IAES International Journal of Artificial Intelligence (IJ-AI)
Vol. 14, No. 1, February 2025, pp. 368~375
ISSN: 2252-8938, DOI: 10.11591/ijai.v14.i1.pp368-375

Prediction of metabolic syndrome in mexicans using machine learning

Zaira Pineda-Rico[1], Diana Luz de los Angeles Rojas Mendoza[1], Ulises Pineda-Rico[2]
[1]Unidad Academica Multidisciplinaria Region Altiplano, Universidad Autónoma de San Luis Potosi, San Luis Potosi, Mexico
[2]Facultad de Ciencias, Universidad Autónoma de San Luis Potosi, San Luis Potosi, Mexico

Article Info

Article history:

Received Nov 30, 2023
Revised Aug 8, 2024
Accepted Aug 30, 2024

Keywords:

Metabolic syndrome
Metabolic syndrome in Mexicans
Prediction of metabolic syndrome
Predictive models of metabolic syndrome
Sedentary work habits

ABSTRACT

Metabolic syndrome (MetS) is a compelling public health issue in Mexico, with high prevalence rates of overweight, obesity, arterial hypertension, diabetes, high triglycerides, low high-density lipoprotein cholesterol, and high total cholesterol. Despite this, predictive models tailored for under-researched professional groups with sedentary habits are scarce. This study introduces a novel predictive model for MetS using data from the National Center for Health Statistics and a unique dataset of higher education staff. By employing and comparing machine learning algorithms such as decision trees, random forest, artificial neural networks, and adaptive boosting, the research provides new insights into gender and race-specific aspects of MetS. The data was labeled using standards from the International Diabetes Federation and the National Cholesterol Education Program Adult Treatment Panel III to create classification models, which were tested on the higher education staff dataset. Model predictions were assessed using F1-score, accuracy and area under the curve - receiver operating characteristic (AUC-ROC), with random forest, decision tree, and adaptive boosting performing best. The key predictive features identified for MetS prediction include triglycerides, glucose, high-density lipoprotein cholesterol, waist-to-height ratio, and body mass index.

Corresponding Author:

Zaira Pineda-Rico
Unidad Academica Multidisciplinaria Region Altiplano, Universidad Autónoma de San Luis Potosi
Carr. Cedral km 5 y 600, Ejido San José de las Trojes, Matehuala, San Luis Potosi, 78700, Mexico
Email: zaira.pineda@uaslp.mx

1. INTRODUCTION

Metabolic syndrome (MetS) is a group of factors that can be used as parameters to identify an individual's increased risk for developing chronic diseases, such as type 2 diabetes, coronary diseases, and stroke [1]. MetS can be defined as a "pathologic condition characterized by abdominal obesity, insulin resistance, hypertension, and hyperlipidemia" [2], [3]. According to the World Health Organization (WHO), there are about 422 million people in the world with diabetes and more than 1 billion people with obesity. Hence, halting the rise in diabetes and obesity by 2025 has become a globally agreed upon target.

The National Institute of Diabetes and Digestive and Kidney Diseases stated that the biomarkers used to diagnose MetS change according to the individual's diet, exercise habits, cultural, and socioeconomic factors [4], [5]. Then, MetS needs to be treated by developing prevention strategies such as detection of risk factors, to reduce the health impact of this condition on the body, as well as on healthcare systems [6]–[8]. Research focused on particular groups has demonstrated a notable correlation between adopting healthy habits and decreasing the risk of developing MetS [9], [10].

In this context, machine learning (ML) has become a tool that allows the study of large amounts of information to find underlying patterns within the data, that contribute to the occurrence of MetS on specific population. ML is a subset of artificial intelligence (AI) technology; it has been described as the science of programming computers to learn from provided information. In recent years, this type of technology has had a significant impact on healthcare, supporting activities such as data analysis and decision-making. These activities include data collection, diagnosis, medical imaging, treatment alternatives and personalized treatments to construct healthcare projects aimed at enhancing medical diagnosis or assisting healthcare personnel in recognizing a patient's condition [11]–[14].

