Measurement 167 (2021) 108173

Contents lists available at ScienceDirect

Measurement

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

Estimation of body fat percentage using hybrid machine learning algorithms

Muhammed Kürşad Uçar [a,*], Zeliha Uçar [b], Fatih Köksal [c], Nihat Daldal [d]

[a] Sakarya University, Faculty of Engineering, Electrical-Electronics Engineering, 54187, Serdivan, Sakarya, Turkey
[b] Istanbul Okan University, Institute of Health Sciences, Nutrition and Dietetics, Mecidiyekoy, Istanbul, Turkey
[c] Bursa Higher Specialization Education and Research Hospital, 16310, Yildirim, Bursa, Turkey
[d] Department of Electrical and Electronics Engineering, Faculty of Engineering, Bolu Abant Izzet Baysal University, 14280 Bolu, Turkey

A R T I C L E I N F O

Article history:
Received 20 March 2020
Received in revised form 3 June 2020
Accepted 29 June 2020
Available online 24 July 2020

Keywords:
Body composition
Body fat percentage calculation
Body fat percentage estimation
Machine learning
Artificial intelligence

A B S T R A C T

Before obesity treatment, body fat percentage (BFP) should be determined. BFP cannot be measured by weighing. The devices developed to produce solutions to this problem are called "Body Analysis Devices". These devices are very costly. Therefore, more practical and cost-effective solutions are needed. This study aims to determine BFP using hybrid machine learning methods with high accuracy rate and minimum parameter. This study uses real data sets, which are 13 anthropometric measurements of individuals. Different feature groups were created with feature selection algorithm. In the next step, 4 different hybrid models were created by using MLFFNN, SVMs, and DT regression models. According to the results, BFP of individuals can be estimated with a correlation value of $R = 0.79$ with one anthropometric measurement. The results show that the developed system can be used to estimate BFP in practice. Besides, the system can calculate BFP with just one anthropometric measurement without device requirement.

1. Introduction

The human body consists of extracellular fluid, bone, fat and muscle cells. The combination of these four groups at a perfect rate achieves a perfect balance of body composition. Fat tissue is divided into white and brown. Brown fat tissue is a tissue that is mostly produced for the production of heat in newborn babies. White fat tissue is the largest energy store. White fat tissues serve for storage of excess energy and for immediate use. Fat tissue mainly consists fat cells called adipocytes, in addition to fibroblasts, preadipocytes, macrophages [1,2].

White fat tissue constitutes 15–20% of the body. It is found in two different places, under the skin and around the internal organs. Brown fat tissue is about 4% of newborn body weight [3]. Fat tissue is considered an endocrine organ with known heat and energy regulation functions as well as secretory feature [1].

Adipose tissue procures many secretion products for body needs and is vital. For example, TNF-α causes insulin resistance in obese people, which if not treated leads to diabetes [2]. Obesity is a disease in which excess nutrients are stored as fat cells and increased in volume, decreasing the quality of life, causing diabetes and cardiovascular diseases [2]. In addition, obesity can lead to psychological disturbances by adversely affecting life and quality of life. It is known that obesity is also effective in the formation of musculoskeletal, respiratory and digestive problems [4].

In obese, the most important stage of treatment is the reduction of excess fat tissue. The major cause of insulin resistance is accumulated fats. Fats are not just energy tanks. They secrete some hormone components and insulin resistance, high blood pressure, type 2 diabetes, unstable cholesterol, cardiovascular disease causes problems such as [3]. Therefore, maintaining a certain balance of body fat is essential for a healthy life. Evaluation of body composition is the most important stage in obesity. The body composition consists of body fat mass (kg) and lean body mass (kg). Lean body mass is composed of bones, muscle and body fluid. Evaluation of body composition is used primarily in areas such as obesity, sports sciences, public health related areas, cardiology, nephrology [2–4]. With all this information is being utilized, it is essential to determine the body composition components with high impeccability, the minimum need for knowledge and least cost systems.

The devices developed to produce solutions to this problem are called "Body Analysis Devices". These devices generally use the Bioelectrical Impedance Analysis (BIA) method [5,6]. X-ray absorptiometry methods are also used for body fat percentage

* Corresponding author.
E-mail address: mucar@sakarya.edu.tr (M.K. Uçar).

https://doi.org/10.1016/j.measurement.2020.108173
0263-2241/

(BFP) measurements. However, this method is relatively expensive and not common for often clinical use [6]. Evaluation of body composition can be done by direct and indirect measurements [7,8]. Direct measurement methods are not suitable for application in the living human body. However, it can be applied on cadavers. Indirect measurement methods can be applied in the living human body. Many methods have been developed in this field [8–12].

Anthropometric measurements are one of the indirect measurements [5]. In this method, height, weight, skin fold thickness, diameter and circumference of various parts of the body are used. The skinfold caliper is used to measure the skinfold thickness. The use of this meter requires considerable knowledge. In order to prevent incorrect measurements, it is necessary to take measurements at certain points of the body. Therefore, it is not possible for a person without technical knowledge to make an accurate calculation with this meter.

It is discovered that when the methods are compared, there are still problems in need of processing. The most significant disadvantages of the best efficient work methods are high costs and technical staff need.

Staff with technical knowledge is required in the Skinfold method, which contains manual tools without technology [7]. Personnel expected to be well educated and highly experienced with measurements. Technical staff begin to measure accurately after completing approximately 1000 measurements therefore educating a technical personnel is highly laboursome.

There are many studies in the literature for estimating BFP [9,13–17]. A large part of these studies are statistical based approaches and tried to find solutions to the problem by producing linear and nonlinear equations specific to the problem. They are simple models and it is very difficult to work decisively. Anthropometric measurements are generally used for equations. In addition, several studies have produced a solution to the problem by machine learning method [9,13]. However, these studies developed a single model and did not carry out detailed analysis. Anthropometric measurements are also used in machine learning based studies. However, the system stability has increased considerably because the method is different [9,13].

Calculating the skinfold thickness and body fat percentage is accepted as the standard method in this field [18,19]. Skinfold thickness and the body fat percentages calculated using BIA (Tanita TBF300) method were compared in a study examined by Kaner and his friends. The results of the study determined that BIA and skin fold thickness method can be used interchangeably.

Several equations have been developed with error rate performance of %2,5 for men and %2.65 for women in a study where equations were developed based on anthropometric measurements [11]. Machine learning methods present a preferable performance than simple regression equations. The fact that general equations can be used to estimate the percentage of body fat is the biggest indication that it can perform very well in machine learning studies. In addition, there are many studies that focus on equation extraction in the literature [10–12].

Body mass index, waist height ratio and abdominal volume index were found to be very effective in calculating body fat percentage in a study that determines which anthropometric parameters are more successful in calculating the percentage of body fat [8]. It also been reported that waist height ratio, body mass index and waist circumference is related to body fat percentage [7,20].

It is determined that body mass index was more strongly associated with cardiometabolic risk than skin fold thickness in a study dated 2017 [21]. There are many studies where body mass index is associated with body fat index [18,19,22,23].

The aim of the study is to calculate BFP which is the most important component of body composition by hybrid machine learning based methods. A real data set was used for this. Anthropometric measurements are available in the database. The working flow diagram is being given in Fig. 1. The machine learning-based BFP was calculated using anthropometric measurements with machine learning. In order to improve system performance, some anthropometric measurements were reduced by using feature selection algorithm and system performance was evaluated.

This study, when compared with other examinations in the litterateur, leads significantly. These can be listed as; (1) Whereas statistic base designs have been taken place often, Hybrid artificial intelligence based design was used in this study. (2) In the litterateur, equations are predominantly used for BFP calculations. The same process is performed with artificial intelligence without the need for an equation in this study. (3) Anthropometric measurements taken from individuals for BFP calculation are used in litterateur. Feature selection algorithms were not used. Features related to the feature selection algorithm was selected and used in this study. Thanks to this advantage, the cost of measurementcan be reduced. This study gains innovation to the litterateur when all these advantages are evaluated.

According to the results, BFP of individuals can be estimated with a correlation value of $R = 0.788$ with one anthropometric measurement. The results show that the developed system can be used to estimate BFP in practice. Besides, the system can calculate BFP with just one anthropometric measurement without device requirement.

2. Material and method

The study is organized according to the flow diagram in Fig. 1. Firstly, the data were selected with feature selection algorithms. As a second step, predictive models were created. All features groups were analyzed with all models. In the feature selection stage, the features are grouped according to their correlation values. In the machine learning phase, hybrid models were created with different machine learning models. All machine learning algorithms are used for each feature group.

2.1. Data collection

In the study, a real database was used for estimating body fat percentage [24]. The data set includes anthropometric measurements and body fat percentage values of 252 individuals. Table 1 summarizes the data collected from individuals. The goal here is to calculate the Y Body Fat Percentage using the minimum X variable. Body fat percentage measurements were obtained using Siri Equations [25].

Statistical information about the data is given in Table 2. Property numbers are numbered 1 for X1, 2 for X2, and 3 for X3.

Table 3 gives the size of the relationship between features and body fat percentage (Y). The relationship is in correlation coefficient.

2.2. Feature selection/sort algorithm

Sperman correlation coefficients and Principal Component Analysis (PCA) methods were used as feature selection algorithm in this study.

