Research Article

A Balanced Machine Learning Approach to Obesity Risk Classification: Comparative Analysis and Feature Importance

***Haydar KOÇ*[1], *Tuba KOÇ*[2]**

Abstract

Obesity is a growing public health concern, particularly among university students who are exposed to lifestyle changes, disordered eating habits, and reduced physical activity. The aim of this study is to classify obesity risk levels among university students using machine learning classification methods and to identify the most influential factors associated with this risk. The study sample consisted of data collected from 445 students studying at Çankırı Karatekin University. In this context, eight machine learning algorithms—Logistic Regression, Random Forest, Extra Trees, Support Vector Machines, K-Nearest Neighbor, Quadratic Discriminant Analysis, Naive Bayes, and Multilayer Perceptron—were compared to classify obesity risk. Class imbalance in the dataset was addressed using the Adaptive Synthetic Sampling (ADASYN) method applied exclusively to the training set. The models were evaluated using standard performance metrics, and the highest accuracy rate (96.26%) was achieved by the Random Forest model, followed by Logistic Regression with 94.77% accuracy. Variable importance analysis indicated that age, internet use scale score, and fast-food consumption frequency were the most influential factors in classification, while the low correlation between variables (|r| < 0.2) suggested that model performance was driven by the combined contribution of multiple features. Overall, the findings demonstrate that the balanced machine learning approach, particularly ensemble-based methods, can classify obesity risk with high accuracy and provide valuable insights for targeted prevention strategies among university students.

Keywords: Adaptive synthetic sampling, machine learning, obesity, young adults.

1. Assoc. Prof. Dr., Çankırı Karatekin University, Faculty of Science, Department of Statistics, haydarkoc@karatekin.edu.tr, ORCID:0000-0002-8568-4717
2. Assoc. Prof. Dr., Çankırı Karatekin University, Faculty of Science, Department of Statistics, tubakoc@karatekin.edu.tr, https://orcid.org/0000-0001-5204-0846

Received : 02.09.2025
Accepted : 01.12.2025

Cite this paper: *Koç, H. & Koç, T. (2025). A Balanced Machine Learning Approach to Obesity Risk Classification: Comparative Analysis and Feature Importance, Eurasian Journal of Health Technology Assessment, 9(2): 90-107.*

1. Introduction

Obesity is an important public health problem on a global scale. Obesity is increasing day by day in both developed and developing countries. Obesity reduces the quality of life of individuals and increases the incidence of chronic diseases such as cardiovascular diseases, hypertension, and diabetes. According to the World Health Organization, approximately 39% of individuals aged 18 and over are overweight and 13% are obese (WHO, 2024). Obesity has negative effects not only on physical health, but also on the psychosocial well-being of individuals (Hruby & Hu, 2015). University age is a critical period in shaping health behaviors, in which individuals experience great changes both physically and behaviorally during the transition from childhood and adolescence to adulthood. During this period, individuals leave the family and step into a more independent life, starting to make individual decisions on many issues from their own dietary preferences to the level of physical activity). At the same time, factors such as course load, exam stress, social adjustment problems, and irregular lifestyle can pave the way for unhealthy diet, low physical activity, and increased sedentary behaviors (Nelson et al., 2008). An increase in fast food consumption, the development of meal skipping habits, and a decrease in exercise frequency is commonly observed among students who have just started university (Alzahrani et al., 2020). These lifestyle changes increase the risk of weight gain in young people and can pave the way for obesity in the long term. In addition, considering that the habits acquired during the university period continue in adulthood, this period offers an important intervention opportunity in terms of changing health behaviors. In the context of obesity research, demographic variables such as age and sex are known to influence energy balance and body composition (Hruby & Hu, 2015). Behavioral factors including dietary habits, fast-food consumption, meal-skipping frequency, and physical activity play a critical role in determining weight status. Moreover, digital lifestyle indicators such as internet usage duration and internet addiction levels have recently emerged as psychosocial determinants of sedentary behavior and irregular eating patterns among young adults. Accordingly, these demographic, behavioral, and psychological dimensions were jointly considered in the model to capture the multifactorial structure of obesity.

While traditional statistical methods evaluate a limited number of variables affecting obesity, machine learning (ML) techniques provide effective predictive frameworks by analyzing numerous, complex, and interactive factors simultaneously. Previous studies have demonstrated the strong potential of ML algorithms in obesity prediction. For example, Zheng and Ruggiero (2017) compared four different machine learning models and reported that, compared to logistic regression, the Advanced Decision Trees (IDT), Weighted K Nearest Neighbor (KNN), and Artificial Neural Networks (ANN) achieved higher classification performance. Similarly, Chatterjee et al. (2020) explored how ML techniques can be employed to identify key risk factors associated with obesity and overweight. Ferdowsy et al. (2021) found that the logistic regression algorithm performed best (97.09% accuracy) among nine algorithms, whereas gradient boosting achieved the lowest (64.08%). Musa &Basaky (2022) observed that the GBoost Classifier yielded the highest accuracy (99.05%) using clinical datasets. In later studies, hybrid and ensemble methods also demonstrated promising results in predicting obesity levels (Dirik, 2023; Helforoush & Sayyad, 2024; Yağmur, 2024). Collectively, these studies provide an important methodological background supporting the application of ML-based classification techniques in obesity risk prediction.

In this study, the risk of obesity was tried to be classified by using the data collected through a comprehensive questionnaire applied to university students.

The main purpose of the study is to compare the performance of different machine learning classification algorithms (Logistic Regression, Random Forest, Extra Trees, Support Vector Machines, K-Nearest Neighbor, Quadratic Discriminant Analysis, Naive Bayes, and Multilayer Perceptron) in predicting obesity risk. Following the identification of an imbalance between the classes in the data set, the data was balanced using the ADASYN (Adaptive Synthetic Sampling) method prior to modelling. In addition to various performance metrics, feature importance analysis for the best-performing model and the correlation matrix of the top 15 variables that stand out in this analysis were examined. Consequently, the model's predictive efficacy and interpretability of the results were enhanced by elucidating the most contributing variables to the classification process and the structural relationships between these variables. In this respect, the study aims to reveal the effectiveness of different algorithms in determining obesity classes and to increase the practical applicability of the findings with variable importance analysis.

In the second part of the article, the data set, the steps involved in its preparation, the classification algorithms and the evaluation metrics utilized in the study are explained in detail. In the third part, the findings obtained from the applied models are presented and comparative analyses are made. In the final section, the results obtained are discussed in the context of extant literature and the practical contributions of the study are outlined, along with suggestions for future research.

