Turkish Journal of Electrical Engineering & Computer Sciences

http://journals.tubitak.gov.tr/elektrik/

Research Article

Turk J Elec Eng & Comp Sci
(2019) 27: 3648 – 3664
© TÜBİTAK
doi:10.3906/elk-1808-138

A robust ensemble feature selector based on rank aggregation for developing new VO_2max prediction models using support vector machines

Fatih ABUT[1], **Mehmet Fatih AKAY**[2,*], **James GEORGE**[3]

[1]Department of Computer Engineering, Faculty of Engineering, Adana Alparslan Türkeş Science and Technology University, Adana, Turkey

[2]Department of Computer Engineering, Faculty of Engineering, Çukurova University, Adana, Turkey

[3]Department of Exercise Sciences, College of Life Sciences, Brigham Young University, Provo, Utah, USA

Received: 17.08.2018 • **Accepted/Published Online:** 19.06.2019 • **Final Version:** 18.09.2019

Abstract: This paper proposes a new ensemble feature selector, called the majority voting feature selector (MVFS), for developing new maximal oxygen uptake (VO_2max) prediction models using a support vector machine (SVM). The approach is based on rank aggregation, which meaningfully utilizes the correlation among the relevance ranks of predictor variables given by three state-of-the-art feature selectors: Relief-F, minimum redundancy maximum relevance (mRMR), and maximum likelihood feature selection (MLFS). By applying the SVM combined with MVFS on a self-created dataset containing maximal and submaximal exercise data from 185 college students, several new hybrid VO_2max prediction models have been created. To compare the performance of the proposed ensemble approach on prediction of VO_2max, SVM-based models with individual combinations of Relief-F, mRMR, and MLFS as well as with other alternative ensemble feature selectors from the literature have also been developed. The results reveal that MVFS outperforms other individual and ensemble feature selectors and yields up to 8.76% increment and 11.15% decrement rates in multiple correlation coefficients (Rs) and root mean square errors ($RMSE$s), respectively. Furthermore, in addition to reconfirming the relevance of sex, age, and maximal heart rate in predicting VO_2max, which were previously reported in the literature, it is revealed that submaximal heart rates and exercise times at 1.5-mile distance are two further discriminative predictors of VO_2max. The results have also been compared to those obtained by a general regression neural network and single decision tree combined with MVFS, and it is shown that the SVM exhibits much better performance than other methods for prediction of VO_2max.

Key words: Ensemble feature selection, rank aggregation, support vector machine, maximal oxygen uptake, prediction

1. Introduction

VO_2max is defined as the maximum volume of oxygen an individual uses while exercising at maximal capacity, and it is usually measured as a relative rate in milliliters of oxygen per kilogram of body mass per minute (mL kg^{-1} min^{-1}). The direct measurement of VO_2max during a maximal exercise test is known as the most proper technique for the evaluation of endurance capacity. However, given that the tests of direct measurement of VO_2max need expensive equipment, a great deal of time, and trained staff with expertise, a large number of researchers have attempted to find indirect and simpler ways of determining VO_2max based on prediction equations. Consequently, dozens of various prediction models to estimate VO_2max of different target audiences ranging from professional athletes to fit adults, teenagers, and children have been proposed in the related

*Correspondence: fakay@cu.edu.tr

literature. This plethora of models shows a wide spectrum of different characteristics, depending on data obtained by maximal, submaximal, nonexercise, or combined "hybrid" tests performed; machine learning or statistical methods applied; or the predictor variables used. A detailed survey of previously conducted studies on prediction of VO_2max along with major conclusions can be found in [1].

In the course of recent years, another active research area in this field has been the application of feature selection for identifying the most relevant predictors of VO_2max, with the final aim of developing more accurate prediction models [2]. Our literature review shows that there are three major studies [3–5] in the literature employing individual feature selectors such as Relief-F, mRMR, and correlation-based feature selection (CFS) to determine the discriminative predictors of VO_2max. Particularly, the study in [3] proposed various Relief-F-based and CFS-based regular VO_2max models using maximal, submaximal, or nonexercise variables only. Support vector machine (SVM), multilayer perceptron (MLP), and multiple linear regression (MLR) were used to build the prediction models. In [4], SVM with individual combinations of Relief-F and mRMR was used to develop new hybrid VO_2max prediction models based on maximal and questionnaire data, with the aim of comparing the performance of Relief-F with mRMR. For comparison purposes, prediction models based on a radial basis function neural network (RBFNN) and decision tree forest (DTF) have also been created. In a follow-up work, the study in [5] proposed new hybrid VO_2max models by combining the maximal, submaximal, and questionnaire variables using SVM combined with the Relief-F algorithm and compared the results to those of MLP-based and tree boost (TB)-based models. To sum up the major conclusions made by these three studies, it has been reported that sex, age, MX-HR, maximal exercise time, submaximal ending speed of the treadmill, and Perceived Functional Ability (Q-PFA) and Activity Code (AC) questionnaires are the essential predictors of VO_2max, and SVM has been reported to be superior to the rest of the utilized regression methods for prediction of VO_2max.

However, the analysis of these studies indicates that every single feature selector strategy has its own shortcomings, and there is no generalizable superiority of one feature selector over the others that consistently performs well over all datasets. To overcome the shortcomings of individual methods, one possible way would be to design ensemble feature selectors based on rank aggregation, which tends to achieve higher accuracy and generalize better than individual methods [6, 7]. A rank aggregation can be described as a collection of individual feature selectors whose results are combined typically by means of weighted or unweighted voting. Although the idea of rank aggregation is not new, it has been applied mostly to classification problems [8–11] rather than regression problems. An exhaustive review of the related literature revealed that, to the best of our knowledge, there is only one major study [12] that introduced an ensemble feature selector, referred to as EFS, also applicable to regression problems. However, EFS independently sums the results of each individual feature selector up to a cumulative ranking after normalizing the individual results to a common scale, with no deeper investigation of consistency relationships among the relevance ranks of variables reported by respective individual feature selectors. A new approach quantifying and evaluating the correlation of the rank results of individual feature selectors may further increase the robustness of an ensemble ranker in determining relevant and irrelevant predictors of VO_2max.

The purpose of this paper is to design a new ensemble feature selector, called the majority voting feature selector (MVFS), which aggregates the consensus properties of Relief-F, mRMR, and MLFS, with the aim of making more robust decisions about the set of relevantly identified VO_2max predictors and creating new accurate VO_2max prediction models. SVM has been used to develop the prediction models, and the dataset on which the proposed approach has been evaluated was created using data of 185 college students who successfully completed a submaximal 1.5-mile endurance test and a maximal graded exercise test. To compare the performance of the proposed approach in prediction of VO_2max, SVM-based models with individual combinations of Relief-F, mRMR, and MLFS have also been developed. Similarly, to draw a comparison with the performance of SVM, prediction models based on two popular types of artificial neural network-based and tree-structured methods including the general regression neural network (GRNN) and single decision tree (SDT), respectively, have also been created.

