JESPAR
Journal of Exercise Sciences
& Physical Activity Reviews

ORIGINAL ARTICLE

Use of Logistic Regression Method in Predicting Obesity Levels with Machine Learning Method

Abdulvahap Pinar[1] | **Fatma Hilal Yagin**[2*] | **Georgian Badicu**[3]

[1]Rectorate Unit, Adıyaman University, Adıyaman, Türkiye
[2]Biostatistics and Medical Informatics, Faculty of Medicine, Inonu University, Malatya, Türkiye
[3]Department of Physical Education and Special Motricity, Transilvania University of Brasov, 500068 Brasov, Romania

ABSTRACT

Obesity is a worldwide health issue due to excessive fat accumulation, especially prevalent in developing countries. It increases risks for diabetes, heart disease, and cancer, affecting multiple body systems. In 2016, 1.9 billion people were overweight, with 650 million classified as obese, emphasizing its global impact on public health. Both rich and developing nations are seeing sharp increases in their obesity rates, while low- and middle-income nations are seeing the biggest increases. This emphasizes how critical it is to create international plans for the administration and avoidance of obesity. This thorough analysis demonstrates the substantial effects of obesity on public health, health systems, and individual health. Public health policy are thus greatly influenced by studies on the causes, effects, and practical management techniques of obesity. The aim of this study is to derive classification metrics for a machine learning(ML) model suitable for classifying obesity levels of individuals and to present the corresponding accurate classification performance metric. Using the logistic regression model, the following classification performance metrics for predicting obesity levels were calculated: Area under ROC curve (AUC) is 0.980, Classification accuracy (CA) is 0.909, F1-Score is 0.911, Precision (Prec) is 0.909, Recall is 0.860, Matthews correlation coefficient (MCC) is 0.992, and Specificity (Spec) is 0.992. Notably, the classification accuracy (CA) of 90.9% indicates a significant achievement in correctly classifying the levels of obesity.This evaluation demonstrates the efficacy of the logistic regression model in distinguishing between different obesity levels, with high values in various performance metrics such as AUC and MCC underscoring the model's robustness and reliability in medical.

Keywords: Obesity, public health, epidemiology, machine learning, classification

***Corresponding:** Fatma Hilal Yagin; hilal.yagin@inonu.edu.tr
Journal home page: www.e-jespar.com
Academic Editor: Dr. Mehmet Gülü
https://doi.org/10.5281/zenodo.12601115

ARTICLE HISTORY
Received: 01 April 2024
Accepted: 27 May 2024
Published: 01 July 2024

INTRODUCTION

Obesity is a long-term medical development and aesthetic disorder characterized by excessive accumulation of fat in the body, resulting from irregular diet and physical inactivity. Obesity is defined by the World Health Organization (WHO) as a body mass index (BMI) of 30 or more. Obesity is becoming increasingly common in both industrialized and developing countries. Genetic factors, dietary habits, lack of physical activity and modern lifestyle are associated with this condition (Organization, 2000). Adopting a healthy lifestyle is essential in the fight against obesity. Obesity can be prevented and treated primarily through three strategies: behavioral changes, regular physical activity and a balanced diet. and a diet that limits calories while providing all the nutrients the body needs. Regular physical activity of at least 150 minutes of moderate-intensity exercise per week or 75 minutes of vigorous exercise per week is one of the factors preventing obesity. Behavioral improvements require people to create long-term plans to increase their levels of physical activity and nutrition. Health policy and community-based programs are also necessary to reduce the prevalence of obesity. These include planning efforts to raise awareness of obesity, promoting physical activity and increasing the availability of healthy foods. Increasing the availability and affordability of fruits, vegetables and whole grain foods, especially in low-income areas, and increasing the accessibility of healthy foods in developing countries can reduce obesity rates (Rippe & Hess, 1998). Improving the environment through the creation of parks, sports fields, walking and cycling paths is one way to encourage physical exercise. Awareness campaigns aim to prevent and increase public knowledge of the negative health effects of obesity through media and educational initiatives.

