ORIGINAL ARTICLE

A Glance at the Relative Importance of ATP III Criteria in Metabolic Syndrome: Deep Learning Approach

Mojtaba Hajihasani[1*], **Rassoul Hajizadeh**[2]

[1] *Faculty of Engineering Modern Technologies, Amol University of Special Modern Technologies, Amol, Iran*

[2] *Machine Learning and Deep Learning Laboratory, Faculty of Engineering Modern Technologies, Amol University of Special Modern Technologies, Amol, Iran*

*Corresponding Author: Mojtaba Hajihasani

Email: mhajihasani@ausmt.ac.ir

Received: 07 September 2025 / Accepted: 30 December 2025

Abstract

Purpose: Metabolic Syndrome (MetS) significantly increases the risk of cardiovascular disease and is characterized by a combination of atherogenic dyslipidemia, central obesity, and hypertension. This study aims to investigate additional factors contributing to MestS and evaluate their relevance by employing advanced analytical techniques.

Materials and Methods: Specifically, we utilized ensemble learning, a supervised learning approach that combines multiple individual models to enhance predictive performance. Input factors were ranked through a univariate feature selection technique, while unsupervised methods were employed to mitigate biases associated with traditional labeling based on textbook ATP III criteria. An autoencoder was trained to reduce dimensionality without requiring prior knowledge of the data.

Results: Our findings demonstrated an accuracy of 0.9836 ±0.0298 with 95% CI: [0.9724, 0.9947] and calibration Brier score 0.0012 ± 0.0027 achieved by the gradient boosting classifier excluding six main ATP III defined factors. Contrary to the expectation of identifying only the main factors, our analysis identified HOMA-IR, BMI, age, and cholesterol as additional key predictors of MetS. Furthermore, the autoencoder's compressed representation revealed a distinctive three-state classification, rather than the conventional binary classification of MetS presence or absence. Tukey HSD post-hot test confirms that the three latent clusters are not only statistically distinct but also clinically ordered: cluster 1 represents a high-risk MetS profile, cluster 0 a low-risk profile, and cluster 2 an intermediate or mixed risk profile.

Conclusion: The results indicate that obesity, followed by high blood sugar and Hypertriglyceridemia, Insulin Resistance (HOMA-IR), hypertension, to a lesser extent, age, and cholesterol are the most essential features of the MetS prediction. Additionally, the discovery of a distinct third cluster provides an alternative insight into MetS, suggesting the potential for developing an intermediate/subtype MetS that helps personalized patient management strategy.

Keywords: Metabolic Syndrome; Adult Treatment Panel III; Ensemble Learning; Feature Selection; Clustering; Autoencoder.

DOI:

1. Introduction

Metabolic Syndrome (MetS) is a cluster of risk factors that increases the risk of cardiovascular disease and type 2 diabetes [1, 2]. It includes elevated blood pressure, dyslipidemia, raised fasting glucose, and central obesity. Various organizations have proposed diagnostic criteria differing mainly in their definitions of central obesity. The International Diabetes Federation and the American Heart Association/National Heart, Lung, and Blood Institute have proposed differing criteria, different waist measurement thresholds [3]. An individual is diagnosed with metabolic syndrome if three or more of the five abnormal findings are present, with waist circumference serving as a useful screening tool. The syndrome is associated with insulin resistance, atherogenic dyslipidemia, central obesity, and hypertension, and if left untreated, it significantly increases the risk of diabetes and cardiovascular diseases [4].

The key risk factors include central obesity, dyslipidemia, hypertension, impaired glucose metabolism, and poor physical fitness - the cluster of which defines metabolic syndrome [5, 6]. Dyslipidemia is characterized by elevated triglycerides and low HDL cholesterol levels. Low physical fitness and reduced cardiovascular or muscular capacity are also associated with an increased risk of metabolic syndrome, particularly in children [7]. Demographic factors such as age, sex, and ethnicity also influence metabolic syndrome risk. The prevalence is higher among males, Hispanics, and older individuals [8].

Furthermore, the clustering of metabolic syndrome components can help identify individuals at risk of developing cardiovascular disease and type 2 diabetes. For instance, one study found that the clustering of metabolic syndrome factors was associated with increased carotid Intima-Media Thickness (cIMT) in children and adolescents [9]. Similarly, another study found the clustering of metabolic syndrome factors was associated with increasing risk of cardiovascular disease and type 2 diabetes in adults [10].

In medical terminology, clustering refers to data analytics techniques such as factor analysis, Cox regression, Fisher exact test, and bootstrap [10]. These techniques utilize to understand the pattern of clustering of the basic variables of the syndromes. Metabolic syndrome clustering is a significant predictor of cardiovascular disease and type 2 diabetes, and its components are influenced by lifestyle factors [11, 12]. Early detection and management of metabolic syndrome clustering can help reducing the risk of these conditions.

In recent years, Artificial Intelligence (AI) has been increasingly applied in the medical context, focusing on improving diagnosis, treatment, and management. AI-based methodologies have been employed to analyze large datasets, identify patterns, and predict the risk of metabolic syndrome [13-17].

AI has many applications in metabolic syndrome, such as: i) Predictive Modeling: Machine learning algorithms have been used to develop predictive models for identifying individuals at risk of metabolic syndrome. These models incorporate clinical, genetic, and factors to accurately predict the risk [15, 18]; ii) Disease diagnosis: AI-based systems have been developed to diagnose metabolic syndrome using parameters such as blood pressure, blood glucose levels, and Body Mass Index (BMI) and can also assess severity and recommend treatment [19]; iii) Personalized medicine: AI has been used to develop personalized treatment plans for patients with metabolic syndrome. This involves analyzing individual patient data and recommending tailored interventions based on their specific needs [16]; iv) Health monitoring: AI-based systems have been designed to monitor patients with metabolic syndrome, tracking their health status and alerting healthcare providers to any changes or potential complications [20].

Therefore, machine learning algorithms have been extensively used in metabolic syndrome research to analyze large datasets, identify patterns, and predict the risk of metabolic syndrome [21, 22]. Commonly used algorithms include: XGBoost, Stacking Ensemble Algorithm [15]; Logistic Regression, Multi-Layer Perceptron Neural Network [23, 24], Naive Bayes, Random Forest, LogitBoost, BayesNet [18], [25]; Extreme Boosting with Random Forest (XGBRF), CatBoost [19, 26]. These algorithms have been applied to predict the occurrence of metabolic syndrome with improved accuracy, enhanced predictive power, early detection, reduced bias, reduced healthcare costs, enhanced decision-making, and enhanced decision-making. In this study, when a supervised model is trained, the input factors are analyzed and ranked based on their relevance to the

MetS definition using univariate feature scores and the receiver operating characteristic area under the curve (ROC AUC) metric [27- 29].