2. Dataset generation

The dataset contains data from 185 college students (115 males and 70 females) from Brigham Young University, aged from 18 to 26 years, where prior to any testing all subjects signed an informed consent document approved by the University Institutional Review Board and completed a brief physical activity questionnaire to filter for possible cardiovascular contraindications.

To produce the dataset, the subjects accomplished a submaximal 1.5-mile endurance test and a maximal graded exercise test (GXT). The submaximal 1.5-mile test was completed on a 229.7-m indoor track, which required the completion of ten and a half laps. A test administrator registered the elapsed exercise time and heart rate via heart rate monitor as the subjects passed the 0.5-mile, 1-mile, and 1.5-mile distances. At least 24 h after the 1.5-mile test, subjects completed a maximal GXT in the laboratory. Upon entering the laboratory, subjects' heights and weights were measured using a standard physician's scale and a wall-mounted stadiometer, respectively. To warm up for the maximal test, subjects began walking on the treadmill and then were instructed to continue walking briskly or to begin jogging/running at a comfortable pace for about 2–3 min. Following the warm-up, subjects continued to exercise at this same self-selected pace while the treadmill grade was enhanced 1.5% every minute. These progressive increases in grade continued until subjects reached volitional fatigue and could not move on despite oral motivation. After the maximal test, a brief 5-min cool-down (walking at a slow pace) was completed. During the maximal GXT, subjects' SM-HR was registered at the end of each stage. VO_2max was computed by averaging the highest four consecutive 15-s scores close to the end of the maximal test.

In addition to physiological variables such as sex (0.62 ± 0.48), age (22.42 ± 3.39), height (1.75 ± 0.10), and weight (74.03 ± 15.42), the dataset contains two further categories of predictor variables. The first category, i.e. maximal variables, contains MX-HR (193.83 ± 8.34) and MX-RER (1.15 ± 0.04), and the second category, i.e. submaximal variables, consists of the exercise times SM-MIN1 (165.71 ± 16.25), SM-MIN2 (166.99 ± 26.56), and SM-MIN3 (171.47 ± 17.05) and heart rates SM-HR1 (4.88 ± 1.38), SM-HR2 (9.72 ± 2.49), and SM-HR3 (14.70 ± 3.56) at the 0.5-mile mark, 1-mile mark, and 1.5-mile mark, respectively.

3. Majority voting feature selector

The MVFS algorithm proposed in this study is an ensemble feature selector consisting of three individual feature selectors including Relief-F, mRMR, and MLFS, the details of which were described in [13], [14], and [15], respectively. The motivation behind designing this ensemble approach was that each individual feature selector uses a specific feature evaluation criterion and may yield different ranking results even when applied to the same dataset. The MVFS algorithm aims to aggregate the consensus properties of Relief-F, mRMR, and MLFS to produce more robust decisions about the set of relevantly identified variables. The necessary and sufficient condition for rank aggregation to outperform its individual members is that the combined rankers are accurate and diverse. Another critical decision when performing ensemble feature selection is the aggregation technique used for combining the resulting variable lists from the multiple feature selectors into a single decision for each variable. Our aggregation technique is based on performing majority voting or calculating correlation scores among variables or, in the worst case, using the priority order among the three individual feature selectors, which is to be predefined based on their average prediction performance on the dataset being considered.

The main steps of the MVFS algorithm can be summarized as follows. The ranking lists of the predictor variables reported by Relief-F, mRMR, and MLFS are passed as input parameters to the ensemble feature selector. For every rank of the lists, three steps are followed until one of the variables reported by the three individual feature selectors is chosen. The first step involves the majority voting strategy where a consensus among the feature selector results is to be reached. Majority voting belongs to the simplest strategies in designing combination classifiers. However, the power of this strategy is comparable to other complex methods [16]. Obviously, the decision criterion for majority voting is fulfilled when at least two of the three feature selectors point to the same variable, or a unique consensus among the three feature selectors for the same variable has been observed. In the other case, where the decision criterion is not fulfilled, i.e. each of the three feature selectors report a different variable for the current rank, the algorithm proceeds with the second step.

The second step involves separate calculation of the so-called correlation score (*CS*) for each variable reported by the three individual feature selectors. In the related literature, similar evaluation metrics correlating the independent estimates of individual feature selectors were already introduced for classification problems and have also been proved as effective solutions for robust decision making in ensemble techniques for more relevant variables [17]. In this study, the *CS* for a variable, as given in Eq. (1), is formed by the addition of two further scores, namely the inconsistency score (*IS*) of a variable across the three feature selector results and the ranking distance of a variable to the first rank (*DS*). The formulas for *CS* and *DS* are shown in Eq. (2) and Eq. (3), respectively:

$$CS_{var}\left(fs,i\right) = IS_{var}\left(fs,i\right) + \quad DS_{var}\left(fs,i\right), \tag{1}$$

$$IS_{var}\left(fs,i\right) = \begin{pmatrix} \left|rank_{var}\left(Relief-F\right) - \ rank_{var}\left(mRMR\right)\right| + \\ \left|rank_{var}\left(Relief-F\right) - \ rank_{var}\left(MLFS\right)\right| + \\ \left|rank_{var}\left(MLFS\right) - \ rank_{var}\left(mRMR\right)\right| \end{pmatrix}, \tag{2}$$

$$DS_{var}\left(fs,i\right) = \begin{pmatrix} \left(rank_{var}\left(Relief-F\right) - 1\right) + \\ \left(rank_{var}\left(mRMR\right) - 1\right) + \\ \left(rank_{var}\left(MLFS\right) - 1\right) \end{pmatrix}, \tag{3}$$

where the *rank* function gives the relevance rank of the variable *var* reported by the corresponding feature selector *fs*. *IS* expresses the inconsistency level of a variable in terms of its rank reported by the three individual feature selectors, whereas *DS* is defined as the absolute difference between the rank of the variable being considered and the first rank. The lower the values of both scores are, the more relevant the variables can be considered to be. Particularly, a lower *IS* for a variable indicates that it has been relatively reported by approximately the same rank by the three feature selectors. Similarly, a lower *DS* for a variable indicates that it is relatively more relevant compared to the other two variables. In conclusion, *CS* is calculated for each variable in the triple candidate set and the variable having the lowest *CS* is selected for that rank. In the other case, where at least two variables have the same *CS*, the algorithm continues with the last step.

For cases where the variable selection process for a rank is completed (i.e. the decision criterion of majority voting for a variable is fulfilled in the first step or a variable with the highest *CS* is chosen in the second step), MVFS subsequently updates the list of variable ranks of the feature selector whose rank is assumed to be incorrect. Particularly, the rank of the variable being corrected is moved up to the same rank of other dominant feature selectors and the rest of the affected variables are rearranged so that they are listed after that reranked variable by holding their previous relative orders.