Several research groups have presented studies based on the prevalence of MetS in specific populations with sociodemographic characteristics, by finding associations between risk factors, lifestyle and nutrition relying on ML techniques. Yang *et al.* [15] analyzed risk factors including age and gender as features, to develop a predictive model with regional characteristics, enabling efficient pre-evaluation for the occurrence of MetS. Park *et al.* [14] studied demographic, anthropometric, biochemical, genetic, nutrient, and lifestyle characteristics of a Korean population to develop a predictive model for insulin resistance. Through their study, the authors determined that several ML algorithms, including logistic regression (LR), random forest (RF), and artificial neural networks (ANN), were effective at identifying risk factors for different diseases and predicting their occurrence in a clinical setting. Panagoulias *et al.* [16] constructed and studied cascaded support vector machines (SVM) based classifiers for automated MetS diagnosis using blood tests to obtain biomarkers. Yu *et al.* [17] analyzed the accuracy of various decision tree (DT) ML algorithms in predicting MetS in self-paid health examination subjects using FibroScan. The use of ML techniques resulted in high predictive accuracy, highlighting their potential for identifying MetS in this population. Esparza *et al.* [18] use correlation-based feature selection (CFS) and chi-squared filter methods to score the variable importance of health parameters, with the aim of predicting MetS in a group of Mexican population of the city of Tlalpan. They remarked the importance to extend such studies to different sets of population with specific characteristics [19].

In Mexico, diabetes is among the five diseases with a major economic impact on the national health system. The 2018 National Health and Nutrition Survey shows that diabetes, cardiovascular diseases (CVD), and obesity account for 16.4% of all causes of consultation. The International Diabetes Federation (IDF) estimates that there will be nine million people with diabetes in Mexico by 2025 [20]. Additionally, almost 50% of Mexican adults are diagnosed with MetS due to sedentary behaviors, unhealthy dietary habits and inadequate sleep. The increase in sitting time (more than 420 minutes/day) among Mexican adults aged 20 to 69 years in Mexico City has resulted in a 5.4% increase in overweight/obesity and a 1.3% rise in the diagnosis of diabetes [21], [22]. The 2020 National Health and Nutrition Survey on Covid-19, stated that 76% of adult women and 72.1% of adult men are overweight or obese. Additionally, there is a prevalence of arterial hypertension of 30.2% among adults over 20 years old, 15.6% of adults have diabetes, 49% have high levels of triglycerides, 28.2% have low high-density lipoprotein cholesterol (HDL-C), and 26.1% have high levels of total cholesterol [23].

The aim of this study is to construct a predictive model for MetS tailored to a Mexican population subgroup. Specifically, the focus is on higher education staff, who often spend extended working hours in office settings, leading to sedentary behaviors that can have a negative influence on their health. In an educational institution, detecting health risks among staff members should be crucial for preemptively addressing the onset of chronic illnesses.

The present work uses the criteria for clinically diagnosing MetS set forth by the IDF and the revised National Cholesterol Education Programs Adult Treatment Panel III (NCEP ATP III). These criteria encompass factors such as glycemic control, fasting blood glucose (GLU), waist circumference (WC), triglycerides, HDL-C, blood pressure, and body mass index (BMI). Given the gender and race-based nature of MetS diagnosis, this study employs a database from the National Center for Health Statistics (NCHS) to train five supervised predictive models implemented with Python and Scikit-Learn. The previously mentioned database contains information gathered during the National Health and Nutrition Examination Survey (NHANES) concerning the Mexican American population. The supervised learning algorithms implemented in this study are DT, ANN, RF, and adaptive boosting (Adaboost). The predictive models were tested on a dataset created from records of individuals employed at a Higher Education Institution in Mexico, with the objective of predicting MetS.

2. METHOD

The process to construct the prediction models is partitioned into the three stages: Extract transform load (ETL), exploratory data analysis (EDA), and results. The diagram of the workflow is shown in Figure 1. In the ETL stage, the data is recovered from the databases and it is cleaned by restoring corrupted elements with the estimated mean, also the feature scaling is performed. In the EDA stage, statistical analysis is conducted using Pearson correlation value (P-value). The data was labeled according to the presence of MetS, considering both the ATP III criteria and the IDF criteria. The following classifiers were implemented: DT, RF, ANN, and AdaBoost. In the results stage, RF is used to obtain variable importance measures (VIM) through shapley additive explanations (SHAP) values. The quality of the classifiers was assesed using the Accuracy metric, the F1 score, and the area under the receiver operating characteristic curve (AUC-ROC).

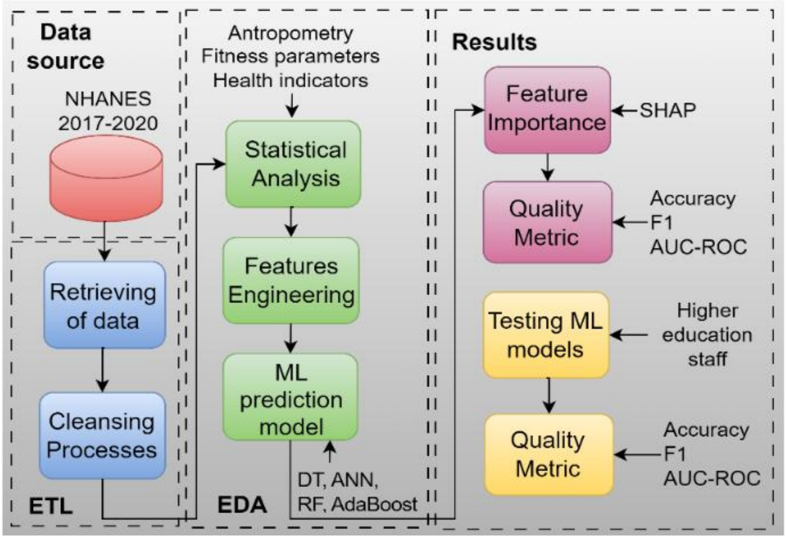


Figure 1. Block diagram of the data analysis describing: data source, ETL, EDA, and results

