DIAGNOSIS OF OBESITY LEVEL BASED ON BAGGING ENSEMBLE CLASSIFIER AND FEATURE SELECTION METHODS

Asaad Alzayed, Waheeda Almayyan and Ahmed Al-Hunaiyyan

Computer Information Department, Collage of Business Studies, PAAET, Kuwait

ABSTRACT

In the current era, the amount of data generated from various device sources and business transactions is rising exponentially, and the current machine learning techniques are not feasible for handling the massive volume of data. Two commonly adopted schemes exist to solve such issues scaling up the data mining algorithms and data reduction. Scaling the data mining algorithms is not the best way, but data reduction is feasible. There are two approaches to reducing datasets selecting an optimal subset of features from the initial dataset or eliminating those that contribute less information. Overweight and obesity are increasing worldwide, and forecasting future overweight or obesity could help intervention. Our primary objective is to find the optimal subset of features to diagnose obesity. This article proposes adapting a bagging algorithm based on filter-based feature selection to improve the prediction accuracy of obesity with a minimal number of feature subsets. We utilized several machine learning algorithms for classifying the obesity classes and several filter feature selection methods to maximize the classifier accuracy. Based on the results of experiments, Pairwise Consistency and Pairwise Correlation techniques are shown to be promising tools for feature selection in respect of the quality of obtained feature subset and computation efficiency. Analyzing the results obtained from the original and modified datasets has improved the classification accuracy and established a relationship between obesity/overweight and common risk factors such as weight, age, and physical activity patterns.

KEYWORDS

Data mining, Obesity, Feature reduction, Filter feature selection, Bagging algorithm.

1. INTRODUCTION

The World Health Organization defines obesity and overweight as excessive fat accumulation in particular body parts that can be dangerous to wellbeing. Obesity among adults, teens, and children has emerged worldwide health problem. The number of people who suffer from obesity has doubled since 1980. The studies estimated that more than 1900 million adults nowadays suffer from weight alteration, and by 2030, lifestyle diseases will contribute to 30% of global deaths [1]. There is an alarming increase in obesity, and overweight is the primary lifestyle disease that leads to other health disorders, such as cardiovascular diseases, chronic obstructive pulmonary disease, cancer, type II diabetes, hypertension, anxiety, and depression [2].

Obesity is a disease with numerous factors, such as uncontrolled weight gain, excessive fat intake, and energy consumption. Some of the reasons for being overweight are the increased consumption of high energy-dense food high in fat and ow frequency of physical activity due to an inactive type of work, the new transport manners, and growing urbanization. Biological hazard factors such as genetic background can also cause obesity, so there can be several kinds of obesity as Monogenic, leptin, polygenic and syndromic. Besides, other risk factors are social,

DOI: 10.5121/ijaia.2022.13203

psychological, and unhealthy eating habits [3]. On the other side, some literature suggests other determining several potential risk factors that contribute to obesity, such as being an only child family conflicts as divorce, depression, and anxiety [4,5]. Identifying risk factors can prevent obesity and overweight with the appropriate behavioral intervention strategies [6].

Data mining is the analytical process of knowledge discovery in large and complex datasets, obtaining new information to support decision-making [7,8,9]. This study had the objective of implementing several data mining techniques to predict overweight/obesity status. Several scholars have looked into this disorder and generated means to figure out the obesity level of an individual; yet, such tools are limited to calculating the BMI, neglecting relevant factors such as the family history of obesity and leisure time dedicated to physical activities [9]. We considered building a competent tool to detect obesity levels more efficiently based on this.

Many lifestyle factors or unhealthy habits contributing to obesity have been reported, ranging from dietary and physical activity patterns [2, 4, 5, 6]—Accordingly, an apparent demand to develop an automated solution to diagnose obesity. Consequently, a large scale of datasets is a significant challenge. The feature selection process minimizes the computational overhead and improves the overall performance by eliminating nonessential and unimportant features from datasets before model implementation [10], a crucial requirement in medical data classification problems. A considerable number of machine learning algorithms have been proposed for medical datasets and related classification problems, including neural networks methods [11,12,19], decision trees algorithms [11,14,19], convolutional neural network algorithm approaches [22,23], SVM model approaches [13,16,17], and k-nearest neighbor classifiers [13,18].

Obesity already exists; this article seeks to explore the critical factors behind this disease from a data-mining point of view. The primary objective is to improve the prediction accuracy of obesity with a minimal number of feature subsets using the dataset gathered by Palechor and de la Hoz Manotas to estimate obesity levels among people from Mexico, Peru, and Colombia [24]. The main contributions of this paper include four aspects.

I. In the current work, we investigate the performance of different feature selection methods to build optimal feature subsets to consider the obesity and overweight risks such as The Inconsistency Metric, mRMR, ReliefF, Pairwise Consistency, Pairwise Correlation, and Tabu search.
II. We added Basal Metabolic Rate, Resting Metabolic Rate, and Body Fat Percentage parameters to the dataset list of features.
III. We built several machine learning classifiers, C4.5, FURIA, RandF, and Bagging, to classify the feature subsets.
IV. We carried out experiments considering the 10-fold cross-validation and several evaluation metrics to test the effectiveness of the suggested classification algorithm in predicting the obesity level.

The remainder of the article is as follows; Section 2 consists of the existing literature on using the machine learning model in overweight and obesity prediction. Then Section 3 represents the approaches used in this article. Section 4 clarifies the methodology that includes the proposed feature selection technique and the classifiers. Section 5 discusses the detail regarding implementing the proposed classification algorithm and details about experiments and results. Finally, conclusions are given in Section 6.

2. PREVIOUS WORK

Recently, data mining has been vastly applied in many fields, including healthcare and medical science. Several data mining techniques have been used to explore the obesity problem among adolescents. In [11], researchers evaluated different multivariate regression methods and multilayer perceptron feedforward neural network models to define people at risk of obesity using the UK Millennium Cohort Study set up to follow the lives of children born at the turn of the new century. With an accuracy of above 90% to predict teenager BMI from prior BMI readings. In a recent research [12] that carried on the experiments over the dataset, they used the Synthetic Minority Oversampling Technique (SMOTE) technique to handle the imbalance issue and ensemble classifier.

Bassam et al. [13] built analytical models on data obtained from the Kuwait Health Network to provide early insight into the future risk of type II diabetes. The study considered several parameters, including age, gender, BMI, pre-existing hypertension, and family history of hypertension. Using a machine-learning algorithm that compares the performance of logistic regression, k-nearest neighbor (kNN), support vector machine (SVM) based on a five-fold cross-validation technique, the kNN classifier has outperformed the other classifiers in terms of AUC values. Meghana et al. [14] developed a machine learning tool using auto-sklearn model to classify two cardiovascular disease datasets and compare the classification accuracies by the opinion of graduate students. The results reported outperformed traditional machine learning classifiers and the graduate student.