The last step is based on the previously defined priority order of the feature selectors to decide the final decision of a variable in the candidate set. To determine the priority ranking among the three individual feature selectors for prediction of VO_2max, the average performances of the three individual feature selectors on the used dataset is considered, i.e. the feature selector producing the lowest average *RMSE* over all respective VO_2max models is assigned with the first priority, whereas the feature selectors occupying the second and third places in terms of the performance regarding the average *RMSE*s receive the second and third priority for prediction of VO_2max, respectively. According to the average results obtained by applying the three individual feature selectors on the utilized dataset, as presented in Section 5.3, their ranking from highest to the lowest priority has been determined as MLFS, mRMR, and Relief-F.

This three-step procedure is repeated until the relevance of all predictor variables reported by the three individual feature selectors has been evaluated. The pseudocode of the proposed ensemble ranker is illustrated in Algorithm.

Algorithm Pseudocode of the MVFS algorithm.

```
Input: ranking lists of predictor variables reported by Relief-F, mRMR, and MLFS
Output: a new MVFS-based ranking list of predictor variables
repeat
        Retrieve the triple of predictor variables at rank i
        Apply the majority of voting principle on the triple of variables
               if (consensus for a variable is reached)
                      then select the corresponding variable by majority of votes
               else
                      Calculate the CS for each variable in the triple candidate set using Eq. 1
                      if (there is a single minimum CS)
                             then select the variable with the lowest CS
                      else
                             Select the predictor variable reported by the feature selector that
                             has been assigned with the highest priority
                      end if
               end if
until the relevance of all predictor variables has been evaluated
```

4. Methodology and prediction models

By running the MVFS algorithm on the dataset, the relevance rank of every predictor variable has been calculated. The relevance order of the variables, from highest to the lowest rank, is illustrated in Table 1. Particularly, the results obtained by the MVFS show that the decisions on the rankings of 7 variables (i.e. age, height, weight, MX-RER, SM-HR1, SM-HR2, and SM-MIN3) have been made by consensus voting; 4 variables (i.e. sex, MX-HR, SM-MIN1, and SM-MIN2) have been ranked using the calculated *CS*s; and finally, only a single variable (i.e. SM-HR3) has been ranked via the priority order of MLFS over the other two individual feature selectors.

The prediction models have been created by iteratively removing the predictor variable with the lowest rank until a single variable with the highest rank remains in the last model. In the same manner, the relevance ranks of the same predictor variables have been determined for Relief-F, mRMR, and MLFS. Additionally, another alternative ensemble feature selector proposed in the literature, i.e. EFS, has also been applied to the dataset. The achieved ranks of predictor variables for Relief-F, mRMR, MLFS, and EFS are illustrated in Table 1. Table 2 shows the MVFS-based VO_2max prediction models along with their predictor variables, and similarly Table 3 through Table 6 illustrate the Relief-F-based, mRMR-based, MLFS-based, and EFS-based prediction models, respectively, that have been created for comparison purposes.

Table 1. Relevance ranks of the variables given by MVFS, Relief-F, mRMR, MLFS and EFS.

Rank	MVFS	Relief-F	mRMR	MLFS	EFS
1	SM-MIN3	Weight	SM-MIN3	SM-MIN3	SM-MIN3
2	Weight	SM-MIN3	Weight	SM-MIN2	SM-HR1
3	SM-MIN2	MX-HR	MX-RER	Sex	Sex
4	Sex	SM-MIN2	Sex	SM-MIN1	SM-MIN2
5	SM-MIN1	Height	SM-MIN1	Weight	Weight
6	MX-HR	Sex	MX-HR	SM-HR3	SM-HR2
7	SM-HR3	SM-HR3	SM-MIN2	MX-RER	SM-MIN1
8	MX-RER	SM-MIN1	Age	MX-HR	SM-HR3
9	Age	Age	SM-HR1	Age	Age
10	SM-HR1	SM-HR2	Height	SM-HR1	MX-HR
11	Height	SM-HR1	SM-HR2	SM-HR2	Height
12	SM-HR2	MX-RER	SM-HR3	Height	MX-RER

Table 2. Overview of MVFS-based VO_2max models along with their predictor variables.

Models	Predictor variables
Model 1	SM-MIN3, weight, SM-MIN2, sex, SM-MIN1, MX-HR, SM-HR3, MX-RER, age, SM-HR1, height, SM-HR2
Model 2	SM-MIN3, weight, SM-MIN2, sex, SM-MIN1, MX-HR, SM-HR3, MX-RER, age, SM-HR1, height
Model 3	SM-MIN3, weight, SM-MIN2, sex, SM-MIN1, MX-HR, SM-HR3, MX-RER, age, SM-HR1
Model 4	SM-MIN3, weight, SM-MIN2, sex, SM-MIN1, MX-HR, SM-HR3, MX-RER, age
Model 5	SM-MIN3, weight, SM-MIN2, sex, SM-MIN1, MX-HR, SM-HR3, MX-RER
Model 6	SM-MIN3, weight, SM-MIN2, sex, SM-MIN1, MX-HR, SM-HR3
Model 7	SM-MIN3, weight, SM-MIN2, sex, SM-MIN1, MX-HR
Model 8	SM-MIN3, weight, SM-MIN2, sex, SM-MIN1
Model 9	SM-MIN3, weight, SM-MIN2, sex
Model 10	SM-MIN3, weight, SM-MIN2
Model 11	SM-MIN3, weight
Model 12	SM-MIN3

Table 3. Overview of Relief-F-based VO_2max models along with their predictor variables.

Models	Predictor variables
Model 13	Weight, SM-MIN3, MX-HR, SM-MIN2, height, sex, SM-HR3, SM-MIN1, age, SM-HR2, SM-HR1, MX-RER
Model 14	Weight, SM-MIN3, MX-HR, SM-MIN2, height, sex, SM-HR3, SM-MIN1, age, SM-HR2, SM-HR1
Model 15	Weight, SM-MIN3, MX-HR, SM-MIN2, height, sex, SM-HR3, SM-MIN1, age, SM-HR2
Model 16	Weight, SM-MIN3, MX-HR, SM-MIN2, height, sex, SM-HR3, SM-MIN1, age
Model 17	Weight, SM-MIN3, MX-HR, SM-MIN2, height, sex, SM-HR3, SM-MIN1
Model 18	Weight, SM-MIN3, MX-HR, SM-MIN2, height, sex, SM-HR3
Model 19	Weight, SM-MIN3, MX-HR, SM-MIN2, height, sex
Model 20	Weight, SM-MIN3, MX-HR, SM-MIN2, height
Model 21	Weight, SM-MIN3, MX-HR, SM-MIN2
Model 22	Weight, SM-MIN3, MX-HR
Model 23	Weight, SM-MIN3
Model 24	Weight

Table 4. Overview of mRMR-based VO_2max models along with their predictor variables.