2.1. The data

The dataset used to train the predictive models was obtained from the results of the National Health and NHANES 2017-2020. NHANES is a program of the NCHS, which is part of the Centers for Disease Control and Prevention (CDC). The program evaluates the health and nutritional status of people in the United States through a combination of interviews and physical examinations. NHANES over-samples individuals aged 60 and older, African Americans, and Hispanics to produce reliable statistics. As a result, the survey provides valuable information about MetS and other health issues among Mexican Americans. The data collected in the survey from 2017 to March 2020 included records of 660 Mexican American individuals, 340 women (51.5%), and 320 men (48.5%), aged 12 to 80 years. The dataset included various cardiovascular risk factors, encompassing clinical and anthropometric measurements, lifestyle habits, and biomedical evaluations. The dataset used in this work to train the predictive models does not include lifestyle habits.

After training, the predictive models were tested in a dataset created from records of 45 individuals employed in a Higher Education Institution located in the city of Matehuala, Mexico. The dataset is comprised by 23 women (51%) and 22 men (49%), aged 26 to 55 years. The recorded data encompasses various anthropometric measurements, including age (AG), weight (WE), height (HE), BMI, WC, waist-to-height ratio (WHtR), arm circumference (AR), hip circumference (HP), systolic blood pressure (BSY), diastolic blood pressure (DSY), and heart rate (HR). Additionally, health indicators such as GLU, triglycerides (TRI), and HDL-C are included, along with functional fitness parameters like muscle (MM), visceral fat (VF), and body fat (BF). The data corresponding to WE, BMI, MM, MA, VF, BF, and BA, was collected with an Omron HBF-514C body monitor. This device passes electrical currents through the individual using handheld electrodes and electrodes placed on the scale's surface, allowing for a general analysis of the user’s body.

To ensure accuracy, participants were instructed to refrain from exercise and fasting (including coffee) on the day of the tests. Blood pressure was measured with an inflatable cuff and a gauge around the arm, both diastolic (BDI) and BSY blood pressure readings were millimeters of mercury (mmhg). The magnitude of WA, HP and AR was obtained using a tape measure in centimeters. HR was acquired at the wrist on the radial artery, in beats per minute. The calculation of WHtR involves dividing the WC by the height, expressed in centimeters. GLU levels were determined using a one touch blood sugar meter using a blood sample from the fingertip, in millimoles per liter (mmol/l). The staff members can be described, in terms of the anthropometric measurements, functional physical fitness indices (obtained through bioelectrical impedance analysis (BIA)) and blood glucose levels, as follows: they are 37.5 years of age, with a height of 163.2 centimeters, a BMI of 26.91 kg/m² (>25), with normal glucose levels (<99), normal HDL-C, borderline high cholesterol levels of 173.97 mmol/l and normal systolic/DSY of 113/76.

2.2. Metabolic syndrome criteria

The diagnostic criteria for MetS proposed by the NCEP ATP III report, are based on a multiple risk factor assessment for CVD. Furthermore, the IDF suggests the inclusion of WC as a necessary component for diagnosing MetS [24]. A detailed description of the risk factors considered for both sets of criteria is provided in Table 1.

Table 1. Diagnostic criteria for MetS proposed by the NCEP ATP III and the IDF [25]

Parameters	Revised NCEP ATP III	IDF
Definition	Any three of the following five features	Increased WC
Elevated WC	≥ 102 cm in men, ≥ 88 cm in women	Men ≥ 90cm, Women ≥ 80cm Along with any 2 of following features
Triglycerides	≥ 1.7mmol/l or TG treatment	
HDL-C	Men < 1.03 mmol/l or women < 1.29 mmol/l or HDL-C treatment	
Blood pressure	Systolic ≥ 130 mmHg or Diastolic ≥ 85mmHg or hypertension treatment or previously diagnosed hypertension	
Fasting blood glucose	≥ 5.6 mmol/l or treatment for elevated glucose or previously diagnosed Type 2 Diabetes.	