Jindal et al. [15] conducted a research R ensemble prediction model and Python interface model for obesity prediction based on age, height, weight, and BMI. The ensemble approach scored an accuracy of 89.68% and outperformed generalized linear model, random forest, and partial least squares. Seyla et al. [16] investigated the impact of dietary and physical activity behaviors on obesity using discriminant analysis, support vector machines (SVM), and neural nets algorithms. As a result, SVM achieved a higher prediction accuracy and validated that dietary and exercise behavior patterns can better explain the prevalence of obesity instead of individual components. Dunstan et al. [17] utilized SVM, Random Forest, and Extreme Gradient Boosting to predict country-level obesity prevalence, based on countrywide food sales of a small subset of food and beverage group categories. Random Forest predicted obesity with an absolute error of 10%. The study indicated that commercial baked goods and flours, followed by cheese and sweet carbonated drinks, were the most appropriate food categories in predicting obesity. Zheng et al. [18] analyzed the 2015 Youth Risk Behavior Surveillance System dataset for the state of Tennessee in order to predict obesity in high school students by focusing on nine health-related behaviors. They have utilized binary logistic regression, improved decision tree (IDT), weighted kNN, and neural network. Results showed that the weighted kNN model achieved an 88.82% accuracy and 93.44% specificity in the classification problem.

DeGregory et al. [19] proposed adapting smart wearable sensor devices, electronic healthcare records, and smartphone apps in obesity-related data gathering to prevent obesity/overweight problems. They studied the behavior of several machine learning methods and topological data analysis on the National Health and Nutrition Examination Survey. Golino et al. [20] used a random tree technique to investigate the prediction of elevated blood pressure using BMI, waist and hip circumference, and waist-hip ratio to study data collected from 400 students. The proposed model outperformed the traditional logistic regression model in predictive power. The proposed model reported a sensitivity of 80.86%, specificity of 81.22% in the training set, and 45.65%, 65.15% in the test sample for women. Moreover, a sensitivity of 72% and specificity of 86.25% in the training set and, respectively, 58.38% 69.70% in the test set for men.

Pleuss et al. [21] led research to deploy a machine learning-based model to process a 3D image to obtain hundreds of anthropometric measurements after analyzing the images obtained from a 3D scanner. They used anthropometric information to estimate BMI, waist circumference, and hip circumference to body fat. Maharana et al. [22] built a regression model based on a convolutional neural network to approximately process 150,000 3-D satellite images from Google Static Maps API in six cities in the US to extract features of the built environment to study connections between the built environment and obesity. They concluded that consistently presents a strong association between obesity prevalence and the built environment indicators, despite varying city and neighborhood values. Pouladzadeh et al. [23] proposed a combination of graph cut segmentation and deep learning neural networks to classify and recognize food items that would run on smartphones to calculate the amount of calorie intake automatically. Results showed that combining the two methods provides 100 % food recognition accuracy. From the related work, we identified a list of machine learning models and risk factors related to obesity/overweight, as described in Table 1.

Table 1. The risk factors related to obesity, overweight according to data mining approaches

Researcher	**Risk Factors**
DeGregory et al. [19]	Inactivity, inappropriate diet
Singh et al. [11,12]	BMI
Bassam et al. [13]	Age, gender, BMI, pre-existing hypertension, family history of hypertension, and diabetes type II
Meghana et al. [14]	Cardiovascular diseases
Seyla et al. [16]	Activity, nutrition
Jindal et al. [15]	Age, height, weight, BMI
Zheng et al. [18]	Inactivity, improper diet
Dunstan et al. [17]	Unhealthy diet
Golino et al. [20]	Blood Pressure, BMI, Waist Circumference, Hip Circumference, Waist–Hip Ratio
Pleuss et al. [21]	BMI, Circumference, Hip Circumference
Maharana et al. [22]	Environment, context
Pouladzadeh et al. [23]	Nutrition

3. METHODOLOGY AND MATERIALS

The primary objective of this research work is to study the performance of various machine learning classifiers with various feature sets. One of the motivations in this research is to improve classification performance by applying several feature selection approaches to build a new feature subset and several machine learning classifiers for prediction. The process can be completed in two steps: feature selection and classification. In the framework of the proposed approach, we utilized the Bagging algorithm to predict the obesity level. The Bagging algorithm has been used as an effective ensemble algorithm to ensure good performance of the base classifiers in this work since the Bagging algorithm can thoroughly handle the core characteristic of base classifiers algorithms, instability [19]. It also utilizes the Bagging algorithm as a reliable method due to its generalized capability to avoid overfitting problems by reducing the variance. Generally, the generalization capability of the classification model is reduced by the existence of redundant features. Therefore, we apply the Filter-based feature selection method to construct a compact feature selection of high-impact features and reduce redundant features.

3.1. Dataset

Overweight and obesity are related to various factors, such as food and dietary practices, energy consumption, genetics, lifestyle, socioeconomic aspects, anxiety, and depression [3]. The deployed dataset was constructed by [24] after searching for literary sources for the main factors or habits contributing to obesity. The sample population included 1043 females and 1068 males between 14 and 61 years old, including Colombia, Mexico, and Perú. The dataset includes 2111 records and 18 variables based on the surveys to identify their obesity level. Information gathered included age, gender, weight, the frequency of physical activity, the frequency of fast food intake, and other aspects that could help describe the nutritional behavior of obese people (see Table 2). To specify the obesity levels, we used the table provided by WHO (Table 3) to categorize correctly the data analyzed based on the BMI.

Most of the literature concerning obesity depends on BMI to assess overweight and obesity, although it is not generic enough to be applied in every context, such as a pregnant lady or an older man. Therefore, we decided to include metabolic and anthropometric measures. We decided to add Basal Metabolic Rate, Resting Metabolic Rate, and Body Fat Percentage parameters to the dataset.

Basal metabolic rate (BMR) defines the rate of energy consumed by the human body. According to the Harris-Benedict equation [25],

$$\mathbf{BMR} = \begin{cases} (13.7\times \text{weight}) + (5\times \text{height}) - (6.8\times \text{age}) + 66 & ,\textit{gender is male} \\ (9.6\times \text{weight}) + (1.8\times \text{cm}) - (4.7\times \text{age}) + 655 & ,\textit{gender is female} \end{cases} \quad 1$$

Resting metabolic rate (RMR) depends on age, weight, height, gender. According to [26],

$$\mathbf{RMR} = \begin{cases} (10\times \text{weight}) + (6.25\times \text{cm}) - (5\times \text{age}) + 5 & ,\textit{gender is male} \\ (10\times \text{weight}) + (6.25\times \text{cm}) - (5\times \text{age}) - 16 & ,\textit{gender is female} \end{cases} \quad 2$$

Body fat percentage (BFP) and according to Deurenberg [27] If the person is a child, then

$$\mathbf{BFR} = \begin{cases} (1.51\times \text{BMI}) - (0.70\times \text{age}) - (3.6\times \text{gender}) + 1.4 & ,\textit{subject is child} \\ (1.20\times \text{BMI}) + (0.23\times \text{age}) - (10.8\times \text{gender}) - 5.4 & ,\textit{subject is adult} \end{cases} \quad 3$$