Models	Predictor variables
Model 25	SM-MIN3, weight, MX-RER, sex, SM-MIN1, MX-HR, SM-MIN2, age, SM-HR1, height, SM-HR2, SM-HR3
Model 26	SM-MIN3, weight, MX-RER, sex, SM-MIN1, MX-HR, SM-MIN2, age, SM-HR1, height, SM-HR2
Model 27	SM-MIN3, weight, MX-RER, sex, SM-MIN1, MX-HR, SM-MIN2, age, SM-HR1, height
Model 28	SM-MIN3, weight, MX-RER, sex, SM-MIN1, MX-HR, SM-MIN2, age, SM-HR1
Model 29	SM-MIN3, weight, MX-RER, sex, SM-MIN1, MX-HR, SM-MIN2, age
Model 30	SM-MIN3, weight, MX-RER, sex, SM-MIN1, MX-HR, SM-MIN2
Model 31	SM-MIN3, weight, MX-RER, sex, SM-MIN1, MX-HR
Model 32	SM-MIN3, weight, MX-RER, sex, SM-MIN1
Model 33	SM-MIN3, weight, MX-RER, sex
Model 34	SM-MIN3, weight, MX-RER
Model 35	SM-MIN3, weight
Model 36	SM-MIN3

Due to fact that SVM [18], in the literature, has been often reported to be superior to other machine learning methods especially in the field of sport physiology [1, 19, 20], it has been chosen to build the VO_2max prediction models. SVM constructs a hyperplane or set of hyperplanes in a high- or infinite-dimensional space to perform a regression analysis. The values of cost (C), epsilon (ϵ) for the ϵ-insensitive loss function, and the type of kernel function are the main parameters influencing the performance of SVM-based models. The radial

Table 5. Overview of MLFS-based VO_2max models along with their predictor variables.

Models	Predictor variables
Model 37	SM-MIN3, SM-MIN2, sex, SM-MIN1, weight, SM-HR3, MX-RER, MX-HR, age, SM-HR1, SM-HR2, height
Model 38	SM-MIN3, SM-MIN2, sex, SM-MIN1, weight, SM-HR3, MX-RER, MX-HR, age, SM-HR1, SM-HR2
Model 39	SM-MIN3, SM-MIN2, sex, SM-MIN1, weight, SM-HR3, MX-RER, MX-HR, age, SM-HR1
Model 40	SM-MIN3, SM-MIN2, sex, SM-MIN1, weight, SM-HR3, MX-RER, MX-HR, age
Model 41	SM-MIN3, SM-MIN2, sex, SM-MIN1, weight, SM-HR3, MX-RER, MX-HR
Model 42	SM-MIN3, SM-MIN2, sex, SM-MIN1, weight, SM-HR3, MX-RER
Model 43	SM-MIN3, SM-MIN2, sex, SM-MIN1, weight, SM-HR3
Model 44	SM-MIN3, SM-MIN2, sex, SM-MIN1, weight
Model 45	SM-MIN3, SM-MIN2, sex, SM-MIN1
Model 46	SM-MIN3, SM-MIN2, sex
Model 47	SM-MIN3, SM-MIN2
Model 48	SM-MIN3

Table 6. Overview of EFS-based VO_2max models along with their predictor variables.

Models	Predictor variables
Model 49	SM-MIN3, SM-HR1, sex, SM-MIN2, weight, SM-HR2, SM-MIN1, SM-HR3, age, MX-HR, height, MX-RER
Model 50	SM-MIN3, SM-HR1, sex, SM-MIN2, weight, SM-HR2, SM-MIN1, SM-HR3, age, MX-HR, height
Model 51	SM-MIN3, SM-HR1, sex, SM-MIN2, weight, SM-HR2, SM-MIN1, SM-HR3, age, MX-HR
Model 52	SM-MIN3, SM-HR1, sex, SM-MIN2, weight, SM-HR2, SM-MIN1, SM-HR3, age
Model 53	SM-MIN3, SM-HR1, sex, SM-MIN2, weight, SM-HR2, SM-MIN1, SM-HR3
Model 54	SM-MIN3, SM-HR1, sex, SM-MIN2, weight, SM-HR2, SM-MIN1
Model 55	SM-MIN3, SM-HR1, sex, SM-MIN2, weight, SM-HR2
Model 56	SM-MIN3, SM-HR1, sex, SM-MIN2, weight
Model 57	SM-MIN3, SM-HR1, sex, SM-MIN2
Model 58	SM-MIN3, SM-HR1, sex
Model 59	SM-MIN3, SM-HR1
Model 60	SM-MIN3

basis function (RBF) has been utilized as the kernel function, which requests the optimization of the function parameter gamma (γ). The optimal values of the three model parameters C, ϵ, and γ have been determined using the well-known grid search technique [21]. The grid search technique works by looking for values of every parameter using predetermined geometric steps across a search range, and the values of C, ϵ, and γ yielding the maximum prediction performance are selected. The lower and upper limit values of the search range for C

and γ have been chosen according to the recommendations made in [22]. Particularly, in [22], it was reported that trying exponentially growing sequences of C and γ is an effective way to determine the optimal values.

To draw a comparison with the performance of SVM, prediction models based on GRNN [23] and SDT [24] have also been developed. GRNN is a popular artificial neural network (ANN)-based method that does not necessitate an iterative training procedure like backpropagation networks. Thus, it is usually much faster to train a GRNN than other backpropagation networks such as MLP or RBFNN, and GRNNs have the potential to often exhibit more satisfactory prediction performance. For a GRNN-based model, the values of minimum and maximum values for sigma and the search step influence the performance of GRNN-based models. The Gaussian function has been chosen as the kernel function. Finally, for the category of tree-structured methods, preference has been given to SDT due to its simplicity, potential effectiveness in making decisions, and negligible training times, which often are in the order of milliseconds. The quality of an SDT-based model is determined by the values of minimum rows in a node, minimum size of node to split, and maximum tree levels. Table 7 and Table 8 show the ranges of values of the utilized parameters for the SVM-based models and the GRNN- and SDT-based models, respectively.

Table 7. Ranges of values of the utilized parameters for the SVM-based models.

Method	Parameter	Range
SVM	Cost (C)	$[2^{-2} - 2^{19}]$
	Epsilon (ϵ)	$[0.001 - 170]$
	Gamma (γ)	$[2^{-10}–2^{12}]$

Table 8. Ranges of values of the utilized parameters for the GRNN-based and SDT-based models.