2.3. Prediction models

One of the primary objectives of applying ML techniques to extensive datasets is to uncover patterns within various features. In this study the data was labeled using both ATP III and IDF criteria. The supervised learning algorithms applied in this work are DT, ANN, RF and Adaboost. All the algorithms were implemented using python Scikit-Learn. DT can be employed for both classification and regression tasks on complex datasets, whereas RF is constructed by amalgamating multiple individual DTs. DT learn the optimal strategy to partition the training dataset into smaller, more precise subsets until the desired prediction is achieved. In the case of RF, the predictions from all individual trees are added to form the final prediction. Within ML algorithms, the importance of each feature is assessed using a scoring system that is implemented after the algorithm has been trained. This approach proves beneficial in comprehending which characteristics hold significance when selecting a subset of features. The quality assessment in the classifiers is made through metrics, such as the F1 score, the confusion matrix, and the AUC-ROC.

3. RESULTS AND DISCUSSION

3.1. Labeled data

The data of the Mexican American group (training set) was labeled according to both the ATP III and the IDF criteria for MetS, the detailed anthropometries and health indicators of this group are described in Table 2. The characteristics of the two groups were compared using the Pearson correlation coefficient (P-value). The P- value varies between 1 and -1, a 0 value means that there is no correlation. Within the Mexican American group, according to the ATP III criteria, 26.5% is considered to have MetS. In comparison, 43.5% of the group have MetS when considering IDF criteria. Within the Mexican staff members (testing set), according to the ATP III criteria, 17.7% of the group is considered to have MetS, with an average age of 43.8 years, height of 168.5 centimeter, BMI of 31.2, WC of 112.8 centimeter, WHtR of 0.6, hip of 110 centimeter, BSY of 144 mmHg, DSY of 86 mmHg, heart rate of 70 beats per minute, glucose of 108.3 mmol/l, triglycerides of 237.1 mmol/l, HDL-C of 180.6 mmol/l. According to the IDF criteria, 25% of the group is considered to have MetsS, with an average age of 40.6 years, height of 165.6 centimeter, BMI of 31.4, WC of 117.5 centimeter, WHtR of 0.7, hip 112 centimeter, BSY of 125 mmHg, DSY of 82 mmHg, heart rate of 71.4 beats per minute, glucose of 109.1 mmol/l, triglycerides of 252 mmol/l, HDL-C of 182.1 mmol/l.

3.2. Classification models

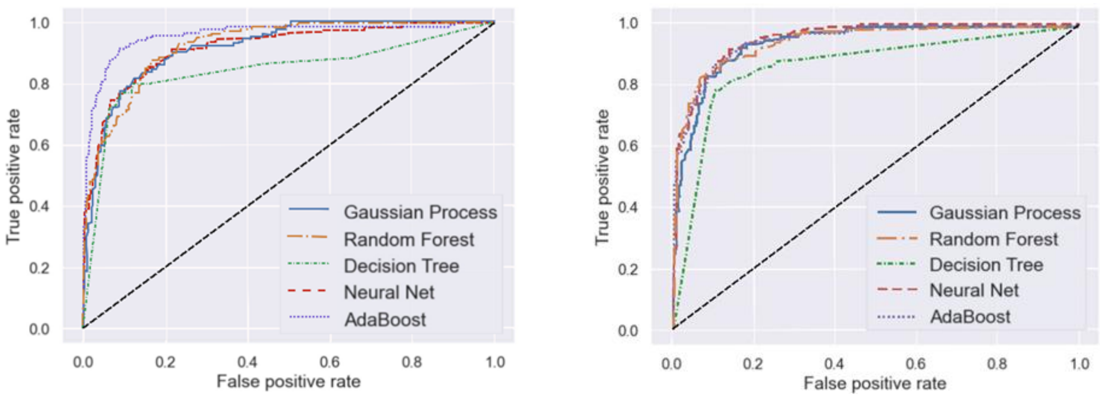
The accuracy of the classification was measured using Scikit-Learn, supported by the estimation of the F1-score and the AUC value when training the models. Also, the AUC-ROC was created for each classification model in order to visualize its performance. The score is normalized, meaning that the best performance exhibits a value close to 1.0. Table 3 summarizes the performance of each classification model and Figure 2 shows the ROC curve for all the classification models implemented considering both a) ATP III criteria as shown in Figure 2(a) and IDF criteria as shown in Figure 2(b), the AUC was above 0.9 in all cases.