Two significant factors that may have a role in the development of obesity are alcohol consumption and genetics. Obesity and weight growth can be exacerbated by lifestyle decisions and genetic predispositions (Bougneres, 2002). Drinking alcohol increases the body's calorie production because it is a high-calorie beverage. While comparing alcohol to protein or carbohydrates, it has roughly 7 more calories per gram. Alcohol use can alter how the body processes fat. Fat buildup may result from the liver's slowing down of fat metabolism during alcohol breakdown. Due to an increase in insulin resistance, alcohol drinking can also raise the risk of obesity. Environmental and lifestyle factors are frequently shared by family members. This covers things like lifestyle decisions, levels of physical exercise, and eating habits. Poor lifestyle choices, like calorie-dense foods and little exercise, are frequently observed in families with a history of obesity. Youngsters take on their parents' food and lifestyle choices. Due to variables including poor eating habits and insufficient physical activity, children who grow up in households where obesity has a history are more susceptible to obesity risk (Osaka ve ark., 2017). Fatigue can result from inadequate water consumption's detrimental effects on metabolic systems. Maintaining metabolic rate, which aids in the body's calorie burning, depends critically on water. Reducing the body's ability to burn fat might result in weight gain if inadequate water intake occurs. Additionally, by enhancing feelings of satiety, water can lessen the urge to overeat (Intakes ve ark., 2005). Because it alters metabolism, smoking can make you gain weight. A balanced diet and regular exercise are less common among smokers, who also tend to have unhealthy lifestyles. Moreover, increased appetite and a slowed metabolic rate upon quitting smoking sometimes result in weight gain. This may make obesity more likely (Kaner ve ark., 2017).

This study aims to obtain and analyze metrics to classify obesity levels using logistic regression, one of the ML methods. These metrics aim to contribute to the development of healthy living interventions and policies by evaluating the effectiveness of the classification model for obesity. The findings of the study aim to provide a scientific basis for strategies to prevent and manage obesity and offer practical guidance in public health and clinical practice.

METHODS

Participant and Data

This study uses data from Mexico, Peru and Colombia to estimate obesity levels based on various eating habits and physical conditions of individuals between the ages of 14 and 61 (Olmedo, 2011).
The dataset used in the study is open access (publicly available) and was obtained from https://archive.ics.uci.edu/dataset/544/estimation+of+obesity+levels+based+on+eating+habits+and+physical+condition.The factors related to the data set and their descriptive statistics are presented in Table 1. The dataset included the following variables related to dietary habits: frequency of consumption of high-calorie foods (FAVC), frequency of vegetable consumption (FCVC), number of main meals (NCP), food consumption between meals (CAEC), daily water consumption (CH20) and alcohol consumption (CALC). Technological time devices used (TUE), frequency of physical activity (FAF), tracking of calorie consumption (SCC) and transportation used (MTRANS) were characteristics associated with physical status. Age, height and weight were quantitative factors and gender was a qualitative variable.

Variables	Categories	Descriptive statistics
Gender	Female	227 (45,6%)
	Male	271 (54,4%)
Age	Numeric value	23 ± 7
Height	Numeric value	1,69 ± 0,10
Weight	Numeric value	69,6 ± 17
Family History has overweight	No	198 (39,8%)
	Yes	300 (60,2%)
Eat High Caloric Food Frequently	No	150 (30,1%)
	Yes	348 (69,9%)
Vegetables Consumption Frequency	Never	32 (6,4%)
	Sometimes	272 (54,6%)
	Always	194 (39,0%)
Number of main meals daily	Between 1 y 2	108 (21,7%)
	Three	0 (0,0%)
	More than three	390 (78,3%)

Consumption of food between meals	No	53 (10,6%)
	Sometimes	136 (27,3%)
	Frequently	289 (58,0%)
	Always	20 (4,0%)
Smoking	No	466 (93,6%)
	Yes	32 (6,4%)
Liquid intake daily	Less than a liter	135 (27,1%)
	Between 1 and 2 L	266 (53,4%)
	More than 2 L	97 (19,5%)
Calories consumption monitoring	No	443 (89,0%)
	Yes	55 (11,0%)
Physical activity	I do not have	162 (32,5%)
	1 or 2 days	158 (31,7%)
	2 or 4 days	113 (22,7%)
	4 or 5 days	65 (13,1%)
Time-using technology devices	0-2 hours	243 (48,8%)
	3-5 hours	181 (36,3%)
	More than 5 hours	74 (14,9%)
Alcohol consumption	No	1 (0,2%)
	Sometimes	45 (9,0%)
	Frequently	273 (54,8%)
	Always	179 (35,9%)
Type of Transportation used	Automobile	99 (19,9%)
	Motorbike	7 (1,4%)
	Bike	11 (2,2%)
	Public Transportation	326 (65,5%)
	Walking	55 (11,0%)
Obesity level	Insufficient Weight	34 (6,8%)
	Normal Weight	287 (57,6%)
	Overweight Level I	47 (9,4%)
	Overweight Level II	11 (2,2%)
	Obesity Type I	3 (0,6%)
	Obesity Type II	58 (11,6%)
	Obesity Type III	58 (11,6%)

