Sinopfbd, 2024; 9(1): 217-239 https://doi.org/10.33484/sinopfbd.1445215

Sinop Üniversitesi Fen Bilimleri Dergisi
Sinop Uni J Nat Sci
E-ISSN: 2564-7873

https://dergipark.org.tr/tr/pub/sinopfbd

Using Artificial Intelligence Techniques for the Analysis of Obesity Status According to the Individuals' Social and Physical Activities

Niğmet KÖKLÜ[1] and Süleyman Alpaslan SULAK[2]

How to cite: Koklu, N., & Sulak, S.A. (2024). Using artificial intelligence techniques for the analysis of obesity status according to the individuals' social and physical activities. *Sinop Üniversitesi Fen Bilimleri Dergisi, 9*(1), 217-239. https://doi.org/10.33484/sinopfbd.1445215

Research Article

Corresponding Author
Niğmet KÖKLÜ
nkoklu@ktun.edu.tr

ORCID of the Authors
N.K: 0000-0001-9563-3473
S.A.S: 0000-0001-9716-9336

Received: 29.02.2024
Accepted: 10.06.2024

Abstract

Obesity is a serious and chronic disease with genetic and environmental interactions. It is defined as an excessive amount of fat tissue in the body that is harmful to health. The main risk factors for obesity include social, psychological, and eating habits. Obesity is a significant health problem for all age groups in the world. Currently, more than 2 billion people worldwide are obese or overweight. Research has shown that obesity can be prevented. In this study, artificial intelligence methods were used to identify individuals at risk of obesity. An online survey was conducted on 1610 individuals to create the obesity dataset. To analyze the survey data, four commonly used artificial intelligence methods in literature, namely Artificial Neural Network, K Nearest Neighbors, Random Forest and Support Vector Machine, were employed after pre-processing. As a result of this analysis, obesity classes were predicted correctly with success rates of 74.96%, 74.03%, 74.03% and 87.82%, respectively. Random Forest was the most successful artificial intelligence method for this dataset and accurately classified obesity with a success rate of 87.82%.

Keywords: Obesity dataset, artificial intelligence methods, artificial neural network, support vector machine, k nearest neighbors, random forest

Kişilerin Sosyal ve Fiziksel Aktivitelerine Göre Obezite Durumunun Analizi için Yapay Zeka Tekniklerinin Kullanımı

[1] Technical Science Vocational High School, Konya Technical University, Konya, Türkiye

[2] Ahmet Keleşoğlu Educational Faculty, Necmettin Erbakan University, Konya, Türkiye

Öz

Obezite, genetik ve çevresel etkileşimlere sahip ciddi ve kronik bir hastalıktır. Sağlığa zararlı olan vücuttaki aşırı miktardaki yağ dokusu olarak tanımlanır. Obezitenin başlıca risk faktörleri, sosyal, psikolojik ve beslenme alışkanlıklarını içerir. Obezite, dünya genelinde tüm yaş grupları için önemli bir sağlık sorunudur. Şu anda dünya genelinde 2 milyardan fazla insan obez veya aşırı kilolu durumdadır. Araştırmalar, obezitenin önlenebileceğini göstermektedir. Bu çalışmada, obezite riski taşıyan bireyleri tanımlamak için yapay zeka yöntemleri kullanıldı. Obezite veri setini oluşturmak için 1610 birey üzerinde çevrimiçi bir anket yapıldı. Anket verilerini analiz etmek için literatürde yaygın olarak kullanılan dört yapay zeka yöntemi olan Yapay Sinir Ağı, K En Yakın Komşu, Rastgele Orman ve Destek Vektör Makinesi, kullanıldı. Bu analizin sonucunda, obezite sınıfları sırasıyla %74.96, %74.03, %74.03 ve %87.82 başarı oranlarıyla doğru bir şekilde tahmin edildi. Rastgele Orman, bu veri seti için en başarılı yapay zeka yöntemi oldu ve obeziteyi %87.82 başarı oranıyla doğru bir şekilde sınıflandırdı.

Anahtar Kelimeler: Obezite veri seti, yapay zeka yöntemleri, yapay sinir ağı, destek vektör makinesi, k- en yakın komşu, rastgele orman

Introduction

Nutrition is a behavior that should be done consciously to protect human health and increase the quality of life to take the nutrients needed by the body in sufficient quantities and at appropriate times. Adequate and balanced nutrition is one of the basic conditions for an individual to live healthily, develop economically and socially, and increase the level of welfare [1]. While scientific, technological, and economic developments cause a decrease in health problems related to malnutrition, they also cause problems related to overnutrition and excess energy intake [2]. Obesity refers to the accumulation of an excessive amount of adipose tissue in specific areas of the body, which may pose health risks [3]. This rise presents a risk to the typical health development in individuals of all ages. Many adults complain weight gain caused by factors associated with high-calorie food, sedentary lifestyles, and modes of transportation [4]. Obesity is caused by the fact that the energy taken by the body with food is more than the energy spent. It is crucial not only to manage the intake of energy but also ensure a balance in energy expenditure to prevent obesity [5]. Obesity is a significant health issue affecting individuals of all ages worldwide [6, 7]. Defined by the World Health Organization (WHO) as a significant determinant of death and disability, obesity stands out as the fourth most common risk factor in terms of non-communicable diseases, following high blood pressure, dietary risks, and tobacco [8, 9]. The issue of obesity has transformed into a global health challenge and is progressively escalating in numerous nations, starting from childhood and adolescence [10, 11]. It is prevalent among individuals with middle and low incomes in developed countries, while in developing nations, it is more prevalent among individuals with middle to high incomes [12]. Over the years, the prevalence of obesity has multiplied several times. As of 2016, 39% of adults were overweight and 13% were obese. Moreover, over 340 million children aged between 5-18 years were either overweight or obese [13]. Today, more than two billion people worldwide are overweight or obese [14]. Many studies have stated that obesity can be prevented [15]. Studies on the spread of diseases have recognized a high body mass index as a risk element for an increasing number of long-term illnesses [16]. Obesity can occur due to hereditary reasons. It can be leptin, monogenic, syndromic, and polygenic [17]. Additional contributing factors to obesity include social, psychological, and eating habits. The risk of obesity is affected by the environmental factor in early life [18]. Obesity can be caused by a combination of environmental, biological, and cultural factors [19, 20]. The incorporation of physical activity into daily life is considered the first step towards a healthy lifestyle and is also recognized as a crucial factor in preventing obesity [21]. Around 25-50% of the energy expended by humans is due to physical activity. As body mass increases, so does the energy used because moving a heavier object requires more energy expenditure [22]. According to a study, the number of overweight adults exceeds one billion, and body mass increases across all regions of the world when considering the entire population distribution [23]. In addition, obesity has adverse effects not only on physical health but also on psychosocial and emotional health [24]. The prejudice and negative social perspective of society against obesity can lead to psychosocial problems such as decreased self-esteem, low body image, difficulty in finding a job, and working with lower wages in individuals who have this problem [25]. Factors such as age, gender, dietary habits, sociocultural factors, and physical inactivity have a significant role in the development of obesity [26]. Research shows that children with low physical activity are more likely to be obese as observed in adults [27]. School interventions involving physical activity have been found to have a greater impact on body weight compared to interventions that do not include physical activity [28, 29]. Therefore, emphasis should be placed on physical activity during school age [30, 31]. Determining the factors affecting the prevalence of obesity is important for solving possible health problems and taking necessary precautions [32]. The primary factors responsible for the development of obesity are: a high-fat diet (HF diet), positive energy imbalances determined as genetic factors, and environmental factors [33]. Deep learning is a machine learning class. Deep learning methods have been used in many problems that are being solved with artificial intelligence and many deep learning approaches have been revealed [34]. In recent years, deep learning techniques have been used in many studies in the field of obesity. In particular, deep learning models have been developed to predict obesity and their performance has been evaluated with various metrics. For example, in the study conducted by [35], the relationships between obesity and the environment and physical activities were investigated using deep learning techniques. Similarly, in another study conducted by Alkhalaf and colleagues, a model was developed for determining and classifying the risk of obesity using deep learning methods [36]. These studies demonstrate the potential of deep learning techniques in predicting and classifying obesity.