Table 2. Characteristics of the Mexican American group as labeled considering the ATP III and the IDF criteria

Code	ATP III Yes (N = 175 (26.5%))	ATP III No (N = 485 (73.34%))	ATP III p value	IDF Yes (N = 287 (43.5%))	IDF No (N = 373 (56.5%))	IDF p value
AG	50.34 (12-80)	36.01 (12-80)	0.0004	45.37(12-80)	35.53 (12-80)	0.0000
Anthropometries						
HE	164.74 (138.3-189.5)	157.79 (max. 192.5)	0.1928	162.77(138.3-192.5)	157.22 (max. 188)	0.0075
BMI	33.4 (21.3-51.8)	27.41 (max. 67)	0.0000	34.52(23.6-67)	24.75 (max. 47.5)	0.0000
WC	107.02 (max. 147.4)	88.57 (max. 178)	0.0230	110.53 (89-178)	80.33 (max. 134.1)	0.0009
WHtR	0.65 (max. 0.95)	0.54 (0.45-1.11)	0.0000	0.68(0.53-1.11)	0.49 (max. 0.78)	0.0000
HP	108.13 (153.9)	96.36 (176.6)	0.0226	113.64(max. 176.6)	88.58 (max. 130.4)	0.0003
BSY	121.74(214)	96.27(182.6)	0.0773	109.11(max. 190)	98.3 (max. 214)	0.0000
BDI	71.88(106.33)	57.69 (95.33)	0.0051	66.2(max. 105.33)	57.8 (max. 100.33)	0.0000
HR	61.60 (104.66)	57.03 (95.33)	0.0000	59.97 (max. 104.66)	56.91 (max. 100.33)	0.0000
Health indicator						
GLU	133.41 (320)	102.39 (421)	0.0151	121.60(max. 325)	102.16 (max. 421)	0.0000
TRI	187.12 (967)	87.13 (461)	0.0662	140.92(max. 967)	51.77 (max. 876)	0.0000
HDL-C	41.82 (94)	50.57 (118)	0.0002	43.67(max. 113)		0.0000

Table 3. Accuracy of the classification models implemented to predict MetS. A value of 1.0 means that the model is accurate

Classification model	ATP III Accuracy	ATP III F1	ATP III AUC	IDF Accuracy	IDF F1	IDF AUC
Gaussian process	0.93	0.91	0.97	0.9	0.81	0.97
RF	0.87	0.81	0.99	0.9	0.9	0.98
DT	0.9	0.87	0.99	0.9	0.89	0.99
ANN	0.96	0.94	0.97	0.91	0.91	0.98
Adaboost	0.96	0.96	0.99	0.93	0.93	0.99



(a) (b)

Figure 2. The ROC curve for the classification model considering: (a) ATP III criteria and (b) IDF criteria

3.3. Prediction of metabolic syndrome

The data corresponding to the Mexican staff group was used to test the classification models in the prediction of MetS, the results are summarized in Table 4. DT, RF, and AdaBoost exhibit the best performance among the prediction models created for the data. In order to identify the importance of each feature and its impact on the prediction, SHAP values were computed. As a result, the features with SHAP values close to 1, meaning that they contributed more to the prediction of MetS in the DT, RF, and AdaBoost classification models, were Triglycerides, Glucose, HDL-C, WHtR, and BMI.

The results align with those documented in previous studies conducted on diverse ethnic populations, where BMI, WC, height, and WHtR serve as indicators of MetS and cardiometabolic risk. Specifically concerning the Mexican population, previous research indicates that WC, height and BMI are crucial biomarkers for identifying obesity. Furthermore, WC, BMI, and WHtR play a significant role in predicting MetS when utilizing ML algorithms based on the ATP III criteria as a reference point. In our study, we consider both the ATP III and IDF criteria, alongside indicators to detect obesity according to WHO standards and the Mexican standard NOM-008-SSA3-2017 guidelines. Additionally, we delve into the use of health indicators derived from BIA, aiming to ascertain their importance in predicting MetS.

Table 4. Accuracy of the classification task corresponding to the Mexican staff group in the prediction of MetS. A value of 1.0 means that the classification is accurate

	ATP III			IDF		
Classification model	AC	F1	AUC	AC	F1	AUC
Gaussian process	0.48	0.47	0.97	0.4	0.39	0.97
RF	0.95	0.91	0.99	0.95	0.93	0.98
DT	0.93	0.89	0.99	0.93	0.89	0.99
ANN	0.8	0.44	0.97	0.91	0.88	0.98
AdaBoost	0.97	0.96	0.99	0.95	0.93	0.99

Regarding the performance of the classification models using the ATP III criteria to label the data: RF, DT, and Adaboost achieved superior results compared to Gaussian process and ANN, when assessing quality based on the accuracy and F1 score metrics. Similarly, RF, DT, ANN, and Adaboost outperformed under the IDF criteria using the same quality metrics. All the classification models studied in this work performed similarly according to the AUC quality metric. Whether using the ATP III or the IDF criteria, the features that highly contributed to predicting MetS in the DT, RF, and AdaBoost classification models were Triglycerides, Glucose, HDL-C, WHtR, and BMI.

