



# Metabolic Syndrome Prediction Using Optimized Support Vector Machines

Case study: under 18 years old people in Birjand

**Mostafa Sabzekar**

Assistant Professor

Department of Computer Engineering, Birjand University of Technology

Birjand, Iran

[sabzekar@birjandut.ac.ir](mailto:sabzekar@birjandut.ac.ir)

**Mahdi Alizadeh Ghouzhd**

B.S Graduated

Department of Computer Engineering, Birjand University of Technology

Birjand, Iran

[aliyzadeh@gmail.com](mailto:aliyzadeh@gmail.com)

## ABSTRACT

In recent years, there has been a growing interest in the use of artificial intelligence techniques, including machine learning and deep learning, in healthcare and medical applications. One of the critical areas in medicine is the diagnosis of metabolic syndrome in individuals. This condition is referred to when multiple disorders such as high blood sugar, high blood pressure, high blood fat, and obesity occur in a person, and its consequences can lead to an increased risk of cardiovascular diseases, fatty liver, and more. Unfortunately, the prevalence of this condition is increasing in Iran. In this paper, we aim to utilize the advantages of Support Vector Machines (SVM) in predicting metabolic syndrome. SVM is a machine learning algorithm that can be useful in classification and regression tasks. One of the challenges in machine learning methods is the selection of free parameters to optimize the model's performance. To address this, we use the Particle Swarm Optimization (PSO) algorithm. One of the key strengths of this paper is the use of real data related to individuals under the age of eighteen in Birjand County. Furthermore, the Synthetic Minority Over-sampling Technique (SMOTE) algorithm is employed to balance the dataset. The results obtained from the proposed method indicate its superiority over other approaches. In fact, it achieved over 98% accuracy, sensitivity, and specificity in all parameters examined. This demonstrates that the proposed approach can effectively contribute to the diagnosis of metabolic syndrome in individuals and potentially be a valuable tool in the field of medical diagnosis and prediction.

**KEYWORDS:** Metabolic syndrome, machine learning, Support Vector Machines, Particle Swarm Optimization, optimization, SMOTE algorithm.

## 1 INTRODUCTION

Metabolic syndrome (MS) is a condition in which multiple disorders, such as obesity, high blood sugar, high cholesterol, and high blood pressure, occur simultaneously in an individual [1]. The complications of this syndrome can lead to the development or increased risk of other diseases such as cardiovascular diseases [2], type 2 diabetes [3], and fatty liver disease [4]. It has also been found that some cancers, including breast cancer [5], endometrial cancer [6 and 7], and prostate cancer [8], are closely

related to the components of MS. The most common cause of metabolic syndrome is abdominal obesity [9], but other factors such as increasing age [10], gender [11], physical activity [12], alcohol consumption [13], and smoking [14] are also associated with the development of this condition. The prevalence of this syndrome is rapidly increasing worldwide. According to reports, the total number of people affected by this syndrome worldwide is close to three million, and Iran contributes around 100,000 individuals to this figure. Therefore, early prediction of this condition for the purpose of treatment is essential, and timely intervention can prevent its transformation into more serious diseases.

To address metabolic syndrome, non-invasive predictive studies have been conducted to identify and prevent it. Non-invasive predictive models are those that do not use invasive information obtained from body or skin penetration, such as blood tests. Most of the proposed predictive models utilize lifestyle-related features of individuals.

In previous studies, various methods and models have been developed to predict metabolic syndrome, which can be classified into clinical and non-clinical categories. The International Diabetes Federation (IDF) and the World Health Organization (WHO) consider the presence of at least three metabolic abnormalities primarily in the same individual, including elevated blood sugar levels, insulin resistance, abdominal obesity, high blood pressure, and an abnormal lipid profile (elevated triglyceride levels and low-density lipoprotein cholesterol levels) as clinical markers for metabolic syndrome. These assessment criteria have been useful from a clinical perspective but have limited application from the standpoint of the general population because they require laboratory screening that is costly and challenging for a specific population. The prevalence of the mentioned factors and the fact that most people are unaware of their health status and do not seek healthcare services until the disease causes health problems have led researchers toward the use of universal tools and models for all population segments.

On the other hand, non-clinical methods can be categorized into statistical methods and machine learning-based methods. Many studies have utilized various statistical analysis methods such as linear regression or logistic regression [16] [17]. However, these methods have not been widely accepted due to limitations, including strict assumptions and issues of multicollinearity [18]. On the other hand, the use of machine learning-based methods is rapidly expanding because they do not rely on specific data assumptions and can discover hidden relationships among the available data. With machine learning, non-invasive indices that do not require blood sampling can be used for predicting metabolic syndrome, enabling early detection of the syndrome even in areas with limited medical resources [19].