2.2.1. Spearman correlation coefficient

Spearman Correlation Coefficient was used as feature selection algorithm. Correlation coefficients are measures that point out information about the power and direction of the relationship between variables. Pearson and Spearman correlation coefficients

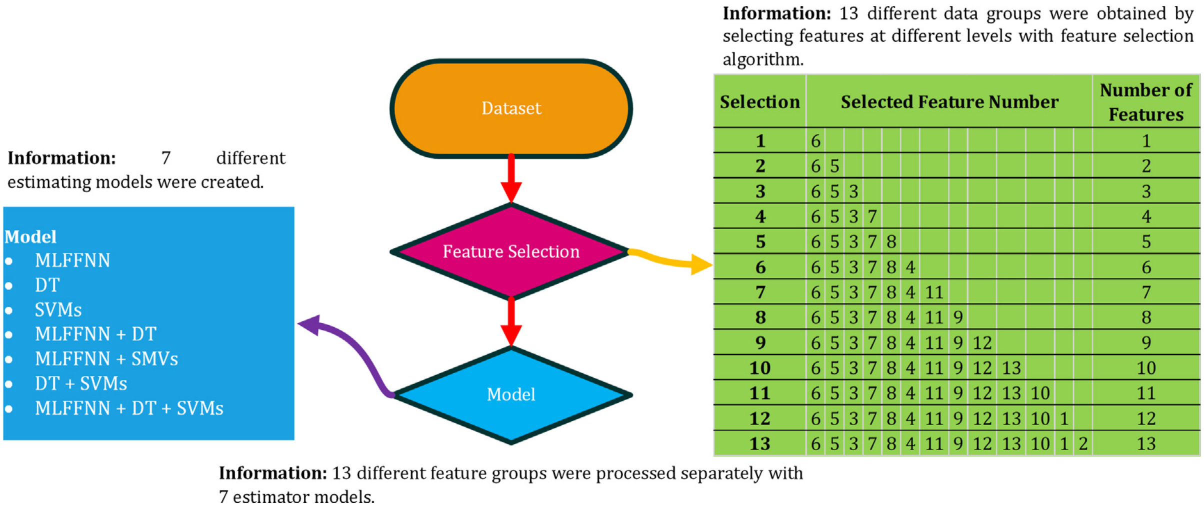

Fig. 1. Flow diagram.

Table 1
Features.

Symbol	Feature Name)
X1	Age (Years)
X2	Height (cm)
X3	Weight (kg)
X4	Neck Circumference (cm)
X5	Chest Circumference (cm)
X6	Abdomen 2 Circumference (cm)
X7	Hip Circumference (cm)
X8	Tigh Circumference (cm)
X9	Knee Circumference (cm)
X10	Ankle Circumference (cm)
X11	Biceps (Extended) Circumference (cm)
X12	Forearm Circumference (cm)
X13	Wrist Circumference (cm)
Y	Body Fat Percentage Siri Equation
D	Density determined from underwater weighing

are used in the case of variables to be examined are numerical variables [26].

It is stated as r_s for spearman correlation coefficient, $d_i$for difference between observation order numbers, $n$for observation order number and r_s for Eq. (1) [26].

$$r_s = 1 - 6\sum_{i=1}^{n} \frac{{d_i}^2}{n \times (n^2 - 1)} \tag{1}$$

If $n > 30$ for statistical decision, $t$value is calculated for the significance of r_s (Eq. (2)). The obtained t statistic is compared with the selected $alpha$($alpha = 0.05$ for this study) $n - 2$ degree of freedom t statistic. In the case of $t < 0.05, r_s$ is meaningful and usable [26].

$$t = \frac{r_s}{\sqrt{\frac{1-{r_s}^2}{n-2}}} \tag{2}$$

Table 3
Relationship between features and body fat percentage.

Feature Number	Relationship Level with BFP
6	0.813
5	0.669
3	0.605
7	0.602
8	0.531
4	0.491
11	0.481
9	0.476
12	0.384
13	0.311
10	0.285
1	0.271
2	−0.015

BFP: Body Fat Percentage.

Table 2
Distribution of data.

Feature Number	Min	Max	Mean	Std	95% CI		R	R^2
					LB	UB		
1	22.00	81.00	44.92	12.67	43.33	46.50	0.27	0.073
2	162.56	196.85	178.59	6.51	177.77	179.40	−0.01	0.000
3	56.70	164.72	81.24	13.15	79.59	82.88	0.61	0.366
4	31.10	51.20	38.02	2.41	37.72	38.32	0.49	0.241
5	83.40	136.20	100.88	8.32	99.84	101.92	0.67	0.448
6	70.40	148.10	92.65	10.68	91.31	93.98	0.81	0.660
7	85.30	147.70	99.91	7.04	99.03	100.80	0.60	0.362
8	49.30	87.30	59.44	5.15	58.79	60.08	0.53	0.281
9	33.00	49.10	38.60	2.37	38.30	38.89	0.48	0.226
10	19.10	33.90	23.12	1.70	22.90	23.33	0.29	0.081
11	25.30	45.00	32.31	2.99	31.94	32.69	0.48	0.231
12	21.00	34.90	28.68	2.01	28.43	28.93	0.38	0.148
13	15.80	21.40	18.24	0.92	18.12	18.35	0.31	0.097

Min Minumum, Max Maximum, Std Standard Deviation, CI 95% - %95 Confidence interval for the mean, LB Lower confidence interval, UB Upper confidence interval, RSpearman Coefficient of Correlation, R^2 Spearman square.

Table 4
Selected features.

Selection	Selected Features Number													Number of Features
1	6													1
2	6	5												2
3	6	5	3											3
4	6	5	3	7										4
5	6	5	3	7	8									5
6	6	5	3	7	8	4								6
7	6	5	3	7	8	4	11							7
8	6	5	3	7	8	4	11	9						8
9	6	5	3	7	8	4	11	9	12					9
10	6	5	3	7	8	4	11	9	12	13				10
11	6	5	3	7	8	4	11	9	12	13	10			11
12	6	5	3	7	8	4	11	9	12	13	10	1		12
13	6	5	3	7	8	4	11	9	12	13	10	1	2	13

The relationship given in Table 3 represents the correlation coefficient values. The data are sorted by relationship level (Table 3). 13 different feature groups were created by selecting 1 feature at a time (Table 4).

2.2.2. Principal component analysis

In general, PCA is a predominantly used method for determining the components/dimensions that compose the total, reducing the number of variables, and ordering the observations objectively [26].

This study was calculated with the principal components as a result of PCA analysis (Table 5). 4 principal components are to explain 95% of all the data regarding obtained parameters therefore, these components were obtained from 13 features in the data set and analyzes were fulfilled.

2.3. Machine learning

In the study, Multilayer Feedforward Neural Networks (MLFFNN), Support Vector Machine Regression Model (SVMs) and Decision Tree Regression (DT) were used as the machine learning algorithm. Each method was used as a stand-alone predictive model and was constructed in hybrid structures (Table 6).

Table 5
Principal component analysis results.

Principal Component No	Voefficient	Cumulative Sum
1	**62.96**	**62.96**
2	**26.25**	**89.21**
3	**5.68**	**94.89**
4	**1.54**	**96.43**
5	1.28	97.71
6	0.57	98.28
7	0.46	98.74
8	0.35	99.09
9	0.28	99.37
10	0.25	99.62
11	0.18	99.81
12	0.16	99.96
13	0.04	100.00

Table 6
Individual and hybrid models.

Individual Model		Hybrid Model	
1	MLFFNN	**1**	MLFFNN + DT
2	SVMs	**2**	MLFFNN + SMVs
3	DT	**3**	DT + SVMs
		4	MLFFNN + DT + SVMs

The reason of the selection these methods can explained as follows: Educational period is short and the accuracy rate is high in the three method [27–30]. Therefore, DT, MLFFNN and SVMs are frequently used methods for many studies in litterateur [31–33]. In addition, DT, MLFFNN and SVM's have a wide application areas in embedded systems [34–36].

Single models were used to create hybrid models. For example, when the MLFFNN and DT hybrid models are formed, a new system is designed by taking the common decision of one or two systems. For example, if MLFFNN gives 10 as the output value and 20 as the DT output value for 1 data, the MLFFNN + DT hybrid model will give (10 + 20)/2 = 15. In other words, the average of the outputs is calculated and accepted as the output of the new hybrid model. After all data were processed in this way, the performance evaluation criteria were recalculated according to the reference values.

For machine learning operations, 252 data are divided into two parts: 70% Training, 30% Test sets.

2.3.1. Multilayer feedforward neural networks

Artificial neural networks are computer architectures created by mathematical modeling of the nerve structures of the human brain [28,37,38]. In this structure, information is transmitted through one-way channels. Usually contains input, hidden and output layer (Fig. 2).

MLFFNN has modifiable parameters. The education algorithm for thisparticular study is "Levenberg–Marquardt". The new structure was created for eachneuron by changing the number of neurons between 1–100. The best performingneuron structure was used in the study.

2.3.2. Support vector machine regression model

SVMs is a machine learning algorithm used in both classification and regression problems [39]. Algorithm aims to fit the curve

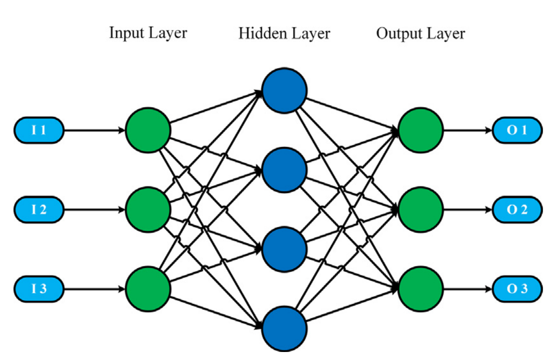

Fig. 2. MLFFNN general network structure.

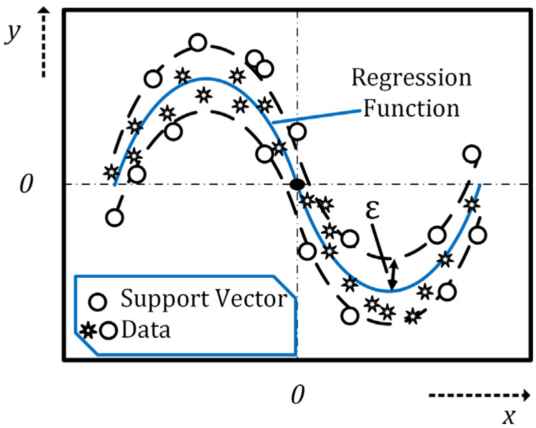

Fig. 3. SVMs general model.

in hyperplane so that it passes over the dataset. It takes advantages of support vectors. Maximum distance between the support vectors is established and a curve is fitted between. This curve is acclaimed as the generalized solution of the dataset (Fig. 3).