Similar findigs have been registered in literature. Yang *et al.* [15] studied the prediction of MetS in a group of 67,730 Chinese population of different ages and genders using an extreme gradient boosting (XGBoost) classifier. They showed that fasting blood glucose and triglycerides contribute the most to the prediction, altogether to both the alterations in health status and the diagnostic markers related to WC and BMI. In a cross-sectional study involving 1001 Spanish adolescents (aged 13.2 ± 1.2 years), Perona *et al.* [25] examined anthropometric measurements to find their correlation with MetS components. They found that WHtR exhibited the most significant predictive value, whereas WC was the most effective predictor of MetS. In review of the use of ML in the prediction of MetS, Kakudi *et al.* [26] listed the five risk factors common to all the clinical definitions of MetS: fasting plasma glucose, WC, triglycerides, HDL-C, and blood pressure. Kim *et al.* [27] used DT, Gaussian naïve Bayes (GNB), K-nearest neighbor (KNN), XGBoost, RF, LR, SVM, multi-layer perceptron (MLP), and 1D convolutional neural network (CNN), with anthropometric, lifestyle, and biochemical factors to predict MetS in a group of 1991 middle-aged Korean population. They identified BMI and WHtR as the most important features for prediction. Shin *et al.* [28] used non-invasive information of 70,370 to predict MetS using ML models based on LR, DT, RF, XGB, and TabNet (TN), they showed that synthetic features based on blood pressure and WC improved the performance of the classification models. Esparza *et al.* [18] created ML algorithms to study several health parameters from 2,942 adults aged 20 years or older, with the objective of finding the most important variables for classifying MetS in a group of Mexican population of the city of Tlalpan. They used CFS and chi-squared filter methods to conclude that WHtR, coupled with the ATP III variables (excluding WC), outperforms WC and BMI as important variables for prediction. Barquera *et al.* [29] studied data of 16,256 individuals to find the predominance of obesity among Mexican adults. Considering physical and sociodemographic factors, the authors created LR models that revealed the relevance of heigh in the identification of obesity in this population, along with WC, and recognized BMI as a risk indicator of comorbidities associated with excessive adipose tissue.

In this work, we use anthropometric measurements and health indicators to find their suitability on the prediction of MetS using ML techniques in staff members of a higher education institution. However, since the indicators used for diagnosing MetS vary based on an individual's diet, physical activity and cultural, and socioeconomic aspects, further prediction models may be studied considering these factors. Our results target specific population; therefore, it is important to identify the implications that adopting healthy behaviors may have in lowering their risk of developing MetS.

The use classification models such as DT, RF, ANN, and AdaBoosting are suitable to predict MetS in a subgroup of Mexican population considering anthropometric measurements and health indicators. The features that contributed to the prediction of MetS using the classification models mentioned before were Tryglicerids, Glucose, HDL-C, WHtR, and BMI. Future work may incorporate individual habits such as diet, sedentary behaviors, exercise, and socioeconomic factors as complementary factors to train the classification models; with the aim of studying the importance of each of these variables in the prediction of MetS in order to create prevention strategies based on habits and lifestyle, suited to this specific population.

4. CONCLUSION

This study identified several key predictive features for MetS using ML techniques. The results align with previous research conducted across various populations, highlighting the reliability of these predictors across diverse ethnicities and regions. This consistency underscores the significance of Triglycerides, Glucose, HDL-C, WHtR, and BMI in predicting MetS within the Mexican population. The results suggest that the importance of specific features may vary among populations and data specific to each population may offer valuable insights. Height emerged as a notable factor in identifying obesity, particularly among Mexican women, in addition to WC. These measurements shed light on adiposity and VF accumulation. Understanding the key predictive features for MetS is crucial for public health interventions and personalized medicine. The identification of high-risk individuals based on these features enables targeted prevention and intervention strategies. This information can inform healthcare decisions, public health policies, and personalized approaches to address MetS effectively.

REFERENCES

[1] S. M. Grundy, H. B. Brewer, J. I. Cleeman, S. C. Smith, and C. Lenfant, "Definition of metabolic syndrome: report of the National Heart, Lung, and Blood Institute/American Heart Association Conference on scientific issues related to definition," *Circulation*, vol. 109, no. 3, pp. 433–438, 2004, doi: 10.1161/01.CIR.0000111245.75752.C6.

[2] M. G. Saklayen, "The global epidemic of the metabolic syndrome," *Current Hypertension Reports*, vol. 20, no. 2, 2018, doi: 10.1007/s11906-018-0812-z.

[3] C. A. A. -Salinas and T. V. -Ruiz, "Recent advances in managing/understanding the metabolic syndrome.," *F1000Research*, vol. 8, 2019, doi: 10.12688/F1000RESEARCH.17122.1.