2. Materials and Methods

2.1. Data Description

The data set utilized in this study was obtained from a face-to-face survey study conducted on Çankırı Karatekin University students in the 2024–2025 academic year. The study was designed to include young people studying at the university, and a simple random sampling method was applied. The target population consisted of all registered undergraduate students in the 2024–2025 academic year (N = 9056). Based on a 95% confidence level and a 5% margin of error, the required sample size was determined as 445 students. The online questionnaire was distributed through the university's student e-mail system, and participation invitations were randomly sent to students from different faculties and departments, ensuring that each student had an equal opportunity to participate voluntarily in the study. Participation was voluntary and anonymous, and informed consent was obtained online before completing the survey. Prior to the initiation of the data collection process, the relevant ethics committee's approval was obtained (Approval No: [44], Date: [23-08-2024]). The questionnaire consisted of various sections, including those pertaining to demographic characteristics, living habits, and health status. The target variable of the study was Body Mass Index (BMI), which was categorized into three distinct groups: underweight (BMI < 18.5; code 1), normal weight (BMI 18.5–24.9; code 2), and obese (BMI ≥ 25; code 3). Although this classification was based on the WHO's thresholds of 18.5 and 25/30, the "overweight" and "obese" groups were combined into a single category for the analyses (Hruby&Hu, 2015). Height and weight information used to calculate BMI was self-reported by participants through an online questionnaire.

The dataset under consideration contains a total of 21 variables, comprising both continuous and categorical variables. The continuous variables incorporated in the study included age (mean = 20.62 years, standard deviation = 1.94, range = 17–33), the duration of daily internet use (mean = 5.45 hours, standard deviation = 3.03, range = 1–18), and the Internet Addiction Scale score (mean = 42.70, standard deviation = 15.40, range = 19–95). Internet Addiction was measured using the Internet Addiction Scale (IAS) developed and adapted to Turkish by Şahin and Korkmaz (2011). The scale consists of 19 items rated on a 5-point Likert scale ranging from 1 (Never) to 5 (Always). The total score ranges from 19 to 95, with higher scores indicating greater levels of internet addiction. In the original study, the scale demonstrated good internal consistency with a Cronbach's alpha coefficient of 0.85. In this study, the total IAS score was used as one of the independent variables representing participants' internet use behavior. Descriptive statistics for these categorical variables, including demographic, behavioral, and health-related characteristics of the participants, are summarized in Table 1.

Table 1. Descriptive Statistics of Categorical Variables

Variable	**Category**	**n**	**%**
Sex	Female	283	63.6
	Male	162	36.4
Mother's education	Primary school	226	50.79
	High school	105	23.6
	University	85	19.1
	Illiterate	20	4.49
	Master's/Doctorate	9	2.02
Father's education	Primary school	185	41.57
	High school	176	39.55
	University	72	16.18
	Illiterate	8	1.8
	Master's/Doctorate	4	0.9
Family type	Nuclear family	339	76.18
	Extended family	81	18.2
	Separated	25	5.62
Income-Expense Perception	Income equals expenses	243	54.61
	Income exceeds expenses	110	24.72
	Income is less than expenses	92	20.67
Smoking Status	No	292	65.62
	Yes	134	30.11
	I quit / Former smoker	19	4.27
Alcohol Use Status	No	352	79.1
	Yes	76	17.08
	I quit / Former user	17	3.82
Do You Exercise Regularly?	No	340	76.4
	Yes	105	23.6
Exercise Frequency	Never	216	48.54
	1-2 days per week	165	37.08
	3-4 days per week	51	11.46
	Every day	13	2.92
Satisfaction with Current Weight	No	241	54.16
	Yes	204	45.84
Do You Have Diabetes?	No	433	97.3
	Yes	12	2.7
Do You Have Hypertension?	No	430	96.63
	Yes	15	3.37
Sleep Pattern	Irregular	294	66.07
	Regular	151	33.93
Meal Skipping Status	I skip meals	345	77.53
	I don't skip meals	100	22.47

Tablo 1. Descriptive Statistics of Categorical Variables (Continued)			
Skipped Meal	Lunch	175	39.33
	Breakfast	162	36.4
	None	85	19.1
	Dinner	23	5.17
Snack/Junk Food Consumption Between Meals	Yes	327	73.48
	No	118	26.52
Fast Food Consumption Frequency	2-4 times a week	168	37.75
	More than once a month	105	23.6
	Less than once a week	77	17.3
	5 or more times a week	43	9.66
	Less than once a month	34	7.64
	Never	18	4.04

2.2. Data Preprocessing

The dataset was subjected to scrutiny for the presence of incomplete observations, yet no instances of missing data were identified. Consequently, the analyses were conducted using the complete data set. However, a significant class imbalance is evident in BMI classes, which are the target variable. The normal weight class constitutes approximately 85% of the sample, while the obese class comprises only 4.5%. The imbalance under scrutiny is visually presented in Figure 1.

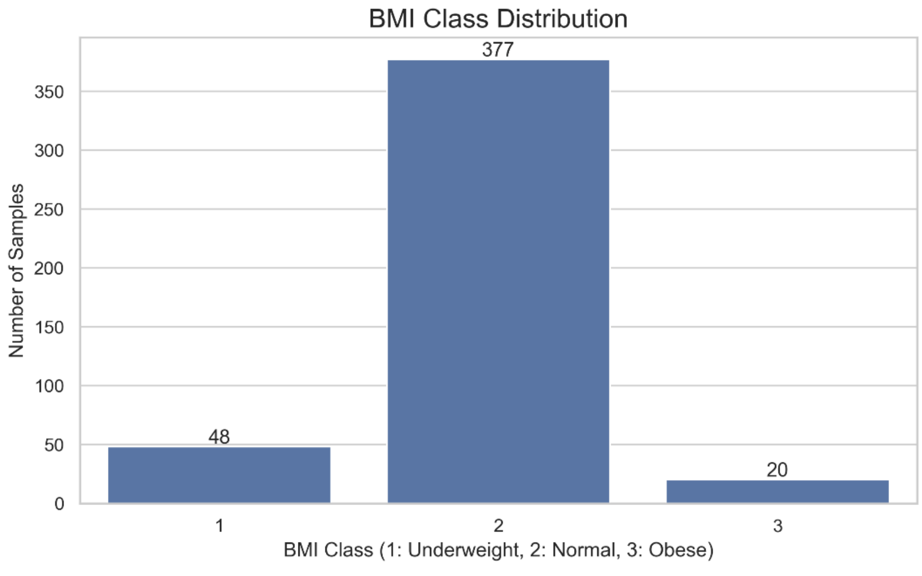

Figure 1. BMI class distribution

The dataset is divided into two sections, designated as training and testing, respectively, to facilitate an objective evaluation of the model's performance. In this separation process, 70% of the data was allocated for model training, with the remaining 30% designated for model evaluation through a testing phase. This partitioning strategy, frequently used in machine learning research, ensures model generalization and objective performance evaluation (Géron, 2022). This approach ensures that an independent set of tests is maintained to evaluate the overall performance of the model, and that the actions performed during training are not reflected in the test data.

Brownlee (2020) emphasizes that preprocessing steps, such as classroom imbalance, should be performed only on the training data; otherwise, the naturalness of the test data will deteriorate, and performance measurements may be misleading. In a similar, Kotsiantis et al. (2006) and Fernández et al. (2018) posit that after the separation of training and testing, the imbalance should be rectified exclusively within the training set.