2.3.3. Decision tree regression

Decision tree algorithms fundamental consist of root, branch, knot and leaves. Each attribute is associated with node when building the tree structure. Between root and node there are branches. It is passed from each node to the other one through branches. Decision in the tree made according to the final leaf [33]. The fundamental logic in creating decision tree structure can be summarized as asking related questions to each node reached, and reaching the final leaf over the branches in the shortest path and time according to the answers given. Thus, it creates decision rules/models regarding to the answers obtained from the questions (see Fig 4).

2.4. Performance evaluation criteria

The implementation process of the performance criterias are as follows: 70% of the current dataset is created using machine learning algorithms. 30% of the data set is asked to the created model. Performance criterias calculated as $y$for the responds of the model and $t$for the real responds. The performance of the models developed in this study was evaluated with seven performance evaluation criteria. These include Mean Absolute Percentage Error (*MAPE*), Root Mean Square Error (*RMSE*), Mean Absolute Difference (*MAD*), Standard Error (*SH*), Correlation Coefficient (*R*), Explanatory Coefficient (R^2) and MSE.

2.4.1. Raw residues (e_i)

Raw Residues (e_i) is the difference between Actual values (t_i) and estimated values (y_i) (Eq. (3)). e_i is requested to be close to zero [26].

$$e_i = t_i - y_i \tag{3}$$

2.4.2. Standard error (SH)

Standard Error (*SH*) indicates the compatibility of the developed method with data [26]. When the correlation is less than 1, the system cannot predict with 100% accuracy. In this case, deviations from the actual values (e_i) occur. The standard error of the developed system *SH*is the standard deviation of these deviations. As e_i decreases, *SH* decreses. System reliability increases as *SH*decreases [26].

$$SH = \sqrt{\frac{\sum_{i=1}^{n}(t_i - y_i)^2}{n-2}} = \sqrt{\frac{\sum_{i=1}^{n}e_i^2}{n-2}} \tag{4}$$

2.4.3. Mean square error

*MSE*refers to the average of the squares of errors (Eq. (5)). e_i in the equation represents errors and is calculated according to Eq. (3).

$$MSE = \frac{1}{n}\sum_{i=1}^{n}e_i^2 \tag{5}$$

2.4.4. Root mean square error

*RMSE*represents the square root of *MSE*. The error rate decreases as *MSE*and *RMSE*approaches 0.

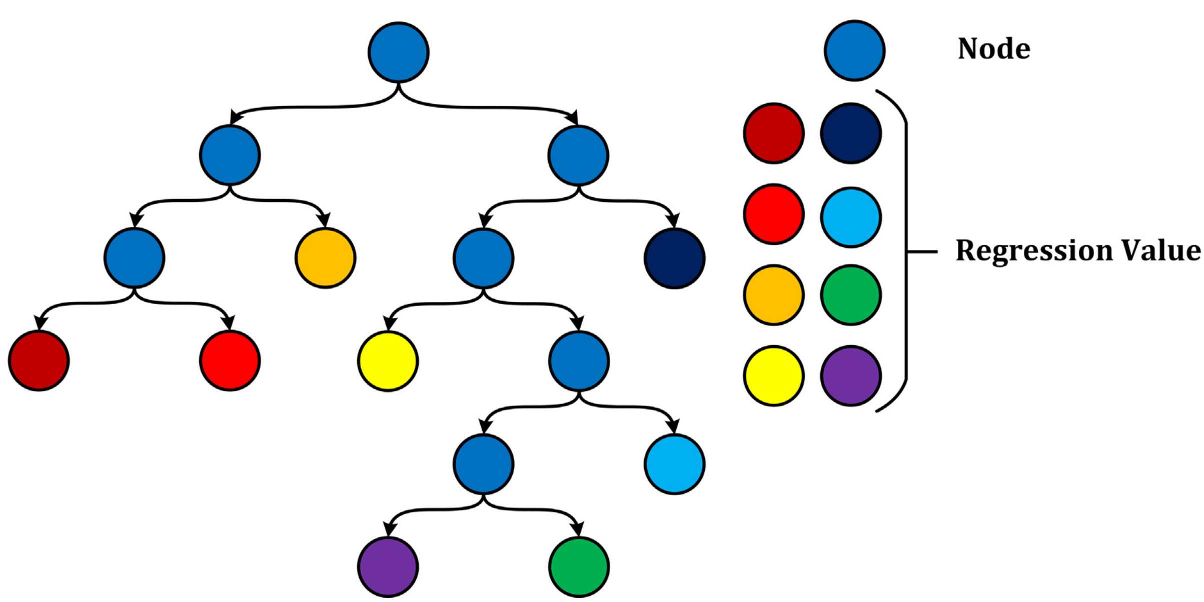

Fig. 4. DT general model.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n} e_i^2} \tag{6}$$

2.4.5. Error rate

Error Rate is the percentage indication of the variation between the actual value and the estimated value (Eq. (7)). In MAD, the equation is used similar to MAPE in 8 and is expected to be close to 0.

$$MAPE = \frac{1}{n}\sum_{i=1}^{n} \frac{|t_i - y_i|}{t_i} \times 100 \tag{7}$$

$$MAD = \frac{1}{n}\sum_{i=1}^{n} |t_i - y_i| \tag{8}$$

2.4.6. Correlation coefficients

Correlation Coefficients are calculated according to the variable types such as the correlation coefficient calculated in the classification process. Since the predicted variables are continuous variables, Pearson (r) or Sperman (r_s) correlation coefficient formulas are used. When selecting these two calculation types, it is checked whether the data is normally distributed or not. If the data is normally distributed, Pearson will be used. If it is not normally distributed, the non-parametric equivalent of the Pearson correlation coefficient, Sperman Correlation Coefficient (r_s) will be used.

2.4.7. Relation between correlation and estimation

If the correlation between the two variables is $|r| < 0.70$, the estimation error rate of the system will be quite high. If $0.5 < |r| < 0.70$, the estimation of the system is low, if $0.7 < |r| < 0.90$, the estimation of the system is medium, and $0.9 < |r|$, the estimation of the system is high [26].

2.4.8. Explanatory coefficient (R^2)

Defines the share of the change that can be explained by the regression model as a percentage of the total change (Eq. (9)) [26]. In other words, what percentage of the total change in the dependent variable R^2 can be explained by the independent variable.

$$R^2 = r^2 = \frac{KT_R}{KT_Y} \tag{9}$$

R^2 descriptive equations: Actual values (t_i), Estimated values (y_i), Number of data (n), Sum of Squares T (KT_T, Sum of Squares R (KT_R), Sum of Squares A (KT_A).

$$KT_T = KT_R + KT_A \tag{10a}$$

$$KT_T = \sum_{i=1}^{n} (t_i - \bar{t})^2 = \sum_{i=1}^{n} t_i^2 - \frac{\left(\sum_{i=1}^{n} t_i\right)^2}{n} \tag{10b}$$

$$KT_R = \sum_{i=1}^{n} (y_i - \bar{t})^2 \tag{10c}$$

$$KT_A = \sum_{i=1}^{n} (t_i - y_i)^2 \tag{10d}$$

Table 7
Performance evaluation 1/5.

	Feature Selection Level 1						
Feature Numbers	**6**						
Model	**Performance Evaluation Criteria**						
	MAPE	***RMSE***	***MAD***	***SH***	***R***	R^2	***MSE***
MLFFNN	58.14	12.57	6.633	12.75	0.709	0.503	158.1
DT	29.85	4.608	3.699	4.672	0.771	0.595	21.23
SVMs	31.63	4.526	3.652	4.588	0.789	0.623	20.48
MLFFNN + DT	42.58	7.474	4.971	7.577	0.729	0.531	55.86
MLFFNN + SMVs	42.94	7.446	4.901	7.549	0.739	0.547	55.44
DT + SVMs	30.22	4.482	3.592	4.544	**0.79**	0.624	20.09
MLFFNN + DT + SVMs	82.4	12.53	10.76	12.71	0.75	0.562	157.1
	Feature Selection Level 2						
Feature Numbers	**6-5**						
Model	**Performance Evaluation Criteria**						
	MAPE	***RMSE***	***MAD***	***SH***	***R***	R^2	***MSE***
MLFFNN	46.91	8.854	6.162	8.976	0.405	0.164	78.4
DT	31.4	4.845	3.933	4.912	0.733	0.538	23.48
SVMs	30.55	4.264	3.417	4.323	0.801	0.641	18.18
MLFFNN + DT	34.67	5.854	4.358	5.935	0.632	0.399	34.27
MLFFNN + SMVs	34.83	5.617	4.195	5.694	0.679	0.461	31.55
DT + SVMs	30.06	4.39	3.537	4.451	**0.802**	0.643	19.27
MLFFNN + DT + SVMs	70.88	9.926	8.739	10.06	0.734	0.538	98.53
	Feature Selection Level 3						
Feature Numbers	**6-5-3**						
Model	**Performance Evaluation Criteria**						
	MAPE	***RMSE***	***MAD***	***SH***	***R***	R^2	***MSE***
MLFFNN	33.85	5.64	4.227	5.717	0.737	0.543	31.8
DT	29.82	4.829	3.905	4.895	0.749	0.561	23.32
SVMs	32.43	4.546	3.674	4.609	0.773	0.597	20.67
MLFFNN + DT	28.83	4.631	3.587	4.695	0.783	0.613	21.45
MLFFNN + SMVs	31.27	4.757	3.641	4.823	0.77	0.593	22.63
DT + SVMs	30.32	4.46	3.652	4.522	0.781	0.61	19.89
MLFFNN + DT + SVMs	74.51	11.34	9.909	11.5	**0.784**	0.615	128.6

In addition to these parameters, the Blant-Altman graph was prepared for 4 models considered to be the best model. Blant-Altman, When a new method against reference is developed with the same units, the Bland–Altman graphical approach allows to look at the subject from a different perspective when examining whether a new method which is thought to measure the same feature makes similar measurements with classical meter or when the compatibility between the evaluators is examined.

