, height, and weight were used as common predictor variables. Second: gender, age, and BMI were used as mutual predictive variables. Other variables were time, speed and HRmax. No significant differences between the two categories were found. SEE and R values were calculated using 10-fold cross-validation. The predictor models with common predictor variables and time resulted in the lowest SEE. In the first category, SVM produced 5.47% and 16.09% lower SEE than MLP and MLR respectively. In the second category, SVM produced 6.51% and 14.60% lower SEE than MLP and MLR respectively.

An ANN model by Beltrame et al., 2016 [26] was used to predict VO_2 max. Data consisted of heart rate and treadmill ergometer inputs from 10 healthy young people (5 males, age 29.8 ± 7.6 yrs, height 178.4 ± 11.2 cm, body mass 75 ± 11.3 kg; 5 females, age 22.8 ± 0.7 yrs, height 165.2 ± 7.5 cm, body mass 62.1 ± 5.8 kg). The measured data were used to train an ANN to predict VO_2 max based on treadmill speed, treadmill grade, gender, exercise time, heart rate and BMI. Low bias and high linear correlation indicated accurate predictions; the ANN showed an R-value of 0.97. The authors trained a separate model using data from 9 participants that included 7 predictor variables, namely gender, BMI, exercise time, recovery time, treadmill grade, speed and HRmax. 10-fold leave one out cross-validation was used. The model consisted of 11 hidden and one output neuron, and resulted in an R-value of 0.98. Due to its simplicity, the authors speculated that the proposed model would work for different populations irrespective of factors such as height and weight, but further tests should be performed to verify this hypothesis and further improve prediction accuracy.

Wearable sensor data from 16 healthy adult males (age 27 ± 7 yrs, height 174 ± 7 cm, body mass 78 ± 14 kg) during walking was used to create a VO_2 max predictor by Beltrame et al., 2017 [8] using a random forest ensemble predictor. In the previous study [26], data from treadmill walking exercises was used to predict VO_2 max using an ANN. Wearable sensor data permitted the prediction of oxygen consumption during everyday living activities and random paced walking. The prediction model resulted in an R-value of 0.87. Nodes with two offspring nodes were found in every tree, starting at the root and continuing to the top. If the heart rate exceeded 50 beats per minute (splitting criteria), the node split into left and right subtrees according to the feature vector evaluation based on the decision value and resulted in the leaf node's prediction value. The process continued until a full tree was formed. The leaf node at the bottom of the tree contained the predicted output for the given feature values. A randomly selected portion of the training data was used to build each regression tree. The final projected value for a specific time was derived from the average of the predictions made by all of the tree's leaves. The study showed that machine learning prediction methods and data from non-intrusive wearable sensors can be used to forecast oxygen consumption patterns with reasonable accuracy. The study consisted of daily exercises limited to light and moderate intensity, which could be evaluated using the model developed in the study. Exercises of higher intensity can be tested using extensive testing techniques. Higher intensity exercises result in complex dynamics with non-linear patterns. The population in this study consisted of healthy men within a defined weight and age range [26], so further validation testing is required before applying this approach to different populations.

In a study by Akay et al., 2017a [27] SVM, Radial Basis Function Network (RBFN), Generalized Regression Neural Networks (GRNN), and DTF (which makes predictions based on several variables unlike a single decision tree) were used to create VO_2 max prediction models. Data from 98 participants (58 males, 40 females, age 20.79 ± 2.12 yrs, height 173.05 ± 1.99 cm, body mass 65.83 ± 1.89 kg) including predictive variables age, gender, weight, height, HRmax, treadmill grade, exercise time, and speed were used to create 15 different models, each using a different set of 4–8 predictor variables. Model accuracy was evaluated using the SEE and R values. The results indicated that the predictor variables treadmill grade, speed, and duration had a significant impact on VO_2 max prediction, as evidenced by a lower SEE. With the lowest SEE of 4.51, GRNN outperformed the other regression methods yielding on average 3.31%, 14.16% and 27.62% lower SEE than SVM, DTF and RBFN based models respectively. According to the results, for GRNN, SVM, DTF and RBF models, SEE decreases and R increases as more predictor variables are used. However, one clear exception can be seen in Models 3 and 10 with predictor variables (gender, age, weight, height, HRmax and speed) for which SEE increases and R decreases. This variation may be due to the effect of predictor variable speed. The authors noted that the accuracy of VO_2 max prediction models can be improved by using different machine learning methods combined with feature selection algorithms.