Table 2. Collected dataset description

Attribute	Abbreviation	Type	Possible values
1. Gender	-	Nominal	Male, Female
2. Age	-	Numeric	Integer Numeric Values
3. Height	-	Numeric	Integer Numeric Values (Mt)
4. Weight	-	Numeric	Integer Numeric Values (Kg)
5. Family with overweight/Obesity	FHWO	Nominal	Yes, No
6. Fast Food Intake	FAVC	Nominal	Yes, No
7. Vegetables Consumption Frequency	FCVC	Numeric	1: Always 2: Sometimes 3: Rarely
8. Number of main meals daily	NCP	Numeric	1 to 2: UD 3: TR More than 3: MT

9. Food intake between meals	CAEC	Nominal	S: Always CS: Usually A: Sometimes CN: Rarely
10. Smoking	SMOKE	Nominal	Yes, No
11. Liquid intake daily	CH2O	Numeric	1: Less than one-liter 2: Between 1- and 2-liters 3: More than 2 liters
12. Calories Consumption Calculation	SCC	Nominal	Yes, No
13. Physical Activity	FAF	Numeric	1: 1 to 2 days 2: 3 to 4 days 3: 5 to 6 days 0: No physical activity
14. Schedule dedicated to technology	TUE	Numeric	0: 0 to 2 hours 1: 3 to 5 hours 2: More than 5 hours
15. Alcohol consumption	CALC	Nominal	NO: No consume of alcohol CF: Rarely S: Weekly D: Daily
16. Type of Transportation used	MTRANS	Nominal	TP: Public transportation MTA: Motorbike BTA: Bike CA: Walking AU: Automobile
17. Class	IMC	Nominal	Based on the WHO Classification: -Underweight -Normal -Overweight -Obesity I -Obesity II -Obesity III

Table 3. Definition of underweight, overweight and obesity according to the WHO reference system for México [24]

WHO Category	**BMI (kg/m)**
Underweight	Less than 18.5
Normal	18.5-24.9
Overweight	25.0-29.9
Obesity Class I	30.0-34.9.
Obesity Class II	35.0-39.9
Obesity Class III	More than 40.0

3.2. Feature selection

The feature selection phase aims to specify essential attributes before constructing a classification model by removing non-essential and redundant attributes. There are two main feature selection paradigms: Filter and Wrapper methods, where each method has a unique feature selection mechanism [28]. The filter methods evaluate the feature's relevancy by utilizing a ranking procedure that withdraws the most minor ranked features accordingly. This method enhances the correlation between features and classes through the evaluation criteria and weakens the correlation between features. Due to that, the Filter method is robust, scalable, computationally efficient, and most importantly, independent of the classifier.

On the other hand, the wrapper model works nearly like the filter methods, but they evaluate the subsets using a predefined classification algorithm rather than an independent measure. Correspondingly, the predictive accuracy is often high, while the generalization ability is inadequate. Therefore, the wrapper methods deliver more satisfactory results, but they tend to be more computationally expensive with large-scale datasets [29].

Overall, we can summarise the difference between the Filter and Wrapper methods. The Filter method has a better-generalized ability, though the estimation performance of the Wrapper method is much better. Moreover, the computational cost of the Wrapper method is better, and it has a more significant potential to handle the overfitting issue. Therefore, several articles have applied feature selection methods before conducting the classification process [12, 18, 20].

1. Tabu Search Technique

Tabu search is an iterative memory-based algorithm proposed by Glover in 1986 to solve combinatorial optimization problems [30]. It contains a local search mechanism conjoined with a tabu mechanism. Tabu search starts with an initial solution X'$\in\Omega$ among neighborhood solutions, where Ω feasible solutions are set. Since then, several researchers have applied the Tabu successfully search in several multiclass classification problems [31].
The algorithm explores and assesses all the possible solutions N(X)$\subseteq \Omega$ to obtain a new one, X' $\in\Omega$ N(X), with a better functional value if the new solution X' is not listed in the tabu list or it satisfies the aspiration criterion [24]. If the candidate solution X' is better than X_{best}, the value of X_{best} is overridden; else, the Tabu search will go uphill to avoid local minima. Moreover, to avoid local minima, Tabu search limits visiting previously visited solutions for a certain number of iterations. Next, the neighborhood search resumes the new solution X' until it meets the stopping criterion.

2. mRMR feature selection method

The mRMR is a heuristic technique proposed by Peng et al. to measure the relevancy and redundancy of features and determine the most informative features [32]. The mRMR feature selection technique aims to simultaneously identify features with maximum relevancy for target classes and minimum redundancy with other features in the dataset. The mRMR method helps determine crucial features and, at the same time, minimize the classification error. The basic concept of the mRMR method is to use two mutual information operations: one to measure the relevancy between classes and each feature and the second to measure redundancy between every feature. The main goal of applying the mRMR feature is to find a subset of features from S with m features, $\{x_i\}$, that jointly has the most considerable dependency on the target class C or have the minimal redundancy on the selected features subset S.

3. ReliefF

ReliefF is a well-known multivariate filter that extends a prior version of Relief [33] based on nearest neighbors. It randomly selects samples and searches for nearest neighbors from the same class (ignoring the rest). The ReliefF works by comparing the values of the selected sample with the hits and misses, and then it updates the relevance score for each feature. A helpful feature should weigh similar examples from the identical class and different instances from the other classes similar to examples from the same class and different from examples from the other classes.

4. The Inconsistency Metric

The idea behind this attribute subset evaluation method is to find a subset of features that can separate the dataset with a highly predominant class [34]. The consistency measure approach finds out the best discrimination set from the original data rather than applying an algorithm to split the classes. The inconsistency measure considers the features inconsistent when two or more samples have the same values but distinct labels.

5. Pairwise Correlation

The Pairwise Correlation technique was developed by Jiménez F. et al. in 2021 and was inspired by the correlation-based feature selection method [35]. The proposed Pairwise Correlation method evaluates an attribute i by using the following function

$$\phi_D^A(i) = \frac{1}{n-1} \cdot \sum_{i \in \{1,\dots,n\}} \phi_D^A(\{i,j\}) \qquad 4$$

where ΦD ({i, j}) is the merit of the subset formed by attributes i and j, for all j = 1, . . ., n,

$$\phi_D(S) = \frac{k \cdot \delta^{c}{}_D}{\sqrt{k + k \cdot (k-1) \cdot \delta^{c}{}_D}} \qquad 5$$

The merit $\Phi^A{}_D$ (i) of an attribute i is the mean of the merits ΦD of the attribute subsets formed by i and each of the other attributes. The method prefers attributes with low correlation to other attributes, and the class is highly correlated.

6. Pairwise Consistency

The Pairwise Consistency is a relatively new feature selection method introduced by Jiménez F. et al. that developed the consistency metric for attribute subsets introduced by Liu and Setiono [35, 36]. The idea behind the consistency measure is to locate attributes that can split the dataset into parts with a favorably predominant class. The proposed Pairwise Consistency method evaluates an attribute i ∈ {1, . . ., n} by using the following function Ψ AD

$$\psi_D^A(i) = \frac{1}{n-1} \cdot \sum_{i \in \{1,\dots,n\}} \psi_D^A(\{i,j\}) \qquad 6$$

where $\Psi D(\{i,j\}) = 1 - ID(\{i,j\})$ is the consistency rate of the subset formed by the attributes i and j, for all j = 1, . . . , n, with j‡i. The merit ΨAD (i) of an attribute i is the mean of the consistency rates of the attribute i and each of the other attributes.