In this direction, the Adaptive Synthetic Sampling (ADASYN) method was applied exclusively to the training data in the solution of the class imbalance problem. ADASYN produces synthetic samples that bear a resemblance to the Synthetic Minority Over-sampling Technique (SMOTE) method. However, it does not adhere to a predetermined rate for this process; rather, it dynamically determines the necessity for sample production between classes. In this respect, ADASYN produces more balanced and learnable examples, especially about the internal structure of the minority class. As posited by He et al. (2008) ADASYN is the optimal solution due to its numerous advantages, including the reduction of bias, the enhancement of classroom boundaries, and the mitigation of overfitting risks.

As demonstrated in Figure 2, the application of the ADASYN method results in a balanced class distribution.

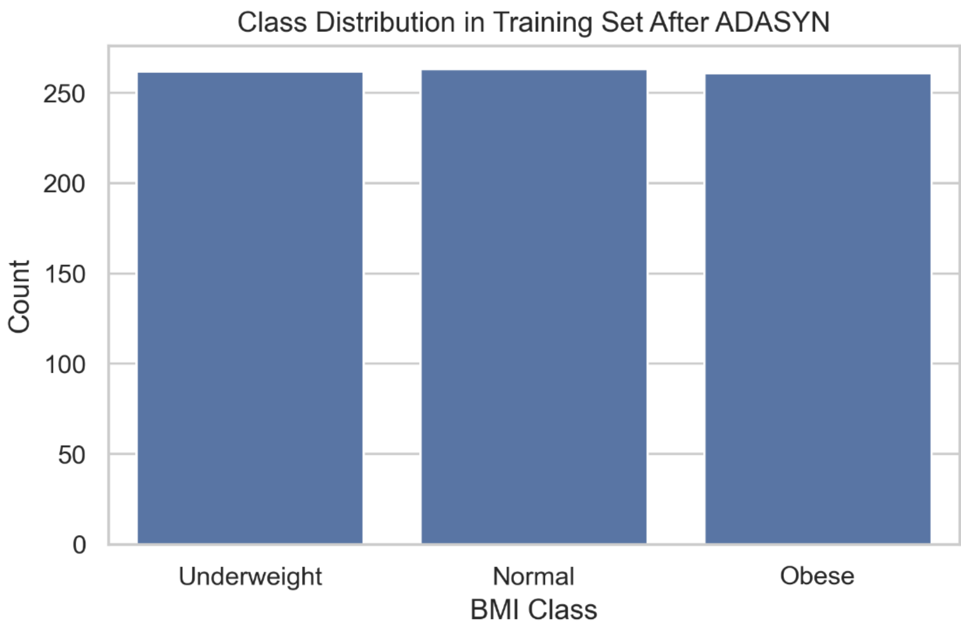

Figure 2. Balanced class distribution obtained using ADASYN

2.3. Model Descriptions

2.3.1. Logistic Regression

Logistic regression (LR) is a probability-based supervised learning algorithm that is widely used in binary and multiclass classification problems (Hosmer et al., 2013). In contradistinction to linear regression, it predicts the probability that an observation belongs to a certain class as opposed to continuous values. Consequently, it is employed as a standard classification method in numerous fields.

The linear combination of the independent variables is converted to the range 0 to 1 using the sigmoid (logistic) function, and the model output is interpreted as the probability that the observation belongs to the corresponding class. The estimation of model parameters is achieved through the minimization of the logarithmic loss function. The coefficients thus obtained are indicative of the effect of the variables on the probability of belonging to the class at the log-odds level.

2.3.2. Random Forest

Random Forest (RF) is a powerful ensemble learning method that is widely used in both regression and classification problems (Breiman, 2001). The method is founded on the principle of independently training a substantial number of decision trees on random (bootstrap) samples of the training data and taking a majority vote in classification and averaging in regression.

The evaluation of only randomly selected variables in each tree has been demonstrated to reduce the variance of the model and provide resistance to overfitting. The RF algorithm is distinguished by its consistent performance in high-dimensional and complex data sets. However, given the potential limitations in interpretability of the model, additional analyses such as variable significance levels are frequently employed. It is evident that RF demonstrates a high level of resistance to overfitting and is reliable, particularly in the context of high-dimensional and complex data sets (Fernández-Delgado et al., 2014).

2.3.3. Extra Trees Classifier

Extra Trees (Extremely Randomized Trees) is a classifier; that is to say, it is an ensemble learning method based on decision trees (Geurts et al., 2006). For each tree, the model generates multiple decision trees using randomly selected subsets of training data and randomly determined split thresholds. The estimation process involves the majority vote for classification problems and the calculation of the means of the predictions for regression problems. A high level of randomization has been shown to reduce the risk of overfitting while increasing the model's ability to generalize. Literature acknowledges Extra Trees as a highly effective method for capturing relationships between variables, especially in high-dimensional and complex data sets.

2.3.4. Support Vector Machines

Support Vector Machines (SVM) represent a powerful supervised machine learning algorithm that is widely used in both classification and regression problems (Cortes& Vapnik, 1995). The primary objective of the SVM is to ascertain a hyperplane that optimally differentiates the various classes, thereby maximizing the distance (margin) between them. This approach enhances the model's capacity for generalization.

In the context of non-linear separable data, the employment of kernel functions is instrumental in the transformation of data into a higher-dimensional space, thereby facilitating the establishment of complex decision bounds. The SVM model is distinguished by its ability to avoid overfitting in high-dimensional and complex data sets (Akın, 2023).

2.3.5. K-Nearest Neighbor

K Nearest Neighbor (KNN) is a non-parametric, heuristic classification algorithm (Cover & Hart, 1967). The determination of the class of a new observation is determined by the algorithm based on measures of similarity or distance (most commonly, Euclidean distance) from the samples in the training set. The fundamental approach entails the identification of the nearest K neighbors to the new observation, with the majority class amongst these neighbors being designated as the prediction.

As the K value increases, the algorithm's noise resistance is enhanced; however, very high K values can lead to a decline in overall performance. The efficacy of the method is contingent upon the judicious selection of the K-value and the distance measure. The simplicity and interpretability of KNN have resulted in its wide application in different disciplines.

2.3.6. Quadratic Discriminant Analysis

Quadratic Discriminant Analysis (QDA) is a statistical method that has been utilized for the purposes of classification and size reduction (Friedman, 2009). QDA, a variant of LDA (Linear Discriminant Analysis), employs distinct covariance matrices for each class, thereby generating quadratic decision boundaries in lieu of linear ones. This approach enables the effective delineation of classes, even in non-linear scenarios. QDA has the capacity to model the differences between groups not only through mean vectors, but also through covariance structures. However, it should be noted that the applicability of this method may be limited in very high-dimensional data sets, since the number of parameters increases with the square of the number of variables. The method is favored on account of its simplicity, flexibility, and capacity to model differences between classes in relatively low-dimensional data sets.