As the prevalence of overweight and obesity continues to rise, there is an increasing need to develop new methods to predict who will benefit from dietary interventions. AI has the potential to solve problems using computer systems integrated with large data sources. In the field of nutrition, AI applications can offer benefits such as effective nutrition planning, interpretation, diet analysis, quality nutrition counseling, and providing in-depth knowledge about the effects of nutrition on health [37, 38]. Apart from the success metric, the data collection method should also be taken into account. Data mining is a field of computer science that investigates algorithms designed to learn from data, and it has been applied to various problems, including examining the efficacy of interventions [4]. With current advancements in technology, there are potentially alternative methods to provide better forecasting. Researchers and experts in this field have employed diverse data mining techniques [39, 40]. Nowadays, data mining is utilized in numerous industries, including the field of medicine, where it is used to enhance decision-making in medical affairs [41]. The surge of data in the medical field has resulted in greater levels of uncertainty, highlighting the growing significance of data mining in managing such uncertainties [42, 43]. The vast number of datasets present in biomedicine leads to a multitude of potential patterns that cannot be easily predicted, but data mining techniques offer a valuable tool for uncovering such patterns [44]. Utilizing data mining techniques can assist medical professionals in shifting their focus from population-based to individual-based approaches [45]. Considering the studies in literature, 4 different artificial intelligence methods were used in this study to determine 4 different obesity states based on 14 features. Concerning literature, this study provides the following contributions.

- Researchers working in this field will have access to a numerical dataset obtained from 1610 individuals.
- Fourteen (14) different variables were examined to determine whether individuals had obesity or not.
- Instead of statistical prediction methods, four different artificial intelligence methods were used: Artificial Neural Network (ANN), K Nearest Neighbor (KNN) Support Vector Machine (SVM), and Random Forest (RF).
- Detailed performance analyses of ANN, SVM, KNN, and RF methods were conducted.

The rest of the article is organized in the following manner: Part two presents an overview of the dataset, methods employed, and the metrics used to assess performance. Part three presents the experimental results. Finally, in part four, the study's findings and recommendations are presented.

Methods and Techniques

This section offers a comprehensive overview of the dataset employed in the study, the artificial intelligence techniques utilized, and the performance criteria used to evaluate these models. The application was developed by applying the programming language Python and its accompanying libraries. The study's general structure and operation are illustrated in Figure 1 through a flowchart.

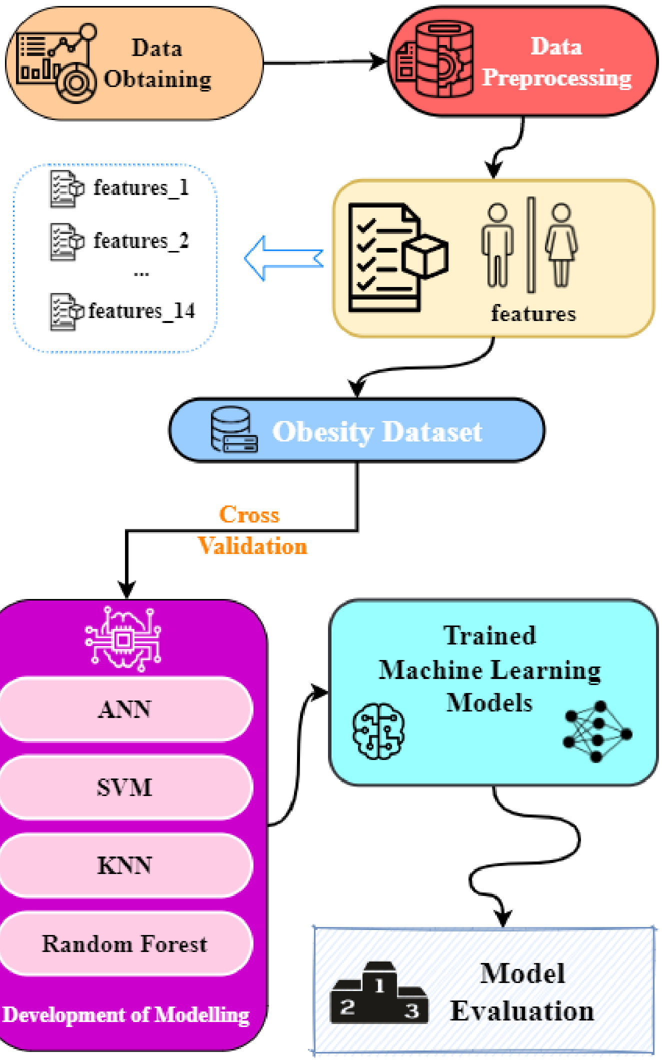

Figure 1. *Flow chart of the study*

Obesity Dataset

To create the Obesity Dataset, first of all, the variables affecting obesity were determined by examining literature. The generated dataset has 14 variables required to determine obesity. The Obesity Dataset was obtained through internet by a questionnaire applied to a total of 1610 people living in Türkiye. The data distributions of the features in the obesity dataset are given in Table 1.

Table 1. *Features and class distribution in the obesity dataset*

Attributes	Features	Values
Sex	1. Male	712
	2. Female	898
Age	Values in integers	
Height	Values in integers (cm)	
Overweight/Obese Families	1. Yes	266
	2. No	1344
Consumption of Fast Food	1. Yes	436
	2. No	1174
Frequency of Consuming Vegetables	1. Rarely	400
	2. Sometimes	708
	3. Always	502
Number of Main Meals Daily	1. 1-2	444
	2. 3	928
	3. 3+	238
Food Intake Between Meals	1. Rarely	346
	2. Sometimes	564
	3. Usually	417
	4. Always	283
Smoking	1. Yes	492
	2. No	118
Liquid Intake Daily	1. Amount smaller than one liter	456
	2. Within the range of 1 to 2 liters	523
	3. In excess of 2 liters	631
Calculation Of Calorie Intake	1. Yes	286
	2. No	1324
Physical Exercise	1. No physical activity	206
	2. In the range of 1-2 days	290
	3. In the range of 3-4 days	370
	4. In the range of 5-6 days	358
	5. 6+ days	386
Schedule Dedicated to Technology	1. Between 0 and 2 hours	382
	2. Between 3 and 5 hours	826
	3. Exceeding five hours	402
Type of Transportation Used	1. Automobile	660
	2. Motorbike	94
	3. Bike	116
	4. Public transportation	602
	5. Walking	138
Class	1. Underweight	73
	2. Normal	658
	3. Overweight	592
	4. Obesity	287

A visualized form of how data is dispersed within the Obesity dataset is described in Figure 2. As seen in Figure 2, the obesity dataset consists of a total of 1610 individuals. Among these, 898 are female and 712 are male. The youngest participant in the dataset is 18 years old, and the oldest participant is 54 years old.

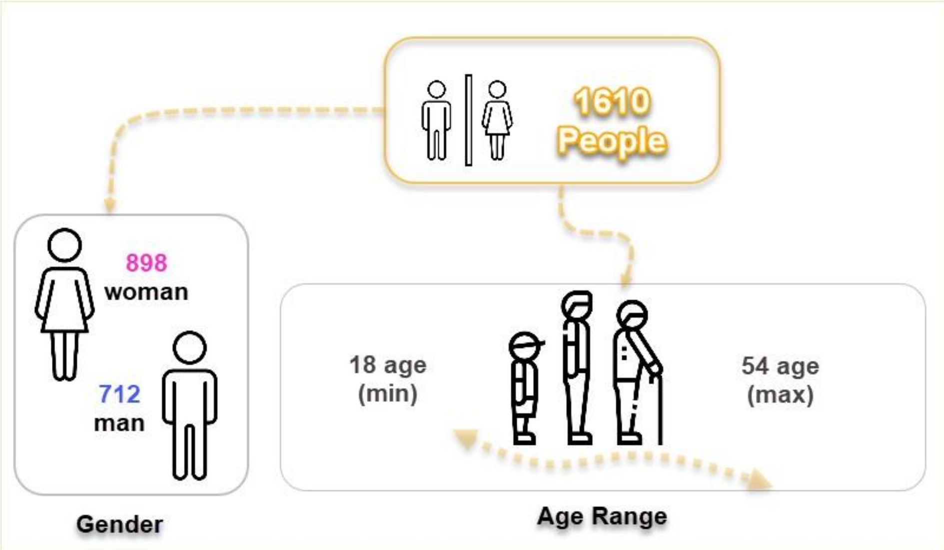

Figure 2. *Information about the study group that created the obesity dataset*

Confusion Matrix and Performance Metrics

A confusion matrix is a tabular representation implemented for evaluating the accuracy of classification in artificial intelligence methods. The table contains four distinct values, and based on these values, the performance indicators of the artificial intelligence model can be processed. Table 2 illustrates a two-class confusion matrix with its corresponding definitions.