Akay et al., 2017b [28]built an MLR based model to predict VO_2 max in college students using physiological and questionnaire data from 62 young students (28 females, 38 males, aged 18–27). VO_2 max was predicted with 7 different models based on the predictor variables gender, age, weight, height, and PFA-1 and PFA-2 scores. VO_2 max was measured using maximal tests conducted on a treadmill. SEE and R-value was used to evaluate model accuracy. The results demonstrated that the MLR model reliably estimated VO_2 max. The addition of PA-R as a predictor variable increased VO_2 max prediction accuracy by 34.33%. The prediction equation:

$$VO_2max = (15.47 * gender) - (0.12 * age) + (0.04 * height) - (0.45 * weight) + (1.74 * PFA1) + (1.45 * PAR) + 49.74$$

gave the lowest SEE of 5.14 and the highest R-value of 0.93. The authors further speculated that feature selection algorithms could help improve VO_2 max estimates.

Akay et al., 2017c [29] developed new equations for predicting VO_2 max based on age, gender, weight, height, BMI, HRmax, and TT using data from 18 young individuals (age 21.06 ± 1.92 yrs, height 175.17 ± 8.73 cm, BMI 21.81 ± 1.08 kg m^{-2}). With the use of MLR, 12 VO_2 max prediction equations were created. SEE and R-values were used to evaluate model performance. The regression equation:

$$VO_2max = -(12.331 * gender) - (0.805 * age) + (0.883 * height) - (1.167 * weight) - (0.052 * HRmax) - (0.158 * TT) + 6.473$$

produced the lowest SEE of 3.49 and the highest R of 0.88. Predictor variables: age, height and weight play a significant role in VO_2 max prediction. As was the case for several of the reviewed studies, the authors suggested that additional research should be conducted with a larger dataset for VO_2 max prediction.

SVM and Relief-F feature selection were used by Yigit et al., 2017 [30] to create a new hybrid decision prediction model with data from 143 individuals (87 males, 56 females). Gender, exam scores, grade point average and high school specialized area were predictor variables, including results from coordination and skill tests, vertical leap, 30-m sprint, and 20-m shuttle run. The dataset was randomly split into training and test sets using 10-fold cross-validation, to ensure the validity of the results, and varied percentage ratios. Nine models were created with different combinations of predictor variables. Classification accuracy, specificity, sensitivity, negative predictive value (NPV) and positive predictive value (PPV) were used to evaluate model performance. The model that included all predictor variables produced the highest accuracy of 97.22%. The model including only a subset of predictor variables (coordination and skill test, vertical leap, sprint test) resulted in an accuracy of 77.78%, whereas the model that included sprint test score as a predictor variable resulted in an accuracy of 72.22%. The findings demonstrated that when the number of predictor variables in the prediction models was reduced, classification accuracy decreased. The approach used in this study may also be useful for predicting VO_2 max on a large dataset. Along with model accuracy, specificity, sensitivity, NPV and PPV can also be used to evaluate prediction models.

In Akay et al., 2018a [31], novel models for predicting school students' physical fitness were developed using MLR. The participants' data were separated according to specific variables, including results from the 30 m speed test, 20 m stage run, balance test, and the handgrip (right/left) test. The prediction models used gender, age, BMI, body fat, and the number of curl-ups and push-ups done in 30 s as predictor variables. Data from 333 students (133 males and 200 females, aged 11–16 yrs) was used. Eight physical fitness prediction models were created with different predictor variables (gender, age and BMI being common to all). Model accuracy was evaluated using SEE. The results revealed that the best MLR model resulted in a SEE of 3.95, and the authors concluded that this is an acceptable approach for physical fitness prediction. The authors also suggested that machine learning methods combined with feature selection could improve the accuracy of the prediction models.

SVM with Relief-F feature selection was used by Akay et al., 2018b [32] to create prediction models using data from 97 individuals (57 males and 40 females, aged 15–33 yrs). Relief-F scores were used to develop 10 models using age, gender, height, weight, HRmax, speed, time, PFA-1, PFA-2 and PA-R for VO_2 max prediction. SEE and R-values were calculated to determine model accuracy. The SVM models were also compared to RBFN models and TB models. The prediction model that included PA-R, weight, PFA-1, PFA-2, gender and HRmax had the lowest SEE of 6.42 and the highest R of 0.79. SVM-based models outperformed RBFN models and TB models, with SEE values that were (on average) 22.91% and 13.34% lower respectively. Again, the authors suggested using feature selection with different machine learning methods to improve the accuracy of VO_2 max prediction but did not apply this approach themselves.