4. CLASSIFICATION

4.1. C4.5 decision tree Classifier

C4.5 is an improved version of the ID3 decision tree algorithm. Its capabilities include approximating discrete-valued functions, robust noisy data, and handling missing attribute values. The C4.5 algorithm chooses the best splitting attributes using the information gain (IG) ratio as the default attribute choice criterion [37]. This algorithm begins visiting each node to select an optimal split based on the IG ratio's maximization and designates it as the root node of a tree. Then, it forms a leaf for all of the potential values. The algorithm's primary mission is to choose the appropriate feature to partition the dataset into several categories.

4.2. RandF (RandF) Machine Learning

RandF algorithm is a tree-based-ensemble machine learning method based on a combination tree of predictors. Each tree uses a random vector sampled independently from the classification input vector. Each tree depends on the values of a random vector sampled independently and with the same distribution for all trees [38]. The algorithm assesses the last decision by compiling the votes from all the trees using a rule-based approach or an iterative error minimization technique by reducing the weights for the correctly classified samples. RandF accelerates the building speed by building a variety of individual base classifiers.

4.3. Fuzzy Unordered Rules Induction (FURIA) classifier

Hühn and Hüllermeier first introduced the FURIA algorithm to improve and extend the state-of-the-art rule learner algorithm RIPPER. FURIA is more accurate than the original RIPPER algorithm and C4.5 classifier [39]. FURIA depends on fuzzy rules and unordered rule sets instead of conventional rules for classification. It also uses the rule stretching technique to deal with uncovered cases to generate an unordered one vs. rest scheme for each class. For that, the order of the rule is not considered because is no default rule as each class is separated from others. FURIA applies a rule stretching strategy if any rule may not cover a new query. Fuzzy intervals, namely fuzzy sets with trapezoidal membership functions, are used to obtain fuzzy rules.

4.4. Bagging classifier

Breiman proposed bootstrap aggregating bagging as a meta-algorithm based on decision trees' ensemble to improve classification and regression models [40]. Although this technique can be applied in any algorithm, it mainly applies to decision tree models. Moreover, bagging helps decrease the variance and reduce the over-fitting of estimation. Bagging diversity is achieved by generating copies of the original training set, where different training datasets are randomly drawn with replacement. Accordingly, a decision tree is built based on the standard approach with each training data replica. Consequently, every tree can be described by a distinctive set of variables, nodes, and leaves. Ultimately, their projections are joint to obtain the outcome [41].

4.5. Evaluation Metrics

The performance of the proposed model is measured using precision, the receiver operator characteristic (ROC), and Cohen's kappa coefficient. We can define precision as

$$\text{precision}=\frac{TN}{(FP+TN)}\times 100\% \qquad 7$$

FP is the false positive rate, and TN is the True negative rate.

The ROC curve is a standard metric for analyzing classifier performance over a range of trade-offs between true positive rate (TP) and false-positive rate (FP) [42]. ROC usually ranges from 0.5 for a perfectly random classifier and 1.0 for a perfect classifier. The area under the ROC curve is a graphical plot for evaluating two-class decision problems.

Kappa error or Cohen's kappa coefficient is a useful measure to compare different classifiers' performance and selected features' quality ranges from 1 to -1 [42].

If the kappa value -according to the successive formula- approaches +1, there is a better opportunity for an agreement, and when it approaches -1, it indicates a poor chance for agreement.

$$\text{Kappa error} = \frac{P(A)-P(E)}{1-P(E)} \qquad 8$$

P(A) is the total agreement probability, and P(E) is the theoretical probability of chance agreement.

5. RESULTS AND DISCUSSION

Since we have defined the theoretical mechanism of the proposed algorithm, it is necessary to examine its practical efficiency as well. Accordingly, we plan to conduct several simulation experiments to verify the reliability of the proposed algorithm. In the simulation experiments, we have chosen to apply the 10-fold cross-validation method for the validation, and separately recorded classification accuracy, AUC, Kappa, and MCC of FURIA, C4.5, RandF, and Bagging classifiers. Initially, we should report the classifications without feature selection to demonstrate the performance of the suggested feature selection techniques. Table 2 reports the performance before performing feature selection. The RandF classifier (95.78%) slightly outperformed the FURIA (95.07%) and C4.5 (93.74%) classifiers. At the same time, the Bagging based on the RandF classifier outperformed the other classifiers by scoring accuracy of 95.878%. In most classifiers ' performances, the results showed good performance in AUC, Kappa, and RMSE metrics. The time to build the model is higher in FURIA and bagging-based FURIA than C4.5 and RandF classifiers.

Table 4. Classification performance without feature selection methods

Classifier	**ACC(%)**	**AUC**	**Kappa**	**RMSE**	**Time(s)**
FURIA	95.07	0.987	0.9425	0.1103	1.85
C4.5	93.74	0.978	0.927	0.1286	0.02
RandF	95.78	0.998	0.9508	0.1143	0.44
Bagging(FURIA)	96.7314	0.998	0.9618	0.0938	10.79
Bagging(C4.5)	95.405	0.997	0.9463	0.1039	0.14
Bagging(RandF)	95.878	0.998	0.9519	0.1202	2.89

Table 4 lists the obtained features using mRMR, ReliefF, Pairwise Consistency, Pairwise Correlation, The Inconsistency Metric, and Tabu search. This step has reduced the feature size from 16 to 11 and 7. The feature selection methods helped reduce the number of attributes of the original datasets, notably the Inconsistency Metric and TabuSearch techniques. Figure ?? shows the resulting Venn diagram, which demonstrates the relationship among Feature selection Techniques. Table 5 illustrates all the possible intersections between Feature selection Techniques using the original dataset. All the feature selection techniques share Weight, CH2O, and TUE attributes. Most of the methods share Age, FAF, and Height attributes.