2.3.7. Naive Bayes

The Naive Bayes (NB) classifier is a probabilistic classification method that is both simple and effective. It is based on Bayes' theorem and hypothesized conditional independence between features (Rish, 2001). The method employed to calculate the probability that an observation belongs to a specific class is as follows:

$$\boldsymbol{P}(\boldsymbol{C}|\boldsymbol{X}) = \frac{\boldsymbol{P}(\boldsymbol{X}|\boldsymbol{C})\boldsymbol{P}(\boldsymbol{C})}{\boldsymbol{P}(\boldsymbol{X})} \quad (1)$$

In this context, P(C) denotes the priori probability, whilst $\boldsymbol{P}(\boldsymbol{X}|\boldsymbol{C})$ signifies the probability of occurrence of observations under the specified class. NB can be trained expeditiously, even on voluminous data sets, can operate with a limited number of training examples, can process both continuous and discrete data, and can be applied to binary or multi-class problems. Despite its simplicity, the system performs competitively in several areas, including text mining, medicine, spam filtering and recommendation systems (Domingos & Pazzani, 1997).

2.3.8. Multi-Layer Perceptron

Multi-Layer Perceptron (MLP) is a feedforward neural network model consisting of input, one or more hidden layers, and output layers (Murtagh, 1991). The data contained within the input layer is processed by nonlinear activation functions through neurons located in the hidden layers and subsequently transferred to the output layer. The model employs a backpropagation algorithm and gradient-based optimization methods to update the weights during the training process. This configuration endows MLP with the capacity to acquire intricate and non-linear separable data patterns. As stated in the literature, machine learning algorithms (MLP) have been shown to provide higher levels of accuracy in comparison to traditional methods when it comes to classification and risk estimation, as evidenced by studies in the field of medical data analysis (Bikku, 2020).

2.4. Evaluation Metrics

Evaluation metrics are defined as measurable performance indicators that are utilized for the evaluation of the efficacy of classification and discriminatory analysis methods in binary and multi-class environments. Naidu et al. (2023) demonstrate, the utilization of disparate metrics enables the highlighting of differing aspects of the classifier's behavior. This, in turn, facilitates fair comparison, threshold selection and hyperparameter adjustment between models. These metrics, which originated from the early formalization of precision and recall in information retrieval, have become standard tools for reporting and interpreting predictive performance in machine learning studies. The following section details the formula for several of the primary evaluation metrics. In this context, TP (true positive) denotes the number of true positives, TN (true negative) signifies the number of true negatives, FP (false positive) represents the number of false positives, and FN (false negative) corresponds to the number of false negatives.

$$Accuracy = (TP + TN) / (TP + TN + FP + FN) \quad (2)$$

$$Precision = TP / (TP + FP) \quad (3)$$

$$Recall = TP / (TP + FN) \quad (4)$$

$$F1 - score = 2 \times (Precision \times Recall) / (Precision + Recall) \quad (5)$$

3. Results

In this section, eight different machine learning algorithms were employed to predict the obesity class of university students. The class imbalance that was evident in the dataset was only eliminated by implementing the ADASYN method on the training data. Consequently, this ensured that the models were trained on a balanced data set. The hold-out validation technique was utilized in the modelling process. A total of 70% of the data was allocated for training purposes, while the remaining 30% was designated as the test set. All models were trained and tested in accordance with this structure. The comparative results obtained are presented in Table 2.

Table 2. Performance Metrics of Machine Learning Models for Obesity Classification

Model	Accuracy	Precision	Recall	F1-Score
Logistic Regression	0.947761	0.948399	0.947761	0.943442
Random Forest	0.962687	0.962687	0.962687	0.962687
Extra Trees	0.902985	0.912916	0.902985	0.880335
SVM	0.850746	0.723769	0.850746	0.782138
KNN	0.850746	0.723769	0.850746	0.782138
QDA	0.873134	0.857045	0.873134	0.850298
Naive Bayes	0.559701	0.905068	0.559701	0.661328
MLP	0.895522	0.882844	0.895522	0.878314

As illustrated in Table 2, the RF model demonstrated the highest level of success among all methods, with an accuracy rate of 96.26%. In the extant literature, RF is frequently cited as one of the most successful methods, due to its resistance to overfitting and its ability to generalize, especially in high-dimensional and complex data sets (Breiman, 2001). The second-best performance was obtained from the LR model, with an accuracy rate of 94.77%. The Extra Trees and MLP algorithms demonstrated moderate performance, while the QDA, SVM, and KNN methods exhibited comparable accuracy values. The lowest success rate was observed in the Naive Bayes model, with an accuracy rate of 55.97%. The findings demonstrate that ensemble methods generally exhibit superior predictive performance in comparison to individual models and can attain favorable outcomes even in the context of imbalanced data sets.

The machine learning algorithms employed in the study have been optimized with specific hyperparameters to ensure that the model functions optimally for the data set. Hyperparameter selections have been demonstrated to enhance the accuracy, generalization ability and learning time of the model in a balanced manner. As illustrated in Table 3, the fundamental hyperparameters employed for each model are presented.

Table 3. Tuned Hyperparameters for the Applied Machine Learning Algorithms

Model	**Key Hyperparameters and Their Values**
Logistic Regression	penalty=l2, C=1.0, solver=lbfgs, max_iter=2000
Random Forest	n_estimators=100, min_samples_split=2, random_state=42
Extra Trees	n_estimators=100, min_samples_split=2, random_state=42
Support Vector Machine	kernel=rbf, C=1.0, gamma=scale, random_state=42
K-Nearest Neighbors	n_neighbors=5, weights=uniform, algorithm=auto
Quadratic Discriminant Analysis	reg_param=1.0
Naive Bayes	(No hyperparameters / Default settings)
Multi-Layer Perceptron	hidden_layer_sizes=(100,), activation=relu, solver=adam, alpha=0.0001, max_iter=2000

To evaluate the discrimination power of the models, the Receiver Operating Characteristic (ROC) curves and the Area Under the Curve (AUC) values were examined. The ROC analysis comparatively demonstrates the sensitivity and 1-specificity rates of classification models at varying threshold values. Figure 3 presents the ROC curves obtained for each model.