Table 2. *Two-class confusion matrix*

Evaluation Criteria	**Definition**
True positive (**TP**)	The number of specimens that are predicted as well as actually classified as positive.
False positive (**FP**)	The number of specimens that are predicted as positive but actually fall under the negative category.
True negative (**TN**)	The number of specimens that are classified as negative as well as predicted as negative.
False negative (**FN**)	The number of specimens that are predicted as negative but actually belong to the positive category.

Artificial intelligence models can be evaluated using a variety of metrics to assess their performance [46, 47]. In this study, commonly used metrics such as Precision (P), Recall (R), F-measure (F), and Accuracy (AC) were utilized to evaluate the performance of artificial intelligence models. The computation of these metrics for a two-class confusion matrix is as follows; **Accuracy (AC):** refers to the total number of correctly classified records by a classifier. It is measured as the percentage of correctly classified test sets based on the model, defining how accurate the classifier is [48].

$$\text{Accuracy (\%)} = \frac{\text{TP + TN}}{\text{TP + FP + FN + TN}} \times 100 \quad (1)$$

Precision (P): The proportion of true positive samples to all samples that are classified as positive [48, 49].

$$\text{Sensitivity (\%)} = \frac{\text{TP}}{\text{TP + FP}} \times 100 \quad (2)$$

Recall (R): As a measure of accurate identification of positive samples, it refers to the true positive rate [48, 50].

$$\text{Recall (\%)} = \frac{\text{TP}}{\text{TP + FN}} \times 100 \quad (3)$$

F1-Score (F): F-measure is calculated by combining precision and recall metrics to evaluate the model's performance [46].

$$\text{F1} - \text{Score} = \frac{2 \times \text{TP}}{2 \times \text{TP+FP + FN}} \times 100 \quad (4)$$

A four-class confusion matrix was utilized in this study due to the presence of four classes in the dataset. Table 3 depicts the definitions and values of the four-class confusion matrix.

Table 3. *Obesity data confusion matrix*

(4 x 4) Multi-Class		**Predicted Class**			
		Underweight	Normal	Overweight	Obesity
Actual Class	Underweight	**T1**	F_{12}	F_{13}	F_{14}
	Normal	F_{21}	**T2**	F_{23}	F_{24}
	Overweight	F_{31}	F_{32}	$\mathbf{T_3}$	F_{34}
	Obesity	F_{41}	F_{42}	F_{43}	$\mathbf{T_4}$

Table 4 illustrates the computation of TP, FP, TN, and FN values utilizing the data from a four-class confusion matrix.

Table 4. *Calculating TP, FP, TN, and FN values through a confusion matrix with four classes*

CLASS	TP	TN	FP	FN
C1	$TP_1 = T_1$	$TN_1 = T_2 + T_3 + T_4 + F_{23} + F_{24} + F_{32} + F_{34} + F_{42} + F_{43}$	$FP_1 = F_{21} + F_{31} + F_{41}$	$FN_1 = F_{12} + F_{13} + F_{14}$
C2	$TP_2 = T_2$	$TN_2 = T_1 + T_3 + T_4 + F_{13} + F_{14} + F_{31} + F_{41} + F_{34} + F_{43}$	$FP_2 = F_{12} + F_{32} + F_{42}$	$FN_2 = F_{21} + F_{23} + F_{24}$
C3	$TP_3 = T_3$	$TN_3 = T_1 + T_2 + T_4 + F_{12} + F_{14} + F_{21} + F_{24} + F_{41} + F_{42}$	$FP_3 = F_{13} + F_{23} + F_{43}$	$FN_3 = F_{31} + F_{32} + F_{34}$
C4	$TP_4 = T_4$	$TN_4 = T_1 + T_2 + T_3 + F_{12} + F_{13} + F_{21} + F_{23} + F_{31} + F_{32}$	$FP_4 = F_{14} + F_{24} + F_{34}$	$FN_4 = F_{41} + F_{42} + F_{43}$

Cross Validation

One of the popular techniques for adjusting model parameters involves sampling data to detect errors in the actual and predicted outcomes. Additionally, the cross-validation approach is utilized to enhance classification reliability. This method randomly divides the dataset into equal clusters based on a specified number, as illustrated in Figure 3. One cluster is designated as the test set while the rest are used for training, and this process is repeated for each cluster, with the test set changing at each iteration. Cross-validation is comparable to the repeated random subsampling method, except that the partitions are created such that there is no overlap between the two test sets [51]. For this research, a cross-validation value of 10 was established, which led to the partitioning of the dataset into 10 distinct portions. The training and testing procedures were executed ten times, with nine portions designated for training and one for testing in each iteration. To obtain the overall performance metrics, the results were averaged arithmetically.

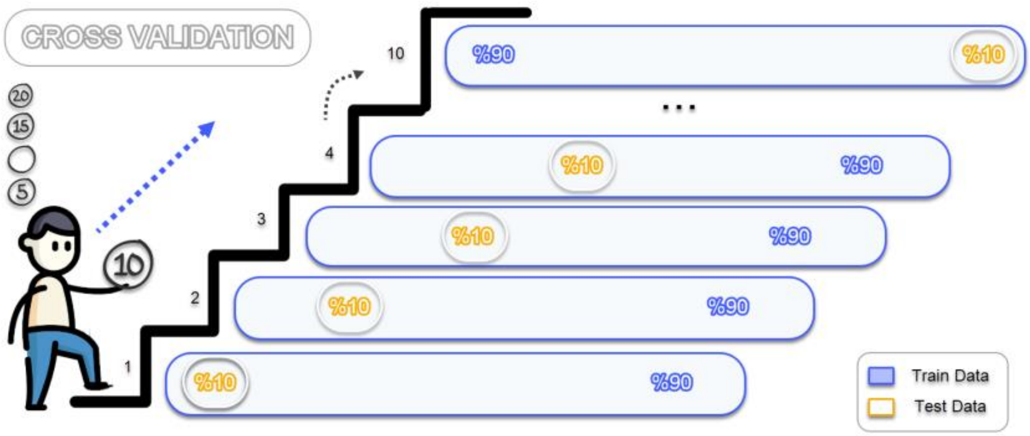

Figure 3. *Cross validation scheme used in the study*

Artificial Intelligence Models

Four distinct artificial intelligence techniques were utilized in this study, and a description of each approach is provided in the subsequent sections.

Artificial neural networks (ANN)

Artificial neural networks (AI), in which basic functions such as the ability to generate new data from data collected by the brain by learning, remembering, and generalizing by imitating the learning path of the human brain, are performed by computer software [52]. The human brain is a sophisticated system with the capacity to store and process information to solve problems. Neurons are the fundamental building blocks of this intricate structure. In recent times, one of the benefits of artificial intelligence (AI) applications is their ability to improve the accuracy and ease of using models for complex natural systems with large inputs. From the past to the present, ANN has been used in many fields and applications [53]. Artificial Neural Networks (ANN) is a popular and effective model utilized in problem-solving and machine learning [54, 55]. ANNs are composed of simple processing units called neurons, which are extensively distributed in parallel and have the innate ability to store and retrieve relevant experimental information. An artificial neuron or processing element is the fundamental unit that constitutes the foundation of the system [56]. Artificial Neural Networks (ANNs) possess numerous characteristics that make them a highly potent and favored tool for modeling, prediction, and performance optimization of diverse systems. These features include high-speed information processing, mapping capabilities, fault tolerance, generalization, and robustness [54]. Neural networks have three primary types of layers: input, hidden, and output layers. When a neural network has multiple hidden layers, it is referred to as a multilayer neural network or deep neural network.

In Figure 4, it can be seen that each input node is linked to all nodes in the subsequent hidden layer. The neural network learns from examples through repeated connections in the subsequent steps, ultimately generating the corresponding outputs [57].