[4] F. S. Sigit *et al.*, "The prevalence of metabolic syndrome and its association with body fat distribution in middle-aged individuals from Indonesia and the Netherlands: A cross-sectional analysis of two population-based studies," *Diabetology and Metabolic Syndrome*, vol. 12, no. 1, 2020, doi: 10.1186/s13098-019-0503-1.

[5] J. Madhusoodanan, "Searching for better biomarkers for metabolic syndrome," *ACS Central Science*, vol. 8, no. 6, pp. 682–685, 2022, doi: 10.1021/acscentsci.2c00629.

[6] J. Myers, P. Kokkinos, and E. Nyelin, "Physical activity, cardiorespiratory fitness, and the metabolic syndrome," *Nutrients*, vol. 11, no. 7, 2019, doi: 10.3390/nu11071652.

[7] S. C. -Barquero, A. M. R. -León, M. S. -Pérez, R. Estruch, and R. Casas, "Dietary strategies for metabolic syndrome: A comprehensive review," *Nutrients*, vol. 12, no. 10, pp. 1–21, 2020, doi: 10.3390/nu12102983.

[8] C. A. Paley and M. I. Johnson, "Abdominal obesity and metabolic syndrome: Exercise as medicine?," *BMC Sports Science, Medicine and Rehabilitation*, vol. 10, no. 1, 2018, doi: 10.1186/s13102-018-0097-1.

[9] M. Garralda-Del-Villar *et al.*, "Healthy lifestyle and incidence of metabolic syndrome in the SUN cohort," *Nutrients*, vol. 11, no. 1, 2019, doi: 10.3390/nu11010065.

[10] E. H. Herningtyas and T. S. Ng, "Prevalence and distribution of metabolic syndrome and its components among provinces and ethnic groups in Indonesia," *BMC Public Health*, vol. 19, no. 1, 2019, doi: 10.1186/s12889-019-6711-7.

[11] M. Javaid, A. Haleem, R. Pratap Singh, R. Suman, and S. Rab, "Significance of machine learning in healthcare: Features, pillars and applications," *International Journal of Intelligent Networks*, vol. 3, no. 4, pp. 58–73, 2022, doi: 10.1016/j.ijin.2022.05.002.

[12] P. Manickam *et al.*, "Artificial intelligence (AI) and internet of medical things (IoMT) assisted biomedical systems for intelligent healthcare," *Biosensors*, vol. 12, no. 8, 2022, doi: 10.3390/bios12080562.

[13] H. H. Wang, D. K. Lee, M. Liu, P. Portincasa, and D. Q. H. Wang, "Novel insights into the pathogenesis and management of the metabolic syndrome," *Pediatric Gastroenterology, Hepatology and Nutrition*, vol. 23, no. 3, 2020, doi: 10.5223/PGHN.2020.23.3.189.

[14] S. Park, C. Kim, and X. Wu, "Development and validation of an insulin resistance predicting model using a machine-learning approach in a population-based cohort in Korea," *Diagnostics*, vol. 12, no. 1, 2022, doi: 10.3390/diagnostics12010212.

[15] H. Yang *et al.*, "Machine learning-aided risk prediction for metabolic syndrome based on 3 years study," *Scientific Reports*, vol. 12, no. 1, 2022, doi: 10.1038/s41598-022-06235-2.

[16] D. P. Panagoulias, D. N. Sotiropoulos, and G. A. Tsihrintzis, "SVM-based blood exam classification for predicting defining factors in metabolic syndrome diagnosis," *Electronics*, vol. 11, no. 6, 2022, doi: 10.3390/electronics11060857.

[17] C. S. Yu *et al.*, "Predicting metabolic syndrome with machine learning models using a decision tree algorithm: Retrospective cohort study," *JMIR Medical Informatics*, vol. 8, no. 3, 2020, doi: 10.2196/17110.

[18] G. O. G. -Esparza, O. I. Vázquez, M. Vallejo, and J. H. -Torruco, "Prediction of metabolic syndrome in a Mexican population applying machine learning algorithms," *Symmetry*, vol. 12, no. 4, 2020, doi: 10.3390/SYM12040581.

[19] P. M. Nilsson, J. Tuomilehto, and L. Rydén, "The metabolic syndrome – What is it and how should it be managed?," *European Journal of Preventive Cardiology*, vol. 26, no. 2, pp. 33–46, 2019, doi: 10.1177/2047487319886404.

[20] T. S. Levy *et al.*, "Encuesta Nacional de Salud y Nutrición 2018-19: Resultados Nacionales," *INEGI*, Cuernavaca, México: Instituto Nacional de Salud Pública, 2020.