Supervised machine learning models work based on labeling derived from textbook criteria such as the ATP III guidelines [30]. Although, it is valuable for all above-mention benefits, it is still biased by the criteria definition that may affect the data mining applications. In fact, the main purpose of supervised learning is model prediction and control. If it comes a problem to analysis the data for new aspects, the labeling of training data influences the investigation through the data. In such a case, we utilize an unsupervised technique, specifically, an autoencoder, to detect anomalies and to reduce dimensions [31].

An autoencoder is a type of artificial neural network that learns to represent data in a compressed, efficient manner [32]. It is trained by minimizing the difference between the original input data and the reconstructed output. This forces the encoder to capture the most important features of the data.

In section 3 ("Material and Method"), we discuss the dataset and the supervised and unsupervised models in detail. Our findings are discussed in the "Result" and "Discussion" sections, respectively. Finally, the study will be concluded.

1.1. Statement of Significance

(1) Problem or Issue	Metabolic syndrome (MetS) significantly increases the risk of cardiovascular disease and is characterized by a combination of atherogenic dyslipidemia, central obesity, and hypertension. This study investigates additional factors contributing to MetS and their relevance by employing advanced analytical techniques.
(2) What is Already Known	We have a few key criteria, such as ATP III, which checks Central obesity, Hypertriglyceridemia, Low HDL-C, hypertension, and fasting plasma glucose to diagnose whether the subject is MetS or non-MetS.
(3) What this Paper Adds	Achieved 98.22% accuracy using gradient boosting supervised learning on 23 extra factors, including those factors that are required in ATP III, and resulted in HOMA-IR, BMI, age, and cholesterol being also significant factors for diagnosing. Additionally, unsupervised autoencoder learning is employed and revealed a novel three-state MetS classification rather than only two classic states of either MetS or non-MetS.
(4) Who would benefit from the new knowledge in this paper	This discovered transitional cluster and four extra effective factors indicates early-stage metabolic syndrome risk. It highlights the potential for AI-driven personalized patient management strategies.

2. Materials and Methods

2.1. Subjects

In this study, a hospital-based dataset was used for evaluation. This hospital-based cross-sectional study included a total of 2555 subjects aged 20-69 years (31% male and 69% female), comprising males to female's ratio of 1:2.2. Since the MetS prevalence in Iran is between 1.3 and 1.6 [18], this indicates a mildly imbalanced dataset. On the other hand, according to the ATP III definition, the prevalence of MetS vs. non-MetS is 1132 vs. 1432 cases (i.e., the ratio of 1:1.2), which is within the balanced dataset. However, utilizing the random oversampling technique for the training dataset would be a shrinking solution for mild imbalance in such medical cases [19, 33]. The participants were registered from those patients who visited the outpatient endocrine clinic at Tehran University General Hospital and underwent health examinations. Subjects with conditions that could affect glucose metabolism, such as use of glucocorticoids or chemotherapy agents, partially treated thyroid disease, malignancies, cirrhosis, renal failure, pregnancy, or lactation, were excluded. In our study, 679 (22.7%) of the patients had a history of antihypertensive drug use. The clinical and laboratory characteristics of the subjects are shown in Tables 1 and 2, which list 22 continuous and categorical factors, including 23 factors in total. The study was approved by the Ethics Review Committee on human experimentation. Written informed consent was obtained from all participants prior to inclusion.

Table 1. Statistics of the continuous-valued factors in the dataset

	Age	Height	Weight	Waist-circ	Hip-circ	SBP	DBP	FBS	Insulin	Chol	HDL	LDL	TG	Cr	C-Peptide	BMI
Mean	45.0	161.4	75.2	94.4	105.4	121.8	78.5	126.9	9.1	197.1	48.0	115.3	167.13	0.9	2.5	28.9
Std	13.0	8.8	14.9	12.6	10.9	18.7	9.5	57.0	5.9	42.5	12.0	34.5	104.3	0.3	1.2	5.2
Min	19	139	41	62	72	80	60	63	1	90	16	15	38	0.4	0.3	17.3
25%	34	155	65	86	99	110	70	89	5	168	40	91	100	0.8	1.8	25.25
50%	46	160	73	94	104	120	80	100	8	193	47	113	145	0.97	2.2	28.1
75%	56	167	83	102	111	130	80	151	11.9	220	55	135	199	1	3	31.63
Max	69	190	180	159	191	220	120	470	100	497	169	361	996	11	13	55.56

Table 2. Count the value of categorical factors in the dataset

Factors	Negative	Positive
DM	1448	1113
Family history DM	1314	1247
HTN	1806	755
Family history of HTN	1560	1001
Family history of CAD	1798	763

2.2. Supervised Models

As mentioned earlier, ranking the contributing factors can provide valuable insights for detecting metabolic syndrome. To model the complex interactions between factors and MetS, we employed ensemble learning methods based on the ATP III criteria.

In this context, an efficient classifier is required to achieve high discrimination between MetS and no MetS subjects. Ensemble learning is a machine learning technique that combines multiple individual models to improve overall predictive performance [34]. By combining the predictions of multiple models, ensemble learning can help reduce variance, increase accuracy, and improve generalization. Two common examples of ensemble methods are gradient-boosted trees and random forests [35]. Moreover, ensemble models can be applied to other non-tree-based algorithms using (bagging) or boosting methods such as AdaBoost [35]. Additionally, the Support Vector Machine (SVM) with radial basis function kernel was used as a baseline model [36].

After creating the model, we analyzed for ranking of input features based on their relevancy of MetS definition [27]. In the context of feature selection, classification models were trained in two scenarios of including and excluding the ATP III-defined features. Then, the feature selection technique SelectKBest utilized f_classif univariate feature score, which calculates the ANOVA F-value between label/feature for classification tasks [28, 29, 37].

These tests assign a score to each feature based on its relationship with the label, and features with higher scores are considered more informative for predicting MetS.

2.3. Unsupervised Clustering Model

Among various clustering techniques, we employed an autoencoder, which is a type of deep neural network that learns to represent data in a compressed, efficient form, making it suitable for visualization and pattern discovery. An autoencoder consists of two main components:

i) **E**ncoder: compresses the input data into a latent-space representation;

ii) **D**ecoder: reconstructs the original data from the compressed latent representation. The Latent-space layer contains the most important features extracted from the data.