2.5. Statistical analysis

The relationship between anthropometric measurements used for the calculation of body fat percentage and body fat percentage was statistically analyzed. Spearman correlation coefficients were used for the examination. Correlation values were agreed as significant If $t < 0.05$. (Section 2.2.1).

3. Results

The aim of the study is to calculate the percentage of body fat by using minimum data, high accuracy rate with machine learning methods. A real database was used for this purpose. The database contains 13 characteristics and percentage of body fat. With the feature selection/sorting algorithm, features are sorted according to level of relevance. After sorting, the data set is divided into 13 different data sets. In this study, 7 different machine learning structures were used. MLFFNN, DT and SVMs are singular estimators. In addition, MLFFNN + DT, MLFFNN + SVMs, DT + SVMs and MLFFNN + DT + SVMs structures were created to estimate body fat percentages. Performance evaluation criteria were calculated after each model. In addition to these analyzes, the relationship between anthropometric measurements and BFP was statistically analyzed. Data reduction was performed with PCA analysis and regression analysis was repeated.

Prediction performance varies by feature selection level (Tables 7–11). At the first feature selection levels (Selection level 1–6) the overall performance of the models has increased. However, there have been some reductions in subsequent election levels.

At the same feature selection level, machine learning models can give different performance (Table 7, Selection level 13). This can affect the performance of hybrid models both positively and negatively (Table 7 - Selection level 1, Table 9 - Selection level 9).

PCA analysis was also examined in order to reduce the number of features (Table 12). PCA reduces features. Four principal components were obtained using PCA. It has been determined that the model with the best performance compared to the 4 principal components is MLFFNN. On the side, MLFFNN + performed well on SVMs. Compared to machine learning methods, PCA is lower than the performance value (DT) achieved by a single feature (Table 7). Despite, PCA present high performance compared to the machine learning methods.

In order to make the tables more understandable, a summary graph in Fig. 5 has been prepared. The graph shows the performance evaluation criteria results for each model and selection level. $R$and R^2(RR) values are expected to be close to 1, while other performance values are expected to be close to 0. It is healthier to consider different parameters together in performance evaluation.

At the 1st feature selection level, the best model for $R$is DT + SVMs, while the best model for *MAPE*is DT (Table 7). Therefore, it

Table 8
Performance evaluation 2/5.

Feature Numbers / Model	*MAPE*	*RMSE*	*MAD*	*SH*	*R*	R^2	*MSE*
Feature Selection Level 4							
Feature Numbers: 6-5-3-7							
MLFFNN	35.17	5.43	4.31	5.505	0.762	0.581	29.49
DT	30.92	4.958	4.009	5.027	0.729	0.531	24.58
SVMs	31.46	4.466	3.592	4.528	0.78	0.608	19.94
MLFFNN + DT	30.79	4.7	3.731	4.765	0.78	0.608	22.09
MLFFNN + SMVs	32.84	4.789	3.839	4.855	0.778	0.605	22.93
DT + SVMs	30.36	4.483	3.654	4.545	0.78	0.609	20.09
MLFFNN + DT + SVMs	79.35	11.91	10.7	12.07	**0.788**	0.62	141.7
Feature Selection Level 5							
Feature Numbers: 6-5-3-7-8							
MLFFNN	43.43	6.869	5.485	6.964	0.608	0.37	47.18
DT	31.84	5.002	4.04	5.071	0.717	0.515	25.02
SVMs	31.93	4.451	3.59	4.513	**0.782**	0.612	19.81
MLFFNN + DT	34.97	5.359	4.249	5.433	0.7	0.49	28.72
MLFFNN + SMVs	35.52	5.25	4.194	5.323	0.714	0.51	27.57
DT + SVMs	31.22	4.553	3.706	4.616	0.772	0.596	20.73
MLFFNN + DT + SVMs	80.74	12.16	10.7	12.33	0.733	0.537	147.9
Feature Selection Level 6							
Feature Numbers: 6-5-3-7-8-4							
MLFFNN	41.11	6.39	5.026	6.479	0.692	0.479	40.84
DT	30.93	4.908	4.111	4.976	0.742	0.55	24.09
SVMs	31.57	4.472	3.596	4.534	0.778	0.605	20
MLFFNN + DT	32.3	4.958	3.842	5.026	0.754	0.569	24.58
MLFFNN + SMVs	33.73	4.941	3.904	5.009	0.749	0.561	24.41
DT + SVMs	30.32	4.453	3.689	4.514	0.78	0.609	19.83
MLFFNN + DT + SVMs	80.64	12.15	10.85	12.31	**0.785**	0.616	147.6

Performance Evaluation Criteria

Table 9
Performance evaluation 3/5.

	Feature Selection Level 7						
Feature Numbers	**6-5-3-7-8-4-11**						
Model	Performance Evaluation Criteria						
	MAPE	***RMSE***	***MAD***	***SH***	***R***	R^2	***MSE***
MLFFNN	57.22	9.18	6.277	9.306	0.535	0.287	84.27
DT	31.58	4.835	4.086	4.901	0.738	0.545	23.37
SVMs	32.01	4.463	3.602	4.525	0.781	0.61	19.92
MLFFNN + DT	40.82	5.969	4.488	6.051	0.668	0.446	35.62
MLFFNN + SMVs	41.71	6.092	4.44	6.176	0.662	0.438	37.11
DT + SVMs	30.57	4.399	3.636	4.46	**0.783**	0.613	19.35
MLFFNN + DT + SVMs	85.58	12.38	10.8	12.55	0.711	0.506	153.3
	Feature Selection Level 8						
Feature Numbers	**6-5-3-7-8-4-11-9**						
Model	Performance Evaluation Criteria						
	MAPE	***RMSE***	***MAD***	***SH***	***R***	R^2	***MSE***
MLFFNN	79.36	13.34	9.586	13.52	0.381	0.145	178
DT	31.74	4.919	4.146	4.987	0.734	0.538	24.2
SVMs	31.14	4.503	3.598	4.565	0.77	0.592	20.27
MLFFNN + DT	49.53	7.629	5.727	7.734	0.611	0.373	58.21
MLFFNN + SMVs	49.62	7.551	5.65	7.655	0.622	0.386	57.02
DT + SVMs	30.34	4.466	3.675	4.527	**0.775**	0.6	19.94
MLFFNN + DT + SVMs	89.86	13.82	11.57	14.01	0.694	0.482	190.9
	Feature Selection Level 9						
Feature Numbers	**6-5-3-7-8-4-11-9-12**						
Model	Performance Evaluation Criteria						
	MAPE	***RMSE***	***MAD***	***SH***	***R***	R^2	***MSE***
MLFFNN	50.19	8.982	6.174	9.106	0.499	0.249	80.67
DT	30.38	4.864	4.059	4.931	0.746	0.557	23.65
SVMs	31.18	4.513	3.609	4.575	0.77	0.593	20.37
MLFFNN + DT	36.71	6.209	4.685	6.294	0.622	0.387	38.55
MLFFNN + SMVs	37.16	6.041	4.54	6.124	0.66	0.435	36.49
DT + SVMs	29.7	4.432	3.656	4.493	**0.78**	0.608	19.64
MLFFNN + DT + SVMs	78.3	12.32	10.35	12.49	0.696	0.485	151.8

will be healthier to determine the best model by considering many parameters (see Fig 6).

As a result of the developed models, it has been found that feature selection levels 3–7-10 and 12 have high predictive performance with MLFFNN. According to the Blant-Altman Graph of these models, the highest level of fit is 12th. Although there are random errors at other levels, the developed system is practical.

Statistical relationship between body fat percentage and anthropometric measurements was examined in the study. Best 9 Correlational values were shown in graphs at the end of the examination (Fig. 7). In the evaluation, $t < 0.05$ was determined for all correlation coefficients. Correlation values change between $0.38424 < R < 0.8126$. These values are an indication that the data are related to body fat percentage.

In Table 13, the highest performing equations for some studies in the litterateur are being shown. These equations are generally developed on a gender basis.

4. Discussion

In the previous section, detailed analyzes of machine learning based models are presented. In this section, the performance of the analyzes based on real data will be compared with the literature.

The number of publications for determining fat percentage is going up with the increase in academic work opportunities and specialists working in these fields. [46–50]. Calculating the percentage of body fat for the body composition is rough and pricy therefore, efficient and low-cost methods are needed [51–53]. If the developed methods are practically usable, a considerable economic expansion can occur. The reliability of measurements without the need for the device is going to save a period of time and make money.

This study is to calculate BFP with the highest accuracy rate with minimum anthropometric measurements and hybrid machine learning algorithms. Studies in the litterateur are mostly statistics based [40–42]. While developing models, regression model is created with the whole data set. However, it is not possible to test the model developed in this case. When the complete data set is used in education, the accuracy rate ($R$and R^2) is quite high and provides misleading information. The data sets used in modeling and testing are given. The model is not created with all data. Therefore, although the accuracy rate seems lower than in the litterateur, it is more credible.

In the hybrid method (DT + SVMs) proposed in this study, the performance value of $MAD = 3.592$ was obtained by 1 anthropometric measurement. In the study using the same data in the literature, the proposed hybrid system performance was determined as $MAD = 3.6974$ [9]. At this level, 13 anthropometric properties were used. The most important aspect of the study is to calculate the same performance with less process load with feature selection algorithms. The ratio of training and test data set used in machine learning is the same in both studies. However, it is not known which data are used in education in the literature [54–56].