Table 1. Descriptive Statistics for the variables in the data set and their characteristics

Biostatistical Data Analysis

Frequency (percent) calculations were used to summarize the qualitative factors. The association between obesity and qualitative factors was investigated using chi-square testing. By computing the mean and standard deviation, quantitative variables were condensed. Statistical significance was defined as $p < 0.05$ for all results. Using the IBM application SPSS 26.0, statistical analyses were carried out. In addition, the dataset was split into training and test sets using the 10-fold cross-validation approach, and the The ML model was validated using the Orange 3.37.0 Data Mining version. Using the 10-fold cross-validation approach, the dataset is partitioned into 10 equal portions. A distinct subset of the dataset is used as test data each time, with the remaining subset being used as training data. In this validation process, a total of 399 (80%) samples of training data and 99 (20%) samples of test data were used.

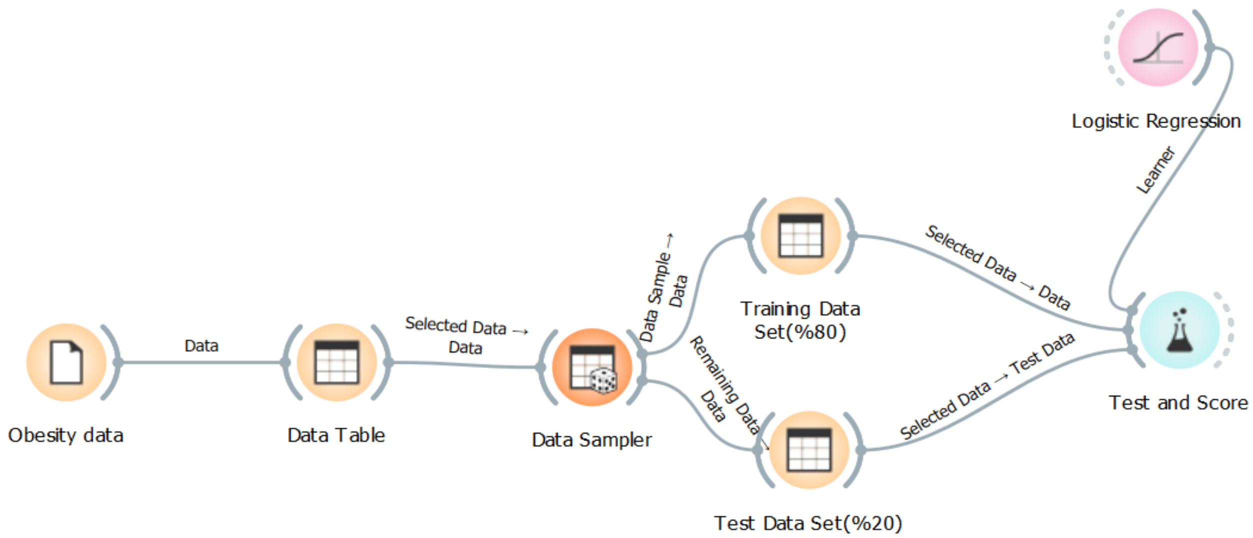

Figure 1. Figure 1. Setup design of the data design model of the Python-based orange3 package program

Orange is a data mining and ML suite for Python scripting-based data processing. The Orange Library is a set of data mining tools arranged hierarchically. Higher-level algorithms, such classification tree learning, are built from the lower-level processes at the base of the hierarchy, like input filtering, probability assessment, and feature scoring (Demšar ve ark., 2013). In Figure 3, in the "Data Sampler" section, the data were divided into training data set (80%) and test data set (20%). The data belonging to obesity levels separated as Training and Test data were connected to the "Test and Score" cell and 10-fold cross validation was performed. Finally, the LR model was connected and performance measures were calculated.