Autoencoders have a wide range of applications, including dimensionality reduction, feature extraction, anomaly detection, and generative modeling [32, 38]. There are many different types of autoencoders that exist, each with its own advantages and limitations. In this study, we utilize stacked autoencoders Figure 1, which are composed of multiple layers that allow the model to learn more complex representations of the data.

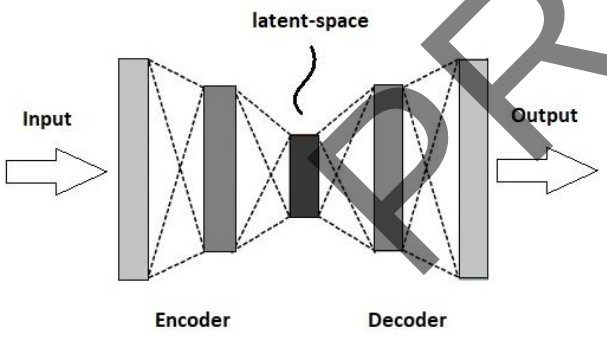

Figure 1. Illustrates the structure of the autoencoder, which includes three decreasing layers in the encoder to compress data into the latent space, followed by two expanding layers in the decoder to reconstruct the output

3. Results

3.1. Modeling the MetS Based on the ATP III Guideline

Initially, according to the ATP III guideline, we label the dataset with six main factors, i.e., WC, TG, HDL-C, systolic BP, diastolic BP, FPG, and/or specific medications/previously diagnosed, including DM and HTN as the input and diagnosing the MetS or non-MetS as the output. Furthermore, the five mentioned classifiers are trained by 23 factors dataset in which the training and test sets are randomly split by the ratio of 67% and 33%, respectively, utilizing 10-fold cross-validation. The binary classification performance reports by accuracy, F1-score, and PR-AUC are accompanied by their 95% confidence interval for comparison. Additionally, the Brier score is calculated, which is a calibration assessment, and the lower is better.

The classification performance is performed in two scenarios to dodge leakage: the former is baseline, which includes the main defining factors plus all-other factors reported in Table 3, and the latter is all-other factors excluding the main ones reported in Table 4.

These two tables compare how well five machine-learning models distinguish MetS from non-MetS using several performance metrics over repeated runs (mean ± standard deviation with 95% confidence intervals). Gradient Boosting is by far the best model in both scenarios reported in Tables 3 and 4: accuracy and F1 are about 0.999 and 0.98, respectively, with extremely tight confidence intervals; PR-AUC is essentially 1.0 and 0.985, respectively, and the Brier score is almost zero and 0.014, respectively.

3.2. Feature Selection

After creating the model, we analyze the input factors and rank the factors based on their relevance to the MetS definition. In the context of feature selection, the univariate score, which calculates the ANOVA F-value, is applied on both models derived by two scenarios to evaluate the performance of a classification model after selecting a subset of features. For this purpose, we utilize the best model from Tables 3 and 4, i.e., the Gradient Boosting classifier by baseline and excluding ATP III defining features.

As it is shown in Figure 3, the distribution of binary classes (MetS or no MetS) against each univariate factor is analyzed by the univariate score for both scenarios depicted in parts a) and b).

Features with higher scores are considered more relevant or informative for predicting the target variable. Table 5 lists the top 10 features selected according to Figure 2.

Table 3. Classification performances of the models of baseline to diagnose MetS or no MetS

Classifiers	Accuracy Mean±STD [CI 95%]	F1-Score	PR-AUC	Brier Score
Gradient Boosting	**0.9988 ± 0.0027 [0.9978, 0.9999]**	**0.9987 ± 0.0031 [0.9975, 0.9998]**	**0.9998 ± 0.0007 [0.9995, 1.000]**	**0.0012 ± 0.0027 [0.0001, 0.0022]**
SVM	0.8555 ± 0.0225 [0.8471, 0.8639]	0.8385 ± 0.0267 [0.8285, 0.8485]	0.9202 ± 0.0214 [0.9122, 0.9282]	0.0997 ± 0.0126 [0.0950, 0.1044]
Random Forest	0.925 ± 0.016 [0.919, 0.931]	0.9145 ± 0.0195 [0.9072, 0.9218]	0.9778 ± 0.0073 [0.9750, 0.9805]	0.061 ± 0.007 [0.058, 0.064]
Bagging	0.856 ± 0.025 [0.847, 0.866]	0.8355 ± 0.0303 [0.8242, 0.8468]	0.9135 ± 0.0287 [0.9027, 0.9242]	0.1016 ± 0.0139 [0.0964, 0.1068]
Ada Boost	0.928 ± 0.013 [0.923, 0.932]	0.9183 ± 0.0151 [0.9127, 0.9240]	0.9778 ± 0.0078 [0.9749, 0.9807]	0.1631 ± 0.0036 [0.1618, 0.1644]

Table 4. Classification performances of the models excluding the main guideline factors to diagnose MetS or no MetS

Classifiers	Accuracy Mean±STD [CI 95%]	F1-Score	PR-AUC	Brier Score
Gradient Boosting	**0.9836 ± 0.0298 [0.9724, 0.9947]**	**0.9816 ± 0.0332 [0.9692, 0.9940]**	**0.9859 ± 0.0290 [0.9751, 0.9967]**	**0.0140 ± 0.0252 [0.0046, 0.0234]**
SVM	0.7864 ± 0.0208 [0.7787, 0.7942]	0.7659 ± 0.0241 [0.7569, 0.7749]	0.8172 ± 0.0275 [0.8069, 0.8274]	0.1491 ± 0.0114 [0.1449, 0.1534]
Random Forest	0.8020 ± 0.0253 [0.7925, 0.8114]	0.7778 ± 0.0293 [0.7669, 0.7887]	0.8514 ± 0.0272 [0.8413, 0.8616]	0.1384 ± 0.0111 [0.1343, 0.1426]
Bagging	0.7578 ± 0.0245 [0.7487, 0.7670]	0.7251 ± 0.0301 [0.7138, 0.7363]	0.7870 ± 0.0367 [0.7733, 0.8007]	0.1604 ± 0.0122 [0.1558, 0.1650]
Ada Boost	0.8073 ± 0.0237 [0.7985, 0.8161]	0.7818 ± 0.0274 [0.7715, 0.7920]	0.8596 ± 0.0234 [0.8509, 0.8683]	0.2021 ± 0.0032 [0.2009, 0.2033]