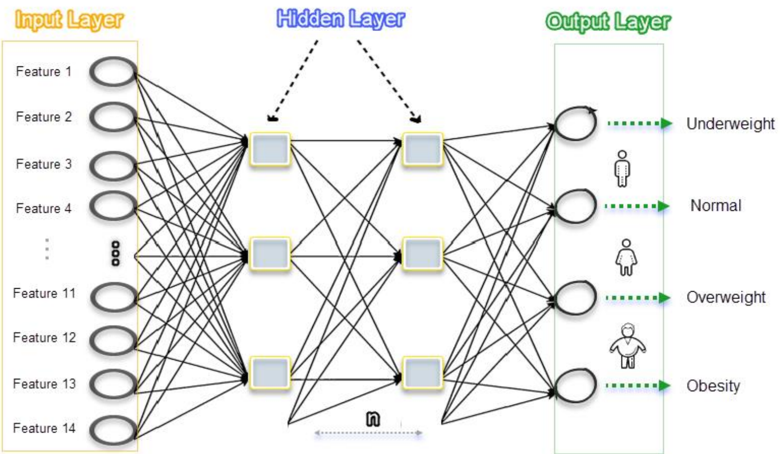

Figure 4. *The general working structure of the ANN model*

Support Vector Machine (SVM)

Support Vector Machines (SVMs) are a widely employed and highly effective machine-learning technique for data classification [56]. SVMs are supervised learning models based on statistical learning theory, which involve learning algorithms that analyze the data for classification and regression analysis. This method transforms the initial input space into a higher-dimensional feature space to enhance the classification process. With SVM, limits can be defined for both linear and nonlinear datasets. Support Vector Machines (SVM) have become widely favored due to their capability to identify optimal hyperplanes that maximize the separation between classes in the feature space [58]. The fundamental principle underlying SVM involves separating classes by drawing margins between them. As illustrated in Figure 5, these margins are calculated in a manner that maximizes the distance between the margin and the classes, thereby minimizing classification error [48, 59-61].

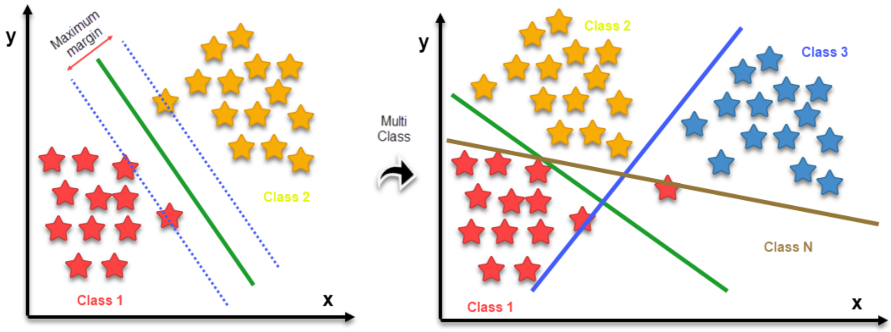

Figure 5. *The general working structure of the SVM model*

Support Vector Machines (SVM) aim to pinpoint the optimal hyperplane that minimizes classification errors while maximizing the margin, a crucial space between data points. This pursuit involves a cost parameter, which balances the desire to maximize accuracy while minimizing misclassifications [62]. Ultimately, SVM strives to delineate a decision boundary that effectively separates different classes of data, ensuring accurate classification of new, unseen data points while guarding against overfitting [63]. To achieve this, SVM employs kernel functions like linear, polynomial, or radial basis function (RBF) to transform input data into higher-dimensional spaces. Within these transformed spaces, SVM constructs a hyperplane that efficiently segregates data points into their respective classes, relying solely on a subset of training data points known as support vectors [64]. This streamlined approach enhances memory efficiency and effectiveness, especially in high-dimensional fields. One key advantage of SVM lies in its adeptness at handling high-dimensional data, rendering it suitable for applications with numerous features. Additionally, SVM boasts a solid theoretical foundation and demonstrates reduced vulnerability to overfitting compared to alternative machine learning methods. However, the performance of SVM hinges heavily on the selection of appropriate kernel functions and parameters, necessitating careful tuning to achieve optimal results [65]. SVM emerges as a versatile and robust technique widely applied across various domains such as text categorization, image recognition, and bioinformatics, owing to its ability to efficiently tackle complex classification tasks while mitigating the risk of overfitting.

K Nearest Neighbors (KNN)

K Nearest Neighbor (KNN) is a non-generalizing or sample-based learning algorithm, also referred to as a "lazy learning" algorithm, that gained popularity in statistical data analysis during the early 1970s. Unlike other algorithms, KNN does not concentrate on constructing an internal model by storing all training data samples in n-dimensional space [66-68]. KNN classifies new data points according to similarity measurements using the data. As seen in Figure 6, this algorithm calculates the classification according to the k value given as a parameter by the vote of the simple majority of its nearest neighbors and classifies accordingly. Noisy training data is also very effective, but the accuracy value varies depending on the quality of the data [68].

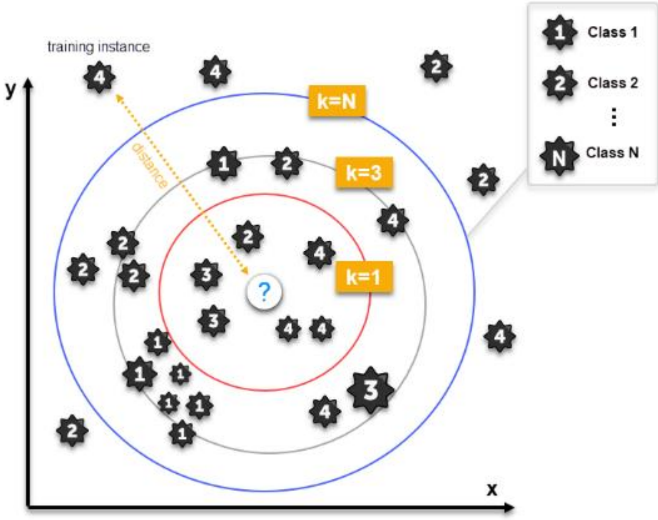

Figure 6. *The general working structure of the KNN model*

Random Forest (RF)

Random Forest is an ensemble learning approach proposed and developed by Breiman to solve classification and regression problems [69]. This algorithm creates many decision trees in general structure and combines them to get the best result [70]. The algorithm is extended by training each tree with its unique randomly generated subset of the training data, utilizing only a subset of the variables for that specific tree. Multiple classifiers are compared using ensemble return values to attain more precise results than a single classifier. To classify a sample, each tree in the forest is provided with an input vector, and a result is generated for each tree. The algorithm then branches each node using the best variable, chosen randomly at each node. The decision trees are in their largest form and are untrimmed. Data that is not utilized in training, known as out-of-bag data, can be used to provide an independent estimate of the overall accuracy of classification [56, 71, 72]. The working scheme of the Random Forest algorithm for the study is shown in Figure 7.

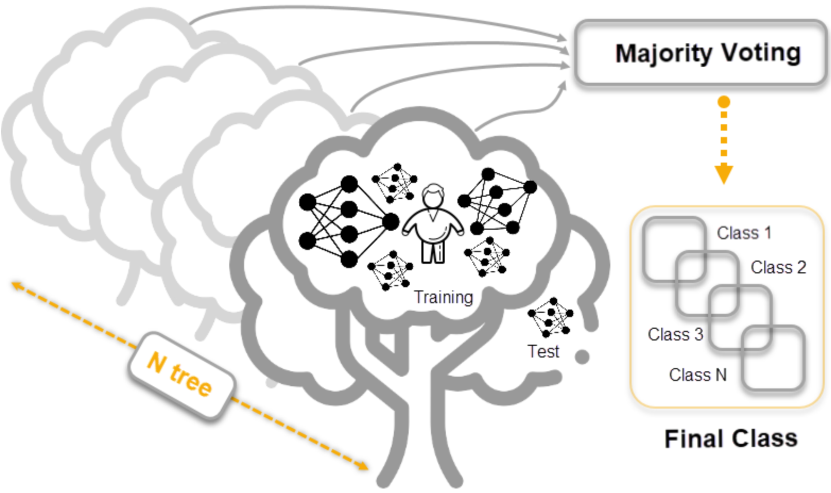

Figure 7. *Random forest general study structure*

Receiver Operating Characteristic (ROC) Curve

Evaluating the performance of artificial intelligence methods is crucial. The ROC curve is a probability curve, and the AUC-ROC curve is a commonly used metric nowadays. AUC stands for "area under the ROC curve," and it measures the extent to which the model can differentiate between classes. The ROC plot offers a means to visually represent, organize, and compare classifiers according to their performance [50, 73] (Fawcett, 2006). Spackman was one of the pioneers in using ROC charts for assessing and contrasting algorithms in machine learning [74]. The analysis of ROC curves is a common method to evaluate and compare the accuracy of classifiers over a spectrum of sensitivity and specificity values. Different parametric regression models have been suggested to model and predict ROC curves [75-77]. It is preferable to have a high AUC value in ROC curve graphs. Figure 8 provides some examples of ROC-AUC graphs. A higher AUC value indicates a better predictive ability of the machine learning model [78] (Narkhede, 2018).