Lojistik Regresyon (LR)

When forecasting the existence or absence of a condition based on a variety of parameters, LR is especially helpful. When used to healthcare, the term "ML method" refers to the process of predicting a patient's likelihood of contracting an illness based on test results, age, gender, and past medical history, among other factors (Çolak, 2001). Classifying or categorizing one or more independent variables (predictors) is done using this strategy, which is ideal for solving classification problems. It is essentially used to calculate the likelihood that the outcome, or dependent variable, falls into one of two categories. These odds are computed using a logistic function, which is a sigmoid function.

One benefit of logistic regression is that it allows one to understand which factors have what kinds of influence on categorization, making the predictions interpretable. Additionally, the model's parameters are typically calculated using maximum likelihood, guaranteeing that the model is near-accurate in its parameter estimates. Many domains, including medicine, biology, economics, the social sciences, and marketing, use logistic regression extensively. Examples of applications for it include analyzing consumer behavior in marketing studies and forecasting illness risk in medical research. Additionally, by increasing the accuracy of logistic regression, different regularization strategies (such as L1 and L2 regularization) might assist lessen overfitting (Bishop & Nasrabadi, 2006; Friedman, 2009).

RESULTS

Table 2 shows the results of the study in which various factors affecting the level of obesity were examined. The table presents the distribution of different categorical variables (family history of obesity, consumption of high-calorie foods, frequency of vegetable consumption, etc.) and their distribution over different obesity levels. For each category, the number of participants below the obesity level (such as Ideal Weight, Type I and Type II obesity) is given in percentages. For example, according to the responses to the question "Do you have a history of obesity?", individuals with a family history of obesity are more likely to be obese. This shows how family history and genetic variables affect obesity risk. Similarly, people with a "habit of frequent consumption of high-calorie foods" tend to be more obese. Table 2 also shows the p-values for each variable. These p-values are used to assess whether the variables have a statistically significant effect on obesity levels. For example, the variables "Smoking habit", "Family History has overweight" and "Time-using technology devices" have p-values of 0.007, 0.001 and 0.038 respectively (p<0.05). From this point of view, smoking, Family History has overweight and Time-using technology devices have a statistically significant effect on obesity levels.

		Obesity Level							
Qualitative Variables	Categories	Insufficient Weight	Normal Weight	OV Level I	OV Level II	Obesity Type I	Obesity Type II	Obesity Type III	p-value
		n (%)	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)	
Family History has overweight	No	18 (52,9%)	132 (46,0%)	7 (14,9%)	1 (9,1%)	0 (0,0%)	23 (39,7%)	17 (29,3%)	<0.001*
	Yes	16 (47,1%)	155 (54,0%)	40 (85,1%)	10 (90,9%)	3 (100,0%)	35 (60,3%)	41 (70,7%)	
Eat High Caloric Food Frequently	No	13 (38,2%)	79 (27,5%)	11 (23,4%)	5 (45,5%)	1 (33,3%)	19 (32,8%)	22 (37,9%)	0.412*
	Yes	21 (61,8%)	208 (72,5%)	36 (76,6%)	6 (54,5%)	2 (66,7%)	39 (67,2%)	36 (62,1%)	
Vegetables Consumption Frequency	Never	3 (61,8%)	18 (72,5%)	3 (76,6%)	1 (54,5%)	0 (66,7%)	4 (67,2%)	4 (61,8%)	0.273*
	Sometimes	12 (35,3%)	155 (54,0%)	31 (66,0%)	5 (45,5%)	0 (0,0%)	33 (56,9%)	36 (62,1%)	
	Always	19 (55,9%)	114 (39,7%)	13 (27,7%)	5 (45,5%)	3 (100,0%)	21 (36,2%)	19 (32,8%)	
Number of main meals daily	Between 1 and 2	5 (14,7%)	52 (18,1%)	13 (27,7%)	3 (27,3%)	0 (0,0%)	17 (29,3%)	18 (31,0%)	0.12*
	More than three	29 (85,3%)	235 (81,9%)	34 (72,3%)	8 (72,7%)	3 (100,0%)	41 (70,7%)	40 (69,0%)	