Przednowek et al., 2018 [33] provided various models for predicting VO_2 max based on the results of the 20 m shuttle run test and anthropometric data. The study identified the best prediction model for estimating VO_2 max in 308 young adults (154 females, 154 males, aged 19–27 yrs) using 23 independent variables, namely gender, distance, HRmax, recovery heart rate, age, weight, height, waist, hip, waist to height ratio, waist to hip ratio, BMI, fat mass index, fat-free mass index, body adiposity index, body surface area, Fat, fat-free percentage, and total body water. The researchers employed MISO model types (multiple input, single output) in their investigation and used MLP, SVM, and ANN with RBF. RMSE was used to assess all models. Leave one out cross-validation was used to choose the best model. After examining different combinations of variables, it was found that for females an RBF-type neural network with 8 neurons in the hidden layer produced the most accurate model, with an RMSE of 4.07. The model for women generated a smaller error than the model for men with an RMSE value of 5.30. The common model (male and female) resulted in an RMSE value of 4.78. A large population was tested in this paper using an RBF based ANN model, which resulted in a smaller error as compared to MLP and SVM models. However, a limitation was the narrow age range of the included subjects (19–27). Age and gender have a significant effect on the prediction of VO_2 max given their strong predictive power for an athlete's aerobic endurance.

ANN was used to predict responses using a submaximal test (cycling at self-selected intensity) in Borror et al., 2019 [34]. 12 healthy adult men (aged 21.1 ± 2.5 yrs, body mass 82.1 ± 11.7 kg, height 179.3 ± 8.9 cm) cycled for 50 min at different intensities, while wearing heart rate monitors. The variables used to train, validate, and test the ANN included heart rate, the time derivative of heart rate, power output, cadence, and body mass. The model's accuracy was tested using a 12-fold hold-out cross-validation procedure. SEE and R-values were used to assess the model's accuracy. The model resulted in an R-value of 0.91 ± 0.04 and SEE of 3.34 ± 1.07. A wide range of exercise intensities and durations were assessed, potentially allowing more robust models to be developed using ANN. The method proved to be less dependent on strict protocols. Less variation was seen in the predicted data and the target data, and the results suggest that ANN could significantly improve energy expenditure estimations. Thus, this kind of simple methodology might improve the practicality of oxygen uptake measurement. The limitations of the study were the small sample size and the narrow age range of participants. Further studies should test the proposed ANN on a larger dataset with varied ages and fitness levels. The algorithm could also be modified for other physical activities such as walking.

Abut et al., 2019 [35] used SVM and Majority Voting Feature Selector (MVFS) to build a novel VO_2 max prediction model. The method relied on rank aggregation and made use of the connection between predictor variable relevance rankings provided by feature selectors. Several hybrid (combination of maximal and submaximal tests) VO_2 max prediction models were built using maximal treadmill test data from 185 college students (18–26 yrs). RMSE and R-values were used to compare the approach's performance with Relief-f, minimum redundancy maximum relevance and maximum likelihood feature selection. The results showed that MVFS outperformed other individual and ensemble feature selectors and delivered an increase in R of 8.76% and a decrease in RMSE of 11.15%. The research also showed that submaximal heart rates and exercise durations over 1.5 miles could serve as distinct predictors of VO_2 max. Using a GRNN and SDT coupled with MVFS as benchmarks, the results demonstrated that SVM outperformed both approaches for predicting VO_2 max.

Zignoli et al., 2020 [36] used a recurrent neural network (RNN) to build VO_2 max prediction models using cardiovascular features from 7 male participants (body mass 76 ± 6.6 kg) using easy to obtain inputs (Intensity Levels, weight, peak power output, HR, and Respiratory frequency) during cycling. An RNN model with 3 hidden layers with 32 neurons, 1 hidden layer with 10 neurons and one output neuron accurately predicted VO_2 max. The authors also showed that a larger dataset can be used to build an accurate model without using complex procedures to prepare the training and testing datasets. The models resulted in a peak R-value of 0.94.

Haneen Alzamer et al., 2021 [9] highlighted some recent developments in oxygen uptake prediction using machine learning in studies published between 2005 and 2020. The study provided a good overview of the main concepts regarding oxygen uptake measurements and kinetics, as well as applications of ML in sport sciences. As also outlined above, several successful predictive models have been built to predict VO_2 max. Complex procedures and health concerns associated

with direct prediction are no longer obstacles because models are typically built using data obtained from exercise, non-exercise, or hybrid methods. The article found that choosing between different machine learning algorithms requires the right balance between high R and low SEE. Further research on advanced machine learning algorithms for VO_2 max prediction is required taking into account the sample size.