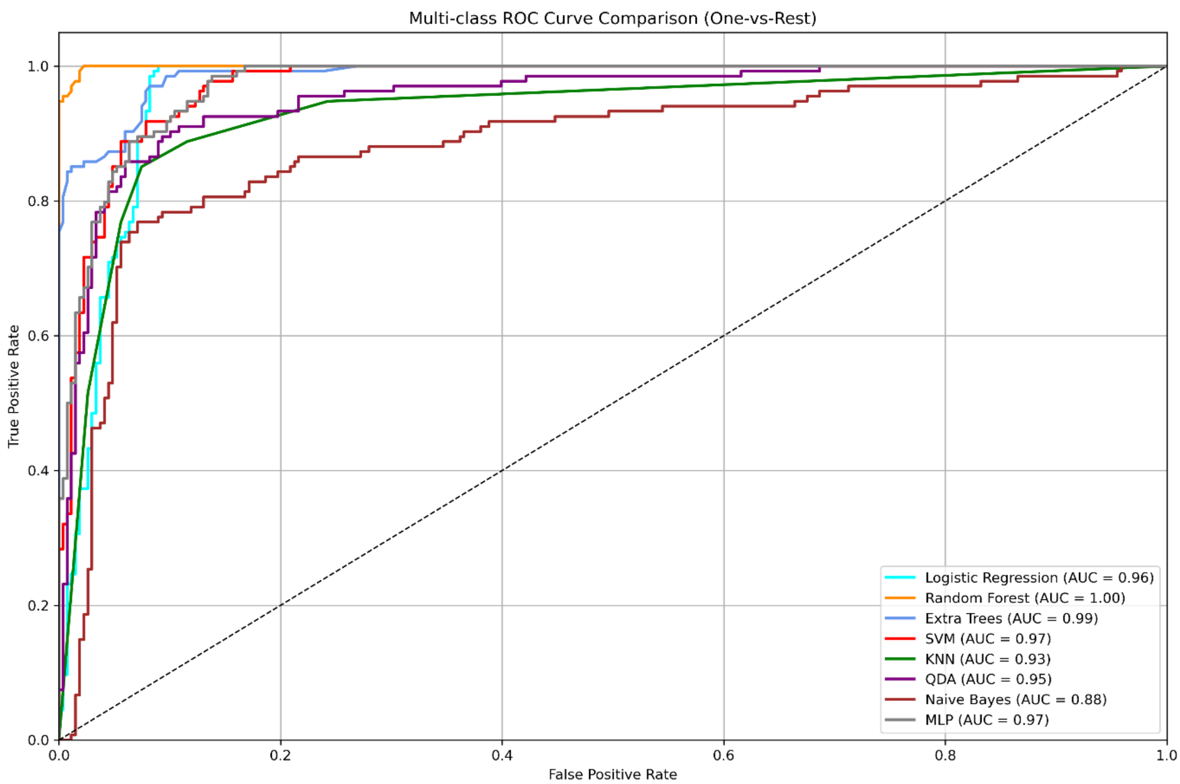

Figure 3. ROC–AUC curves of machine learning algorithms

Figure 3 presents the ROC–AUC curves obtained by the One-vs-Rest (OvR) method for the multiclass classification problem. The Random Forest model demonstrated the highest level of discriminatory strength with an Area Under the Curve (AUC) of 1.00, while Extra Trees (0.99), SVM and MLP (0.97) also exhibited high performance. Logistic regression (0.96) and QDA (0.95) have been shown to achieve successful results, while KNN (0.93) and Naive Bayes (0.88) have been demonstrated to possess lower discriminating power. An AUC of 1.00 indicates that the model ranks all positive samples with higher probability scores than negatives; however, it is likely that other metrics (e.g. accuracy, precision, sensitivity, F1-score) will be lower when a certain threshold value is used (Bishop &Nasrabadi, 2006). The observation that the curves of all models are situated above the random classification line suggests a significant degree of performance. The confusion matrices of all machine learning algorithms employed in Figure 4 are provided.

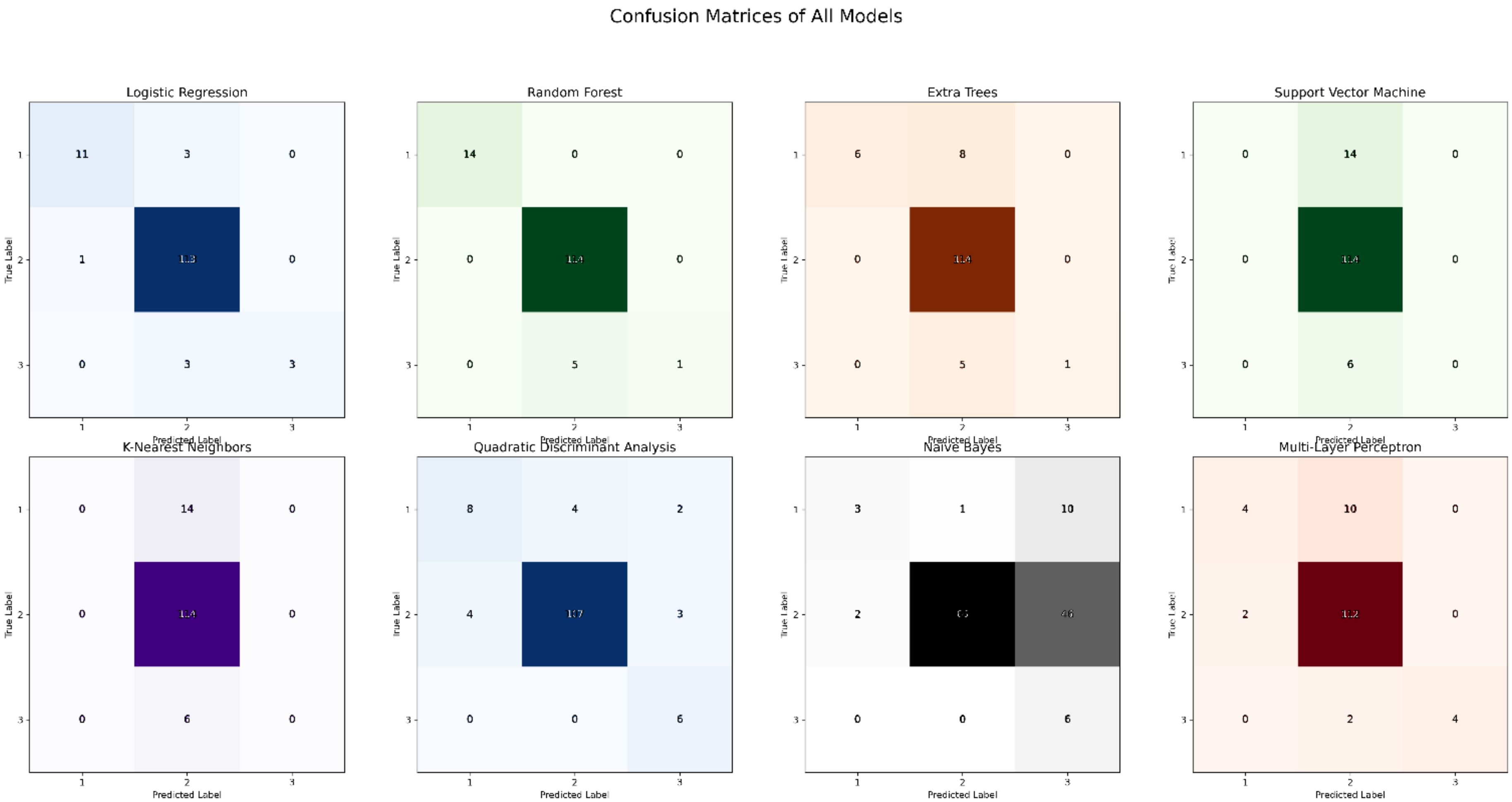

Figure 4. Confusion Matrices for Multi-Class Classification of Machine Learning Models

In multiclass classification problems, confusion matrix analysis provides detailed information on the extent to which the models correctly predict each class, and which classes are subject to confusion. A thorough examination of the confusion matrix results presented in Figure 4 reveals that the RF and SVM models demonstrate an impeccable capacity to accurately classify both class 1 and class 2 categories, exhibiting minimal misclassification in class 3. The efficacy of ensemble-based methods and kernel-based models is well-documented, with numerous studies demonstrating their superiority in scenarios where the distinction between classes is pronounced (Breiman, 2001). Despite the LR model demonstrating robust discrimination capability, as evidenced by its elevated AUC values in the ROC analysis, the outcomes of the confusion matrix analysis suggest a degree of ambiguity between Class 3 and Class 2. This finding indicates that, while the probability estimates of the model demonstrate a high degree of success, the implementation of decision thresholds can impede the accurate classification of data within specific categories (Fawcett, 2006). The simpler hypothetical models, Naive Bayes and QDA, exhibited marked confusion between classes, indicating that feature distributions do not conform to model assumptions and are insufficient to learn complex decision limits (Friedman, 2009).