Consumption of food between meals	No	2 (5,9%)	35 (12,2%)	6 (12,8%)	2 (18,2%)	0 (0,0%)	5 (8,6%)	3 (5,2%)	0.113*
	Sometimes	16 (47,1%)	83 (28,9%)	6 (12,8%)	1 (9,1%)	1 (33,3%)	13 (22,4%)	16 (27,6%)	
	Always	3 (8,8%)	10 (3,5%)	1 (2,1%)	1 (9,1%)	0 (0,0%)	4 (6,9%)	1 (1,7%)	
Smoking	No	33 (97,1%)	274 (95,5%)	41 (87,2%)	8 (72,7%)	2 (66,7%)	55 (94,8%)	53 (91,4%)	0.007*
	Yes	1 (2,9%)	13 (4,5%)	6 (12,8%)	3 (27,3%)	1 (33,3%)	3 (5,2%)	5 (8,6%)	
Liquid intake daily	Less than a liter	10 (29,4%)	83 (28,9%)	11 (23,4%)	5 (45,5%)	1 (33,3%)	12 (20,7%)	13 (22,4%)	0.085*
	Between 1 and 2 L	16 (47,1%)	164 (57,1%)	20 (42,6%)	4 (36,4%)	1 (33,3%)	30 (51,7%)	31 (53,4%)	
	More than 2 L	8 (23,5%)	40 (13,9%)	16 (34,0%)	2 (18,2%)	1 (33,3%)	16 (27,6%)	14 (24,1%)	
Calories consumption monitoring	No	28 (82,4%)	257 (89,5%)	45 (95,7%)	10 (90,9%)	3 (100,0%)	46 (79,3%)	54 (93,1%)	0.097*
	Yes	6 (17,6%)	30 (10,5%)	2 (4,3%)	1 (9,1%)	0 (0,0%)	12 (20,7%)	4 (6,9%)	
Physical activity	I do not have	10 (29,4%)	80 (27,9%)	20 (42,6%)	6 (54,5%)	2 (66,7%)	20 (34,5%)	24 (41,4%)	0.159*
	1 or 2 days	6 (17,6%)	97 (33,8%)	13 (27,7%)	2 (18,2%)	0 (0,0%)	20 (34,5%)	20 (34,5%)	
	2 or 4 days	14 (41,2%)	69 (24,0%)	8 (17,0%)	3 (27,3%)	1 (33,3%)	10 (17,2%)	8 (13,8%)	
	4 or 5 days	4 (11,8%)	41 (14,3%)	6 (12,8%)	0 (0,0%)	0 (0,0%)	8 (13,8%)	6 (10,3%)	
Time-using technology devices	0-2 hours	13 (38,2%)	129 (44,9%)	25 (53,2%)	7 (63,6%)	1 (33,3%)	36 (62,1%)	32 (55,2%)	0.038*
	3-5 hours	13 (38,2%)	122 (42,5%)	12 (25,5%)	2 (18,2%)	2 (66,7%)	11 (19,0%)	19 (32,8%)	
	More than 5 hours	8 (23,5%)	36 (12,5%)	10 (21,3%)	2 (18,2%)	0 (0,0%)	11 (19,0%)	7 (12,1%)	
Alcohol consumption	No	0 (000%)	1 (000%)	0 (000%)	0 (000%)	0 (000%)	0 (000%)	0 (000%)	0.53*
	Sometimes	1 (2,9%)	18 (6,3%)	7 (14,9%)	2 (18,2%)	0 (0,0%)	7 (12,1%)	10 (17,2%)	
	Frequently	19 (55,9%)	161 (56,1%)	22 (46,8%)	6 (54,5%)	2 (66,7%)	36 (62,1%)	27 (46,6%)	
	Always	14 (41,2%)	107 (37,3%)	18 (38,3%)	3 (27,3%)	1 (33,3%)	15 (25,9%)	21 (36,2%)	
Type of Transportation used	Automobile	3 (8,8%)	45 (15,7%)	15 (31,9%)	3 (27,3%)	1 (33,3%)	12 (20,7%)	20 (34,5%)	0.051*
	Motorbike	0 (0,0%)	4 (1,4%)	0 (0,0%)	1 (9,1%)	0 (0,0%)	2 (3,4%)	0 (0,0%)	
	Bike	0 (0,0%)	6 (2,1%)	3 (6,4%)	0 (0,0%)	0 (0,0%)	1 (1,7%)	1 (1,7%)	
	Public Transportation	25 (73,5%)	200 (69,7%)	27 (57,4%)	6 (54,5%)	2 (66,7%)	34 (58,6%)	32 (55,2%)	
	Walking	6 (17,6%)	32 (11,1%)	2 (4,3%)	1 (9,1%)	0 (0,0%)	9 (15,5%)	5 (8,6%)	
Quantitative Variables	**Mean ± SD**								
Age	23 ± 7								
Height	1,69 ± 0,10								
Weight	69,6 ± 17								