In Shandhi et al., 2021 [37], combinations of machine learning algorithms and seismocardiogram data were used to predict VO_2 max in indoor and outdoor environments using data from 17 healthy adults (8 males and 9 females, aged 26.8 ± 4.1yrs) on a treadmill. Linear and nonlinear regression models with feature extraction were used. Heartbeat outliers were removed before training the models. To assess model generalizability, a simple linear regression model with leave-one-subject-out cross-validation was trained to estimate heart rate, and model performance was measured using RMSE and R-value. The prediction model resulted in an RMSE of 4.3 and R of 0.64 in the outdoor environment. This study suggested that environmental variables such as humidity and pressure could also be used as predictive variables in future studies.

4. Results

Table 1 briefly summarizes the studies conducted between 2016 and 2021 that were included in this review. Machine learning techniques and predictor variables used for each study are listed.

Table 2 summarizes the characteristics of the models listed in Table 1. The parameters used for models stated in Table 1 are also listed. No detailed information about the models was given in the papers, so only specific parameters are listed in the table.

Table 3 lists the performance values (R, SEE, RMSE) for all studies listed in this paper for comparison purposes. Blank boxes indicate that data were not presented in the relevant studies. In general, the values in Table 3 show that ANN-based models yield lower SEE/RMSE values and higher R values than other machine learning models. However, it is important to note that the values are not directly comparable because of the diverse datasets and experiments in the different studies and therefore only provide a rough idea of algorithm performance. Hence, it would be advisable for future research to test different models with a single, extensive dataset, which allows R, SEE and RMSE values to be calculated for all models, enabling reliable comparisons between different machine learning approaches.

5. Discussion

Various machine learning and statistical approaches have been used in combination with different predictor variables. Gender, age, BMI, HRmax, and test duration were the most commonly used predictor variables, indicating their strong predictive power for VO_2 max.

Several conclusions may be drawn from the findings of this research. Firstly, the majority of studies used MLR to predict VO_2 max. However, VO_2 max predictions were generally more accurate when using intelligent data processing techniques like SVM and MLP compared to MLR-based models. SVM-based models also performed better than other regression approaches on average. Overall, ANN models generally showed the best accuracy of all machine learning algorithms, as evidenced by smaller SEE and higher R-values [3,26,34].

Secondly, the specific predictor variables used to construct a model have a large effect on the accuracy of VO_2 max predictions. SEE and RMSE are generally the lowest for prediction models that contain more variables [27,28]. Conversely, decreased specificity, sensitivity, PPV, and NPV are often associated with a lower number of predictor variables [30]. Physiological factors like gender and age have been shown to be essential when it comes to creating accurate prediction models. As a result, variables such as gender can be found in nearly every published model.

Feature selection is a relatively recent approach that can potentially

Table 1
Summary of studies conducted between 2016 and 2021. The output variable for each of the studies was VO_2 max and is therefore omitted from the table.

Study	Year	ML Model	Input - Predictor Variables	Model Performance Metric
Abut et al. [3]	2016	SVM with Relief-F MLP Neural Network TB	Gender, age, MX-HR, SM-ES, Q-PFA	RMSE, R
Dincer et al. [23]	2016	MLR	Gender, age, height, weight, BMI, HRmax, TT, PFA-1, PFA-2, PA-R	SEE, R
Kaya et al. [24]	2016	SVM MLP SDT	Height, weight, BMI, TT, HRmax	SEE, R
Ozciloglu et al. [25]	2016	SVM MLP MLR	Gender, age and BMI, HRmax, speed, time	SEE
Beltrame et al. [26]	2016	ANN	Speed, treadmill grade, HR, time, body mass, gender	R
Akay et al. [27]	2017a	SVM GRNN RBFN DTF	Gender, age, height, weight, HRmax, treadmill grade, speed, and exercise time	SEE, R
Akay et al. [28]	2017b	MLR	Gender, age, weight, height, PFA-1, PFA-2, PA-R	SEE, R
Beltrame et al. [8]	2017	Random forest	HR	R
Akay et al. [29]	2017c	MLR	Gender, age, height, weight, BMI, HRmax and TT	SEE, R
Akay et al. [31]	2018a	MLR	Gender, age, BMI, body fat, the number of curl-ups and push-ups performed in 30 s	SEE
Akay et al. [32]	2018b	SVM with Relief-F RBF TB	Gender, age, weight, height, HRmax, time, speed, PFA-1 and PFA-2 and PA-R	SEE, R
Przednowek et al. [33]	2018	MLP ANN with RBF	gender, distance, HRmax, recovery heart rate, age, weight, height, waist1, waist2, hip, waist to height ratio, waist to hip ratio, BMI, fat mass index, fat-free mass index, body adiposity index, body surface area, Fat, fat-free percentage, total body water	RMSE
Borror et al. [34]	2019	ANN	HR, time derivate of HR, Power output, cadence, body mass	SEE, R
Abut et al. [35]	2019	Feature selection with SVM GRNN SDT	Gender, age, height, weight, HRmax, time, HR	RMSE, R
Zignoli et al. [36]	2020	RNN	Intensity Levels, weight, peak power output, HR, Respiratory frequency	R
Shandhi et al. [37]	2021	Simple Linear Regression	ECG, SCG, Pressure	R, RMSE

Table 2
Description of models.