Method	Parameter	Range
GRNN	Minimum sigma	[0.0001, 11]
	Maximum sigma	[1, 65]
	Search step	[1, 59]
SDT	Minimum rows in a node	[1, 20]
	Minimum size of node to split	[3, 10]
	Maximum tree levels	[5, 40]

The prediction errors of the models have been assessed by calculating the Rs and $RMSE$s, whose equations are given in Eq. (4) and Eq. (5), respectively:

$$R = \sqrt{1 - \frac{\sum_{i=1}^{n}(Y - Y')^2}{\sum_{i=1}^{n}\left(Y - \overline{Y}\right)^2}}, \tag{4}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Y - Y')^2}, \tag{5}$$

where Y and $Y^{'}$ represent the measured and predicted VO_2max values, respectively. $\bar{Y}$ gives the mean of the measured VO_2max values and N is the number of samples in a test subset. The evaluation of the generalization errors of the models, on the other hand, has been conducted using 10-fold cross-validation. Consequently, for each fold, the training data included 162 samples, while the test data included 19 samples. The cross-validation process is repeated 10 times using different sets of training and testing data, and the results from the folds have been averaged to increase the confidence in the prediction results and reduce the bias over a single fold.

5. Results and discussion

This section includes four subsections. In the first subsection, numerical results for *R*s and *RMSE*s of all MVFS-based, Relief-F-based, mRMR-based, MLFS-based, and EFS-based prediction models developed by using SVM, GRNN, and SDT are illustrated. The discussion regarding the results and relevant variables affecting the quality of the MVFS-based prediction models are presented in the second subsection. In the third subsection, the results of MVFS-based models are compared with the ones of models created by individually using the Relief-F, mRMR, and MLFS algorithms. Finally, in the fourth subsection, the results of MVFS-based models are compared with the ones of models created by using EFS, which, in contrast to the rest of individual feature selectors, falls into the same class of ensemble feature selectors as MVFS, thus enabling an additional direct comparison with the proposed approach.

5.1. Results

Table 9 presents the *R*s and *RMSE*s of MVFS-based VO_2max models developed by using SVM, GRNN, and SDT. For comparison purposes, Table 10 through Table 13 show the *R*s and *RMSE*s of VO_2max models created by using the relevance ranks of variables reported by Relief-F, mRMR, MLFS, and EFS algorithms, respectively.

Table 9. *R*s and *RMSE*s for MVFS-based VO_2max prediction models.

	SVM		GRNN		SDT	
Models	*R*	*RMSE*	*R*	*RMSE*	*R*	*RMSE*
Model 1	0.89	2.58	0.85	3.06	0.69	4.21
Model 2	0.90	2.55	0.86	3.04	0.71	4.16
Model 3	0.90	2.54	0.86	3.00	0.71	4.14
Model 4	0.90	2.53	0.86	2.98	0.71	4.11
Model 5	0.89	2.62	0.85	3.04	0.70	4.19
Model 6	0.89	2.63	0.85	3.11	0.70	4.19
Model 7	0.88	2.71	0.84	3.22	0.69	4.26
Model 8	0.87	2.83	0.82	3.29	0.66	4.43
Model 9	0.87	2.88	0.82	3.32	0.65	4.46
Model 10	0.82	3.29	0.77	3.76	0.60	4.84
Model 11	0.75	3.89	0.68	4.32	0.46	5.52
Model 12	0.74	3.92	0.67	4.39	0.45	5.59

5.2. Discussion

Based on the relevance ranks of predictor variables reported by the MVFS algorithm, it is seen that the submaximal variable SM-MIN3 is the most relevant predictor of VO_2max. In contrast, height and the submaximal variables SM-HR1 and SM-HR2 have been ranked within the last three orders, suggesting that they are rela-

Table 10. *R*s and *RMSE*s for Relief-F-based VO_2max prediction models.

	SVM		GRNN		SDT	
Models	*R*	*RMSE*	*R*	*RMSE*	*R*	*RMSE*
Model 13	0.89	2.58	0.85	3.06	0.69	4.21
Model 14	0.89	2.59	0.85	3.09	0.69	4.28
Model 15	0.90	2.54	0.85	3.03	0.69	4.23
Model 16	0.90	2.49	0.86	2.99	0.73	3.97
Model 17	0.89	2.61	0.84	3.13	0.67	4.34
Model 18	0.89	2.60	0.85	3.08	0.67	4.31
Model 19	0.85	3.03	0.79	3.57	0.58	4.75
Model 20	0.77	3.73	0.71	4.13	0.55	5.32
Model 21	0.77	3.78	0.70	4.19	0.54	5.40
Model 22	0.75	3.84	0.69	4.23	0.54	5.42
Model 23	0.75	3.89	0.68	4.32	0.46	5.52
Model 24	0.22	5.67	0.21	5.71	0.14	5.84

Table 11. *R*s and *RMSE*s for mRMR-based VO_2max prediction models.

	SVM		GRNN		SDT	
Models	*R*	*RMSE*	*R*	*RMSE*	*R*	*RMSE*
Model 25	0.89	2.58	0.85	3.06	0.69	4.21
Model 26	0.87	2.85	0.82	3.33	0.68	4.50
Model 27	0.87	2.90	0.82	3.35	0.67	4.57
Model 28	0.87	2.84	0.82	3.31	0.68	4.45
Model 29	0.86	2.97	0.82	3.47	0.66	4.68
Model 30	0.85	3.06	0.79	3.55	0.65	4.81
Model 31	0.87	2.88	0.82	3.36	0.67	4.46
Model 32	0.85	3.10	0.80	3.51	0.66	4.58
Model 33	0.85	3.02	0.81	3.47	0.67	4.49
Model 34	0.75	3.86	0.69	4.29	0.47	5.43
Model 35	0.75	3.89	0.68	4.32	0.46	5.52
Model 36	0.74	3.92	0.67	4.39	0.45	5.59

tively the most irrelevant predictors of VO_2max. Similar observations regarding the two submaximal variables also apply for Relief-F, mRMR, and MLFS, for which SM-MIN3 has been ranked within the first two orders, while SM-HR1 and SM-HR2 have consistently been ranked within the last four variables.

Among the set of MVFS-based prediction models, it is seen that Model 4 containing the physiological variables sex, age, and weight; the maximal variables MX-HR and MX-RER; and the submaximal variables SM-MIN1, SM-MIN2, SM-MIN3, and SM-HR3 yields the lowest *RMSE*s for prediction of VO_2max, no matter whether SVM, GRNN, or SDT has been employed for model development. In more detail, SVM has been found to deliver the lowest *RMSE* value with 2.53 mL kg^{-1} min^{-1}; building the same prediction model with

Table 12. *R*s and *RMSE*s for MLFS-based VO_2max prediction models.