[21] N. Macías *et al.*, "Screen-based sedentary behaviors and their association with metabolic syndrome components among adults in Mexico," *Preventing Chronic Disease*, vol. 18, pp. 1–12, 2021, doi: 10.5888/PCD18.210041.

[22] Y. Li, L. Zhao, D. Yu, Z. Wang, and G. Ding, "Metabolic syndrome prevalence and its risk factors among adults in China: A nationally representative cross-sectional study," *PLoS ONE*, vol. 13, no. 6, 2018, doi: 10.1371/journal.pone.0199293.

[23] T. S. Levy *et al.*, "Encuesta Nacional de Salud y Nutrición 2021 sobre Covid-19. Resultados nacionales," *SALUD*, Cuernavaca, México: Instituto Nacional de Salud Pública, 2022.

[24] P. Dobrowolski *et al.*, "Metabolic syndrome a new definition and management guidelines," *Archives of Medical Science*, vol. 18, no. 5, pp. 1133–1156, 2022, doi: 10.5114/aoms/152921.

[25] J. S. Perona, J. S. -RioValle, B. R. -Medina, M. C. -Rodríguez, and E. G. -Jiménez, "Waist circumference shows the highest predictive value for metabolic syndrome, and waist-to-hip ratio for its components, in Spanish adolescents," *Nutrition Research*, vol. 45, pp. 38–45, 2017, doi: 10.1016/j.nutres.2017.06.007.

[26] H. A. Kakudi, L. C. Kiong, F. M. Moy, L. C. Kau, and K. Pasupa, "Diagnosis of metabolic syndrome using machine learning, statistical and risk quantification techniques: a systematic literature review," *Malaysian Journal of Computer Science*, vol. 34, no. 3, pp. 221–241, Jul. 2021, doi: 10.22452/mjcs.vol34no3.1.

[27] J. Kim, S. Mun, S. Lee, K. Jeong, and Y. Baek, "Prediction of metabolic and pre-metabolic syndromes using machine learning models with anthropometric, lifestyle, and biochemical factors from a middle-aged population in Korea," *BMC Public Health*, vol. 22, no. 1, 2022, doi: 10.1186/s12889-022-13131-x.

[28] H. Shin, S. Shim, and S. Oh, "Machine learning-based predictive model for prevention of metabolic syndrome," *PLoS ONE*, vol. 18, no. 6 June, 2023, doi: 10.1371/journal.pone.0286635.
[29] S. Barquera *et al.*, "Obesity and central adiposity in mexican adults: Results from the mexican national health and nutrition survey 2006," *Salud Publica de Mexico*, vol. 51, no. 4, 2009, doi: 10.1590/s0036-36342009001000014.

BIOGRAPHIES OF AUTHORS

Zaira Pineda-Rico received the B.S. degree in Electronics Engineering from the Autonomous University of San Luis Potosi in San Luis Potosi in 2005, the M.S. degree in Electronic Systems from Monterrey Institute of Technology and Higher Education in Monterrey, in Mexico in 2007 and the Ph.D. degree in Engineering Research from University of Leicester, Leicestershire, UK, in 2014. Since 2007, she has been a full time Academic and Researcher in Autonomous University of San Luis Potosi. Her research interests include applied engineering to electronics, bioelectronics, biomechanical, robotics, and energy systems, in addition to the implementation of machine learning algorithms to data and signal processing. She can be contacted at email: zaira.pineda@uaslp.mx.

Diana Luz de los Angeles Rojas Mendoza received the B.S. in Nursing from the Autonomous University of Tamaulipas in Tamaulipas, Mexico in 2004, the M.S. degree in Nursing Sciences from the Autonomous University of Nuevo Leon in Nuevo Leon, Mexico in 2010 and the Ph.D. degree in Education Sciences from Tec Tampico University in Tamaulipas, Mexico in 2014. Since 2010, she has been a full time Academic and Researcher in Autonomous University of San Luis Potosi. Her research interests include healthy lifestyle, the effect of maternal perception on nutritional status and obesity in children, risk factors for overweight and obesity in students, perception of body image in students in relation to body mass index, healthy behaviors, alcohol and tobacco consumption, and associated factors of metabolic syndrome in healthcare professionals. She can be contacted at email: diana.rojas@uaslp.mx.

Ulises Pineda-Rico obtained his Ph.D. degree from The University of Manchester, United Kingdom in 2009. His main research interests are oriented but not limited to: smart antenna arrays, linear and non-linear MIMO precoding, beyond 5G mobile systems, TeraHertz wireless communications, unmanned aerial vehicles, use of wireless technologies for agricultural and health applications, among others demanded by the industry, government, and society. He can be contacted at email: ulises.pineda@uaslp.mx.