Study	ML Model	Parameters	
Abut et al., 2016 [3]	SVM with Relief-F	Cost	$[2^{-2}, 2^{18}]$
		Epsilon	[0.001,160]
		Gamma	$[2^{-10},2^{9}]$
	MLP Neural Network	Number of neurons	[3, 11]
		Hidden layer activation function	Logistic
		Output layer activation function	Linear
	TB	Maximum number of trees used in series	[100,450]
		Minimum size node to split	[2, 28]
		Depth of individual trees	[5, 16]
Dincer et al., 2016 [23]	MLR	VO_2 max = - (7.42 x gender) + (4.26 x age) - (1.44 x BMI) + (4.31 x HRmax) + (3.64 x TT) - (0.16 x PFA-1) + (0.75 x PFA-2) + (0.61 x PAR) - 895.26	
Kaya et al., 2016 [24]	SVM	Cost	[0.1–6000]
		Epsilon	[0.001–150]
		Gamma	[0.0001–100]
	MLP	Number of neurons	[1−24]
		Learning Rate	[0–1]
		Momentum	[0–1]
	SDT	Minimum rows	[2−16]
		Minimum size	[5−20]
		Maximum Tree Level	10
Ozciloglu et al., 2016 [25]	SVM	Different combinations of predictor variables	
	MLP	10 Fold Cross-Validation	
	MLR	No further information available on the used model.	
Beltrame et al., 2016 [26]	ANN	Number of neurons	11
		Activation Function	Linear
		Power	80%
		Significance Level	5%
		10-fold leave one out cross-validation	
Akay et al., 2017a [27]	SVM	No description of the model available	
	GRNN	No description of the model available	
	RBFN	No description of the model available	
	DTF	No description of the model available	
Akay et al., 2017b [28]	MLR	VO_2 max = (15.47 x gender) - (0.12 x age) + (0.04 x height) - (0.45 x weight) + (1.74 x PFA-1) + (1.45 x PAR) + 49.74	
Beltrame et al., 2017 [8]	Random forest	Average output (VO_2 max) of Binary trees	
		Leave one out cross-validation	
Akay et al., 2017c [29]	MLR	VO_2 max = - (12.331 x gender) - (0.805 x age) + (0.883 x height) - (1.167 x weight) - (0.052 x HR max) - (0.158 x TT) + 6.473	
Akay et al., 2018a [31]	MLR	No description of the model available	
Akay et al., 2018b [32]	SVM with Relief-F	Cost	[1–100]
		Gamma	[0.00–50]
		Kernel Function	RBF
	RBF	Maximum Neurons	8
		Radius	[0.001–400]
		Lamda	[10−100]
	TB	Maximum Trees	[200−550]
		Depth	[5−7]
		Minimum size node	[8−23]
Przednowek et al., 2018 [33]	MLP	y = 72.31 + 1.94 x gender + 0.01 x distance − 0.21 x height − 0.23 x fat	
	ANN with RBF	Number of neurons	5
		Leave one out cross-validation	
Borror et al., 2019 [34]	ANN	ANN Function (MATLAB)	fitnet
		Number of Neurons	9
		12-fold hold out cross-validation	
Abut et al., 2019 [35]	Feature selection with SVM	Input: Ranking List of Predictor Variables	
		Output: MVFS Based Ranking List	
		Cost	$[2^{-2} - 2^{19}]$
		Epsilon	[0.001–170]
		Gamma	$[2^{-10}–2^{12}]$
	GRNN	Max sigma	[0.0001–11]
		Min Sigma	[1–65]
		Step	[1–59]
	SDT	Min Rows	[1 - 20]
		Min Node Size	[3 - 10]
		Max Tree Levels	[5–40]
Zignoli et al., 2020 [36]	RNN	Three LSTM Layers	32 neurons each
		Hidden Layer	10 neurons
		Training Method	Stochastic Gradient descent (adagrad)
Shandhi et al., 2021 [37]	Simple Linear Regression	Learning Rate	0.05
		Depth	10
		Gamma	0.3
		Estimators	100

Table 3
Performance evaluation of machine learning models.