Table 5. Optimal feature subsets

Feature Selection Method	No. of Features	Features
1. mRMR	10	Gender, Age, Weight, FHWO, FAVC, FCVC, NCP, CAEC, CH2O, TUE
2. ReliefF	11	Gender, Weight, FHWO, FCVC, CALC, CAEC, TUE, MTRANS, Height, NCP, CH2O
3. Pairwise Consistency	9	Weight, Age, FCVC, FAF, Height, CH2O, TUE, NCP, Gender
4. Pairwise Correlation	8	Weight, FCVC, Age, TUE, NCP, Gender, CH2O, FAF
5. The Inconsistency Metric	7	Age, Height, Weight, FAVC, CH2O, FAF, TUE
6. TabuSearch	7	Age, Height, Weight, CAEC, CH2O, FAF, TUE

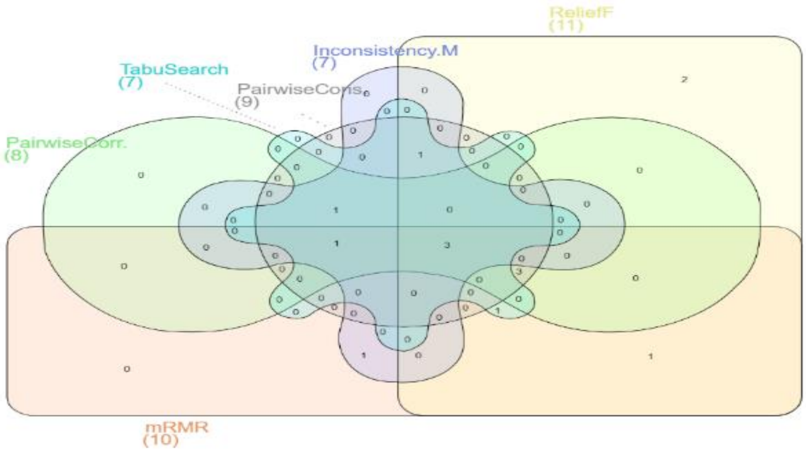

Figure 1. Venn diagram demonstrating the cross-links between the feature selection techniques used over the original dataset

Table 6. Intersecting Features among Feature selection Techniques using the original dataset

Feature Selection Techniques	Intersecting Features
[mRMR] and [ReliefF] and [PairwiseCons.] and [PairwiseCorr.] and [Inconsistency.M] and [TabuSearch]	Weight, CH2O, TUE
[mRMR] and [PairwiseCons.] and [PairwiseCorr.] and [Inconsistency.M] and [TabuSearch]	Age
[PairwiseCons.] and [PairwiseCorr.] and [Inconsistency.M] and [TabuSearch]	FAF
[ReliefF] and [PairwiseCons.] and [Inconsistency.M] and [TabuSearch]	Height
[mRMR] and [ReliefF] and [TabuSearch]	CAEC
[mRMR] and [Inconsistency.M]	FAVC
[mRMR] and [ReliefF] and [PairwiseCons.] and [PairwiseCorr.]	Gender, FCVC, NCP

Table 6 reports the classification performance after performing feature selection. The RandF classifier outperformed the other classifiers achieving an accuracy of 96.77% using the Pairwise Consistency that yielded nine features. At the same time, the bagging based on the FURIA algorithm outperformed the other classifiers by scoring an accuracy of 96.73% while using the Pairwise Consistency technique. As the feature selection methods reduced the features of datasets, it also increased the overall performance accuracy, as with the ReliefF, Pairwise Consistency, The Inconsistency Metric, and TabuSearch techniques. Although the number of features chosen by both the Inconsistency Metric and TabuSearch techniques is less than the other techniques such as mRMR and Pairwise Correlation, the classification results were better. The Bagging algorithm enhanced the prediction accuracy. Overall, the performance of the Bagging algorithm is more promising than that of the base algorithms independently. It is worth noting that the construction time for the model is reduced, chiefly for ReliefF, Pairwise Consistency, Inconsistency Metric, and TabuSearch techniques.

All in all, it indicates that the feature selection step successfully enhances the classification accuracy. Meanwhile, the Inconsistency Metric and Tabu search helped improve the classification performance with a limited number of features. All the suggested models had AUC values above 0.95. Kappa values were between 0.85 and 0.96. the values of RMSE slightly above 0.1. These results demonstrate that the models were successful in predicting obesity. However, the time taken to build the model of FURIA and FURIA bagging-based classifiers is longer than the other classifiers.

Table 7. Classification performance after feature selection

1. **mRMR:10**					
Classifier	**ACC**	**AUC**	**Kappa**	**RMSE**	**Time**
FURIA	87.873	0.962	0.8583	0.1656	1.69
C4.5	87.8257	0.962	0.8583	0.1656	0.08
RandF	**92.3259**	**0.993**	**0.9104**	0.133	0.42
Bagging(FURIA)	91.568	0.991	0.9015	0.1372	13.31
Bagging(C4.5)	88.9626	0.986	0.8711	0.1497	0.16
Bagging(RandF)	**91.9469**	**0.994**	**0.9059**	0.1381	2.56
2. ReliefF:11					
Classifier	**ACC**	**AUC**	**Kappa**	**RMSE**	**Time**
FURIA	94.5523	0.985	0.9364	0.1142	1.02
C4.5	94.3155	0.980	0.9336	0.1231	0.01
RandF	**95.5471**	**0.998**	**0.948**	0.1098	0.28
Bagging(FURIA)	**96.684**	**0.998**	**0.9613**	0.094	8.32
Bagging(C4.5)	95.5945	0.997	0.9485	0.0981	0.11
Bagging(RandF)	95.5945	**0.998**	0.9485	0.1167	2.42
3. Pairwise Consistency:9					
Classifier	**ACC**	**AUC**	**Kappa**	**RMSE**	**Time**
FURIA	95.3103	0.988	0.9452	0.1075	1.12
C4.5	94.5997	0.980	0.9369	0.1203	0.03
RandF	**96.7788**	**0.999**	**0.9624**	0.0998	0.48
Bagging(FURIA)	**96.7314**	**0.998**	**0.9618**	0.096	10.36
Bagging(C4.5)	95.5945	**0.998**	0.9485	0.0974	0.13
Bagging(RandF)	94.0313	0.973	0.9303	0.1284	0.17
4. Pairwise Correlation:8					
Classifier	**ACC**	**AUC**	**Kappa**	**RMSE**	**Time**
FURIA	86.8783	0.955	0.8467	0.1748	1.19
C4.5	86.5467	0.948	0.8428	0.1865	0.02
RandF	**92.0417**	**0.994**	**0.907**	0.1345	0.36
Bagging(FURIA)	90.3837	0.990	0.8877	0.1436	13.84
Bagging(C4.5)	88.9626	0.985	0.8711	0.1531	0.16
Bagging(RandF)	**91.9469**	**0.994**	**0.9059**	0.1402	3.18
5. The Inconsistency Metric :7					
Classifier	**ACC**	**AUC**	**Kappa**	**RMSE**	**Time**

FURIA	94.8839	0.986	0.9402	0.1094	1.03
C4.5	94.3155	0.978	0.9336	0.124	0.01
RandF	**96.5419**	**0.999**	**0.9596**	0.1033	0.3
Bagging(FURIA)	**96.5419**	0.998	**0.9596**	0.0934	9.4
Bagging(C4.5)	95.7366	0.997	0.9502	0.0988	0.11
Bagging(RandF)	96.3998	**0.999**	0.9579	0.1106	2.65
6. TabuSearch:7					
Classifier	**ACC**	**AUC**	**Kappa**	**RMSE**	**Time**
FURIA	94.3629	0.984	0.9342	0.1139	1.59
C4.5	94.2681	0.980	0.9331	0.1229	0.08
RandF	**95.9261**	**0.998**	**0.9524**	0.1092	0.43
Bagging(FURIA)	**96.5419**	**0.998**	**0.9596**	0.0964	9.82
Bagging(C4.5)	95.3103	**0.998**	0.9452	0.1004	0.13
Bagging(RandF)	95.784	**0.998**	0.9508	0.1168	2.63

Table 7 reports the performance after altering the original dataset. The RandF classifier (97.6788%) outperformed the FURIA (94.5997%) and C4.5 (93.9839%) classifiers. At the same time, the Bagging based on the RandF classifier outperformed the other classifiers by scoring accuracy of 97.5367% while using 19 features.