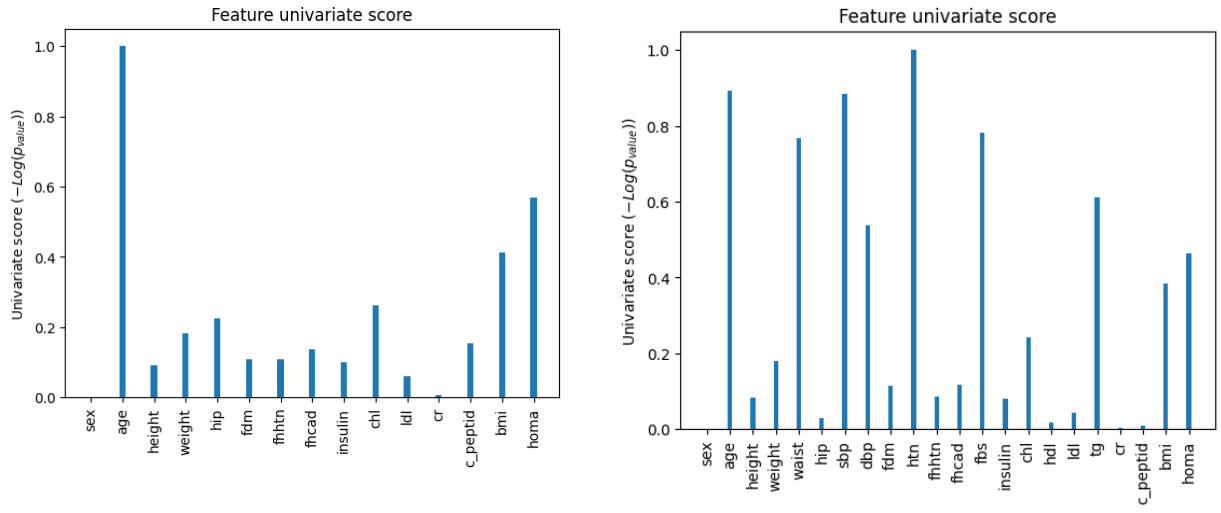

Figure 2. Feature selection techniques: a) excluding defining features; b) including defining features

3.3. Autoencoder Unsupervised Learning

In this section, the factors are restricted to those only used in the ATP III guideline, i.e., six main defined factors (WC, TG, HDL-C, systolic BP, diastolic BP, FPG). Therefore, six factors are under investigation. Thereafter, to find the best autoencoder structure, hyperparameter tuning is utilized, and the number of neurons in each layer is chosen as 16, 8, 3, 8, and 16. The activation functions adjust rectified linear units (ReLU) except for the output layer, which is a sigmoid function. In the training step, the Adam optimizer tries to minimize the Mean Absolute Error (MAE) loss function in which it reaches to 0.05 in early stopping.

The primary objective of setting the latent space of the autoencoder to 3 is to effectively represent and visualize the data. The output of the encoder is plotted in Figure 3. As it is shown, there are three distinct clusters in Figure 3a, accompanied by its projection in a 2D plane as in Figure 3b, which is marked by a red asterisk (*) as their corresponding centroids. As it is clear from Figure 3, Cluster number three was chosen a priori to explore potential subgroups beyond the binary ATP III definition (MetS vs non-MetS) and to allow an intermediate/subtype cluster. Internal validation of the k-means clustering in latent space yielded a Silhouette coefficient of 0.24, a Davies–Bouldin index of 2.44, and a Calinski–Harabasz score of 249.54, indicating a reasonable separation between the three clusters.

To have a better insight into these three distinct clusters, we need to investigate their properties and distribution on factors. It postpones and will discuss in the next (discussion) section.

Table 5. Top best features with the highest relevance of Mets. Based on the ANOVA f-statistic

Ranking	Exclude ATP III Defining Features (out of 1.0)	Include ATP III Defining Features (out of 1.0)
1	Age (1.0),	HTN (1.0)
2	HOMA-IR (0.56),	SBP (0.89)
3	BMI (0.41),	Age (0.88)
4	CHOL (0.26),	FBS (0.78)
5	Hip (0.22),	Waist (0.76)
6	Weight (0.18),	TG (0.61)
7	C_Peptid (0.15),	DBP (0.53)
8	Family History CAD (0.13),	HOMA-IR (0.46)
9	Family DM (0.109),	BMI (0.38)
10	Family History HTN (0.108),	CHOL (0.24)

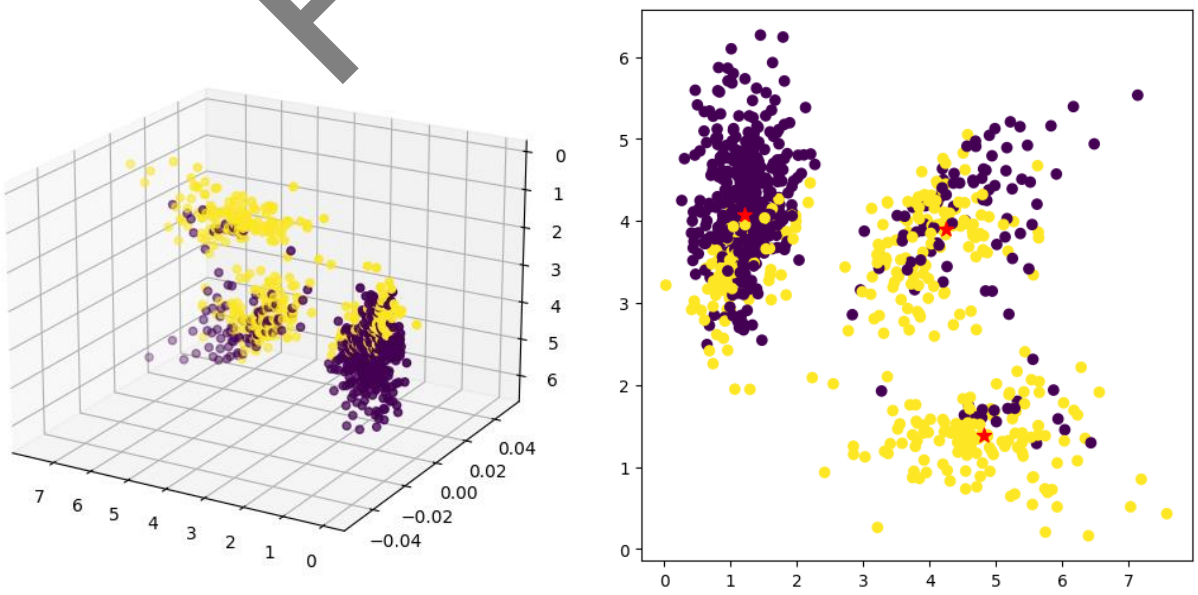

Figure 3. a) compressed output of the encoder in three dimensions and b) its projection in the 2D plane with its centroids

4. Discussion

In this study, a hospital dataset includes 2555 instances, is recognized as MetS vs non-MetS according to the ATP III guideline. Based on feature selection techniques it is found the relationship of 23 factors with MetS. Because the ATP III criteria for diagnosing MetS are related to the pre-diabetic phase, the patients who are prone to diabetes are detected more or less late by these criteria. Therefore, we tried to find out more related factors in order to predict the disease as soon as possible.