Many machine learning models have been used in previous studies to address this issue. In [18], six machine learning methods, including k-Nearest Neighbours, Bayesian networks, Random Forests, Decision Trees, Multilayer Perceptrons (MLP), and Support Vector Machines (SVM), were employed to predict the disease. The dataset consisted of 2871 individuals, with 750 of them having metabolic syndrome. Features such as age, gender, education level, marital status, body mass index, stress, physical activity, alcohol consumption, and smoking status were extracted for each individual. According to the results of the experiments, Support Vector Machines and MLP showed better performance than other methods.

In [20], the relationship between MS and parameters derived from blood tests in 13,978 adults in South Korea was examined. For this purpose, the performance of a simple Bayesian classifier was compared to a binary logistic regression algorithm. The findings showed that multiple factors were associated with MS, and the potential use of machine learning models for detecting MS was suggested.

In another study [15], besides machine learning methods such as Random Forests and C4.5 decision trees, deep learning methods were also used. The data used in this study included 2289 individuals in Mexico, with features like waist circumference, age, body mass index, waist-to-height ratio, height, sleep-related snoring, dietary habits, coffee consumption habits, soft drink and milk consumption, extracted for each person. According to the results, Random Forests outperformed other methods.

Authors in [21] used five classification methods, including Naive Bayes, MLP networks, Random Forests, Decision Trees, and Support Vector Machines, to predict the likelihood of individuals developing MS in non-obese populations. An interesting aspect of this research is that researchers used clinical information such as age, gender, body mass index, smoking status, alcohol consumption status, and exercise status for both training and testing the model in one step. Then, they added the individual's genetic

information to the existing data, which improved the models' performance. Lastly, the suggestion was made for persuading individuals to undergo health assessments and adopt preventive lifestyle behaviours.

A predictive model for MS using only non-invasive information was presented in [3]. In constructing this model, three features that required blood tests, such as triglycerides, blood sugar, and HDL cholesterol, were excluded. The primary innovation in this research is the use of three new artificial features from four main and effective features in the onset of metabolic syndrome, namely waist circumference, systolic and diastolic blood pressure, and gender. Furthermore, several machine learning algorithms were used to evaluate the proposed approach, with the decision tree method demonstrating better performance.

The study closest to our article utilized Random Forests to predict metabolic syndrome in children and adolescents under eighteen in Birjand County [22]. The dataset was imbalanced, so the SMOTE algorithm was used to balance it. One of the disadvantages of the proposed approach is that it does not compare the results with other studies. It's worth noting that in this article, we also used the same dataset from the aforementioned research and will compare the results.

Reviewing the conducted research, we conclude that the methods presented so far have mainly employed common data classification methods, and the only difference is the use of different datasets with various features. This article utilizes real data collected from children and adolescents under eighteen in Birjand. Therefore, its results can provide valuable insights to healthcare professionals. The collected data is imbalanced, so the SMOTE algorithm was used to balance it. Subsequently, the popular Support Vector Machines algorithm was used for model training and testing. The reason for using this method is its high performance and generalization capabilities. To determine the optimal free parameters of this method, the Particle Swarm Optimization (PSO) algorithm was used. The results of the proposed method demonstrate its high performance.

The structure of the article is as follows: in the next chapter, we will briefly explain the foundations and structure of the methods used in the article. The third chapter explains the proposed method, and the results of the proposed method will be presented in the fourth chapter. Finally, the last chapter will provide a summary of the article.

## 2 BACKGROUND

### 2.1 Support Vector Machines

The Support Vector Machine (SVM) [23] was introduced by Dr. Vapnik in 1995 and quickly became a powerful tool for data classification. It has found numerous applications in the medical field, including the prediction of diseases such as metabolic syndrome, which were introduced in the previous chapter. This method was invented to separate two classes of data from each other. Some of these studies have been introduced in the previous chapter for predicting MS. One of the important features of this method is its strong mathematical foundation. In this method, the classifier is the result of solving an optimization problem with a convex objective function and several constraints. The number of these constraints is as much as the number of learning samples. The result of solving this problem is to find the optimal hyperplane with the maximum margin (OSH). Suppose we have the training dataset  $X = \{(x_i, y_i)\}_{i=1}^n$ , where  $n$  is the number of system training samples. Each  $x_i \in \mathcal{R}^m$  is a learning sample with  $m$  features in the input space, and  $y_i \in \{-1, 1\}$  is the label of the class of the  $x_i$  data. The problem of finding OSH in this case is primarily solving the following optimization problem:

$$\begin{aligned} & \text{Minimize } \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n \xi_i \\ & \text{subject to } y_i(w^T x_i + b) \geq 1 - \xi_i \end{aligned} \quad (1)$$