Study	Year	ML Model	R	SEE $ml\,kg^{-1}min^{-1}$	RMSE $ml\,kg^{-1}min^{-1}$
Abut et al. [3]	2016	SVM with Relief-F	0.94		2.92
		MLP Neural Network	0.93		3.14
		TB	0.92		3.38
Dincer et al. [23]	2016	MLR	0.79		4.22
Kaya et al. [24]	2016	SVM	0.72	8.03	
		MLP	0.56	9.58	
		SDT	0.38	10.67	
Beltrame et al. [26]	2016	Neural Network	0.97		
Akay et al. [27]	2017a	SVM	0.77	4.87	
		GRNN	0.81	4.51	
		RBFN	0.51	7.24	
		DTF	0.70	5.62	
Akay et al. [28]	2017b	MLR	0.93	5.14	
Beltrame et al. [8]	2017	Random forest	0.87		
Akay et al. [29]	2017c	MLR	0.88	3.49	
Akay et al. [31]	2018a	MLR		3.95	
Akay et al. [32]	2018b	SVM with Relief-F	0.785	6.415	
		RBF	0.661	7.740	
		TB	0.662	7.771	
Przednowek et al. [33]	2018	MLP			4.78
		ANN with RBF			4.07
Borror et al. [34]	2019	ANN	0.91	3.34	
Abut et al. [35]	2019	Feature selection with SVM	0.86		2.91
		GRNN	0.81		3.37
		SDT	0.64		4.51
Zignoli et al. [36]	2020	RNN	0.94		
Shandhi et al. [37]	2021	Simple Linear Regression	0.64		4.3

improve VO_2 max prediction models. Feature selection helps remove outliers and unnecessary or collinear features, and can thus result in higher predictive model accuracy. Feature importance value should also be calculated in future research to help improve the efficacy of feature selection.

6. Conclusion

This article provides a review of VO_2 max prediction experiments published between 2016 and 2021. Comparison of the previously conducted studies shows that ANN can accurately predict VO_2 max. The reliability of the algorithms can also be increased by training the models with data from a large, heterogeneous group of test subjects. Advanced machine learning techniques like deep learning can be used for more accurate VO_2 max predictions.

6.1. Limitations

The review discusses different machine learning models used in studies between 2016 and 2021 for VO_2 max prediction. However, the use of diverse datasets and parameters, as well as large variations in sample size between studies, hinders accurate comparison of all proposed methods. To allow a more accurate comparison of the various methods, the models should be compared with the same dataset and parameters. Another limitation is that most studies examined only healthy people or college-aged students. It would be of interest to also test the prediction capabilities of the proposed methods for elderly people and people with disabilities, especially since age seems to have strong predictive power for VO_2 max. Finally, several of the examined studies included small sample sizes, which limits the generalizability of the models that are developed.

6.2. Future recommendations

In future studies, neural networks and deep learning methods should be further investigated for VO_2 max prediction due to their generally superior predictive performance. A single dataset should be analyzed with different machine learning models and architectures to allow meaningful comparisons. The small number of participants in most studies increases the risk of overfitting, but only a few studies used cross-validation to mitigate the risk. In future research cross-validation should always be used to produce generalizable results. Feature selection could also be further investigated to improve the accuracy of VO_2 max prediction, and the value could be used as a reliable metric of model predictive accuracy. In addition, studies should aim to include other human populations, such as elderly people and those with clinical disorders. Finally, the release of open-source datasets could help to accelerate the development of more robust models for VO_2 max prediction, and enable benchmarking of new algorithms. Developments in this area could reduce the need to perform expensive exercise testing in order to determine VO_2 max, allowing it to be estimated simply, quickly and non-invasively.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported in part by the Academy of Finland, grants 323472 and 323473 (under consortium "GaitMaven: Machine learning for gait analysis and performance prediction").