It has been found that the best fit model for MetS is through five different classifiers, including ensemble methods and SVM. The best accuracy performance is 98.22% and 99.8%, respectively, for baseline and excluding defined features scenarios, which are obtained by a gradient boosting classifier.

Subsequently, a univariate feature selection technique based on the ANOVA F-statistic was utilized to analyze the factors presented in Figure 2 and Table 5. These scores help identify the most relevant features for the classification task. By applying these feature selection techniques, the key factors influencing the classification performance were identified and analyzed.

Since the ATP III guideline is used as our labels, it is expected to achieve those six main factors on top. However, you can find HOMA-IR, BMI, Age, and Cholesterol among other main factors on top. Even though the HDL-C is among the main factors in the ATP III guideline, it is ranked out of the top ten. It shows that it is possible to pay more attention to these new factors aside the definitions of MetS, which need well-defined criteria, especially between different countries, races, and ethnics [39]. As a result, it's better to make a more precise, flexible, and comprehensive definition of MetS by gathering all related factors. In another point of view, by using the ATP III guideline, the risk factors have all equal values; while Table 5 shows that the DBP and WC are of lower importance compared to others, and also SBP and FBS may be of higher priority in diagnosis.

In our research, with a distribution of 31% male and 69% female, our analysis reveals age and sex among the top, which means age has a significant influence on the MetS definition than sex (Figure 2). The significance of age in our study aligns with the findings of Ma *et al.* [40].

Nonetheless, the factors of age, HOMA-IR, BMI, Cholesterol, Hip circumference, Weight, C_Peptid appear on the top according to the model of excluding ATP III defined factors, as illustrated in Table 5. Our findings in terms of central obesity are in support of Kahn's suggestion and confirm IDF's new definition of MetS, which insists on the importance of central obesity [41].

Although both SBP and DBP are two main criteria in ATP III, our study reveals that they are the 2nd and 7th related factors, respectively. Perhaps higher-ranking factors such as BMI or HOMA-IR may be more closely associated with stress than DBP.

Servais *et al.* have shown in their study that plasma creatinine is not significantly different in metabolic syndrome patients compared with the control group [42]. Our assessments also show that creatinine does not have any influence on diagnosing MetS.

HOMA-IR factor is the simplest insulin resistance measurement. Although in our study, the insulin-receptor sensitivity has a significant relevance (8th), insulin level (16th), and C-peptide level (20th) according to Figure 2b in blood have much lower relevance than HOMA-IR. Esteghamati *et al.* have shown that HOMA-IR is correlated and associated with MetS [43]. While this study confirms that HOMA-IR holds significant relevance to MetS, ranking as the 4th factor among 23 factors.

Anthropometric factors, including BMI, weight, and height are the 9th, 11th, and 14th by feature selection score. This discovery suggests that the combined effect of weight and height is more influential than each factor individually. Esteghamati *et al.* have also shown the correlation of BMI and Mets [44].

In the unsupervised scheme, factor relevance or finding the best model for MetS are not probed; instead, it is tried to visualize instances in three dimensions and transform into another axis to get a different point of view and insight about MetS. We excluded the other factors from the dataset and retained only the primary 6 factors outlined in the ATP III guidelines. We refrained from labeling any instances in the dataset according to the guidelines to ensure that our unsupervised learning approach

remains impartial to predefined symptoms or diseases. The rationale behind focusing solely on the main factors from the ATP III guidelines is to assess whether unsupervised data mining can effectively segregate instances into clusters that align with the MetS guidelines or unveil alternative perspectives.

Remarkably, Figure 3 illustrates three distinct clusters. To delve deeper into these clusters, we will analyze their characteristics. As depicted in Figure 4, Clusters #1 and #3 exhibit counting values correlating with non-MetS and MetS, respectively. Cluster #2 represents a distinct cluster that showcases characteristics of instances on the MetS boundary margin. It is shown in Figure 4 that almost half-half are labeled as MetS vs no-MetS based on the ATP III guideline.

To assess whether the latent clusters differed in cardiometabolic profile, the six ATP III components were compared across clusters using a one-way ANOVA test. Several factors showed statistically significant differences between clusters, as tabulated in Table 6, indicating that the three unsupervised clusters are clinically distinct in terms of MetS severity and composition.

Tukey HSD post-hoc test after one-way ANOVA test is applied on all main factors for pairwise cluster comparison. This analysis of six factors is reported in Table 7. These post-hoc results show that the three clusters differ meaningfully on almost all six metabolic factors; only one pairwise comparison for fasting blood sugar is not significant.

Table 6. The six ATP III components were compared across clusters using a one-way ANOVA test

Factor	F-statistic	p-value
WC	105.14	2.93e-41
SBP	1756.81	0.000e+00
DBP	3730.07	0.000e+00
FBS	32.76	2.11e-14
HDL	118.43	1.40e-45
TG	130.90	0.000e+00