The confusion matrix delineated the prediction accuracy of the model between classes and the types of errors. When this evaluation is taken together with other performance metrics, it is seen that the most successful results are obtained with the Random Forest model. Following the assessment of the model's performance, the identification of the variables influencing classification success enhances the interpretability of the model and facilitates the translation of findings into practical applications. For this purpose, a feature importance analysis was performed on the RF model, which exhibited the optimal performance, thereby revealing the most critical factors in obesity class estimation. In the context of complex data structures, feature importance has been identified as a significant tool for the identification of prominent variables. Furthermore, it has been demonstrated that this tool can be utilized to prioritize future intervention or policy recommendations by demonstrating which variables are more decisive in the decision mechanism of the model (Breiman, 2001). As illustrated in Figure 5, the feature importance of the top 15 variables with the highest severity is demonstrated.

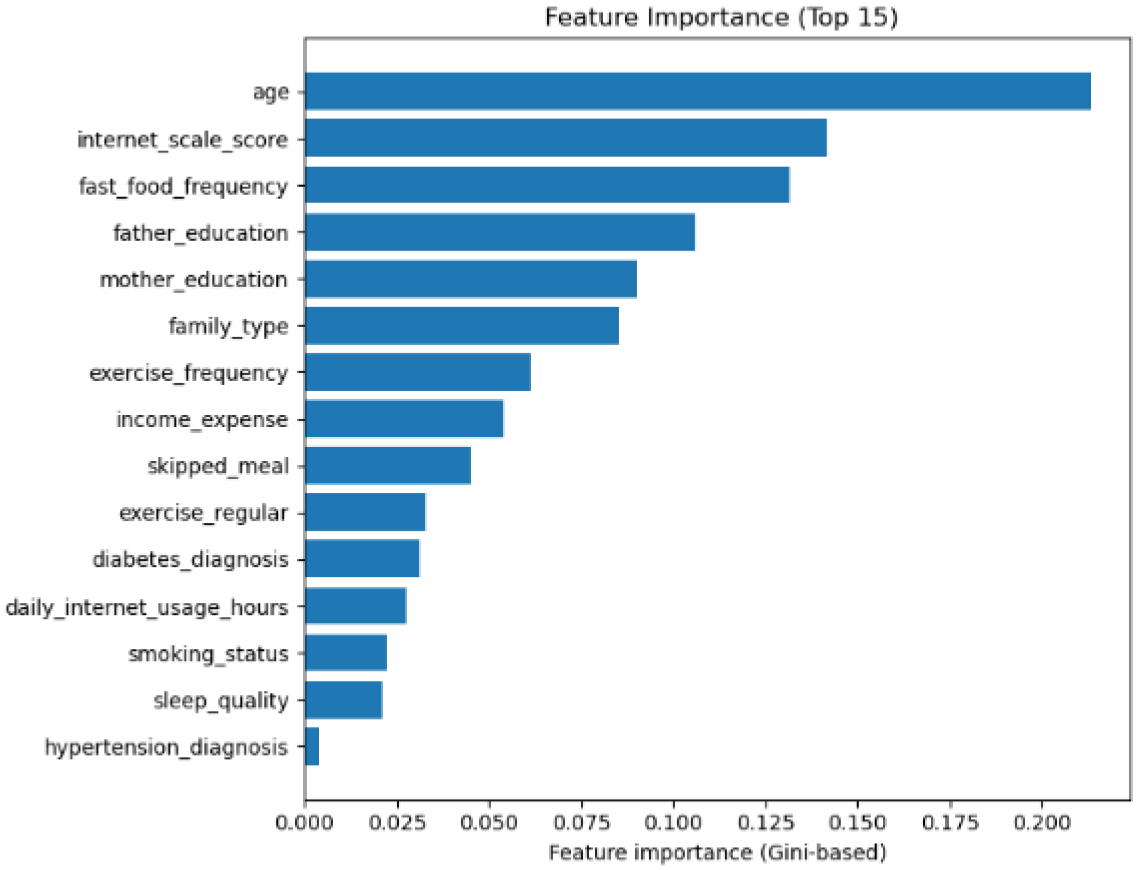

Figure 5. Random Forest Feature Importance (Top 15 Variables, Gini-based)

As demonstrated in Figure 5, it is evident that the variables of age, internet scale score and fast-food frequency exert the most significant influence on the model's classification efficacy. The findings of this study indicate that factors such as age, internet usage habits, and dietary patterns significantly influence the categorization of individuals into distinct obesity classes. Furthermore, variables such as parental education level, family type and exercise frequency were found to be significant contributors to the model, though their contributions were found to be less substantial than those of the initial three variables. The findings indicate that the model employs multidimensional lifestyle indicators with a high degree of effectiveness in distinguishing between obesity classes. This finding aligns with the conclusions of previous studies that have underscored the multifactorial nature of obesity, in which analogous variables have been identified (Alzahrani et al., 2020).

The correlation matrix of the top 15 variables with the highest importance, as determined according to the feature importance analysis, is presented in Figure 6.