All anthropometric measurements were used in the same study [9]. In the methods proposed in this study, hybrid models with different feature levels were created and their performance was evaluated. In this regard, the study includes more comprehensive analyzes. The best model proposed by Shoa was able to reach

Table 10
Performance evaluation 4/5.

	Feature Selection Level 10						
Feature Numbers	**6-5-3-7-8-4-11-9-12-13**						
Model	**Performance Evaluation Criteria**						
	MAPE	***RMSE***	***MAD***	***SH***	***R***	R^2	***MSE***
MLFFNN	51.93	8.136	6.736	8.248	0.707	0.5	66.19
DT	33.04	4.907	3.949	4.975	0.703	0.494	24.08
SVMs	36.12	4.756	3.884	4.821	0.76	0.578	22.62
MLFFNN + DT	39.67	5.892	4.826	5.974	0.736	0.541	34.72
MLFFNN + SMVs	40.62	5.812	4.677	5.892	0.734	0.539	33.78
DT + SVMs	33.63	4.618	3.76	4.682	**0.765**	0.585	21.33
MLFFNN + DT + SVMs	92.55	13.72	12.39	13.91	0.753	0.568	188.3
	Feature Selection Level 11						
Feature Numbers	**6-5-3-7-8-4-11-9-12-13-10**						
Model	**Performance Evaluation Criteria**						
	MAPE	***RMSE***	***MAD***	***SH***	***R***	R^2	***MSE***
MLFFNN	66.21	9.466	7.412	9.597	0.268	0.072	89.61
DT	31.35	4.936	4.147	5.005	0.736	0.541	24.37
SVMs	32.12	4.635	3.683	4.699	0.76	0.578	21.49
MLFFNN + DT	43.55	6.005	4.844	6.088	0.588	0.345	36.06
MLFFNN + SMVs	45.8	6.132	4.894	6.217	0.556	0.309	37.6
DT + SVMs	30.44	4.515	3.718	4.577	**0.769**	0.592	20.39
MLFFNN + DT + SVMs	84	11.91	10.31	12.08	0.696	0.485	141.9
	Feature Selection Level 12						
Feature Numbers	**6-5-3-7-8-4-11-9-12-13-10-1**						
Model	**Performance Evaluation Criteria**						
	MAPE	***RMSE***	***MAD***	***SH***	***R***	R^2	***MSE***
MLFFNN	74.03	14.75	8.265	14.95	0.488	0.238	217.4
DT	29.42	4.818	3.833	4.884	0.756	0.571	23.21
SVMs	27.69	4.272	3.415	4.331	0.787	0.619	18.25
MLFFNN + DT	47.61	8.498	5.407	8.616	0.668	0.447	72.22
MLFFNN + SMVs	47.61	8.491	5.307	8.608	0.663	0.44	72.09
DT + SVMs	27.11	4.293	3.413	4.352	**0.793**	0.629	18.43
MLFFNN + DT + SVMs	83.24	13.76	10.77	13.95	0.711	0.505	189.3

Table 11
Performance evaluation 5/5.

	Feature Selection Level 13						
Feature Numbers	**6-5-3-7-8-4-11-9-12-13-10-1-2**						
	Performance Evaluation Criteria						
Model	***MAPE***	***RMSE***	***MAD***	***SH***	***R***	R^2	***MSE***
MLFFNN	58.5	8.535	6.751	8.653	0.451	0.203	72.85
DT	33.51	5.092	4.013	5.162	0.69	0.477	25.93
SVMs	41.46	5.32	4.31	5.393	0.702	0.493	28.3
MLFFNN + DT	39.51	5.832	4.599	5.912	0.656	0.43	34.01
MLFFNN + SMVs	39.92	5.662	4.478	5.74	0.631	0.398	32.05
DT + SVMs	35.58	4.817	3.836	4.883	**0.74**	0.547	23.2
MLFFNN + DT + SVMs	78.42	11.33	9.584	11.48	0.717	0.514	128.3

Table 12
Regression results for 4 principal components (PCA).

	PCA						
Test	**4 temel bileşen**						
Model	**Performance Evalutaion Criteria**						
	MAPE	***RMSE***	***MAD***	***SH***	***R***	R^2	***MSE***
MLFFNN	30.08	4.87	3.87	4.93	**0.73**	0.54	23.69
DT	39.23	6.32	4.84	6.41	0.60	0.36	39.99
SVMs	36.50	5.16	4.17	5.23	0.69	0.47	26.61
MLFFNN + DT	32.28	5.13	3.92	5.21	0.71	0.50	26.36
MLFFNN + SMVs	31.87	4.83	3.85	4.90	**0.72**	0.52	23.36
DT + SVMs	35.65	5.43	4.12	5.51	0.68	0.46	29.52
MLFFNN + DT + SVMs	76.69	11.34	9.82	11.50	0.71	0.51	128.70

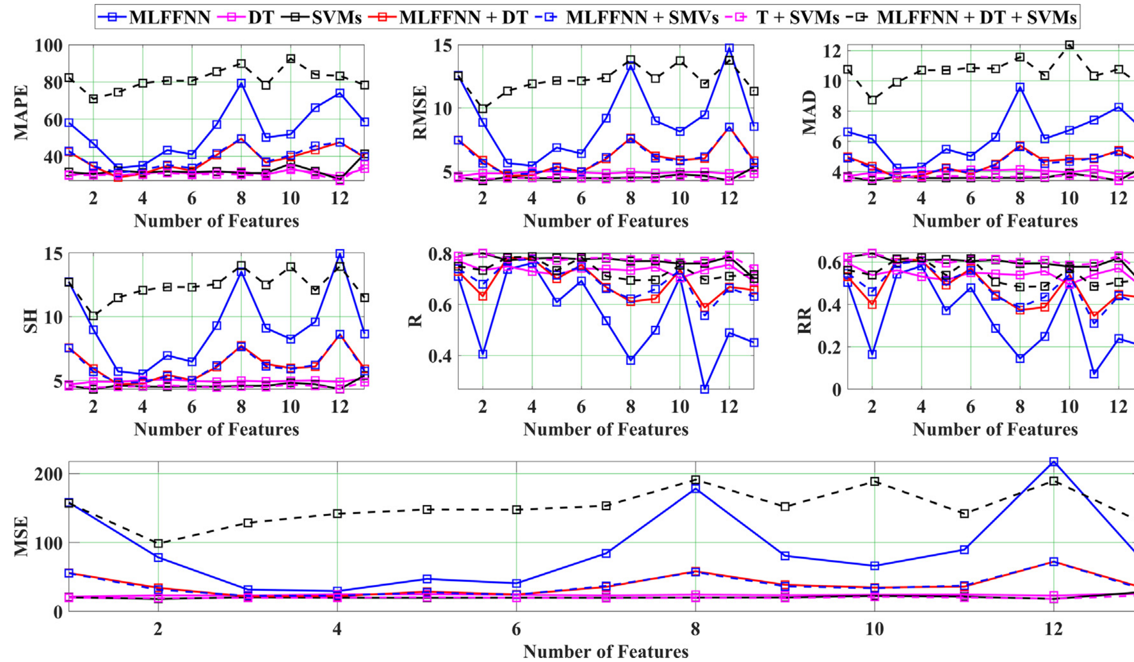

Fig. 5. Summary graphical performance evaluation.

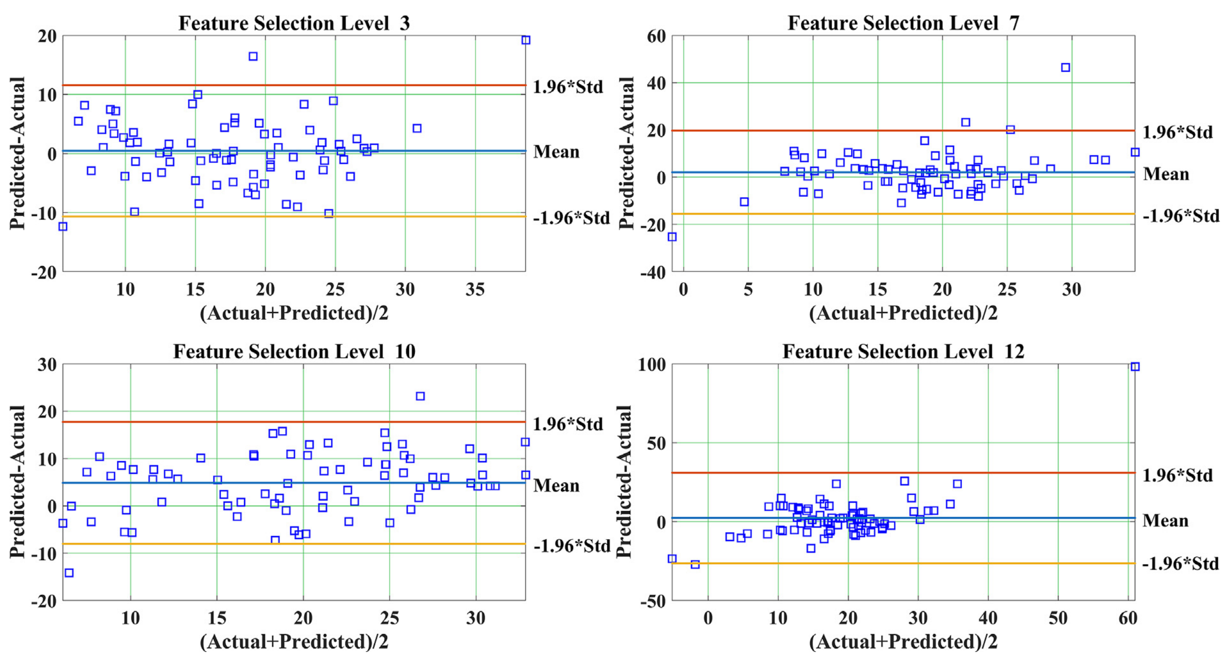

Fig. 6. Blant Altman chart for MLFFNN performance evaluation.