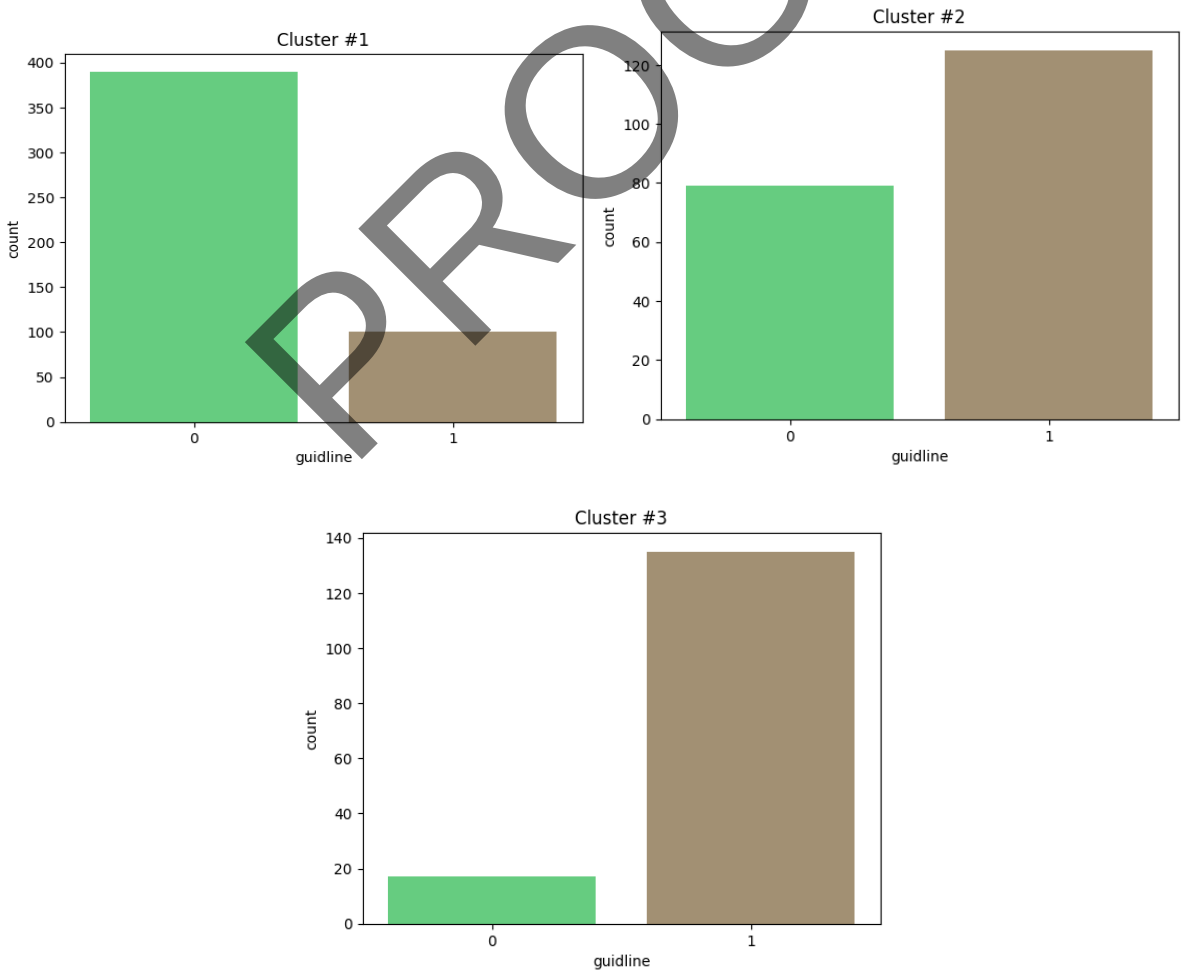

Figure 4. Count of MetS vs no-MetS based on the ATP III guideline in three derived separate clusters from the autoencoder

Table 7. Multiple Comparison of Means - Tukey HSD

--- Factor: Waist ---						
Group1	Group2	Mean diff	p-adj	95% Confidence Interval Lower	Upper	Reject
0	1	-12.27	0.0	-14.40	-10.14	True
0	2	-4.70	0.0	-6.64	-2.76	True
1	2	7.56	0.0	5.97	9.15	True
--- Factor: SBP ---						
Group1	Group2	Mean diff	p-adj	95% Confidence Interval Lower	Upper	Reject
0	1	-40.38	0.0	-42.03	-38.72	True
0	2	-18.74	0.0	-20.25	-17.23	True
1	2	21.63	0.0	20.40	22.86	True
--- Factor: DBP ---						
Group1	Group2	Mean diff	p-adj	95% Confidence Interval Lower	Upper	Reject
0	1	-26.89	0.0	-27.64	-26.15	True
0	2	-13.94	0.0	-14.62	-13.27	True
1	2	12.94	0.0	12.39	13.50	True
--- Factor: FBS ---						
Group1	Group2	Mean diff	p-adj	95% Confidence Interval Lower	Upper	Reject
0	1	-36.74	0.0	-49.34	-24.14	True
0	2	-8.33	0.204	-19.81	3.14	False
1	2	28.40	0.0	19.02	37.79	True
--- Factor: HDL ---						
Group1	Group2	Mean diff	p-adj	95% Confidence Interval Lower	Upper	Reject
0	1	3.26	0.0	2.70	3.81	True
0	2	1.03	0.0	0.53	1.54	True
1	2	-2.22	0.0	-2.64	-1.81	True
--- Factor: TG ---						
Group1	Group2	Mean diff	p-adj	95% Confidence Interval Lower	Upper	Reject
0	1	-50.32	0.0	-58.25	-42.40	True
0	2	-18.29	0.0	-25.50	-11.07	True
1	2	32.03	0.0	26.13	37.94	True

Waist Circumference

- Cluster 1 has a significantly higher waist circumference than cluster 0 (mean difference ≈ 12.27, $p<0.001$) and cluster 2 (≈ 7.57, $p<0.001$), and cluster 2 is also significantly higher than cluster 0 (≈ 4.71, $p<0.001$).
- Thus, central obesity is lowest in cluster 0, intermediate in cluster 2, and highest in cluster 1.

Systolic Blood Pressure (SBP)

- SBP in cluster 1 is dramatically higher than in cluster 0 (≈ 40.38 mmHg, $p<0.001$) and higher than in cluster 2 (≈ 21.64 mmHg, $p<0.001$); cluster 2 is also higher than cluster 0 (≈ 18.75 mmHg, $p<0.001$).
- This indicates a clear gradient of systolic hypertension: cluster 1 > cluster 2 > cluster 0.

Diastolic Blood Pressure (DBP)

- DBP in cluster 1 is significantly higher than in cluster 0 (≈ 26.90 mmHg, $p<0.001$) and cluster 2 (≈ 12.95 mmHg, $p<0.001$); cluster 2 is higher than cluster 0 (≈ 13.95 mmHg, $p<0.001$).
- Again, a consistent severity pattern is observed: cluster 1 has the highest diastolic pressure, cluster 0 the lowest, with cluster 2 in between.

Fasting Blood Sugar (FBS)

- FBS is significantly higher in cluster 1 compared with cluster 0 (≈ 36.74 mg/dL, $p<0.001$) and cluster 2 (≈ 28.41 mg/dL, $p<0.001$).

- The difference between clusters 0 and 2 is not statistically significant (p=0.204), suggesting that hyperglycemia is mainly concentrated in cluster 1, while clusters 0 and 2 are similar in FBS.