	SVM		GRNN		SDT	
Models	*R*	*RMSE*	*R*	*RMSE*	*R*	*RMSE*
Model 37	0.89	2.58	0.85	3.06	0.69	4.21
Model 38	0.90	2.55	0.85	3.01	0.71	4.15
Model 39	0.90	2.54	0.86	3.00	0.71	4.14
Model 40	0.90	2.53	0.86	2.98	0.71	4.11
Model 41	0.89	2.62	0.85	3.04	0.70	4.19
Model 42	0.88	2.70	0.84	3.14	0.67	4.34
Model 43	0.88	2.71	0.84	3.14	0.67	4.37
Model 44	0.87	2.83	0.82	3.29	0.66	4.43
Model 45	0.80	3.51	0.73	3.98	0.57	5.17
Model 46	0.78	3.66	0.72	4.08	0.56	5.23
Model 47	0.76	3.81	0.71	4.16	0.55	5.36
Model 48	0.74	3.92	0.67	4.39	0.45	5.59

Table 13. *R*s and *RMSE*s for EFS-based VO_2max prediction models.

	SVM		GRNN		SDT	
Models	*R*	*RMSE*	*R*	*RMSE*	*R*	*RMSE*
Model 49	0.89	2.58	0.85	3.06	0.69	4.21
Model 50	0.89	2.59	0.85	3.09	0.69	4.28
Model 51	0.90	2.49	0.83	2.97	0.71	4.07
Model 52	0.89	2.58	0.85	3.02	0.70	4.12
Model 53	0.88	2.72	0.84	3.12	0.67	4.30
Model 54	0.87	2.82	0.83	3.21	0.66	4.42
Model 55	0.87	2.80	0.84	3.11	0.67	4.39
Model 56	0.87	2.78	0.84	3.10	0.70	4.17
Model 57	0.79	3.53	0.77	3.65	0.58	4.78
Model 58	0.58	4.74	0.56	4.80	0.30	4.92
Model 59	0.28	5.61	0.24	5.65	0.24	5.66
Model 60	0.74	3.92	0.67	4.39	0.45	5.59

GRNN gives the second most accurate prediction with an *RMSE* value of 2.98 mL kg^{-1} min^{-1}; and finally the development of the model with SDT occupies the last place with an *RMSE* value of 4.11 mL kg^{-1} min^{-1}. In contrast to Model 1 including the full set of predictor variables, the best performing model (i.e. Model 4) including 9 variables yields 1.93%, 2.61%, and 2.38% decrement rates in *RMSE*s for SVM, GRNN, and SDT, respectively. In other words, although Model 4 consists of relatively fewer predictor variables, it yields lower *RMSE*s compared to Model 1.

Among the set of regression methods, SVM unexceptionally exhibits the best performance by yielding the lowest *RMSE*s for all prediction models. Particularly, compared with the *RMSE*s obtained by GRNN-based and SDT-based models, the average percentage decrement rates in *RMSE*s obtained by SVM-based models are 13.63% and 35.50% for prediction of VO_2max, respectively. These results complement the previous studies in

[3–5], where SVM was also reported to be superior to MLP, RBFNN, DTF, TB, and MLR for prediction of VO_2max. GRNN-based models, in turn, outperform the SDT-based models, which comparatively show the worst performance for prediction of VO_2max.

Among the set of predictor variables, it turns out that in addition to SM-MIN3 ranked with the highest importance by the MVFS, the inclusion of the sex, age, MX-HR, and SM-HR3 variables in VO_2max models causes consistent reductions in *RMSE*s, independent of the utilized regression methods. In more detail, the outcomes show that inclusion of sex, age, MX-HR, and SM-HR3 variables into the respective SVM-based prediction models yields 12.46%, 3.43%, 4.24%, and 2.95% lower *RMSE*s, respectively. This observation also applies to the rest of the feature selectors and machine learning methods used in this study. Furthermore, leaving out the MX-RER variable in prediction models has a negligible effect on prediction accuracy, leading to slightly lower *RMSE*s. The rest of the predictor variables, on the other hand, can have an improving, negligible, or deteriorating effect on prediction of VO_2max, depending on with which other variables they are combined to form the prediction model.

As the individual inclusion of sex, age, MX-HR, and SM-HR3 variables in VO_2max models causes consistent reductions in *RMSE*s, it would be interesting to investigate how the special combination of these variables affects the prediction accuracy. To this end, we constructed and evaluated a new model using these variables in an exclusive way. The results show that the model produces *RMSE* values of 3.31 mL kg^{-1} min^{-1}, 3.35 mL kg^{-1} min^{-1}, and 3.86 mL kg^{-1} min^{-1} for SVM, GRNN, and SDT, respectively. Consequently, it can be concluded that although the variables individually have a consistent positive effect on prediction of VO_2max, combining these variables in an exclusive model does not outperform the *RMSE* of the SVM-based Model 4, which exhibits the lowest *RMSE* with 2.53 mL kg^{-1} min^{-1} for MVFS.

Prediction models developed by using the combination of physiological, maximal, and submaximal variables (i.e. Model 1 through Model 6) lead to lower *RMSE*s than the ones of models developed by using the physiological variables in combination with maximal or submaximal variables only.

The most accurate prediction models formed by the MVFS (i.e. Model 4) do not include the SM-HR1, SM-HR2, and height variables, which consistently have been ranked with the lowest three scores according to the MVFS algorithm.

Finally, it is observed that prediction models giving the lowest *RMSE*s for all four feature selectors include at least one predictor variable from each category of physiological, maximal, and submaximal variables. This, in turn, confirms the effectiveness of combining the relevant predictors of VO_2max from different categories in a hybrid model to improve the accuracy over regular models.

5.3. Results for comparing MVFS with individual feature selectors

In this subsection, the average results of MVFS-based models are compared with the average ones of Relief-F-based, mRMR-based, and MLFS-based models, and it is shown that the performance gain obtained by using the MVFS algorithm compared to individual feature selectors on prediction of VO_2max is statistically significant.

Table 14 shows the average *R*s and *RMSE*s of MVFS-based, Relief-F-based, mRMR-based, and MLFS-based models developed by using SVM, GRNN, and SDT. According to these results, the following comments can be made:

- Compared with the average *RMSE*s of the Relief-F-based models, the average *RMSE*s of the MVFS-based models are 11.13%, 9.28%, and 6.06% lower for SVM, GRNN, and SDT, respectively.

Table 14. Comparing the average results of MVFS-based models with the average ones of models based on individual feature selectors using SVM, GRNN, and SDT.

	MVFS		Relief-F		mRMR		MLFS	
	R	*RMSE*	*R*	*RMSE*	*R*	*RMSE*	*R*	*RMSE*
SVM	0.86	2.91	0.79	3.28	0.84	3.16	0.85	3.00
GRNN	0.81	3.37	0.74	3.71	0.78	3.62	0.80	3.44
SDT	0.64	4.51	0.58	4.80	0.62	4.77	0.64	4.61

- Compared with the average *RMSE*s of the mRMR-based models, the *RMSE*s of the MVFS-based models are 7.66%, 6.96%, and 5.57% lower for SVM, GRNN, and SDT, respectively.
- Compared with the average *RMSE*s of the MLFS-based models, the *RMSE*s of the MVFS-based models are 2.75%, 2.13%, and 2.15% lower for SVM, GRNN, and SDT, respectively.