By using the slack variable  $\xi = (\xi_1, \xi_2, \dots, \xi_n)$ , which corresponds to each data point, errors in classification are allowed. The purpose of adding the term  $C \sum_{i=1}^n \xi_i$  is to control the number of misclassified points, and

it is referred to as the penalty coefficient. In fact,  $C$  is a tuning parameter that is initially assigned a value. Adjusting this parameter can strike a balance between maximizing the area between the two classes and the classification error. The larger  $C$  is chosen, the  $\xi_i$  values are constrained, and as a result, the area between the two classes does not expand. Choosing a small value for  $C$  allows for more classification errors. Furthermore, for non-linear classification, data is mapped to a feature space using a kernel function and separated by a hyperplane in the new space. Different kernel functions are used for this purpose, including linear, polynomial, and RBF.

One of the challenging issues in using any data classification method is parameter tuning. The chosen values for these variables affect the model's performance in learning and testing. Therefore, many methods have been proposed to determine the values of these parameters, with one of the most important ones being the use of evolutionary algorithms. The free parameters in the SVM method are the kernel function variable, as well as the parameter  $C$ . In this paper, we will calculate the optimal values of these two parameters for prediction problem using the PSO evolutionary algorithm.

## 2.2 Particle Swarm Optimization

The Particle Swarm Optimization (PSO) algorithm [24] is inspired by the movement of birds in two-dimensional space. The position and velocity of each bird (particle) are represented by a point in a two-dimensional space. A flock of birds attempts to optimize a specific objective function. Each particle knows its best value so far (*pbest*) and its current position. In addition, each particle is aware of the best value searched in the flock so far (*gbest*). Therefore, each particle tries to change its position using the concept of velocity and move towards the optimal solution. The velocity of each particle is updated using Equation (2).

$$v_i^{k+1} = \omega v_i^k + c_1 rand_1 \times (pbest_i - s_i^k) + c_2 rand_2 \times (gbest_i - s_i^k), \quad (2)$$

where  $v_i^k$  is the velocity vector of particle  $i$  at iteration  $k$ ,  $\omega$  is the weight factor,  $rand_1$  and  $rand_2$  are random numbers between zero and one, and  $s_i^k$  is the current position of particle  $i$  at iteration  $k$ . Also, the next position of particle  $i$  is calculated using Equation (3).

$$s_i^{k+1} = s_i^k + v_i^{k+1} \quad (3)$$

The use of the PSO algorithm in various applications across different fields and its numerous developments demonstrate its high efficiency and popularity.

## 2.3 SMOTE Algorithm

Imbalanced data is generally referred to as data where the number of samples in one class is much larger than the other classes. As a result, the model faces difficulties during training on the minority class. To address this issue, data needs to be pre-processed and balanced using various techniques. The Synthetic Minority Over-sampling Technique (SMOTE) [25] is an oversampling method widely used in the medical field to deal with imbalanced class data [26]. This method increases the number of data samples in the minority class by generating synthetic data points from the nearest neighbours using the Euclidean distance. The new samples are similar to the original data points based on their initial features.

## 3 THE PROPOSED METHOD

In this chapter, the proposed PSO-SVM method is presented for predicting metabolic syndrome. As the title suggests, it utilizes the Support Vector Machine (SVM) method for classification. To determine

the optimal parameters in the SVM method, the Particle Swarm Optimization (PSO) technique is employed. Figure 1 illustrates the pseudocode for the proposed method.