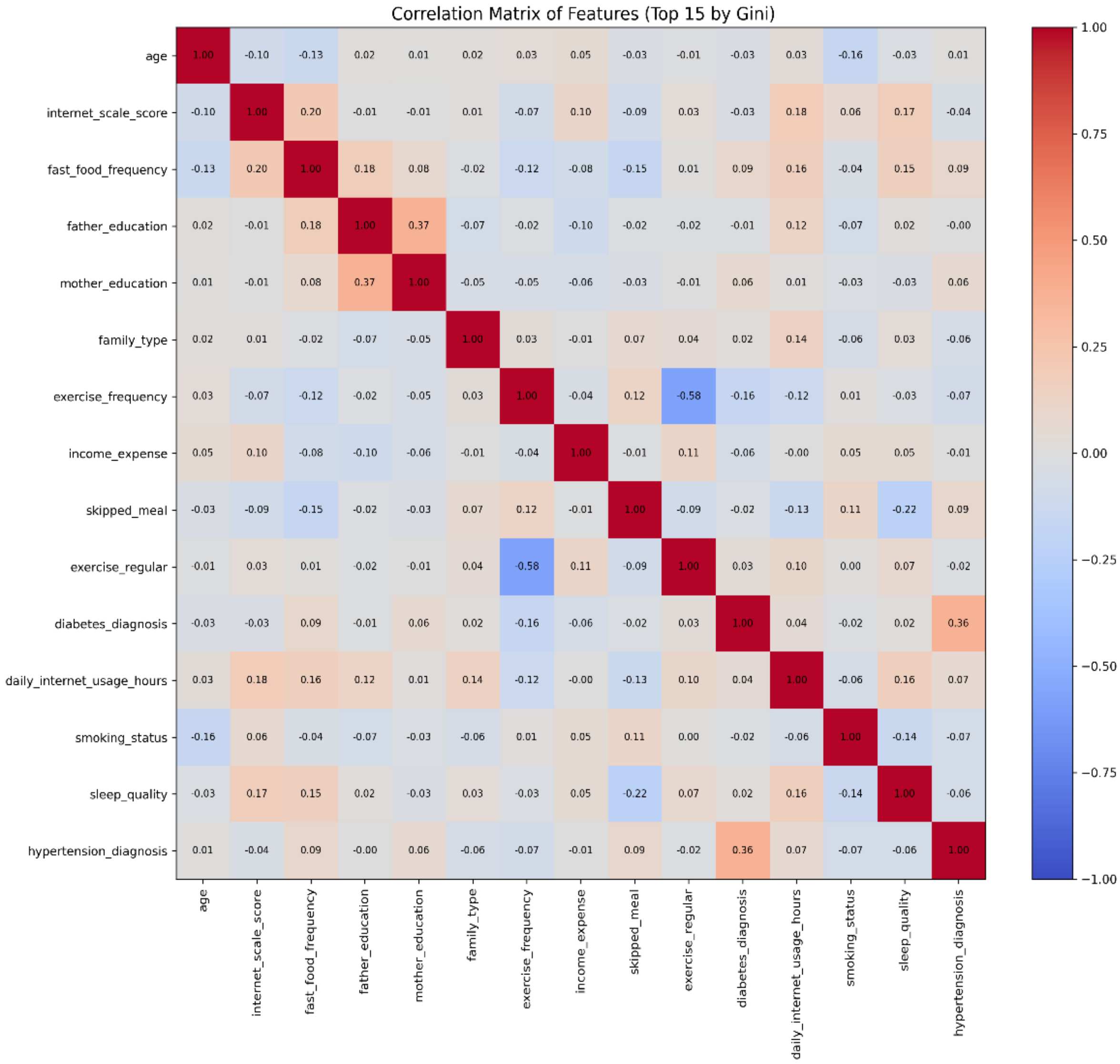

Figure 6. Correlation Matrix of the Top 15 Variables by Gini Importance

As illustrated in Figure 6, the Spearman correlation matrix of the top 15 most significant variables, as determined by the Gini coefficient, is presented. The observation that most correlation coefficients are below ±0.2 suggests that the model's predictive capability is derived from the collective influence of numerous low-intermediate correlated variables, rather than being driven by a single dominant variable. This finding indicates that the risk of multicollinearity is minimal and that the model can effectively utilize independent signals between features (Dormann et al., 2013). However, two exceptions merit particular attention: a high positive correlation between meal skipping and skipped meal (r = 0.73) and a moderate-high negative correlation between exercise regular and exercise frequency (r = –0.58). The initial finding indicates a robust correlation between the frequency of skipping meals and the occurrence of skipping meals, a phenomenon that has been frequently documented in research on eating behaviors. Furthermore, survey questions are often interpreted in accordance with respondents' expectations. The second finding suggests that the inverse relationship between the definition of regular exercise and exercise frequency may be since some participants perceive the phrase "regular exercise" as 1–2 days per week, and that similar measurement uncertainties have been reported in physical activity studies

(Al-Hazzaa et al., 2014). Furthermore, the correlations between demographic variables (e.g., age) and lifestyle indicators remain low (|r| < 0.2), indicating that the model will consider the characteristics in this group as independent signals. This finding is also consistent with the results of a study conducted on young adults in Turkey, which revealed that age and sex have an effect independent of other risk factors (Şengul et al. 2020). In conclusion, the correlation matrix both confirms the low risk of multiple links and shows that the relationships between nutrition and physical activity habits are consistent with findings in academic literature. This approach provides a robust statistical foundation for the interpretability and reliability of the classification model.

4. Discussion and Conclusion

Obesity is widely recognized as a major global public health concern, characterized by its adverse effects on individual well-being and its steadily increasing prevalence, which imposes a substantial burden on healthcare systems (WHO, 2024). The university period represents a critical transitional stage in which individuals begin to reshape their lifestyle preferences, often experiencing changes in dietary habits and a reduction in physical activity levels. During this period, the transition from family dependence to independent living, combined with academic pressure, altered social environments, and increased access to fast food, places young adults at a heightened risk for obesity. Therefore, accurately identifying obesity risk factors has become essential for developing effective preventive strategies and public health interventions.

In this study, the performance of eight different machine learning algorithms—Logistic Regression, Random Forest, Extra Trees, SVM, KNN, QDA, Naive Bayes, and MLP—was compared to classify obesity risk among university students. To address the class imbalance problem inherent in the data, the Adaptive Synthetic Sampling (ADASYN) method was applied exclusively to the training set, while the models were evaluated using standard classification metrics such as Accuracy, Precision, Recall, and F1-Score. The Random Forest algorithm achieved the highest accuracy rate (96.26%), followed by Logistic Regression (94.77%), which aligns with the literature emphasizing the robust generalization capacity and resistance to overfitting of ensemble-based methods (Breiman, 2001).

Feature importance analysis for the Random Forest model revealed that age, Internet Addiction Scale score, and fast-food consumption frequency were the most influential predictors in obesity classification. These findings indicate that behavioral and psychological factors, alongside demographic attributes, play a key role in determining obesity risk. Furthermore, variables such as parental education, family type, and exercise frequency contributed meaningfully to the model, albeit with relatively smaller effects. Correlation analysis also demonstrated a strong positive association between meal skipping and skipped meal type, supporting previous studies that have identified meal-skipping behavior as a significant predictor of obesity and poor nutritional outcomes among college students (Olagunju et al., 2024; Pendergast et al., 2016). Thus, the model effectively captures complex relationships between nutritional behaviors and lifestyle patterns, enhancing its interpretive power in classifying obesity risk. Although the sample consisted of university students, age remained a significant predictor of obesity classification. Even within this relatively homogeneous group, small age differences may reflect distinct lifestyle habits, activity levels, and metabolic variations that contribute to obesity risk. Furthermore, while daily internet usage hours were not a significant predictor, the Internet Addiction Scale, which measures behavioral and psychological dependences, showed a stronger relationship with obesity classification. This suggests that the qualitative nature of digital behavior may be more relevant than quantitative exposure in predicting obesity risk. From a methodological perspective, although ADASYN effectively reduced class imbalance, synthetic resampling may introduce slight variations in model consistency and generalizability, which should be acknowledged as a limitation. Moreover, as height and weight data were self-reported, potential reporting bias and measurement inaccuracies must be considered.

Overall, this study contributes to the growing body of research that integrates machine learning approaches into obesity risk prediction and prevention. The results highlight the multifactorial structure of obesity and demonstrate that digital behavior, dietary habits, and lifestyle characteristics are significant predictors even within a relatively homogeneous university population. These insights can guide policymakers and university administrators in developing evidence-based prevention programs. Campus-based initiatives that encourage healthy eating, regular exercise, and controlled internet use can help reduce obesity risk among young adults. Additionally, digital well-being campaigns focusing on the behavioral consequences of excessive internet use could contribute to healthier lifestyle patterns.