Table 8 lists the obtained features using mRMR, ReliefF, Pairwise Consistency, Pairwise Correlation, The Inconsistency Metric, and Tabu search. This step has reduced the feature size from 19 to 11 and 6 features only. Figure ?? shows the resulting Venn diagram, which demonstrates the relationship among Feature selection Techniques. Table 9 illustrates all the possible intersections between Feature selection Techniques using the original dataset. Worth noting that all the feature selection techniques share Weight and BFP attributes. Most methods share Age, Gender, FHWO, FCVC, and CAEC features.

Table 10 reports the classification performance after performing feature selection over the modified dataset. The RandF classifier outperformed the other classifiers achieving an accuracy of 97.63% using the Pairwise Consistency that yielded 15 features. At the same time, the Bagging based on the RandF algorithm outperformed the other classifiers by scoring an accuracy of 97.58% while using the Pairwise Consistency technique. As the feature selection methods reduced the features of datasets, it also increased the overall performance accuracy, as with the Pairwise Consistency and TabuSearch techniques.

Although the number of features chosen by the Pairwise Consistency, Pairwise Correlation, and Inconsistency Metric techniques is less than the initial dataset, the classification results were better. The Inconsistency Metric, Pairwise Consistency, and Pairwise Correlation were the best feature selection techniques. The Bagging algorithm enhanced the prediction accuracy. Overall, the performance of the Bagging algorithm is more promising than that of the base algorithms independently. It is worth noting that the construction time for the model is reduced, especially for The Inconsistency Metric, Pairwise Consistency, and Pairwise Correlation techniques. The recent results indicate that the feature reduction stage successfully helps in improving classification accuracy. As the filter-based feature selection method reduces the features of datasets, it also decreases the time taken to build the model and increases the overall performance. Based on the results of experiments, Pairwise Consistency and Pairwise Correlation techniques are shown to be promising tools for feature selection in respect of the quality of obtained feature subset and computation efficiency.

Table 8. Classification performance without feature selection methods for the modified dataset

Classifier	ACC(%)	AUC	Kappa	RMSE	Time(s)
FURIA	94.5997	0.984	0.9369	0.1106	1.58
C4.5	93.9839	0.976	0.9297	0.1288	0.11
RandF	**97.6788**	**0.999**	**0.9729**	0.088	0.45
Bagging(FURIA)	97.2525	0.998	0.9679	0.086	8.7
Bagging(C4.5)	95.0261	0.995	0.9419	0.1066	0.19
Bagging(RandF)	**97.5367**	**0.999**	**0.9712**	0.0933	2.9

Table 9. Optimal feature subsets the modified dataset

Feature Selection Method	No. of Features	Features
1. mRMR	11	Gender, Age, Weight, FHWO, FAVC, FCVC, NCP, CAEC, CH2O, TUE, BFP
2. ReliefF	11	BFP, Gender, Weight, RMR, BMR, FHWO, FCVC, CALC, CAEC, Height, FAF
3. Pairwise Consistency	15	BFP, Weight, BMR, RMR, Age, FCVC, FAF, Height, CH2O, TUE, NCP, Gender, CAEC, CALC, FHWO
4. Pairwise Correlation	14	BFP, Weight, RMR, BMR, FCVC, Age, TUE, NCP, Gender, CH2O, FAF, FHWO, CAEC, Height
5. The Inconsistency Metric	6	Age, Height, Weight, FAF, CALC, BFP
6.Tabu	11	Gender, Age, Weight, FHWO, FAVC, FCVC, NCP, CAEC, CH2O, TUE, BFP

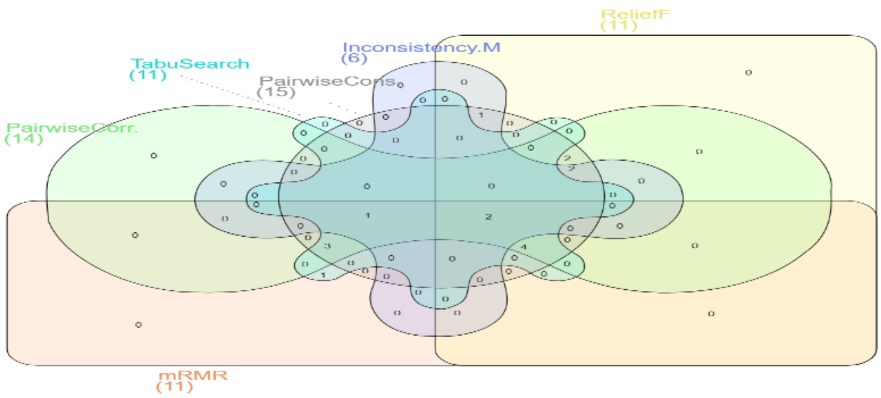

Figure 2. Venn diagram demonstrating the cross-links between the feature selection techniques used over the modified dataset

Table 10. Intersecting Features among Feature selection Techniques using the modified dataset

Feature Selection Techniques	Intersecting Features
[mRMR] and [ReliefF] and [PairwiseCons.] and [PairwiseCorr.] and [Inconsistency.M] and [TabuSearch]:	Weight BFP
[mRMR] and [PairwiseCons.] and [PairwiseCorr.] and [Inconsistency.M] and [TabuSearch]	Age

[mRMR] and [ReliefF] and [PairwiseCons.] and [PairwiseCorr.] and [TabuSearch]:	Gender FHWO FCVC CAEC
[mRMR] and [PairwiseCons.] and [PairwiseCorr.] and [TabuSearch]:	NCP CH2O TUE
[ReliefF] and [PairwiseCons.] and [Inconsistency.M]	CALC
[ReliefF] and [PairwiseCons.] and [PairwiseCorr.]:	RMR BMR
[ReliefF] and [PairwiseCons.] and [PairwiseCorr.] and [Inconsistency.M]	Height FAF
[mRMR] and [TabuSearch]	FAVC

Table 11. Classification performance after feature selection for the modified dataset