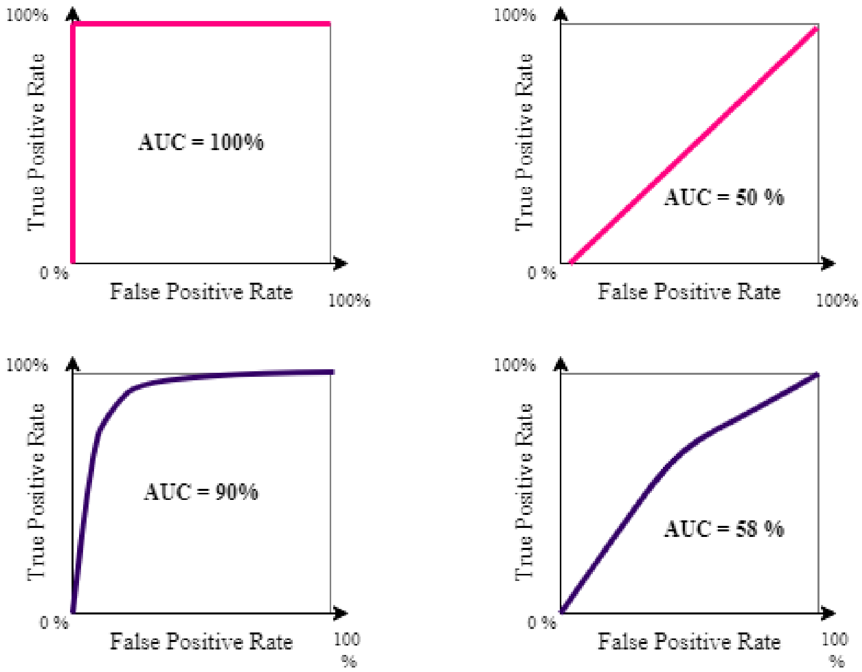

Figure 8. *ROC-AUC performance examples*

Results

This section presents experimental results obtained by training and testing artificial intelligence methods (ANN, SVM, KNN, and RF). In the subsequent section, the classification outcomes of the models employed in the investigation are outlined.

ANN Architecture-Based Classification Results

Table 5 presents the classification values obtained from the confusion matrix of the ANN model. Based on these values, the accuracy rates for the underweight, normal, overweight, and obesity categories were calculated. The results indicate an overall classification success of 74.96%, as shown in Figure 9. However, the confusion matrix also reveals that the overweight and obesity categories were less accurately differentiated.

Table 5. *ANN confusion matrix*

ANN		Predicted Class			
(4 x 4)		Underweight	Normal	Overweight	Obesity
Actual Class	Underweight	**40**	28	2	3
	Normal	15	**536**	89	18
	Overweight	2	86	**445**	59
	Obesity	2	15	84	**186**

When Table 5 is examined, it is seen that 40 data in the Underweight class, 536 in the Normal class, 445 in the Overweight class, and 186 in the Obesity class are correctly classified. When TP, TN, FP, and FN data are examined, the most successful class is observed to be Underweight. It has been determined that the Normal, Overweight, and Obesity classes are highly mixed with each other. The reason is that, the training of the models is not fully realized due to the data imbalance between the classes. Classes are confused with each other since the data of the classes are very close to each other.

SVM Architecture-Based Classification Results

The classification results for the SVM model are presented in this section, with the corresponding confusion matrix given in Table 6. The accuracy rates of the underweight, normal, overweight, and obesity classes were calculated using this matrix, and it was found that the model achieved a classification success rate of 74.03%, as shown in Figure 9. However, the confusion matrix in Table 6 indicates that the SVM model had difficulty differentiating between the overweight and obesity classes.

Table 6. *SVM Confusion matrix*

SVM		Predicted Class			
(4 x 4)		Underweight	Normal	Overweight	Obesity
Actual Class	Underweight	**30**	37	4	2
	Normal	6	**563**	78	11
	Overweight	0	98	**426**	68
	Obesity	4	15	95	**173**

Upon analyzing Table 6, it is observed that the SVM model accurately classified 30 instances of the Underweight class, 563 instances of the Normal class, 426 instances of the Overweight class, and 173 instances of the Obesity class. A closer look at the TP, TN, FP, and FN data indicates that the Normal class had the highest classification performance. However, due to limited data availability, the Underweight class demonstrated the lowest classification success.

KNN Architecture-Based Classification Results

Upon analyzing the classification values in Table 7, the accuracy rates of the underweight, normal, overweight, and obesity classes were computed for the KNN model. Figure 9 demonstrates a classification success of 74.03%. However, it is evident from the confusion matrix in Table 7 that the differentiation performance of the overweight and obesity classes is inadequate.

Table 7. *KNN confusion matrix*

KNN		**Predicted Class**			
(4 x 4)		Underweight	Normal	Overweight	Obesity
Actual Class	Underweight	**55**	17	1	0
	Normal	15	**551**	84	8
	Overweight	3	90	**459**	40
	Obesity	0	3	51	**233**

Upon examining Table 7, it can be observed that the number of correctly classified data is 55 in the Underweight class, 551 in the Normal class, 459 in the Overweight class, and 233 in the Obesity class for the KNN model. Analysis of the TP, TN, FP, and FN data reveals that the most successful class is Obesity. However, it is evident that the classification success of the Underweight class is the lowest due to the scarcity of data.

RF Architecture-Based Classification Results

After analyzing the classification results of the ANN model, as shown in Table 8, the accuracy rates of underweight, normal, overweight, and obesity classes were calculated. Figure 9 illustrates that the ANN model achieved the highest classification success rate of 87.82% accuracy. However, the confusion matrix in Table 8 indicates that the normal and overweight classes had low performance in terms of differentiation. Upon reviewing Table 8, it can be observed that 49 data in the Underweight class, 607 in the Normal class, 520 in the Overweight class, and 238 in the Obesity class are classified accurately. Upon analyzing the TP, TN, FP, and FN data, it can be deduced that the Normal class has the highest classification success. Due to the limited amount of data, the Underweight class has the lowest classification success.

Table 8. *RF confusion matrix*

RF		**Predicted Class**			
(4 x 4)		Underweight	Normal	Overweight	Obesity
Actual Class	Underweight	**49**	20	3	1
	Normal	2	**607**	46	3
	Overweight	0	57	**520**	15
	Obesity	0	3	46	**238**

Results of All Classification Models

The accuracy, precision, recall, and F1 Score values of each model were obtained using the confusion matrix data, and the results and graph for all models are shown in Figure 9.

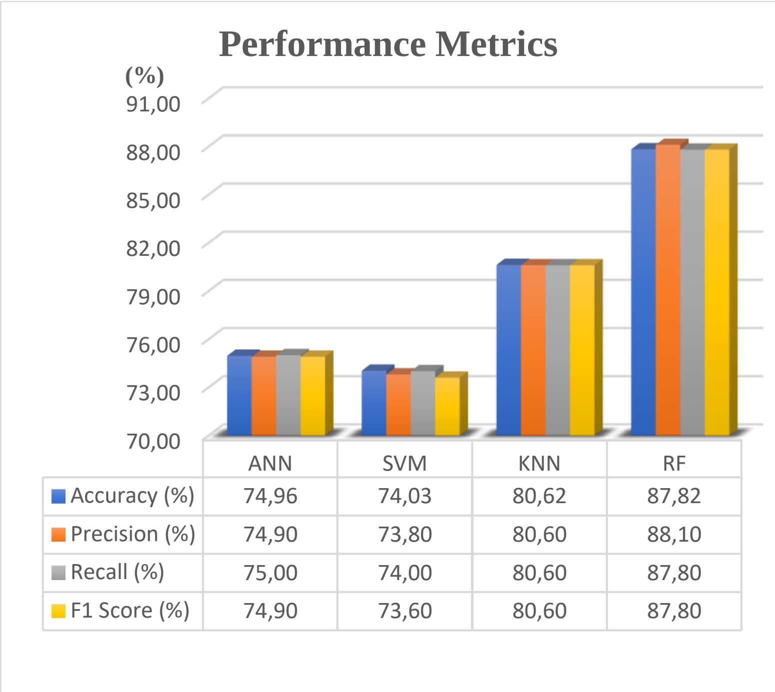

Figure 9. *Performance metrics were obtained for ANN, SVM, KNN, and RF*

Upon analyzing the outcomes presented in Figure 9, it was revealed that the RF model has the highest classification success rate (87.12%), whereas the SVM model has the lowest (74.03%). In a similar vein, the RF model exhibits the highest metric values beyond classification success, whereas the SVM model displays the lowest values. ROC curves were obtained to conduct a thorough analysis of the model's performance, and their results are shown in Figure 10 for all models.