Future research should aim to extend the present findings by incorporating larger and more diverse samples from multiple universities or regions to enhance generalizability. Integrating objectively measured anthropometric indicators (e.g., height, weight, waist circumference) and biochemical data would improve the accuracy of obesity risk estimation. Moreover, longitudinal designs could help reveal causal pathways linking digital behavior, dietary practices, and obesity over time. Finally, employing hybrid deep learning and explainable AI methods could provide more comprehensive insights into the complex mechanisms underlying obesity among young adults.

References

1. Akın, P. (2023). A new hybrid approach based on genetic algorithm and support vector machine methods for hyperparameter optimization in synthetic minority over-sampling technique (SMOTE). *AIMS Mathematics*, *8*(6), 9400–9415.
2. Alzahrani, S. H., Saeedi, A. A., Baamer, M. K., Shalabi, A. F., & Alzahrani, A. M. (2020). Eating habits among medical students at king abdulaziz university, Jeddah, Saudi Arabia. International journal of general medicine, 77-88.
3. Bikku, T. (2020). Multi-layered deep learning perceptron approach for health risk prediction. Journal of Big Data, 7(1), 50.
4. Bishop, C. M., & Nasrabadi, N. M. (2006). Pattern recognition and machine learning (Vol. 4, No. 4, p. 738). New York: springer.
5. Breiman, L. (2001). Random forests. Machine learning, 45(1), 5-32.
6. Brownlee, J. (2020). Imbalanced classification with Python: better metrics, balance skewed classes, cost-sensitive learning. Machine Learning Mastery.
7. Chatterjee, A., Gerdes, M. W., & Martinez, S. G. (2020). Identification of risk factors associated with obesity and overweight—a machine learning overview. Sensors, 20(9), 2734.
8. Choudhuri, A. (2022). A hybrid machine learning model for estimation of obesity levels. In *Data management, analytics and innovation conference* (pp. 257–266). Springer. https://doi.org/10.1007/978-981-19-2600-6_22
9. Cortes, C., & Vapnik, V. (1995). Support-vector networks. Machine learning, 20(3), 273-297.
10. Cover, T., & Hart, P. (1967). Nearest neighbor pattern classification. IEEE transactions on information theory, 13(1), 21-27.
11. Dirik, M. (2023). Application of machine learning techniques for obesity prediction: a comparative study. Journal of complexity in Health Sciences, 6(2), 16-34.
12. Domingos, P., & Pazzani, M. (1997). On the optimality of the simple Bayesian classifier under zero-one loss. Machine learning, 29(2), 103-130.

13. Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., ... & Lautenbach, S. (2013). Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. Ecography, 36(1), 27-46.
14. Ferdowsy, F., Rahi, K. S. A., Jabiullah, M. I., & Habib, M. T. (2021). A machine learning approach for obesity risk prediction. Current Research in Behavioral Sciences, 2, 100053.
15. Fernández, A., García, S., Galar, M., Prati, R. C., Krawczyk, B., & Herrera, F. (2018). Learning from imbalanced data sets (Vol. 10, No. 2018, p. 4). Cham: Springer.
16. Fernández-Delgado, M., Cernadas, E., Barro, S., & Amorim, D. (2014). Do we need hundreds of classifiers to solve real world classification problems?. The journal of machine learning research, 15(1), 3133-3181.
17. Géron, A. (2022). Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow. " O'Reilly Media, Inc.".
18. Geurts, P., Ernst, D., & Wehenkel, L. (2006). Extremely randomized trees. Machine learning, 63(1), 3-42.
19. Friedman, J. (2009). The elements of statistical learning: Data mining, inference, and prediction. (No Title).
20. He, H., Bai, Y., Garcia, E. A., & Li, S. (2008, June). ADASYN: Adaptive synthetic sampling approach for imbalanced learning. In 2008 IEEE international joint conference on neural networks (IEEE world congress on computational intelligence) (pp. 1322-1328). Ieee.
21. Helforoush, Z., & Sayyad, H. (2024). Prediction and classification of obesity risk based on a hybrid metaheuristic machine learning approach. Frontiers in big Data, 7, 1469981.
22. Hosmer Jr, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). Applied logistic regression. John Wiley & Sons.
23. Hruby, A., & Hu, F. B. (2015). The epidemiology of obesity: a big picture. Pharmacoeconomics, 33(7), 673-689.
24. Kotsiantis, S., Kanellopoulos, D., & Pintelas, P. (2006). Handling imbalanced datasets: A review. GESTS international transactions on computer science and engineering, 30(1), 25-36.
25. Musa, F., & Basaky, F. (2022). Obesity prediction using machine learning techniques. *Journal of Applied Artificial Intelligence*, *3*(1), 24–33.
26. Murtagh, F. (1991). Multilayer perceptrons for classification and regression. Neurocomputing, 2(5-6), 183-197.
27. Naidu, G., Zuva, T., Sibanda, E.M. (2023). A Review of Evaluation Metrics in Machine Learning Algorithms. In: Silhavy, R., Silhavy, P. (eds) Artificial Intelligence Application in Networks and Systems. CSOC 2023. Lecture Notes in Networks and Systems, vol 724. Springer, Cham. https://doi.org/10.1007/978-3-031-35314-7_2
28. Nelson, M. C., Story, M., Larson, N. I., Neumark-Sztainer, D., & Lytle, L. A. (2008). Emerging adulthood and college-aged youth: an overlooked age for weight-related behavior change. Obesity.
29. Olagunju, M. T., Aleru, E. O., Abodunrin, O. R., Adedini, C. B., Ola, O. M., Abel, C., ... & Akinsolu, F. T. (2024). Association between meal skipping and the double burden of malnutrition among university students. North African Journal of Food and Nutrition Research, 8(17), 167-177.
30. Şengul, S., Lopcu, K., & Cam, S. (2020). Determinants of the obesity of adults in Turkey: An empirical study. Review of applied socio-economic research, 20(2), 60-71.
31. Pendergast, F. J., Livingstone, K. M., Worsley, A., & McNaughton, S. A. (2016). Correlates of meal skipping in young adults: a systematic review. International Journal of Behavioral Nutrition and Physical Activity, 13(1), 125.
32. Rish, I. (2001, August). An empirical study of the naive Bayes classifier. In IJCAI 2001 workshop on empirical methods in artificial intelligence (Vol. 3, No. 22, pp. 41-46).
33. Şahin, C., & Korkmaz, Ö. (2011). İnternet bağımlılığı ölçeğinin Türkçeye uyarlanması. Selçuk Üniversitesi Ahmet Keleşoğlu Eğitim Fakültesi Dergisi, 32(1), 101-115.
34. World Health Organization. (2024). Obesity and overweight. https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight
35. Yağmur, N. (2024). A hybrid approach to obesity level determination with decision tree and pelican optimization algorithm. *Journal of Scientific Reports-A*, *57*, 97–109. https://doi.org/10.59313/jsr-a.1447814