$RMSE = 4.6384$ with 13 anthropometric measurements. In the model proposed in this article, following performance values obtained, $RMSE = 0.482$ with 1 feature and DT + SVMs model, $RMSE = 4.39$ with 2 features and DT + SVMs model, $RMSE = 4.46$ with 3 features and DT + SVMs model. The performance values obtained at each level are better than the best performance value of the hybrid model proposed by Shao [9].

Apart from anthropometric measurements, there are many studies in the literature that parameters such as age, body mass index (BMI), height and weight are used in body fat percentage [57]. The system performance in these studies has been reported to be between $R^2 = \%60 - 80$ [58–62]. However, these are linear or nonlinear statistics based systems. Not machine learning based systems.

According to the studies in the litterateur, it was revealed that there are different parameters that are effective in the calculation of BPF. The most obvious parameter is gender [51,63–65]. However, in some studies, the common equation for men and women has been drawn [40,50]. The data set used in this study does not contain gender information. Therefore, gender-based analysis could not be performed. This deficiency of the study can be eliminated in future studies.

The maximum Robtained in this study is about 0.8. This ratio is an acceptable value in medical systems [26].

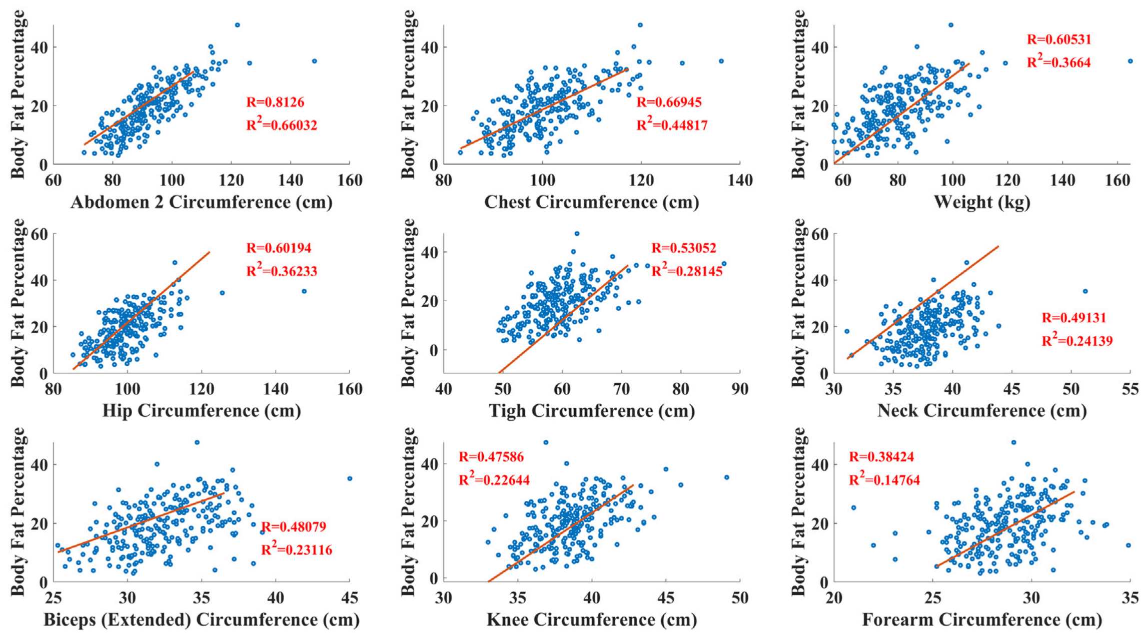

Fig. 7. Statistical analysis of the relationship between body fat percentage and anthropometric measurements.

Table 13
Some equations with the best performance.

Ref	Gender	Equation	R	R^2
[40]	E/K	$\%BFP = -5 \times 10^{-5} \times (Age)^2 + 0.1523 \times (Age) - 20.732$	0.9993	0.9986
[41]	K	$BF/Height^2 (kg/m^2) = -9.3 + 0.714 \times (BMI)$	0.9985	0.9970
[41]	E	$BF/Height^2 (kg/m^2) = -10.5 + 0.642 \times (BMI)$	0.9960	0.9920
[42]	E	$\%BFP = 0.360 \times (WC) + 0.221 \times (HC) - 17.502 \times (WHR) - 0.136 \times (Xc) - 0.198 \times (Ht2/R)$	0.9910	0.9820
[43]	E	$\%BFP = -0.00258 \times (\sum 3SF)^2 + 0.558 \times \sum 3SF + 0.118 \times (Age) + 0.282 \times WC - 2.100 \times FB\text{- } 2.34$	0.9695	0.9400
[44]	E	$\%BFP = 0.465 + 0.180 \times (\sum S7SF) - 0.0002406 \times (S7SF)^2 + 0.06619 \times (Age)$	0.9487	0.9000
[45]	K	$\%BFP = -4.054 + 0.16 \times (Triceps) + 0.154 \times IC + 0.281$ $\times (Biceps) + 0.263 \times (HC) + 0.229 \times ST - 3.249 \times HBB + 0.517 \times RAC + 0.125 \times (WC)$	0.9354	0.8750

Abbreviations
% BFP Body Fat Percentage (%)
BF Body Fat (kg)
WC Waist Circumference (cm)
HC Hip Circumference (cm)
WHR Waist-to-Hip Ratio.
Xc Reactance
Ht^2/R Dividing the length squares by resistance
FB Femur Breadth
BMI Body Mass Index.
HBB Humerus Bone Breadth
ST Skinfold Thickness
IC Iliac Crest
RAC Relaxed Arm Circumference
$\sum 3SF$ Triceps, subscapular, bicep widths total
$\sum 7SF$ Chest, midaxillary, triceps, subscapular, abdominal, suprailiac and thigh sum
Ref Referance.

The Blant-Altman graph of the proposed systems shows that there are random errors. This is an indication that the system is functioning [26].

In PCA analysis to decrease the number of features, the performance rate has increased compared to other data groups. However, all anthropometric measurements must be prepared for PCA analysis. In this case the measurement costs rise. In the analysis made by selecting a feature, the measurement cost has been reduced since analysis with certain features has been made. The performance value obtained by an anthropometric measurement is better compared to PCA analysis. Therefore, the system can be implemented without PCA analysis.

According to the results obtained in the study, 1 anthropometric measurement can estimate the percentage of body fat with the proposed hybrid models at a rate of approximately $R = \%0.788$. These results indicate that the model is very practical in practice.

5. Conclusion

Determination of BPF is very significant in the treatment of obesity. Various methods have been developed to measure PF. Although the gold standard method is Dual Energy X-ray Absorptiometry Measurement, it is very laborious and expensive. Different methods are available in the litterateur which can be seen as a solution of this problem. The first of these methods are statistical based demographic information and regression methods based on anthropometric measurements. However, more practical solutions are needed because anthropometric measurements take time and statistical methods are troublesome. Alternatively, there is a need for systems based on artificial intelligence that can only calculate BPF with some anthropometric measurements.

In this study, hybrid machine learning based system was developed by using 13 anthropometric measurements. As the result of the analysis, a system which allow BPF estimation with anthropometric measurement of only "Abdomen 2 Circumference (cm)" has been processed. There are many benefits in the proposed hybrid model. These can be listed as; (1) Reducing the cost of measurement, (2) Reducing the loss of time, (3) Eliminating the cost of the device, (4) Reducing the economic cost..

The average price of professional devices for BPF measurement is on average $15000. Since the devices are not able to measure remotely, it is essential to locate them in places to be measured. However, with a simple and anthropometric measurement based on hybrid machine learning software, BPF can be measured with high accrurancy rate without the need for this cost and time loss. The correlation value obtained in this study is approximately 0.8. This value shows the usability of the proposed model and is compatible with the studies in the litterateur.

6. Future works

This study can be extended in the future to more data (feature), feature selection algorithms and machine learning algorithms can be repeated. The database can be expanded. In the field of health, it may be advantageous to work with large data. There are very few studies in this field in Turkey. Large collection of data for developing models based on Turkish people can make an important contribution to Turkey.

Future studies can be listed as follows.

- Further collection of features such as anthropometric
- Extend work with different machine learning algorithms
- Developing less and higher performance systems with different feature selection algorithms
- Gender based BPF calculation

CRediT authorship contribution statement

Muhammed Kürşad Uçar: Conceptualization, Methodology, Software, Writing - original draft, Writing - review & editing, Supervision. **Zeliha Uçar:** Data curation, Resources. **Fatih Köksal:** Validation, Investigation, Visualization. **Nihat Daldal:** Validation, Formal analysis, Investigation.

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.

References

[1] M. Uçar, Z. Uçar, Body fat amount/percentage calculation methods: systematic review, J. Inst. Sci. Technol. 10 (2) (2020) 930–943, https://doi.org/10.21597/jist.650481.

[2] A. Sümer, Topiramatin Yag Hücresi Farklilasmasi ve Bazi Karbonik Anhidraz Izoenzimleri Gen Ekspresyonu Üzerine Etkisinin In Vitro Incelenmesi, Ph.D. thesis, Karadeniz Teknik Üniversitesi, 2014.

[3] I.G. POLAT, Effect of Er Stress and Sik2 Reciprocal Relationship on Human Precursor Fat Cell (LiSa-2) Differentiation, Ph.D. thesis, Gebze Technical University, 2017.

[4] M.N. Akgül, The effect of 6-weeks competition period training on body composition of boxers, Ph.D. thesis, Selçuk University - Health Sciences Institute, 2016.