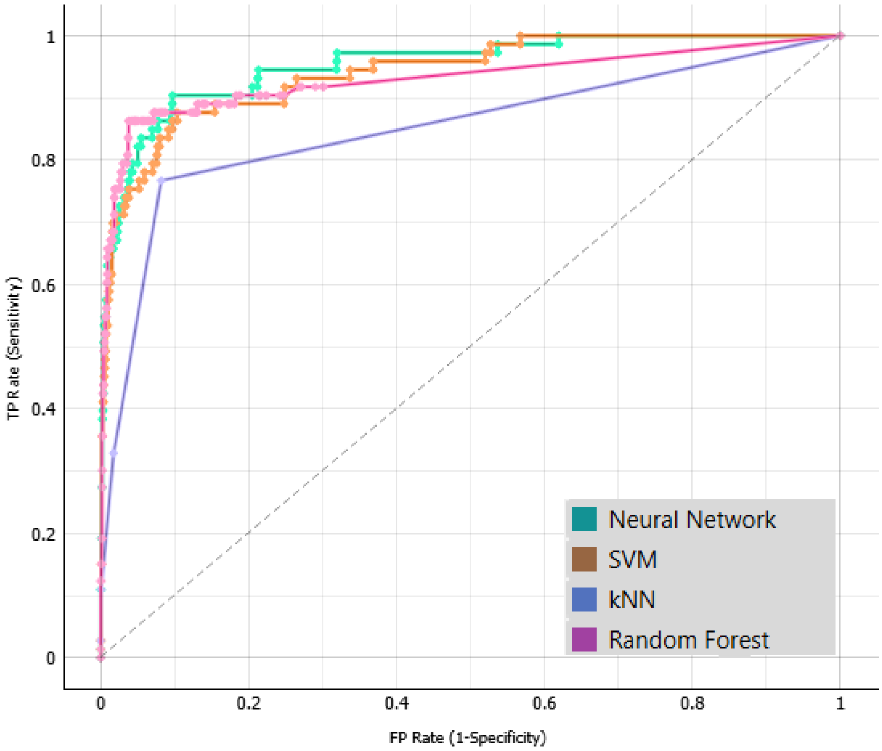

Figure 10. *ROC curves of all models*

Upon examining the ROC curves presented in Figure 10, it can be observed that the RF model achieved the most successful learning performance while the SVM model had the least successful learning performance.

Conclusions and Discussion

Obesity has recently become a rapidly increasing disease on all continents. Many diseases can occur in humans due to obesity. For this reason, it is necessary to determine the risk of obesity with regular controls and take the necessary precautions. The need for artificial intelligence methods has recently increased to accurately and quickly perform these operations. In this study, various machine learning techniques were employed to determine the obesity status of individuals using a dataset consisting of 14 features and four classes. The classification processes were conducted using ANN, SVM, KNN, and RF methods on data collected from 1610 individuals. Cross-validation was used for training and testing purposes. The RF model yielded the highest classification accuracy of 87.82%, while the SVM model had the lowest accuracy of 74.03%. The ANN model achieved a classification success of 74.96%, while the KNN model achieved a classification success of 80.62%. To achieve higher classification success, the amount of data in the dataset should be increased. In addition, equal distribution of data between classes will increase the success of classification. Classification success can be increased by using different machine learning methods. In addition, the selection of features that are effective in classification can be determined by optimization or feature selection methods, which can increase classification success and speed.

Acknowledgement -

Funding/Financial Disclosure This research received no external funding.

Ethics Committee Approval and Permissions This research was conducted with respect to participant rights and in accordance with principles of scientific integrity. Participant privacy and security were prioritized throughout the study. The ethics committee document of the research was received with decision number 2023/201 at the meeting numbered 06 of Necmettin Erbakan University Social and Human Sciences Scientific Research Ethics Committee dated 12/05/2023

Conflicts of 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.

Authors Contribution Conceptualization, Niğmet Köklü, and Süleyman Alpaslan Sulak; methodology, Niğmet Köklü, and Süleyman Alpaslan Sulak; software, Niğmet Köklü, and Süleyman Alpaslan Sulak; validation, Niğmet Köklü, and Süleyman Alpaslan Sulak; formal analysis, Niğmet Köklü; investigation, Niğmet Köklü; resources, Süleyman Alpaslan Sulak; data curation, Niğmet Köklü; writing-original draft preparation, Niğmet Köklü, and Süleyman Alpaslan Sulak; writing-review and editing, Niğmet Köklü, and Süleyman Alpaslan Sulak; visualization, Süleyman Alpaslan Sulak; supervision, Süleyman Alpaslan Sulak; All authors have read and agreed to the published version of the manuscript.

References

[1] Yetkin F. (2008). Konya il merkezinde özel hastanelere başvuran 18-60 yaş grubu kadınların obezite prevalansı and bunu etkileyen etmenler üzerine bir araştırma. Yayınlanmamış [Yüksek Lisans Tezi]. Konya. *s.* 66.

[2] Lakdawalla D & Philipson T. (2009). The growth of obesity and technological change. *Economics & Human Biology, 7*:283-293. https://doi.org/10.1016/j.ehb.2009.08.001

[3] Tan, K. C. B. (2004). Appropriate body-mass index for Asian populations and its implications for policy and intervention strategies. *The lancet*. http://dx.doi.org/10.1016/S0140-6736(03)15268-3

[4] Cervantes, R. C & Palacio, U. M. (2020). Estimation of obesity levels based on computational intelligence. *Informatics in Medicine Unlocked*, *21*, 100472. https://doi.org/10.1016/j.imu.2020.100472

[5] Hill, J. O., Wyatt, H. R & Peters, J. C. (2012). Energy balance and obesity. *Circulation*, *126*, 126-132. https://doi.org/10.1161/CIRCULATIONAHA.111.087213

[6] Kopelman, P. G. (2000). Obesity as a medical problem. *Nature, 404*, 635-643. https://doi.org/10.1038/35007508

[7] Deckelbaum, R. J., & Williams, C. L. (2001). Childhood obesity: the health issue. *Obesity Research. 9*, 239-243. https://doi.org/10.1038/oby.2001.125

[8] Turan, T. (2024). Optimize edilmiş denetimli öğrenme algoritmaları ile obezite analizi ve tahmini. *Mehmet Akif Ersoy Üniversitesi Fen Bilimleri Enstitüsü Dergisi*, *14*(2), 301-312. https://doi.org/10.29048/makufebed.1372323

[9] Vizmanos, B., Cascales, A. I., Rodríguez-Martín, M., Salme-rón, D., Morales, E., Aragón-Alonso, A., Garaulet, M. (2023). Lifestyle mediators of associations among siestas, obesity, and metabolic health. *Obesity, 31*(5): 1227-1239. https://doi.org/10.1002/oby.23765

[10] Ogden, C. L., Carroll, M. D., Curtin, L. R., McDowell, M. A., Tabak, C. J., & Flegal, K. M. (2006). Prevalence of overweight and obesity in the United States, 1999-2004. *Jama*, *295*(13), 1549-1555. https://doi.org/10.1001/jama.295.13.1549

[11] Ng, M., Fleming, T., Robinson, M., Thomson, B., Graetz, N., Margono, C., ... & Gakidou, E. (2014). Global, regional, and national prevalence of overweight and obesity in children and adults during 1980–2013: a systematic analysis for the Global Burden of Disease Study 2013. *The Lancet*, *384*(9945), 766-781. https://doi.org/10.1016/S0140-6736(14)60460-8

[12] Dinsa, G. D, Goryakin, Y., Fumagalli, E., & Suhrcke, M. (2012). Obesity and socioeconomic status in developing countries: a systematic review. *Obesity Reviews, 13,* 1067-1079. https://doi.org/10.1111/j.1467-789X.2012.01017.x