HDL Cholesterol

- HDL is significantly higher in cluster 0 than in clusters 1 (≈ 3.26 units, p<0.001) and 2 (≈ 1.04 units, p<0.001), and cluster 2 has higher HDL than cluster 1 (≈ 2.23 units, p<0.001).
- Thus, cluster 0 shows the most favorable HDL profile, cluster 1 the least favorable, and cluster 2 an intermediate pattern.

Triglycerides (TG)

- TG levels are significantly higher in cluster 1 than in cluster 0 (≈ 50.33 mg/dL, p<0.001) and cluster 2 (≈ 32.04 mg/dL, p<0.001); cluster 2 is also higher than cluster 0 (≈ 18.29 mg/dL, p<0.001).
- This indicates that cluster 1 has the most severe hypertriglyceridemia, cluster 0 the lowest TG, and cluster 2 an intermediate level.

Overall, Tukey HSD confirms that the three latent clusters are not only statistically distinct but also clinically ordered: cluster 1 represents a high-risk MetS profile, cluster 0 a low-risk profile, and cluster 2 an intermediate or mixed-risk profile.

5. Conclusion

In this study, we explored the efficacy of supervised ensemble learning models in predicting and selecting features for Metabolic Syndrome (MetS) by analyzing a dataset of 23 clinical and laboratory factors. Among five trained models, the gradient boosting classifier, excluding six main ATP III defined factors, demonstrated an accuracy of 0.9836 ±0.0298 with 95% CI: [0.9724, 0.9947] and calibration Brier score 0.0012 ± 0.0027. In the context of feature selection and preventing leakage, the gradient boosting classifier is trained in two scenarios of including and excluding the ATP III-defined features. Then, the univariate feature score, which calculates the ANOVA F-value between label/feature for classification tasks, is used for feature importance. Our analysis identified HOMA-IR, BMI, age, and cholesterol as additional key predictors of MetS. In other words, our findings confirm that obesity, high blood sugar, hypertriglyceridemia, and insulin resistance (HOMA-IR) are most influential predictors of MetS, alongside age and cholesterol levels.

Furthermore, the use of unsupervised learning through autoencoders revealed a distinctive third cluster that lies between the MetS and non-MetS groups. One-way ANOVA test showed that statistically significant differences between clusters with p-values < 0.001 for all six factors.

Then, after, Tukey HSD post-hot test confirms that the three latent clusters are not only statistically distinct but also clinically ordered: cluster 1 represents a high-risk MetS profile, cluster 0 a low-risk profile, and cluster 2 an intermediate or mixed-risk profile.

This finding suggests the existence of an intermediate/subtype state that may provide insights into the early identification of individuals at risk for MetS, potentially leading to more personalized and effective management strategies.

Overall, our research contributes to the growing body of evidence supporting the integration of artificial intelligence in healthcare, particularly in the early detection and data mining of metabolic syndrome. Future studies should focus on validating these findings in diverse populations and confirm on the other datasets. The autoencoder is a choice for compressibility and dimensionality reduction in an unsupervised scheme; however, it is required to apply the variation of this model for more comprehensive comparisons. It is required to explore the underlying mechanisms that characterize the proposed third cluster.

Acknowledgment

The participants were registered from those patients who were referred to the outpatient endocrine clinic of the Tehran University General Hospital and given health examinations. The study was approved by the Ethics Review Committee of Amol University of Special Modern Technology and provided the code of IR.ASMT.REC.1404.009.

Additionally, in this paper, the research work has been supported by a research grant from the Amol

University of Special Modern Technologies, Mazandaran, Iran.

References

1- &. A. Abdullah DM, "Machine learning applications based on SVM classification: a review." *Qubahan Academic Journal,* Vol. 1 (No. 2), pp. 81–90, (2021).

2- A. M. Abegaz TM, "Application of Machine Learning Algorithms to Predict Uncontrolled Diabetes Using the All of Us Research Program Data." *Healthcare (Basel),* Vol. 11 (No. 8), p. 1138, (2023).

3- A. H. Abid-Althaqafi NR, " The Effect of Feature Selection on the Accuracy of X-Platform User Credibility Detection with Supervised Machine Learning." *Electronics.,* Vol. 13 (No. 1), p. 205, (2024).

4- E. R.-C. Alberti K G M M, & Prevention, I. D., "Harmonizing the metabolic syndrome: a joint interim statement of the International Diabetes Federation Task Force on Epidemiology and Prevention; National Heart, Lung, and Blood Institute; American Heart Association; World Heart Federation." *Circulation,* Vol. 120 (No. 16), pp. 1640–5, (2009).

5- Z. P. Alberti KG, "Metabolic syndrome--a new worldwide definition. A Consensus Statement from the International Diabetes Federation." *Diabet Med,* Vol. 23 (No. 5), pp. 469–80, (2006).

6- S. F. Boitor O, "Automated Machine Learning to Develop Predictive Models of Metabolic Syndrome in Patients with Periodontal Disease." *Diagnostics (Basel).* Vol. 13 (No. 24), p. 3631, (2023).

7- G. Brown, "Ensemble Learning. Encyclopedia of machine learning." Vol. 312pp. 15–19, (2010).

8- R. N. Chweidan H, "Statistical Methods and Machine Learning Algorithms for Investigating Metabolic Syndrome in Temporomandibular Disorders: A Nationwide Study." *Bioengineering (Basel).* Vol. 11 (No. 2), p. 134, (2024).

9- Y. Z. DONG X, "A survey on ensemble learning." *Front. Comput. Sci.,* Vol. 14 (No. 2), pp. 241–58, (2020).

10- K. O. Esteghamati A, "HOMA-estimated insulin resistance is associated with hypertension in Iranian diabetic and non-diabetic subjects." *Clin Exp Hypertens,* Vol. 30pp. 297–307, (2008).

11- K. O. Esteghamati A., "Metabolic syndrome and insulin resistance significantly correlate with body mass index." *Arch Med Res,* Vol. 39pp. 803–08, (2008).

12- A. L. Fahed G, "Metabolic Syndrome: Updates on Pathophysiology and Management in 2021." *Int J Mol Sci.,* Vol. 23 (No. 2), p. 786, (2022).

13- T. T. Faisal, " Neural network diagnostic system for dengue patients' risk classification." *J Med Syst,* Vol. 36pp. 661–76, (2012).

14- P. Flach, "ROC Analysis.," in *In &. G. C. Sammut, Encyclopedia of Machine Learning and Data Mining.*: *Springer*, (2016), pp. 1–8.