References

[1] Coulson M, Archer D. Practical Fitness Testing. Bloomsbury Publishing Plc; 2009.
[2] Bundy M, Leaver A. A Guide to Sports and Injury Management E-Book. Elsevier Health Sciences; 2012.
[3] Abut F, Akay MF, George J. Developing new VO2max prediction models from maximal, submaximal and questionnaire variables using support vector machines combined with feature selection. Comput Biol Med 2016;79:182–92. https://doi.org/10.1016/j.compbiomed.2016.10.018.
[4] Koutlianos N, Dimitros E, Metaxas T, Cansiz M, Deligiannis AS, Kouidi E. Indirect estimation of VO2max in athletes by ACSM's equation: valid or not? Hippokratia 2013;17(2):136–40.
[5] Akay MF, Shokrollahi N, Aktürk E, George JD. Development Of New VO2max Prediction Models By Using Artificial Neural Networks. Kusadasi, Turkey: Intl. Symp. on Computing in Science and Engineering; 2011. p. 39–44.
[6] Abut F, Akay MF. Machine learning and statistical methods for the prediction of maximal oxygen uptake: recent advances. Med Dev (Auckland, N.Z.) 2015;8:369–79. https://doi.org/10.2147/MDER.S57281.
[7] Akay MF, Inan C, Bradshaw DI, George JD. Support vector regression and multilayer feed forward neural networks for non-exercise prediction of VO2max. Expert Syst Appl 2009;36(6):10112–9.
[8] Beltrame T, Amelard R, Wong A, Hughson RL. Prediction of oxygen uptake dynamics by machine learning analysis of wearable sensors during activities of daily living. Sci Rep 2017;7(1):45738. https://doi.org/10.1038/srep45738.
[9] Alzamer Haneen, Abuhmed Tamer, Hamad Kotiba. A short review on the machine learning-guided oxygen uptake prediction for sport science applications. Electronics (Basel) 2021;10:1956. https://doi.org/10.3390/electronics10161956. 1956.
[10] Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JPA, Clarke M, Devereaux PJ, Kleijnen J, Moher D. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. PLoS Med 2009;6(7):e1000100. https://doi.org/10.1371/journal.pmed.1000100.
[11] Karasu S, Saraç Z. Classification of power quality disturbances by 2D-Riesz Transform, multi-objective grey wolf optimizer and machine learning methods. Digit Signal Process 2020;101:102711. https://doi.org/10.1016/j.dsp.2020.102711.
[12] Demidova L, Klyueva I, Sokolova Y, Stepanov N, Tyart N. Intellectual approaches to improvement of the classification decisions quality on the base of the SVM classifier. Procedia Comput Sci 2017;103:222–30. https://doi.org/10.1016/j.procs.2017.01.070.
[13] Przybyła-Kasperek M, Aning S. Bagging And Single Decision Tree Approaches To Dispersed Data. Computational Science – ICCS 2021; 2021. p. 420–7. https://doi.org/10.1007/978-3-030-77967-2_35.
[14] Ge Y, Wu H. Prediction of corn price fluctuation based on multiple linear regression analysis model under big data. Neural Comput Appl 2019;32(22):16843–55. https://doi.org/10.1007/s00521-018-03970-4.
[15] Brette R. Philosophy of the spike: rate-based vs. Spike-based theories of the brain. Front Syst Neurosci 2015;9. https://doi.org/10.3389/fnsys.2015.00151.
[16] Chandramouli, Dutt S, Das A. Machine Learning. first ed. Pearson Education India; 2018.
[17] Altan A, Karasu S, Zio E. A new hybrid model for wind speed forecasting combining long short-term memory neural network, decomposition methods and grey wolf optimizer. Appl Soft Comput 2021;100:106996. https://doi.org/10.1016/j.asoc.2020.106996.
[18] Rather S, Bala P. A hybrid constriction coefficient-based particle swarm optimization and gravitational search algorithm for training multi-layer perceptron. Int J Intell Comput Cybernet 2020;13(2):129–65. https://doi.org/10.1108/ijicc-09-2019-0105.
[19] Karasu S, Saraç Z. Investigation of power quality disturbances by using 2D discrete orthonormal S-transform, machine learning and multi-objective evolutionary algorithms. Swarm Evolut Comput 2019;44:1060–72. https://doi.org/10.1016/j.swevo.2018.11.002.
[20] Cai J, Luo J, Wang S, Yang S. Feature selection in machine learning: a new perspective. Neurocomputing 2018;300:70–9. https://doi.org/10.1016/j.neucom.2017.11.077.
[21] Majdisova Z, Skala V. Radial basis function approximations: comparison and applications. Appl Math Model 2017;51:728–43. https://doi.org/10.1016/j.apm.2017.07.033.
[22] Chawla D, Joshi H. Consumer perspectives about mobile banking adoption in India – a cluster analysis. Int J Bank Market 2017;35(4):616–36. https://doi.org/10.1108/ijbm-03-2016-0037.
[23] Dincer OF, Akay MF, Cetin E, Yarim I, Daneshvar S. New prediction equations for estimating the maximal oxygen consumption of college-aged students using hybrid data. In: Proceedings of the 1st International Mediterranean Science and Engineering Congress, Adana, Turkey, 26-28 Oct 2016; 2016. 1-1.
[24] Kaya K, Akay MF, Cetin E, Yarim I. Development of new prediction models for maximal oxygen uptake using artificial intelligence methods. In: Proceedings of the International Conference on Natural Science and Engineering, Kilis, Turkey, 19-20 March 2016; 2016. p. 986–8.
[25] Ozcioglu MM, Akay MF, Cetin E, Yarım I. Development of New Maximum Oxygen Uptake Prediction Models For Turkish College Students Using Support Vector Machines and Submaximal Data. Proc. Fourth International Symposium on Engineering, Artificial Intelligence and Applications (ISEAIA 2016). Nov 2016. p. 19–20.
[26] Beltrame T, Amelard R, Villar R, Shafiee MJ, Wong A, Hughson RL. Estimating oxygen uptake and energy expenditure during treadmill walking by neural network analysis of easy-to-obtain inputs. 1985 J Appl Physiol 2016;121(5):1226–33. https://doi.org/10.1152/japplphysiol.00600.2016.
[27] Akay MF, Cetin E, Yarim I, Bozkurt O, Ozcioglu MM. Development of novel maximal oxygen uptake prediction models for Turkish college students using machine learning and exercise data. 2017. https://doi.org/10.1109/