Figure represents the average PDRs in *RMSE*s of VO_2max prediction for MVFS-based models compared to *RMSE*s obtained by Relief-F-based, mRMR-based, and MLFS-based models using SVM, GRNN, and SDT.

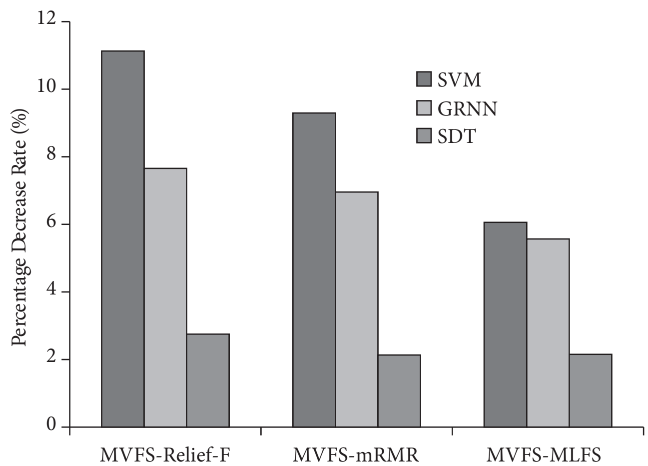

Figure Percentage decrease rates in *RMSE*s of VO_2max for MVFS-based models compared to *RMSE*s obtained by Relief-F-based, mRMR-based, and MLFS-based models using SVM, GRNN, and SDT.

The statistical significance of the performance gain obtained by using the MVFS algorithm compared to individual feature selectors on prediction of VO_2max has been determined using the well-known Wilcoxon signed-rank test [25]. More specifically, the test has been applied on the average *RMSE*s of all SVM-based, GRNN-based, and SDT-based models of the three pairs including (MVFS, Relief-F), (MVFS, mRMR), and (MVFS, MLFS). The sample size of the test case equals nine ($n = 9$), and the two-sided level of significance, i.e. α, is set to 0.05. The test statistic for the Wilcoxon signed-rank test is *W*, defined as the smaller of *W+* and *W-*, which are the sums of the positive and negative ranks, respectively. It is to be checked whether the observed test statistic W supports the null or research hypothesis. This check is performed using the critical value of *W*, which can be found using a predefined and well-known table of critical values. The calculated value of W in each case equals zero, and the critical value of W for $n = 9$ at $\alpha = 0.05$ is 5. Since W is less than the critical value, the null hypothesis is rejected and it can be concluded that the performance gain obtained by using the MVFS algorithm compared to individual feature selectors is statistically significant at $\alpha = 0.05$ for prediction of VO_2max.

5.4. Results for comparing MVFS with ensemble feature selector

In the literature, there exist many ensemble feature selectors exclusively designed for classification problems, making a direct comparison with our approach for regression problems unsuitable. After an exhaustive search of the related literature, we were able to only find one major study that proposed an alternative ensemble feature selector, i.e. EFS, which is also applicable for regression problems.

The rationale behind EFS can be summarized in three steps: (i) calculate the results of the n feature selection methods to be used; (ii) transform the results of each individual feature selector to a common scale (i.e. values between 0 and 1/n); and finally (iii) sum them up to a cumulative ranking (i.e. values between 0 and 1). Currently, EFS incorporates, among others, logistic regression (LogReg), the error-rate (ER)-based and the area under the curve (AUC)-based variable importance measures from the conditional random forest algorithm (ER-RF and AUC-RF, respectively), Pearson correlation (PC), and Spearman correlation (SC) that can be applied to regression problems, and the user can decide which feature selection methods should be used. There is no weighting among the methods, so to obtain the ensemble ranks of the features, the relevance scores for each individual feature reported by respective feature selectors are summed up.

There are three major differences between MVFS and EFS. First, in contrast to EFS, which normalizes the relevance ranks of variables to a common scale and makes a cumulative ranking, MVFS follows a three-step procedure, each of which evaluates any variable by taking into accounting (a) the majority voting principle; (b) the correlation among the variables; and finally (c) the average superiority of one feature selector over the others previously defined by the user. Secondly, MVFS is made up of Relief-F, mRMR, and MLFS, which are currently not included in EFS. Differently from the rest of the individual feature selectors implemented in EFS, Relief-F, mRMR, and MLFS have the merit that in previous studies [4, 5, 26] they have, in general, been shown to perform with satisfactory performance in creating accurate VO_2max models. Finally, in EFS there is no weighting among the feature selectors. In principle, MVFS also starts with no weighting among the methods; however, in case the rank decision of a variable cannot be made by the first two steps (i.e. after performing majority voting and correlation score evaluation), MVFS allows users to prioritize one method over the others for a variable selection in the last step.

To draw a comparison with the performance of MVFS, we constructed several ensemble feature selectors using different combinations of individual feature selectors implemented in EFS. With the help of our extensive experiments, the set of feature selectors leading to the most accurate results for prediction of VO_2max has been determined as the triple ensemble of LogReg, AUC-RF, and PC rankers.

Table 15 shows the average *R*s and *RMSE*s of MVFS-based and EFS-based models developed by using SVM, GRNN, and SDT. It is seen that considering the results obtained by the three utilized machine learning methods, MVFS, on average, achieves a PDR of 6.09% in terms of *RMSE* compared to the EFS algorithm. Particularly, MVFS-based models developed with SVM have comparatively been found to deliver the highest PDR with 10.69%. The second place is occupied by GRNN, which yields a PDR of 6.11%. Finally, the lowest PDR is obtained by SDT, which slightly improves the average accuracy by producing a PDR of 1.47%.

Finally, we compared and evaluated MVFS and EFS in terms of their run times. It is to be noted that the required run time strongly depends on the selection of individual feature selectors making up the ensemble rankers. For example, the run time required for MVFS on a commodity computer with Intel Duo Core 3.00 GhZ

Table 15. Comparing the average results of MVFS-based models with the average ones of EFS-based models using SVM, GRNN, and SDT.

	MVFS		EFS	
	R	$RMSE$	R	$RMSE$
SVM	0.86	2.91	0.78	3.26
GRNN	0.81	3.37	0.74	3.59
SDT	0.64	4.51	0.58	4.57

CPU and 4 GB memory depends on MLFS taking a relatively much longer time than Relief-F and mRMR, lasting approximately 3 min. Similarly, if EFS, consisting of LogReg, AUC-RF, and PC, is applied on the same dataset, the feature ranking process lasts about 6 min. Nonetheless, given the fact that such feature selection processes are executed only once for importance measures of variables before model design, execution times in such magnitudes can be neglected.