*: Chi-square testing; OV;overweight

Table 2. Explanation of factors according to obesity levels

In Table 3, various performance metrics are calculated to measure the performance of the LR model in predicting obesity levels. Classification accuracy indicates the proportion of all data points that the model correctly classifies. That is, your model correctly classifies approximately 90.9% of the dataset. A high accuracy value generally indicates good model performance. An F1-Score of 99.1% indicates that the classification performance of your model is generally balanced. A high specificity value indicates the model's ability to accurately discriminate negatives. A Specificity value of 99.2% indicates that the model predicts negative outcomes quite well and is reliable in this regard. These results suggest that the LR model not only has a high Specificity, i.e. it correctly identifies a significant proportion of obesity levels.

	Performance metrics						
Model	AUC	CA	F1-Score	Prec	Recall	MCC	Spec
Logistic Regression	0.98	0.909	0.911	0.919	0.909	0.86	0.992

AUC: Area under ROC curve; CA: Classification accuracy;Prec: Precision; MCC: Matthews correlation coefficient; Spec: Specificity

Table 3. Results based on the model's performance measures for predicting obesity levels

DISCUSSION

In this work, a machine learning (ML) algorithm was used to create a classification model that assessed obesity levels based on food and physical activity patterns. Several categorization measures were used to assess the model's performance. The area under the curve (AUC), classification accuracy (CA), recall, precision, F1-score, Matthews correlation coefficient (MCC), and specificity are among these measurements. When these measures were combined, the model's predictive ability for obesity levels was thoroughly assessed. While the CA number reflects the model's overall accuracy, the AUC value represents the model's overall discriminative ability. How successfully the model predicts positive classes is measured by F1-Score, Precision, and Recall. Specificity showed the percentage of incorrectly categorized negative groups, whereas MCC offered the general stability of the classification.

Being sedentary and not getting enough exercise are key causes to obesity, especially in young and middle-aged adults. Studies on the subject have shown that eating slower prevents the onset of obesity, using a generalized prediction model. Additionally, sticking to a regular eating schedule is essential for preventing obesity (Siddarth, 2013). One of the main behavioral variables for childhood and adolescent overweight and obesity is overconsumption of high-energy foods. Compared to kids who are not obese, obese kids typically have higher levels of food addiction. For this, it is essential to understand the mechanisms underlying the eating practices that lead to obesity or overweightness (Al-Dalaeen & Al-Domi, 2017; Gülü ve ark., 2022).

The factors used to determine obesity levels include: Gender, Age, Height, Weight, Family History of Overweight, Frequent Consumption of High-Caloric Food, Frequency of Vegetable Consumption, Number of Main Meals Daily, Consumption of Food Between Meals, Smoking, Daily Liquid Intake, Monitoring of Calorie Consumption, Physical Activity, Time Spent Using Technology Devices, Alcohol Consumption, and Type of Transportation Used. The statistical inference between obesity levels and these factors indicates that "smoking habits," "family history of overweight," and "time spent using technology devices" have p-values of 0.007, 0.001, and 0.038, respectively ($p<0.05$). From this perspective, it can be observed that smoking habits, family history of overweight, and time spent using technological devices have a statistically significant effect on obesity levels. On the other hand, according to the Logistic Regression (LR) model obtained with Omega 3.37.0 version, the classification performance for correctly predicting obesity levels was calculated as 90.9%.