CICN.2017.8319382. Paper presented at the 186-189, https://ieeexplore.ieee.org/document/8319382.

[28] Akay MF, Cetin E, Yarım I, Ozciloglu MM. New prediction models for the maximal oxygen uptake of collegeaged students using non-exercise data. New Trends Issues Proc Human Soc Sci 2017;4(4):1–5. https://doi.org/10.18844/prosoc.v4i4.2587.

[29] Akay MF, Abut F, Kaya K, Cetin E, Yarim I. New regression equations for estimating the maximal oxygen uptake of college of physical education and sports students in Turkey. New Trends Issues Proc Human Soc Sci 2017;3(3):11–7. https://doi.org/10.18844/gjhss.v3i3.1505.

[30] Yigit GO, Akay MF, Alak H. Development of new hybrid admission decision prediction models using support vector machines combined with feature selection. New Trends Issues Proc Human Soc Sci 2017;3(3):1–10. https://doi.org/10.18844/gjhss.v3i3.1502.

[31] Akay MF, Bozkurt O, Cetin E, Yarim I. Multiple linear regression-based physical fitness prediction models for Turkish secondary school students. New Trends Issues Proc Human Soc Sci 2018;5(4):58–64. https://doi.org/10.18844/prosoc.v5i4.3704.

[32] Akay MF, Ozciloglu MM, Cetin E, Yarim I, Daneshvar S. Estimating the maximal oxygen uptake with new prediction models for college-aged students using feature selection algorithm. New Trends Issues Proc Human Soc Sci 2018;5(4):52–7. https://doi.org/10.18844/prosoc.v5i4.3703.

[33] Przednowek K, Barabasz Z, Zadarko-Domaradzka M, Przednowek K, Nizioł-Babiarz E, Huzarski M, Sibiga K, Dziadek B, Zadarko E. Predictive modeling of VO2max based on 20 m shuttle run test for young healthy people. Appl Sci 2018;8(11):2213. https://doi.org/10.3390/app8112213.

[34] Borror A, Mazzoleni M, Coppock J, Jensen BC, Wood WA, Mann B, Battaglini CL. Predicting oxygen uptake responses during cycling at varied intensities using an artificial neural network. Biomed Hum Kinet 2019;11(1):60–8. https://doi.org/10.2478/bhk-2019-0008.

[35] Abut F, Akay MF, George J. A robust ensemble feature selector based on rank aggregation for developing new VO2max prediction models using support vector machines. Turk J Electr Eng Comput Sci 2019;27(5):3648–64. https://doi.org/10.3906/elk-1808-138.

[36] Zignoli A, Fornasiero A, Ragni M, Pellegrini B, Schena F, Biral F, Laursen PB. Estimating an individual's oxygen uptake during cycling exercise with a recurrent neural network trained from easy-to-obtain inputs: a pilot study. PLoS One 2020;15(3):e0229466. https://doi.org/10.1371/journal.pone.0229466.

[37] Shandhi MMH, Bartlett WH, Heller JA, Etemadi M, Young A, Plotz T, Inan OT. Estimation of instantaneous oxygen uptake during exercise and daily activities using a wearable cardio-electromechanical and environmental sensor. IEEE J Biomed Health Informat 2021;25(3):634–46. https://doi.org/10.1109/JBHI.2020.3009903.