[13] Stavridou, A., Kapsali, E., Panagouli, E., Thirios, A., Polychronis, K., Bacopoulou, F., Psaltopoulou, T., Tsolia, M., Sergentanis, T. N., & Tsitsika, A. (2021). Obesity in children and adolescents during COVID-19 pandemic. *Children, 8*(2), 135. https://doi.org/10.3390/children8020135

[14] Ryan, D., Barquera, S., Barata Cavalcanti, O., & Ralston, J. (2021). The global pandemic of overweight and obesity: Addressing a twenty-First century multifactorial disease. In *Handbook of global health* (pp. 739-773). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-45009-0_39

[15] Fock, K. M., & Khoo, J. (2013). Diet and exercise in management of obesity and overweight. *Journal of Gastroenterology and Hepatology*, *28*, 59-63. https://doi.org/10.1111/jgh.12407

[16] The GBD 2015 Obesity Collaborators. (2017). Health effects of overweight and obesity in 195 countries over 25 years. *New England Journal of Medicine*, *377*(1), 13-27. https://doi.org/10.1056/NEJMoa1614362

[17] Hainerová, I. A., & Lebl, J. (2013). Treatment options for children with monogenic forms of obesity. *Nutrition and Growth*, *106*, 105-112. https://doi.org/10.1159/000342556

[18] Reilly, J. J., Armstrong, J., Dorosty, A. R., Emmett, P. M., Ness, A., Rogers, I., Steer, C., Sherriff, A. & Avon Longitudinal Study of Parents and Children Study Team (2005). Early life risk factors for obesity in childhood: cohort study. *The BMJ*, *330*(7504), 1357. https://doi.org/10.1136/bmj.38470.670903.E0

[19] Lopez, R. P. (2007). Neighborhood risk factors for obesity. *Obesity*, *15*(8), 2111-2119. https://doi.org/10.1038/oby.2007.251

[20] Komurcu, A. & Derin, D. O. (2024). Sosyal medya kullanımının beden algısı ve yeme tutumuna etkisi. *Beslenme Bilimleri Alanında Uluslararası Araştırmalar I, 57*.

[21] Yazıcı-Gulay, M., Korkmaz, Z., Erten, Z. K., & Gürbüz, K. (2021). Çocukların fiziksel aktivite, obezite düzeylerinin incelenmesi: Kayseri ili örneği. *Genel Sağlık Bilimleri Dergisi*, *3*(3), 228-238. https://doi.org/10.51123/jgehes.2021.32

[22] Prentice, A. M., Black, A. E., Coward, W. A., & Cole, T. J. (1996). Energy expenditure in overweight and obese adults in affluent societies: an analysis of 319 doubly-labelled water measurements. *European Journal of Clinical Nutrition*, *50*(2), 93-97.

[23] Finucane, M. M, Stevens, G. A., Cowan, M. J, Danaei, G., Lin, J. K., Paciorek, C. J., Singh, G. M., Gutierrez, H. R., Lu, Y., & Bahalim, A. N. (2011). National, regional, and global trends in body-mass index since 1980: systematic analysis of health examination surveys and epidemiological studies with 960 country-years and 9· 1 million participants. *The Lancet*, *377*, 557-567. https://doi.org/10.1016/S0140-6736(10)62037-5

[24] Reinehr, T. (2010). Obesity and thyroid function. *Molecular and Cellular Endocrinology, 316*, 165-171. https://doi.org/10.1016/j.mce.2009.06.005

[25] Friedman, K. E, Reichmann, S. K., Costanzo, P. R & Musante, G. J. (2002). Body image partially mediates the relationship between obesity and psychological distress. *Obesity Research, 10,* 33-41. https://doi.org/10.1038/oby.2002.5

[26] Bakhshi, E., Eshraghian, M. R., Mohammad, K., Foroushani, A. R., Zeraati, H., Fotouhi, A., Siassi, F., & Seifi, B. (2008). Sociodemographic and smoking associated with obesity in adult women in Iran: results from the National Health Survey. *Journal of Public Health, 30,* 429-435. https://doi.org/10.1093/pubmed/fdn024

[27] Hills, A. P., Andersen, L. B., & Byrne, N. M. (2011). Physical activity and obesity in children. *British Journal of Sports Medicine, 45*(11), 866-870. https://doi.org/10.1136/bjsports-2011-090199

[28] Summerbell, C. D., Waters, E., Edmunds, L., Kelly, S. A., Brown, T., & Campbell, K. J. (2005). Interventions for preventing obesity in children. *Cochrane Database of Systematic Reviews*, (3). https://doi.org/10.1002/14651858.CD001871.pub2

[29] Jurić, P., Jurak, G., Morrison, S. A., Starc, G., & Sorić, M. (2023). Effectiveness of a population-scaled, school-based physical activity intervention for the prevention of childhood obesity. Obesity, 31(3), 811-822. https://doi.org/10.1002/oby.23695

[30] Strong, W. B., Malina, R. M., Blimkie, C. J., Daniels, S. R., Dishman, R. K., Gutin, B., Hergenroeder, A. C., Must, A., Nixon, P. A , Pivarnik, J M., Rowland, T., Trost, S., & Trudeau, F. (2005). Evidence based physical activity for school-age youth. *The Journal of Pediatrics*, *146*(6), 732-737. https://doi.org/10.1016/j.jpeds.2005.01.055

[31] Sember, V., Jurak, G., Kovač, M., Morrison, S. A., & Starc, G. (2020). Children's physical activity, academic performance, and cognitive functioning: a systematic review and meta-analysis. *Frontiers in Public Health*, *8*, 307. https://doi.org/10.3389/fpubh.2020.00307

[32] Canoy, D., & Buchan, I. (2007). Challenges in obesity epidemiology. *Obesity Reviews*, *8*, 1-11. https://doi.org/10.1111/j.1467-789X.2007.00310.x

[33] Moreno, L. A., & Rodriguez, G. (2007). Dietary risk factors for development of childhood obesity. *Current Opinion in Clinical Nutrition & Metabolic Care*, *10*(3), 336-341. https://doi.org/10.1097/MCO.0b013e3280a94f59

[34] Akın, E., & Şahin, M. E. (2024). Derin öğrenme ve yapay sinir ağı modelleri üzerine bir inceleme. *EMO Bilimsel Dergi*, *14*(1), 27-38

[35] Maharana, A., & Nsoesie, E. O. (2018). Use of deep learning to examine the association of the built environment with prevalence of neighborhood adult obesity. *JAMA Network Open, 1*(4), 181535-181535. https://doi.org/10.1001/jamanetworkopen.2018.1535

[36] Alkhalaf, M., Yu, P., Shen, J., & Deng, C. (2022). A review of the application of machine learning in adult obesity studies. *Applied Computing and Intelligence, 2*(1), 32-48. https://doi.org/10.3934/aci.2022002

[37] Uribe, A. L. M., & Patterson, J. (2023). Are nutrition professionals ready for artificial intelligence? *Journal of Nutrition Education and Behavior*, *55*(9), 623. https://doi.org/10.1016/j.jneb.2023.07.007

[38] Atasoy, Z. B. K., Avcı, E., Beydoğan, R., Ozdemir, E., & Göktaş, P. (2024). Yapay Zeka ve Beslenme. In Göç, Ö. (Ed). *Sağlık&Bilim 2023 Yeni Nesil Teknolojiler.* Efeakademi Yayınları. https://doi.org/10.59617/efepub202367

[39] Masethe, H. D & Masethe, M. A. (2014, 22-24 October). *Prediction of heart disease using classification algorithms.* Proceedings of the world Congress on Engineering and computer Science. San Francisco, USA

[40] Tekin, N. (2023). Eğitimde yapay zekâ: türkiye kaynaklı araştırmaların eğilimleri üzerine bir içerik analizi. *Necmettin Erbakan Üniversitesi Ereğli Eğitim Fakültesi Dergisi, 5*(Özel Sayı), 387-411. https://doi.org/10.51119/ereegf.2023.49

[41] Islam, M. S., Hasan, M. M., Wang, X., Germack, H. D., & Noor-E-Alam, M. (2018). A systematic review on healthcare analytics: application and theoretical perspective of data mining. *Healthcare*, *6*(2). https://doi.org/10.3390/healthcare6020054

[42] Lucas P. (2004). Bayesian analysis, pattern analysis, and data mining in health care. *Current Opinion in Critical Care, 10,* 399-403. *https://doi.org/*10.1097/01.ccx.0000141546.74590.d6

[43] Jacob, S. G & Ramani, R. G. (2012). Data mining in clinical data sets: a review. *International Journal of Applied Information Systems*, *4*(6), 15-26.