CONCLUSIONS

By predicting obesity levels based on eating patterns and physical activity levels, the logistic regression model proved to be a valuable tool. As evidence of the model's applicability and dependability, the categorization accuracy rate reached 90.9%. Area under ROC curve (AUC) 98%, Classification accuracy (CA) 90.9%, F1-Score 91.1%, Precision (Prec) 90.9%, Recall 86%, Matthews correlation coefficient (MCC) 9.2%, and Specificity (Spec) 99.2% were the other performance indicators. Important information for obesity prevention and intervention plan formulation is provided by the measurements taken. The analysis reveals that the model performs well overall, as evidenced by its classification accuracy (CA) of 90.9%. We believe it will make a significant scientific contribution to the creation of laws that encourage healthy living and increase public knowledge of the risks associated with obesity, which will lead to significant advancements in the field of public health.

Author Contributions

Conceptualization, A.P.,F.H.Y. methodology, F.H.Y, A.P., G.B.; formal analysis, G.B.; investigation, G.B.; data curation, F.Y.H.; writing–original draft preparation, A.P., F.H.Y, G.B.; writing–review and editing, A.P., F.H.Y, G.B.

Informed Consent Statement:

The research was conducted in line with the Declaration of Helsinki.

Acknowledgments:

We would like to thank all participants who took part in the research.

Funding:

This research was not funded by any institution or organization.

Conflicts of Interest:

The authors declare that no conflicts interest.

REFERENCES

Al-Dalaeen, A., veAl-Domi, H. (2017). Factors Associated with Obesity among School Children in Amman, Jordan. *Malaysian Journal of Nutrition*, *23*(2).

Bishop, C. M., veNasrabadi, N. M. (2006). *Pattern recognition and machine learning* (Vol. 4). Springer.

Bougneres, P. (2002). Genetics of obesity and type 2 diabetes: tracking pathogenic traits during the predisease period. *Diabetes*, *51*(suppl_3), S295-S303.

Çolak, C. (2001). *Lojistik regresyon analizi ve sağlık bilimlerinde bir uygulama* Yüksek Lisans Tezi, Ġnönü Üniversitesi Sağlık Bilimleri Enstitüsü, Malatya].

Demšar, J., Curk, T., Erjavec, A., Gorup, Č., Hočevar, T., Milutinovič, M., Možina, M., Polajnar, M., Toplak, M., ve Starič, A. (2013). Orange: data mining toolbox in Python. *the Journal of machine Learning research*, *14*(1), 2349-2353.

Friedman, J. (2009). The elements of statistical learning: Data mining, inference, and prediction. *(No Title)*.

Gülü, M., Yapici, H., Mainer-Pardos, E., Alves, A. R., ve Nobari, H. (2022). Investigation of obesity, eating behaviors and physical activity levels living in rural and urban areas during the covid-19 pandemic era: a study of Turkish adolescent. *BMC pediatrics*, *22*(1), 405.

Intakes, S. C. o. t. S. E. o. D. R., Electrolytes, P. o. D. R. I. f., ve Water. (2005). *Dietary reference intakes for water, potassium, sodium, chloride, and sulfate*. National Academies Press.

Kaner, G., Kürklü, N. S., ADIGÜZEL, K. T., ve Koyu, E. B. (2017). İzmir'de beslenme ve diyet polikliniğine başvuran kadınlarda obezite prevalansı ve ilişkili risk faktörlerinin belirlenmesi. *Pamukkale Medical Journal*, *10*(3), 250.

Olmedo, M. (2011). La obesidad: un problema de salud pública. *Revista de divulgació científica y tecnológica de la Universidad Veracruzana*.

Organization, W. H. (2000). Obesity: preventing and managing the global epidemic: report of a WHO consultation.

Osaka, T., Moriyama, E., Arai, S., Date, Y., Yagi, J., Kikuchi, J., ve Tsuneda, S. (2017). Meta-analysis of fecal microbiota and metabolites in experimental colitic mice during the inflammatory and healing phases. *Nutrients*, *9*(12), 1329.

Rippe, J. M., veHess, S. (1998). The role of physical activity in the prevention and management of obesity. *Journal of the American Dietetic Association*, *98*(10), S31-S38.

Siddarth, D. (2013). Risk factors for obesity in children and adults. *Journal of investigative medicine*, *61*(6), 1039-1042.