15- C. S. Ghosh J, " Application of Machine Learning Algorithm and Artificial Intelligence in Improving Metabolic Syndrome-Related Complications: A Review. " *International Journal of Advancement in Life Sciences Research,* Vol. 7 (No. 2), pp. 2581–4877, (2024).

16- C. J. Grundy SM, "Diagnosis and management of the metabolic syndrome: an American Heart Association/National Heart, Lung, and Blood Institute scientific statement." *Curr Opin Cardiol.,* Vol. 21 (No. 1), pp. 1–6, (2006).

17- T. H. Horvat, "Predicting Metabolic Syndrome Risks Using Machine Learning Methods." in *IEEE 21st Jubilee International Symposium on Intelligent Systems and Informatics (SISY)*, Jubilee (2023), pp. 000017–22.

18- C. K. Hsu NW, " Building a model for predicting metabolic syndrome using artificial intelligence based on an investigation of whole-genome sequencing." *J Transl Med.,* Vol. 20 (No. 1), p. 190, (2022).

19- M. C. Hu R, "Effect of metabolic syndrome on prognosis and clinical characteristics of revascularization in patients with coronary artery disease." *Chin Med J (Engl),* Vol. 119pp. 1871–76, (2006).

20- D. K. Ivanović, "ANN Prediction of Metabolic Syndrome: A Complex Puzzle that will be Completed." *J Med Syst,* Vol. 24p. 264, (2016).

21- M. W. Jan Ehrhardt, "Autoencoders and variational autoencoders in medical image analysis.," in *In D. S. Ninon Burgos, Biomedical Image Synthesis and Simulation*: *Academic Press*, (2022), pp. 129–62.

22- Y. Z. Jiang X, ""Big Data" Approaches for Prevention of the Metabolic Syndrome." *Front Genet.,* Vol. 13p. 810152, (2022).

23- K. J. Johnson WD, "Prevalence of risk factors for metabolic syndrome in adolescents: National Health and Nutrition Examination Survey (NHANES), 2001-2006." *Arch Pediatr Adolesc Med.,* Vol. 163 (No. 4), pp. 371–7, (2009).

24- F. A. Kasper DL, Metabolic Syndrome. In Harrison's Manual of Medicine. 19e ed. ed. *McGraw-Hill Education*, (2016).

25- W. J. Kelliny C, "Metabolic syndrome according to different definitions in a rapidly developing country of the African region." *Cardiovasc Diabetol,* Vol. 18p. 7:27, (2008).

26- S. Y. Lv W, "Deep Learning Algorithm for Automated Detection of Polycystic Ovary Syndrome Using Scleral Images." *Front. Endocrinol.,* Vol. 12p. 789878, (2022).

27- F. J. Ma J, " Study on the risk of age-related diabetes mellitus among 8280 cases with metabolic syndrome patients and normal persons in Beijing." *Zhonghua Liu Xing Bing Xue Za Zhi,* Vol. 31 (No. 3), pp. 241–4, (2010).

28- G. E. Magge SN, "The Metabolic Syndrome in Children and Adolescents: Shifting the Focus to Cardiometabolic Risk Factor Clustering." *Pediatrics,* Vol. 140 (No. 2), p. e20171603, (2017).

29- G. V. Marhl M, "Diabetes and metabolic syndrome as risk factors for COVID-19." *Diabetes Metab Syndr.,* Vol. 14 (No. 4), pp. 671–77, (2020).

30- H. P. Nilsson G, "Inflammation and the metabolic syndrome: clustering and impact on survival in a Swedish community-based cohort of 75-year-olds." *Metab Syndr Relat Disord.,* Vol. 11 (No. 2), pp. 92–101, (2013).

31- S. I. Pilnenskiy N, "Feature Selection Algorithms as One of the Python Data Analytical Tools." *Future Internet,* Vol. 12 (No. 3), p. 54, (2020).

32- V. S.-D. Pinaya WHL, "Autoencoders." in *In V. S. Mechelli A, Machine learning*: *Academic Press*, (2020), pp. 193–208.

33- C. R. Pombo A, "Effect of Motor Competence and Health-Related Fitness in the Prevention of Metabolic Syndrome Risk Factors." *Res Q Exerc Sport.,* Vol. 95 (No. 1), pp. 110–17, (2024).

34- J. B. R. Kahn, "The metabolic syndrome: time for a critical appraisal: joint statement from the American Diabetes Association and the European Association for the Study of Diabetes." *Diabetes Care,* Vol. 25pp. 2289–304, (2005).

35- A. S. Rana S, "Metabolic syndrome and underlying genetic determinants: A systematic review." *J Diabetes Metab Disord.,* Vol. 21 (No. 1), pp. 1095–104, (2022).

36- G. P. Servais A, "Is serum cystatin-C a reliable marker for metabolic syndrome?" *Am J Med.,* Vol. 121 (No. 5), pp. 426–32, (2008).

37- D. E. Trigka M, "Predicting the Occurrence of Metabolic Syndrome Using Machine Learning Models." *Computation,* Vol. 11 (No. 9), p. 179, (2023).

38- J. L. Weng, "An autoencoder and machine learning model to predict suicidal ideation with brain structural imaging." *Journal of Clinical Medicine,* Vol. 9 (No. 3), p. 658, (2020).

39- A. F. N. Wilensky, "Periodontitis and Metabolic Syndrome: Statistical and Machine Learning Analytics of a Nationwide Study." *Bioengineering (Basel),* Vol. 10 (No. 12), p. 1384, (2023).

40- E. A. Yang C, "Impact of random oversampling and random undersampling on the performance of prediction models developed using observational health data." *Journal of Big Data,* Vol. 11 (No. 7), (2024).

41- J. X. Young Mee Choi Durham, "A184 Metabolic Syndrome and its Risk Factors in Sexual Minority Women." *Surgery for Obesity and Related Diseases,* Vol. 19 (No. 6), pp. S60–S61, (2023).

42- L. J. Zhang Y, "Identification of co-diagnostic effect genes for aortic dissection and metabolic syndrome by multiple machine learning algorithms." *Sci Rep.,* Vol. 13 (No. 1), p. 14794, (2023).

43- C. C.-B.-T.-P. Zhao M, "International Childhood Vascular Structure Evaluation Consortium. Metabolic syndrome, clustering of cardiovascular risk factors, and high carotid intima-media thickness in children and adolescents." *J Hypertens.,* Vol. 38 (No. 4), pp. 618–24, (2020).

44- Z. Q. Zou G, "Predictive analysis of metabolic syndrome based on 5-year continuous physical examination data." *Scientific Reports,* Vol. 13p. 9132, (2023).

.