[44] Milovic, B., & Milovic, M. (2012). Prediction and decision making in health care using data mining. *Kuwait Chapter of the Arabian Journal of Business and Management Review*,*1*(12), 126-136.

[45] Abdullah, F. S., Manan, N. S. A., Ahmad, A., Wafa, S.W., Shahril, M. R., Zulaily, N., Amin, R.M., & Ahmed, A. (2017). Data mining techniques for classification of childhood obesity among year 6 school children. In: Herawan, T., Ghazali, R., Nawi, N.M., Deris, M.M. (eds) Recent Advances on Soft Computing and Data Mining. SCDM 2016. Advances in Intelligent Systems and Computing, vol 549. Springer, Cham. DOI: https://doi.org/10.1007/978-3-319-51281-5_47

[46] Tang, T. A., Mhamdi, L., McLernon, D., Zaidi, S. A. R., & Ghogho, M. (2018). Deep recurrent neural network for intrusion detection in sdn-based networks. 2018 IEEE International Conference on Network Softwarization (NetSoft 2018)- Technical Sessions. 202-206. https://doi.org/10.1109/NETSOFT.2018.8460090

[47] Taspinar, Y. S., Cinar, I., & Koklu, M. (2021). Prediction of computer type using benchmark scores of hardware units. *Selcuk University Journal of Engineering Sciences, 20*, 11-17.

[48] Vapnik, V. N. (1999). *The Nature of Statistical Learning Theory*. Springer Science & Business media.

[49] Dwivedi, A. K. (2018). Performance evaluation of different machine learning techniques for prediction of heart disease. *Neural Computing and Applications, 29,* 685-693. https://doi.org/10.1007/s00521-016-2604-1

[50] Unal, Y., Taspinar, Y. S., Cinar, I., Kursun, R., & Koklu, M. (2022). Application of pre-trained deep convolutional neural networks for coffee beans species detection. *Food Analytical Methods, 15*, 3232-3243. https://doi.org/10.1007/s12161-022-02362-8

[51] Arlot, S., & Celisse, A. (2010). A survey of cross-validation procedures for model selection. *Statistics Surveys*, *4*, 40-79. https://doi.org/10.1214/09-SS054

[52] Şeker, A., Diri, B., & Balık, H. H. (2017). Derin öğrenme yöntemleri ve uygulamaları hakkında bir inceleme. *Gazi Mühendislik Bilimleri Dergisi*, *3*(3), 47-64.

[53] Keskenler, M. F., & Keskenler, E. F. (2017). Geçmişten günümüze yapay sinir ağları ve tarihçesi. *Takvim-I Vekayi*, *5*(2), 8-18.

[54] Haykin, S. (2009). *Neural Networks and Learning Machines, 3/E*: Pearson Education India.

[55] Tosunoğlu, E., Yılmaz, R., Özeren, E., & Sağlam, Z. (2021). Eğitimde makine öğrenmesi: araştırmalardaki güncel eğilimler üzerine inceleme. *Ahmet Keleşoğlu Eğitim Fakültesi Dergisi, 3*(2), 178-199. https://doi.org/10.38151/akef.2021.16

[56] Ozkan, I. A., Koklu, M & Sert, I. U. (2018). Diagnosis of urinary tract infection based on artificial intelligence methods. *Computer Methods and Programs in Biomedicine, 166*, 51-59. https://doi.org/10.1016/j.cmpb.2018.10.007

[57] Kim P. (2017). *Matlab Deep Learning.* Springer.

[58] Atman Uslu, N., & Onan, A. (2023) Investigating computational ıdentity and empowerment of the students studying programming: A text mining study. *Necmettin Erbakan Üniversitesi Ereğli Eğitim Fakültesi Dergisi, 5*(1), 29-45. https://doi.org/10.51119/ereegf.2023.29

[59] Chen, W., Pourghasemi, H. R., & Naghibi, S. A. (2018). A comparative study of landslide susceptibility maps produced using support vector machine with different kernel functions and entropy data mining models in China. *Bulletin of Engineering Geology and the Environment*, *77*(2), 647-664.

[60] Mahesh, B. (2020). Machine learning algorithms-a review. *International Journal of Science and Research (IJSR). [Internet].* 9:381-386. https://doi.org/10.21275/ART20203995

[61] Tien Bui D, Tuan, T. A., Klempe, H., Pradhan, B., & Revhaug, I. (2016). Spatial prediction models for shallow landslide hazards: a comparative assessment of the efficacy of support vector machines, artificial neural networks, kernel logistic regression, and logistic model tree. *Landslides*, *13*, 361-378. https://doi.org/10.1007/s10346-015-0557-6

[62] Jakkula, V. (2006). Jakkula, V. (2006). Tutorial on support vector machine (svm). School of EECS, Washington State University, 37(2.5), 3.

[63] Patle, A., & Chouhan, D. S. (2013, 23-25 January). *SVM kernel functions for classification.* International Conference on Advances in Technology and Engineering (ICATE, 2013). Mumbai, India. https://doi.org/10.1109/ICAdTE.2013.6524743

[64] Yu, H., & Kim, S. (2012). SVM Tutorial-Classification, Regression and Ranking.In Rozenberg, G., Back, T., & Kok, J. N. (Eds), *Handbook of Natural Computing*, (pp. 479-506). Springer. https://doi.org/10.1007/s10462-018-9614-6

[65] Chauhan, V. K., Dahiya, K., & Sharma, A. (2019). Problem formulations and solvers in linear SVM: a review. *Artificial Intelligence Review*, *52*(2), 803-855. https://doi.org/10.1007/s10462-018-9614-6

[66] Aha, D. W., Kibler, D., & Albert, M. K. (1991). Instance-based learning algorithms. *Machine Learning, 6*, 37-66. https://doi.org/10.1007/BF00153759

[67] Zhang, M. L., & Zhou, Z. H., (2007). ML-KNN: A lazy learning approach to multi-label learning. *Pattern Recognition*, *40*, 2038-2048. https://doi.org/10.1016/j.patcog.2006.12.019

[68] Sarker, I. H. (2021). Machine learning: Algorithms, real-world applications and research directions. *SN Computer Science, 2*, 1-21. https://doi.org/10.1007/s42979-021-00592-x

[69] Breiman L. (2001). Random forests. *Machine Learning, 45,* 5-32. https://doi.org/10.1023/A:1010933404324

[70] Mohan, S., Thirumalai, C., & Srivastava, G. (2019). Effective heart disease prediction using hybrid machine learning techniques. *IEEE Access, 7,* 81542-81554. https://doi.org/10.1109/ACCESS.2019.2923707

[71] Archer, K. J., & Kimes, R. V. (2008). Empirical characterization of random forest variable importance measures. *Computational Statistics & Data Analysis,* 52, 2249-2260. https://doi.org/10.1016/j.csda.2007.08.015

[72] Maxwell, A. E., Warner, T. A., & Fang, F. (2018). Implementation of machine-learning classification in remote sensing: An applied review. *International Journal of Remote Sensing, 39*:2784-2817. https://doi.org/10.1080/01431161.2018.1433343

[73] Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters, 27,* 861-874. https://doi.org/10.1016/j.patrec.2005.10.010

[74] Spackman, K. A. (1989). Signal detection theory: Valuable tools for evaluating inductive learning. *Proceedings of the Sixth International Workshop on Machine Learning,* 160-163. https://doi.org/10.1016/B978-1-55860-036-2.50047-3

[75] Pepe, M. S. (1997). A regression modelling framework for receiver operating characteristic curves in medical diagnostic testing. *Biometrika, 84,* 595-608. https://doi.org/10.1093/biomet/84.3.595

[76] Pepe, M. S. (2003). *The Statistical Evaluation of Medical Tests For Classification and Prediction*: Oxford University Press, USA.

[77] Luckett, D. J., Laber, E. B., El-Kamary, S.S., Fan, C., Jhaveri, R., Perou, C. M., Shebl, F. M & Kosorok, M. R. (2021). Receiver operating characteristic curves and confidence bands for support vector machines, *Biometrics, 77,* 1422-1430. https://doi.org/10.1111/biom.13365

[78] Narkhede, S. (2018). Understanding auc-roc curve. *Towards Data Science, 26*, 220-227.