1. mRMR:11					
Classifier	**ACC**	**AUC**	**Kappa**	**RMSE**	**Time**
FURIA	94.6471	0.986	0.9375	0.1095	1.41
C4.5	93.7944	0.975	0.9275	0.1281	**0.08**
RandF	**97.2051**	0.999	**0.9674**	0.0937	0.45
Bagging(FURIA)	96.3051	0.998	0.9568	0.0907	7.09
Bagging(C4.5)	94.3629	0.995	0.9342	0.1084	0.13
Bagging(RandF)	**96.8261**	**0.999**	0.9629	0.0988	2.38
2. ReliefF:11					
Classifier	**ACC**	**AUC**	**Kappa**	**RMSE**	**Time**
FURIA	94.8839	0.984	0.9402	0.1139	0.86
C4.5	93.747	0.974	0.927	0.1307	**0.02**
RandF	**97.063**	**0.999**	0.9657	0.0872	0.26
Bagging(FURIA)	96.0682	**0.997**	**0.9541**	0.0947	7.26
Bagging(C4.5)	95.1208	0.997	0.943	0.1043	0.13
Bagging(RandF)	**96.8735**	**0.999**	0.9635	0.0919	2.35
3. Pairwise Consistency:15					
Classifier	**ACC**	**AUC**	**Kappa**	**RMSE**	**Time**
FURIA	95.0261	0.985	0.9419	0.1083	0.98
C4.5	94.2207	0.975	0.9325	0.1261	**0.02**
RandF	**97.6315**	**0.999**	0.9723	0.0888	0.29
Bagging(FURIA)	96.684	0.999	0.9613	0.0878	8.24
Bagging(C4.5)	95.2629	0.997	0.9447	0.1042	0.18
Bagging(RandF)	**97.5841**	0.999	0.9718	0.0939	2.69
4. Pairwise Correlation:14					
Classifier	**ACC**	**AUC**	**Kappa**	**RMSE**	**Time**
FURIA	94.8839	0.986	0.9402	0.1078	1.00
C4.5	94.126	0.975	0.9314	0.1269	**0.02**
RandF	**97.4893**	**0.999**	0.9707	0.0875	0.31
Bagging(FURIA)	96.5893	0.999	0.9602	0.0881	7.91
Bagging(C4.5)	95.1208	0.996	0.943	0.1047	0.18
Bagging(RandF)	**97.5367**	**0.999**	0.9712	0.0934	2.78
5. The Inconsistency Metric :6					
Classifier	**ACC**	**AUC**	**Kappa**	**RMSE**	**Time**
FURIA	96.0208	0.987	0.9535	0.1001	0.7
C4.5	94.5523	0.976	0.9364	0.1221	**0.01**

RandF	**97.5367**	**0.999**	**0.9712**	0.0792	0.26
Bagging(FURIA)	97.0156	.998	0.9651	0.0862	6.69
Bagging(C4.5)	95.1208	0.997	0.943	0.1015	0.1
Bagging(RandF)	**97.2999**	**0.999**	0.9685	0.0854	2.27
6. TabuSearch:11					
Classifier	**ACC**	**AUC**	**Kappa**	**RMSE**	**Time**
FURIA	94.6471	0.986	0.9375	0.1095	0.95
C4.5	93.7944	0.975	0.9275	0.1281	**0.02**
RandF	**97.2051**	0.999	0.9674	0.0937	0.27
Bagging(FURIA)	96.3051	0.998	0.9568	0.0907	7.24
Bagging(C4.5)	94.3629	0.995	0.9342	0.1084	0.13
Bagging(RandF)	**96.8261**	**0.999**	0.9629	0.0988	2.43

Table 12. Summary of intersecting features among Inconsistency Metric, Pairwise Consistency, and Pairwise Correlation feature selection techniques

Dataset	**Intersecting Features**
Original dataset	Weight Age FAF CH2O TUE
Modified dataset	BFP Weight Age FAF Height
[original] and [modified]	Weight Age FAF

The analysis of the results obtained from the original and modified datasets has established a relationship between obesity/overweight and common risk factors such as weight, age, and physical activity patterns. The results from the original dataset show that the techniques mentioned above have weight, age, FAF, CH2O, and TUE features in common. While in the modified dataset, the techniques share the weight, age, Height, FAF, and BFP features. Table 10 lists intersecting features among Inconsistency Metric, Pairwise Consistency, and Pairwise Correlation feature selection techniques. In sum, our findings are consistent with the findings from studies of [13,15,16,17,18,19] mentioned earlier in Section 2.

CONCLUSION

Obesity exists; this article seeks to explore the risk factors behind this disease from a data-mining point of view. The primary objective is to improve the prediction accuracy of obesity with a minimal number of feature subsets using the dataset gathered previously. For that reason, we adapted a bagging algorithm based on filter-based feature selection to improve the prediction accuracy of obesity with a minimal number of feature subsets. We utilized several machine learning algorithms for classifying the obesity classes and several filter feature selection methods to maximize the classifier accuracy. The proposed work improves the classification accuracy compared to the previous work from the experimental result. Experiments on the original and modified dataset proved that our proposed method could reduce the number of features by almost 97% and obtain satisfactory results.

REFERENCES

[1] WHO Page. Available online: https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight; https://www.who.int/nmh/publications/ncd_report_chapter1.pdf (accessed on 18 March 2020).

[2] Ward, Z.J.; Bleich, S.N.; Cradock, A.L.; Barrett, J.L.; Giles, C.M.; Flax, C.; Long, M.W.; Gortmaker, S.L. Projected US State-Level Prevalence of Adult Obesity and Severe Obesity. N. Engl. J. Med. 2019, 381, 2440–2450.

[3] Csige, I.; Ujvárosy, D.; Szabó, Z.; Lo˝rincz, I.; Paragh, G.; Harangi, M.; Somodi, S. The impact of obesity on the cardiovascular system. J. Diabetes Res. 2018, 2018, 3407306.

[4] Braet, C., 2007. Childhood obesity. Eating disorders in children and adolescents, pp.182-192.

[5] Lindberg, L., Hagman, E., Danielsson, P., Marcus, C. and Persson, M., 2020. Anxiety and depression in children and adolescents with obesity: a nationwide study in Sweden. BMC medicine, 18(1), pp.1-9.

[6] Oppert, J.M., Bellicha, A., van Baak, M.A., Battista, F., Beaulieu, K., Blundell, J.E., Carraça, E.V., Encantado, J., Ermolao, A., Pramono, A. and Farpour-Lambert, N., 2021. Exercise training in the management of overweight and obesity in adults: Synthesis of the evidence and recommendations from the European Association for the Study of Obesity Physical Activity Working Group. Obesity Reviews, 22, p.e13273.

[7] Dogan, A. and Birant, D., 2021. Machine learning and data mining in manufacturing. Expert Systems with Applications, 166, p.114060.

[8] dos Santos, B.S., Steiner, M.T.A., Fenerich, A.T. and Lima, R.H.P., 2019. Data mining and machine learning techniques applied to public health problems: A bibliometric analysis from 2009 to 2018. Computers & Industrial Engineering, 138, p.106120.

[9] dos Santos, B.S., Steiner, M.T.A., Fenerich, A.T. and Lima, R.H.P., 2019. Data mining and machine learning techniques applied to public health problems: A bibliometric analysis from 2009 to 2018. Computers & Industrial Engineering, 138, p.106120.

[10] Li, Y., Li, T. and Liu, H., 2017. Recent advances in feature selection and its applications. Knowledge and Information Systems, 53(3), pp.551-577.

[11] Singh, B.; Tawfik, H. A Machine Learning Approach for Predicting Weight Gain Risks in Young Adults. In Proceedings of the 10th IEEE International Conference on Dependable Systems, Services and Technologies (DESSERT), Leeds, UK, 5–7 June 2019; pp. 231–234.