6. Conclusion

In this paper, it has been shown that the ensemble nature of the proposed feature selector improves the performance of model generation on average. Particularly, compared with the results of models created by individually using the Relief-F, mRMR, and MLFS algorithms, the proposed MVFS-based models resulted in better performance in terms of the average Rs and $RMSE$s achieved by SVM, GRNN, and SDT. The performance of the MVFS algorithm has also been compared with EFS, another alternative ensemble feature selector from the literature, which was constructed from the LogReg, AUC-RF, and PC rankers. Again, depending on the utilized machine learning method, the proposed hybrid approach shows its advantages by achieving up to 10.69% performance gain compared to the EFS algorithm for prediction of VO_2max. Furthermore, in addition to reconfirming the relevance of sex, age, and MX-HR in predicting VO_2max, which were previously reported in the literature, two additional variables, namely submaximal heart rates and exercise times at 1.5-mile distance, have also been found to have an improving effect on the performance of VO_2max prediction. Finally, regarding the ranking among the machine learning methods, SVM unexceptionally exhibits the best performance by yielding the highest Rs and the lowest $RMSE$s for all prediction models. GRNN-based models, in turn, occupy the second place in terms of performance regarding the Rs and $RMSE$s, whereas the SDT-based models comparatively show the worst performance for prediction of VO_2max.

References

[1] Abut F, Akay MF. Machine learning and statistical methods for the prediction of maximal oxygen uptake: recent advances. Medical Devices: Evidence and Research 2015; 8: 369-379. doi: 10.2147/MDER.S57281

[2] Chandrashekar G, Sahin F. A survey on feature selection methods. Computers and Electrical Engineering 2014; 40 (1): 16-28. doi: 10.1016/j.compeleceng.2013.11.024

[3] Aktürk E. Prediction of maximal oxygen uptake using machine learning methods combined with feature selection. MSc, Çukurova University, Adana, Turkey, 2014.

[4] Abut F, Akay MF, George JD. Support vector machines with individual combinations of Relief-F and mRMR for predicting VO_2max from maximal and questionnaire data. International Journal of Advances in Electronics and Computer Science 2016; 3 (10): 63-67.

[5] Abut F, Akay MF, George J. Developing new VO_2max prediction models from maximal, submaximal and questionnaire variables using support vector machines combined with feature selection. Computers in Biology and Medicine 2016; 79: 182-192. doi: 10.1016/j.compbiomed.2016.10.018

[6] Dietterich TG. Ensemble methods in machine learning. Multiple Classifier Systems 2000; 1857: 1–15. doi: 10.1007/3-540-45014-9_1

[7] Sarkar C, Cooley S, Srivastava J. Robust feature selection technique using rank aggregation. Applied Artificial Intelligence 2014; 28 (3): 243–257. doi: 10.1080/08839514.2014.883903

[8] Giacinto G, Roli F. An approach to the automatic design of multiple classifier systems. Pattern Recognition Letters 2001; 22 (1): 25–33. doi: 10.1016/S0167-8655(00)00096-9

[9] Rokach L. Ensemble methods for classifiers. In: Maimon O, Rokach L (editors). Data Mining and Knowledge Discovery Handbook. Boston, MA, USA: Springer, 2005, pp. 957–980.

[10] Xu J, Sun L, Gao Y, Xu T. An ensemble feature selection technique for cancer recognition. Biomedical Materials and Engineering 2014; 24 (1): 1001–1008. doi: doi: 10.3233/BME-130897

[11] Zhang Z, Yang P. An ensemble of classifiers with genetic algorithm based feature selection. IEEE Intelligent Informatics Bulletin 2008; 9 (1): 18–24.

[12] Neumann U, Genze N, Heider D. EFS: An ensemble feature selection tool implemented as R-package and web-application. BioData Mining 2017; 10 (1): 21. doi: 10.1186/s13040-017-0142-8

[13] Robnik-Šikonja M, Kononenko I. Theoretical and empirical analysis of relief-f and r-relief-f. Machine Learning 2003; 53 (1-2): 23-69. doi: 10.1023/A:1025667309714

[14] Peng H, Long F, Ding C. Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy. IEEE Transactions on Pattern Analysis and Machine Intelligence 2005; 27 (8): 1226-1238. doi: 10.1109/TPAMI.2005.159

[15] Suzuki T, Sugiyama M, Sese J, Kanamori T. Approximating mutual information by maximum likelihood density ratio estimation. In: JMLR Workshop and Conference Proceedings; Whistler, Canada; 2008. pp. 5-20.

[16] Lam L, Suen SY. Application of majority voting to pattern recognition: an analysis of its behavior and performance. IEEE Transactions on Systems, Man, and Cybernetics - Part A 1997; 27 (5): 553–568. doi: 10.1109/3468.618255

[17] Zhang Z, Yang P. An ensemble of classifiers with genetic algorithm based feature selection. IEEE Intelligent Informatics Bulletin 2008; 9 (1): 18-24.

[18] Hearst MA. Support vector machines. IEEE Intelligent Systems 1998; 13 (4): 18-28. doi: 10.1109/5254.708428

[19] Akay MF, Abut F, Daneshvar S, Heil D. Prediction of upper body power of cross-country skiers using support vector machines. Arabian Journal for Science and Engineering 2015; 40 (4): 1045–1055. doi: 10.1007/s13369-015-1588-y

[20] Acikkar M, Akay MF, Ozgunen KT, Aydin K, Kurdak SS. Support vector machines for aerobic fitness prediction of athletes. Expert Systems with Applications 2009; 36 (2): 3596–3602. doi: 10.1016/j.eswa.2008.02.002

[21] Friedrichs F, Igel C. Evolutionary tuning of multiple SVM parameters. Neurocomputing 2005; 64: 107–117. doi: 10.1016/j.neucom.2004.11.022

[22] Hsu CW, Chang CC, Lin CJ. A Practical Guide to Support Vector Classification. Technical Report. Taipei, Taiwan: Department of Computer Science, National Taiwan University, 2003.

[23] Specht DF. A general regression neural network. IEEE Transactions on Neural Networks 1991; 2 (6): 568–576. doi: 10.1109/72.97934

[24] Ali J, Khan R, Ahmad N, Maqsood I. Random forests and decision trees. International Journal of Computer Science Issues 2012; 9 (5): 272–278.

[25] Taheri SM, Hesamian G. A generalization of the Wilcoxon signed-rank test and its applications. Statistical Papers 2012; 54 (2): 457–470. doi: 10.1007/s00362-012-0443-4

[26] Abut F, Akay MF, George J. Maximum-likelihood feature selection for prediction of maximal oxygen uptake using support vector machines. In: International Conference on Educational Research; Kyrenia, North Cyprus; 2017. p. 48.