[5] G. Kaner, G. Pekcan, G. Pamuk, B.Ö. Pamuk, Skinfold thickness versus bioimpedance analysis: body fat prediction in adults, J. Nutr. Dietet. 43 (2) (2015) 111–118, URL https://beslenmevediyetdergisi.org/index.php/bdd/article/view/134v.

[6] J. Ravindranath, P.P. Pillai, S. Parameswaran, S.K. Kamalanathan, G.K. Pal, Body Fat analysis in predialysis chronic kidney disease: multifrequency bioimpedance assay and anthropometry compared with dual-Energy X-ray absorptiometry, J. Renal Nutr. 26 (5) (2016) 315–319, https://doi.org/10.1053/j.jrn.2016.04.002.

[7] L. Grigollo, A. Pelegrini, E.L. Petroski, D.A.S. Silva, J.M.F. de Lima Silva, Anthropometric indicators of obesity in the prediction of high body fat in adolescents, Revista Paulista de Pediatria (English Edition) 33 (1) (2015) 56–62, https://doi.org/10.1016/s2359-3482(15)30031-2.

[8] E. Ehrampoush, P. Arasteh, R. Homayounfar, M. Cheraghpour, M. Alipour, M.M. Naghizadeh, M. Hadibarhaghtalab, S.H. Davoodi, A. Askari, J.M. Razaz, New anthropometric indices or old ones: Which is the better predictor of body fat?, Diab Metabolic Syndrome: Clin. Res. Rev. 11 (4) (2017) 257–263, https://doi.org/10.1016/j.dsx.2016.08.027.

[9] Y.E. Shao, Body fat percentage prediction using intelligent hybrid approaches, Sci. World J. 2014 (2014) 383910. doi:10.1155/2014/383910. URL http://www.ncbi.nlm.nih.gov/pubmed/24723804 http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3958757.

[10] E. Sukić, A. Katić, E. Stokić, A. Kupusinac, O. Rankov, What kind of relationship is between body mass index and body fat percentage?, J. Med. Syst. 41(1) (2016). doi:10.1007/s10916-016-0636-9.

[11] L. Ortiz-Hernández, A.V. Vega López, N. Ramos-Ibáñez, L.J. Cázares Lara, R.J. Medina Gómez, D. Pérez-Salgado, Equations based on anthropometry to predict body fat measured by absorptiometry in schoolchildren and adolescents, J. Pediatria 93 (4) (2017) 365–373, https://doi.org/10.1016/j.jped.2016.08.008.

[12] D. Gallagher, S.B. Heymsfield, M. Heo, S.A. Jebb, P.R. Murgatroyd, Y. Sakamoto, Healthy percentage body fat ranges: an approach for developing guidelines based on body mass index, Am. J. Clin. Nutr. 72 (3) (2000) 694–701, https://doi.org/10.1093/ajcn/72.3.694.

[13] T. Ferenci, L. Kovács, Predicting body fat percentage from anthropometric and laboratory measurements using artificial neural networks, Appl. Soft Comput. 67 (2018) 834–839, https://doi.org/10.1016/J.ASOC.2017.05.063, URL https://www.sciencedirect.com/science/article/pii/S1568494617303411.

[14] J. Hastuti, M. Kagawa, N.M. Byrne, A.P. Hills, Anthropometry to assess body fat in Indonesian adults, Asia Pacific J. Clin. Nutr. 27 (3) (2018) 592–598, https://doi.org/10.6133/apjcn.092017.02, URL http://www.ncbi.nlm.nih.gov/pubmed/29737806.

[15] A.C.C. Salamunes, A.M.W. Stadnik, E.B. Neves, Estimation of female body fat percentage based on body circumferences, Revista Brasileira de Medicina do Esporte 24 (2) (2018) 97–101, https://doi.org/10.1590/1517-869220182402181175, URL http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1517-86922018000200097&lng=en&tlng=en.

[16] J.C. Aristizabal, A. Estrada Restrepo, A. Giraldo García, Development and validation of anthropometric equations to estimate body composition in adult women, Colombia Médica 49 (2) (2018) 154–159. doi:10.25100/cm.v49i2.3643. URL http://www.ncbi.nlm.nih.gov/pubmed/30104807 http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC6084924 http://colombiamedica.univalle.edu.co/index.php/comedica/article/view/3643.

[17] C.J. Henry, S. D/O Ponnalagu, X. Bi, S.-Y. Tan, New equations to predict body fat in asian-chinese adults using age, height, skinfold thickness, and waist circumference, J. Acad. Nutr. Dietet. 118 (7) (2018) 1263–1269, https://doi.org/10.1016/J.JAND.2018.02.019, URL https://www.sciencedirect.com/science/article/pii/S2212267218302703.

[18] R. Arabaci, S. Erden, N. Hasil Korkmaz, C. Çankaya, F. Korkmaz, Relationship between physical activity, nutritin habits and body composition of university students, Nigde Univ. J. Phys. Educ. Sport Sci. 6 (3) (2012) 234–243.

[19] A.T. Ergür, T. Onarlioglu, K. Marakoglu, H. Sezer, B. Suzen, A. Öktem, Comparison of body composition parameters in children and adolescents, using skinfold and bioelectrical impedance methods, Turkish J. Pediatr. Dis. 6 (3) (2012) 133–138.

[20] K.M. Flegal, J.A. Shepherd, A.C. Looker, B.I. Graubard, L.G. Borrud, C.L. Ogden, T.B. Harris, J.E. Everhart, N. Schenker, Comparisons of percentage body fat, body mass index, waist circumference, and waist-stature ratio in adults, Am. J. Clin. Nutr. 89 (3) (2009) 500–508, https://doi.org/10.3945/ajcn.2008.26847.INTRODUCTION.

[21] A.F. Dias, V.B. Lemes, A.R. Gaya, A.C.A. Gaya, J. Mota, C. Brand, Obesity anthropometric indicators associated with cardiometabolic risk in Portuguese children and adolescents, Preventive Med. Rep. 8 (September) (2017) 158–162, https://doi.org/10.1016/j.pmedr.2017.10.002.

[22] O. Akar, A. Genç, S. Çelik, M. Ünlü, K. Üçok, T. Koyuncu, Ü. Sener, Association analyses of oxidative stress, aerobic capacity, daily physical activity, and body composition parameters in patients with mild to moderate COPD, Turkish J. Med. Sci. 44 (2014) 972–979, https://doi.org/10.3906/sag-1308-65.

[23] A. Kupusinac, E. Stokić, R. Doroslovački, Predicting body fat percentage based on gender, age and BMI by using artificial neural networks, Comput. Methods Programs Biomed. 113 (2014) 610–619, https://doi.org/10.1016/j.cmpb.2013.10.013.