[12] Singh, B. and Tawfik, H., 2020, June. Machine learning approach for the early prediction of the risk of overweight and obesity in young people. In International Conference on Computational Science (pp. 523-535). Springer, Cham.

[13] Farran, B.; AlWotayan, R.; Alkandari, H.; Al-Abdulrazzaq, D.; Channanath, A.; Thangavel, A.T. Use of Non-invasive Parameters and Machine-Learning Algorithms for Predicting Future Risk of Type 2 Diabetes: A Retrospective Cohort Study of Health Data From Kuwait. Front. Endocrinol. 2019, 10, 624.

[14] Padmanabhan, M.; Yuan, P.; Chada, G.; van Nguyen, H. Physician-friendly machine learning: A case study with cardiovascular disease risk prediction. J. Clin. Med. 2019, 8, 1050.

[15] Jindal, K.; Baliyan, N.; Rana, P.S. Obesity Prediction Using Ensemble Machine Learning Approaches. In Recent Findings in Intelligent Computing Techniques; Springer: Singapore, 2018; pp. 355–362.

[16] Selya, A.S.; Anshutz, D. Machine Learning for the Classification of Obesity from Dietary and Physical Activity Patterns. In Advanced Data Analytics in Health; Springer: Cham, Switzerland, 2018; pp. 77–97.

[17] Dunstan, J.; Aguirre, M.; Bastías, M.; Nau, C.; Glass, T.A.; Tobar, F. Predicting nationwide obesity from food sales using machine learning. Health Inform. J. 2019. Available online: https://journals.sagepub.com/doi/full/ 10.1177/1460458219845959 (accessed on 18 March 2020).

[18] Zheng, Z.; Ruggiero, K. Using machine learning to predict obesity in high school students. In Proceedings of the 2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Kansas City, MO, USA, 13–16 November 2017; pp. 2132–2138.

[19] DeGregory, K.W.; Kuiper, P.; DeSilvio, T.; Pleuss, J.D.; Miller, R.; Roginski, J.W.; Fisher, C.B.; Harness, D.; Viswanath, S.; Heymsfield, S.B.; et al. A review of machine learning in obesity. Obes. Rev. 2018, 19, 668–685.

[20] Golino, H.F.; Amaral, L.S.D.B.; Duarte, S.F.P.; Gomes, C.M.A.; Soares, T.D.J.; Reis, L.A.D.; Santos, J. Predicting increased blood pressure using machine learning. J. Obes. 2014, 2014, 637635.
[21] Pleuss, J.D.; Talty, K.; Morse, S.; Kuiper, P.; Scioletti, M.; Heymsfield, S.B.; Thomas, D.M. A machine learning approach relating 3D body scans to body composition in humans. Eur. J. Clin. Nutr. 2019, 73, 200–208.
[22] Maharana, A.; Nsoesie, E.O. Use of deep learning to examine the association of the built environment with prevalence of n eighborhood adult obesity. JAMA Netw. Open 2018, 1, e181535.
[23] Pouladzadeh, P.; Kuhad, P.; Peddi, S.V.B.; Yassine, A.; Shirmohammadi, S. Food calorie measurement using deep learning neural network. In Proceedings of the 2016 IEEE International Instrumentation and Measurement Technology Conference, Taipei, Taiwan, 23–26 May 2016; pp. 1–6.
[24] De-La-Hoz-Correa, E., Mendoza Palechor, F., De-La-Hoz-Manotas, A., Morales Ortega, R. and Sánchez Hernández, A.B., 2019. Obesity level estimation software based on decision trees. J Comput Sci, 15 (Issue 1) (2019), pp. 67-77, 10.3844/jcssp.2019.67.77 .
[25] Hsu, P.H., Lee, C.H., Kuo, L.K., Kung, Y.C., Chen, W.J. and Tzeng, M.S., 2018. Determination of the energy requirements in mechanically ventilated critically ill elderly patients in different BMI groups using the Harris–Benedict equation. Journal of the Formosan Medical Association, 117(4), pp.301-307.
[26] Wu, G., 2016. Dietary protein intake and human health. Food & function, 7(3), pp.1251-1265.
[27] Deurenberg, P., Weststrate, J.A. and Seidell, J.C., 1991. Body mass index as a measure of body fatness: age-and sex-specific prediction formulas. British journal of nutrition, 65(2), pp.105-114.
[28] Wan, C., 2019. Feature selection paradigms. In Hierarchical Feature Selection for Knowledge Discovery (pp. 17-23). Springer, Cham.
[29] Lee, S.J., Xu, Z., Li, T. and Yang, Y., 2018. A novel bagging C4. 5 algorithm based on wrapper feature selection for supporting wise clinical decision making. Journal of biomedical informatics, 78, pp.144-155.
[30] Hedar, A.R., Wang, J. and Fukushima, M., 2008. Tabu search for attribute reduction in rough set theory. Soft Computing, 12(9), pp.909-918.
[31] Prajapati, V.K., Jain, M. and Chouhan, L., 2020, February. Tabu search algorithm (TSA): A comprehensive survey. In 2020 3rd International Conference on Emerging Technologies in Computer Engineering: Machine Learning and Internet of Things (ICETCE) (pp. 1-8). IEEE.
[32] Peng, H., Long, F. and Ding, C., 2005. Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. IEEE Transactions on pattern analysis and machine intelligence, 27(8), pp.1226-1238
[33] Urbanowicz, R.J., Meeker, M., La Cava, W., Olson, R.S. and Moore, J.H., 2018. Relief-based feature selection: Introduction and review. Journal of biomedical informatics, 85, pp.189-203.
[34] Liu, J.C., 1996. Optimality and duality for generalized fractional programming involving nonsmooth (F, ρ)-convex functions. Computers & Mathematics with Applications, 32(2), pp.91-102.
[35] Jiménez, F., Palma, G.S.J., Miralles-Pechuán, L. and Botía, J., 2021. Multivariate feature ranking of gene expression data. arXiv preprint arXiv:2111.02357.
[36] Liu, H. and Setiono, R., 1996, July. A probabilistic approach to feature selection-a filter solution. In ICML (Vol. 96, pp. 319-327).
[37] Quinlan, J.R., 2014. C4. 5: programs for machine learning. Elsevier.
[38] Breiman, L., 2001. Random forests. Machine learning, 45(1), pp.5-32.
[39] Hühn, J. and Hüllermeier, E., 2009. FURIA: an algorithm for unordered fuzzy rule induction. Data Mining and Knowledge Discovery, 19(3), pp.293-319.
[40] Breiman, L., 1996. Bagging predictors. Machine learning, 24(2), pp.123-140.
[41] Lee, S.J., Xu, Z., Li, T. and Yang, Y., 2018. A novel bagging C4.5 algorithm based on wrapper feature selection for supporting wise clinical decision making. Journal of biomedical informatics, 78, pp.144-155.
[42] Fernández, A., García, S., Galar, M., Prati, R.C., Krawczyk, B. and Herrera, F., 2018. Performance measures. In Learning from Imbalanced Data Sets (pp. 47-61). Springer, Cham.