[24] R.W. Johnson, Fitting percentage of body fat to simple body measurements, J. Stat. Educ. 4(1) (1996). doi:10.1080/10691898.1996.11910505. URL https://www.tandfonline.com/doi/full/10.1080/10691898.1996.11910505.
[25] William E. Siri, Body Composition From Fluid Spaces and Density: Analysis of Methods, U.S. Atomic Energy Commission (1956) 1–33. URL https://escholarship.org/uc/item/6mh9f4nf.
[26] R. Alpar, Applied Statistic and Validation - Reliability, Detay Publishing, 2010. URL https://books.google.com.tr/books/about/Uygulamali_istatistik_ve_gecerlik_guv.html?id=ITk1MwEACAAJ&pgis=1.
[27] R.U. Rasool, U. Ashraf, K. Ahmed, H. Wang, W. Rafique, Z. Anwar, Cyberpulse: A machine learning based link flooding attack mitigation system for software defined networks, IEEE Access 7 (2019) 34885–34899, https://doi.org/10.1109/ACCESS.2019.2904236.
[28] M. Uçar, M. Bozkurt, C. Bilgin, K. Polat, Automatic detection of respiratory arrests in OSA patients using PPG and machine learning techniques, Neural Comput. Appl. 28(10) (2017). doi:10.1007/s00521-016-2617-9.
[29] M.K. Uçar, M. Nour, H. Sindi, K. Polat, The effect of training and testing process on machine learning in biomedical datasets, Mathe. Probl. Eng. 2020 (2020) 1–17, https://doi.org/10.1155/2020/2836236.
[30] A. Ozdemir, K. Polat, Deep learning applications for hyperspectral imaging: a systematic review, J. Inst. Electron. Comput. 2 (1) (2020) 39–56, https://doi.org/10.33969/jiec.2020.21004.
[31] I.B. Aydilek, A. Arslan, A hybrid method for imputation of missing values using optimized fuzzy c-means with support vector regression and a genetic algorithm, Inf. Sci. 233 (2013) 25–35, https://doi.org/10.1016/j.ins.2013.01.021.
[32] I.B. Aydilek, A. Arslan, A novel hybrid approach to estimating missing values in databases using K-nearest neighbors and neural networks, Int. J. Innovative Comput. Informat. Control 8 (7 A) (2012) 4705–4717.
[33] M. Arican, K. Polat, Binary particle swarm optimization (BPSO) based channel selection in the EEG signals and its application to speller systems, J. Artif. Intell. Syst. 2 (1) (2020) 27–37, https://doi.org/10.33969/ais.2020.21003.
[34] R. Santos, E.D. Moreno, C. Estombelo-Montesco, A comparison of two embedded systems to detect electrical disturbances using decision tree algorithm, in: Proceedings - 32nd Symposium on Integrated Circuits and Systems Design, SBCCI , 2019, 2019., https://doi.org/10.1145/3338852.3339878, URL https://0212390v0-y-https-ieeexplore-ieee-org.proxy.sakarya.deep-knowledge.net/document/8862019.
[35] D. Saguil, A. Azim, Time-efficient offloading for machine learning tasks between embedded systems and fog nodes, in: Proceedings - 2019 IEEE 22nd International Symposium on Real-Time Distributed Computing, ISORC 2019, Institute of Electrical and Electronics Engineers Inc., 2019, pp. 79–82. doi:10.1109/ISORC.2019.00022.
[36] Y. Maret, D. Oberson, M. Gavrilova, Real-Time Embedded System for Gesture Recognition, in: Proceedings - 2018 IEEE International Conference on Systems, Man, and Cybernetics, SMC 2018, Institute of Electrical and Electronics Engineers Inc., 2019, pp. 30–34. doi:10.1109/SMC.2018.00014.
[37] K. Polat, K. Onur Koc, Detection of skin diseases from dermoscopy image using the combination of convolutional neural network and one-versus-all, J. Artif. Intell. Syst. 2 (1) (2020) 80–97, https://doi.org/10.33969/ais.2020.21006.
[38] N. Daldal, M. Nour, K. Polat, A novel demodulation structure for quadrate modulation signals using the segmentary neural network modelling, Appl. Acoust. 164 (2020), https://doi.org/10.1016/j.apacoust.2020.107251, 107251.
[39] M. Vogt, V. Kecman, Active-Set Methods for Support Vector Machines, in: Support Vector Machines: Theory and Applications, Vol. 177, Springer, Berlin, Heidelberg, 2005, pp. 133–158. doi:10.1007/10984697_6.
[40] Z.G. Fthenakis, D. Balaska, V. Zafiropulos, Uncovering the FUTREX-6100XL prediction equation for the percentage body fat, J. Med. Eng. Technol. 36 (7) (2012) 351–357, https://doi.org/10.3109/03091902.2012.708382.
[41] D.C. Frankenfield, W.A. Rowe, R.N. Cooney, J.S. Smith, D. Becker, Limits of body mass index to detect obesity and predict body composition, Nutrition 17 (2001) 26–30.
[42] J. Company, S. Ball, Body composition comparison: Bioelectric impedance analysis with dual-energy X-ray absorptiometry in adult athletes, Meas. Phys. Educ. Exerc. Sci. 14 (2010) 186–201, https://doi.org/10.1080/1091367X.2010.497449.
[43] G.E. van der Ploeg, S.M. Gunn, R.T. Withers, A.C. Modra, Use of anthropometric variables to predict relative body fat determined by a four-compartment body composition model, Eur. J. Clin. Nutr. 57 (2003) 1009–1016, https://doi.org/10.1038/sj.ejcn.1601636.
[44] S.D. Ball, T.S. Altena, P.D. Swan, Comparison of anthropometry to DXA: A new prediction equation for men, Eur. J. Clin. Nutr. 58 (2004) 1525–1531, https://doi.org/10.1038/sj.ejcn.1602003.
[45] M. Kagawa, C. Kuroiwa, K. Uenishi, M. Mori, A.P. Hills, C.W. Binns, New percentage body fat prediction equations for japanese females, J. Physiol. Anthropol. 26 (2007) 23–29, https://doi.org/10.2114/jpa2.26.23.
[46] C.J. Henry, S. D/O Ponnalagu, X. Bi, S.-Y. Tan, New equations to predict body fat in Asian-Chinese adults using age, height, skinfold thickness, and waist circumference, J. Acad. Nutr. Dietet. 118 (7) (2018) 1263–1269, https://doi.org/10.1016/j.jand.2018.02.019.
[47] A.C.C. Salamunes, A.M.W. Stadnik, E.B. Neves, Estimation of female body fat percentage based on body circumferences, Revista Brasileira de Medicina do Esporte 24 (2) (2018) 97–101, https://doi.org/10.3969/j.issn.1001-1935.2018.03.006.
[48] J. Hastuti, M. Kagawa, N.M. Byrne, A.P. Hills, Anthropometry to assess body fat in Indonesian adults, Asia Pacific J. Clin. Nutr. 27 (3) (2018) 592–598, https://doi.org/10.6133/apjcn.092017.02.
[49] J.C. Aristizabal, A. Estrada-Restrepo, A.G. García, Development and validation of anthropometric equations to estimate body composition in adult women, Colombia Medica 49 (2) (2018) 154–159, https://doi.org/10.25100/cm.v49i2.3643.
[50] T. Ferenci, L. Kovács, Predicting body fat percentage from anthropometric and laboratory measurements using artificial neural networks, Appl. Soft Comput. J. 67 (2018) 834–839, https://doi.org/10.1016/j.asoc.2017.05.063.
[51] J. Stevens, K.P. Truesdale, J. Cai, F.S. Ou, K.R. Reynolds, S.B. Heymsfield, Nationally representative equations that include resistance and reactance for the prediction of percent body fat in Americans, Int. J. Obesity 41 (2017) 1669–1675, https://doi.org/10.1038/ijo.2017.167.
[52] M.G. Swainson, A.M. Batterham, C. Tsakirides, Z.H. Rutherford, K. Hind, Prediction of whole-body fat percentage and visceral adipose tissue mass from five anthropometric variables, PLoS One 12 (5) (2017) 1–13, https://doi.org/10.1371/journal.pone.0177175.
[53] H. Sung, J. Mun, Development and cross-validation of equation for estimating percent body fat of Korean adults according to body mass index, J. Obesity Metabolic Syndrome 26 (2017) 122–129, https://doi.org/10.7570/jomes.2017.26.2.122.
[54] J.C. Aristizabal, A. Estrada Restrepo, A. Giraldo García, Development and validation of anthropometric equations to estimate body composition in adult women, Colombia Médica 49 (2) (2018) 154–159, https://doi.org/10.25100/cm.v49i2.3643, URL http://colombiamedica.univalle.edu.co/index.php/comedica/article/view/3643.
[55] A.C.C. Salamunes, A.M.W. Stadnik, E.B. Neves, Estimation of female body fat percentage based on body circumferences, Revista Brasileira de Medicina do Esporte 24 (2) (2018) 97–101, https://doi.org/10.1590/1517-869220182402181175, URL http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1517-86922018000200097&lng=en&tlng=en.
[56] C.J. Henry, S. D/O Ponnalagu, X. Bi, S.Y. Tan, New equations to predict body fat in asian-Chinese adults using age, height, skinfold thickness, and waist circumference, J. Acad. Nutr. Dietet. 118 (7) (2018) 1263–1269, https://doi.org/10.1016/J.JAND.2018.02.019, URL https://www.sciencedirect.com/science/article/pii/S2212267218302703.
[57] H. Sung, J. Mun, Development and cross-validation of equation for estimating percent body fat of Korean adults according to body mass index, J. Obesity Metabolic Syndrome 26 (2) (2017) 122–129, https://doi.org/10.7570/jomes.2017.26.2.122, URL http://www.ncbi.nlm.nih.gov/pubmed/31089506 http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC6484901, http://www.jomes.org/journal/view.html?doi=10.7570/jomes.2017.26.2.122.
[58] P. Deurenberg, J.A. Weststrate, J.C. Seidell, Body mass index as a measure of body fatness: age- and sex-specific prediction formulas, Br. J. Nutr. 65 (2) (1991) 105–114, https://doi.org/10.1079/BJN19910073, URL https://www.cambridge.org/core/product/identifier/S0007114591000193/type/journal_article.
[59] A. Jackson, P. Stanforth, J. Gagnon, T. Rankinen, A. Leon, D. Rao, J. Skinner, C. Bouchard, J. Wilmore, The effect of sex, age and race on estimating percentage body fat from body mass index: The Heritage Family Study, Int. J. Obesity 26 (6) (2002) 789–796, https://doi.org/10.1038/sj.ijo.0802006, URL http://www.nature.com/articles/0802006.
[60] S. Meeuwsen, G. Horgan, M. Elia, The relationship between BMI and percent body fat, measured by bioelectrical impedance, in a large adult sample is curvilinear and influenced by age and sex, Clin. Nutr. 29 (5) (2010) 560–566, https://doi.org/10.1016/J.CLNU.2009.12.011, URL https://www.sciencedirect.com/science/article/pii/S026156141000004X.
[61] P. Deurenberg, A. Andreoli, P. Borg, K. Kukkonen-Harjula, A. de Lorenzo, W. van Marken Lichtenbelt, G. Testolin, R. Vigano, N. Vollaard, The validity of predicted body fat percentage from body mass index and from impedance in samples of five European populations, Eur. J. Clin. Nutr. 55 (11) (2001) 973–979, https://doi.org/10.1038/sj.ejcn.1601254, URL http://www.nature.com/articles/1601254.
[62] D. Gallagher, M. Visser, D. Sepulveda, R.N. Pierson, T. Harris, S.B. Heymsfield, How useful is body mass index for comparison of body fatness across age, sex, and ethnic groups?, Am J. Epidemiol. 143 (3) (1996) 228–239, https://doi.org/10.1093/oxfordjournals.aje.a008733, URL https://academic.oup.com/aje/article-lookup/doi/10.1093/oxfordjournals.aje.a008733.
[63] S. Leahy, C. O'Neill, R. Sohun, C. Toomey, P. Jakeman, Generalised equations for the prediction of percentage body fat by anthropometry in adult men and women aged 18–81 years, Br. J. Nutr. 109 (2013) 678–685, https://doi.org/10.1017/s0007114512001870.
[64] P. Deurenberg, M. Deurenberg-Yap, J. Wang, F.P. Lin, G. Schmidt, Prediction of percentage body fat from anthropometry and bioelectrical impedance in Singaporean and Beijing Chinese, Asia Pacific J. Clin. Nutr. 9 (2) (2000) 93–98, https://doi.org/10.1046/j.1440-6047.2000.00149.x.
[65] R.M. Bielemann, M.C. Gonzalez, T.G. Barbosa-Silva, S.P. Orlandi, M.O. Xavier, R. B. Bergmann, M.C. Formoso Assunção, Estimation of body fat in adults using a portable A-mode ultrasound, Nutrition 32 (2016) 441–446, https://doi.org/10.1016/j.nut.2015.10.009.