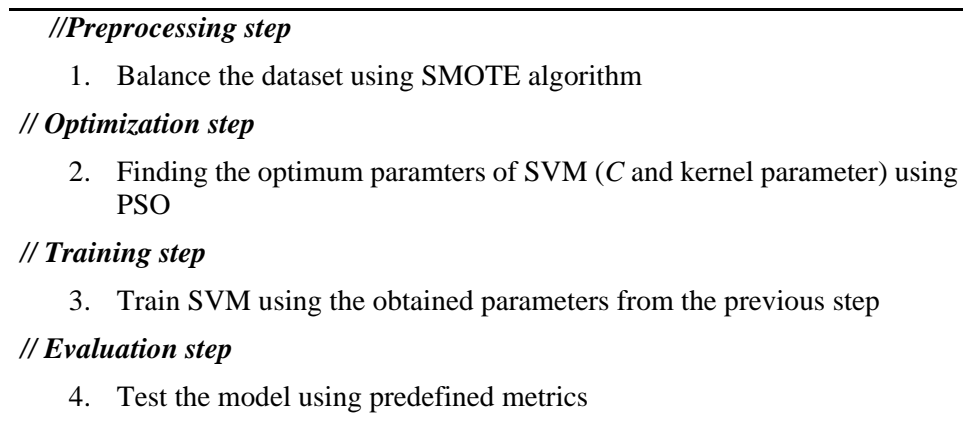


Figure 1: Pseudocode of the proposed PSO-SVM method

As indicated in the algorithm presented in Figure 1, the proposed method consists of four steps. In the first step, the data needs to be balanced. It's worth noting that the available data consists of information on 3,819 individuals, of which 3,577 are normal, and 242 are non-normal. Table 1 shows the information related to the distribution of the available data by gender and age.

Table 1: Distribution of the data used by gender and age

Data Distribution		Number of data with the normal class label	The number of data with the abnormal class label
Gender	Boy	1600	139
	Girl	1977	103
Age	6 to 11 years	1564	92
	12 to 18 years	2013	150

As evident, the ratio of minority to majority data is less than 0.07. Therefore, the data should be balanced before any action. To do this, we use the SMOTE algorithm. After applying the SMOTE algorithm to the data, the number of individuals with the abnormal label (minority) is adjusted to match the majority class.

In the second step, the optimal parameters for the SVM method are calculated using the PSO algorithm. The free parameters of the SVM method include the balancing parameter C and the kernel function parameter. In this article, since the RBF kernel function is used as per equation (4), the optimal value of the parameter  $\sigma$  should be calculated. For this purpose, the parameters of the PSO algorithm are initialized with  $\omega=0.7$ ,  $c_1=1.5$ , and  $c_2=1.5$ . Additionally, it is worth noting that the fitness function used in the PSO method considers the accuracy of the SVM classifier.

$$K(x, y) = \exp\left(-\frac{\|x-y\|^2}{2\sigma^2}\right) \quad (4)$$

In the third and fourth steps, using the optimized parameters obtained for SVM in the previous step, we proceed with the training and testing of the model.

#### 4 EXPERIMENTAL RESULTS

In this section, the results of the proposed method for predicting MS disease are presented. The optimized SVM algorithm using PSO was tested on four subsets (females, males, age groups 6-11 years, and 12-18 years). The performance of this algorithm is compared with the methods presented in [22] as well as basic methods like KNN and SVM. The comparison metrics in this article include accuracy, sensitivity, specificity, and area under the ROC curve (AUC) (equations (5) to (8)).

$$Sensitivity = \frac{TP}{TP + FN} \quad (1)$$

$$Specificity = \frac{TN}{TN + FP} \quad (2)$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (3)$$

The test results of the proposed method are presented. Table 2 displays the comparison results for different methods, categorized by gender. For a better understanding, Figure 2 shows the comparison results for the male group, and Figure 3 shows the results for the female group. It's worth noting that in the charts presented in Figures 2 and 3, the AUC metric has been multiplied by 100 for better visualization.

Table 2: Comparison results of different methods by gender

		<b>CART</b>	<b>CHAID</b>	<b>QUEST</b>	<b>KNN</b>	<b>SVM</b>	<b>SVM- PSO</b>
<b>Male</b>	<b>Sensitivity</b>	93.846	88.846	87.692	<b>100</b>	99.016	97.71
	<b>Specificity</b>	90.759	91.749	86.755	92.54	93.731	<b>100</b>
	<b>Accuracy</b>	92.180	90.410	87.030	96.03	96.250	<b>98.91</b>
	<b>AUC</b>	0.976	0.955	0.949	0.962	0.963	<b>0.996</b>
<b>Female</b>	<b>Sensitivity</b>	98.701	91.775	92.208	<b>100</b>	99.248	98.95
	<b>Specificity</b>	93.684	92.895	91.316	91.836	93.877	<b>100</b>
	<b>Accuracy</b>	95.580	92.470	91.650	95.954	96.586	<b>99.49</b>
	<b>AUC</b>	0.982	0.975	0.953	0.959	0.965	<b>0.997</b>

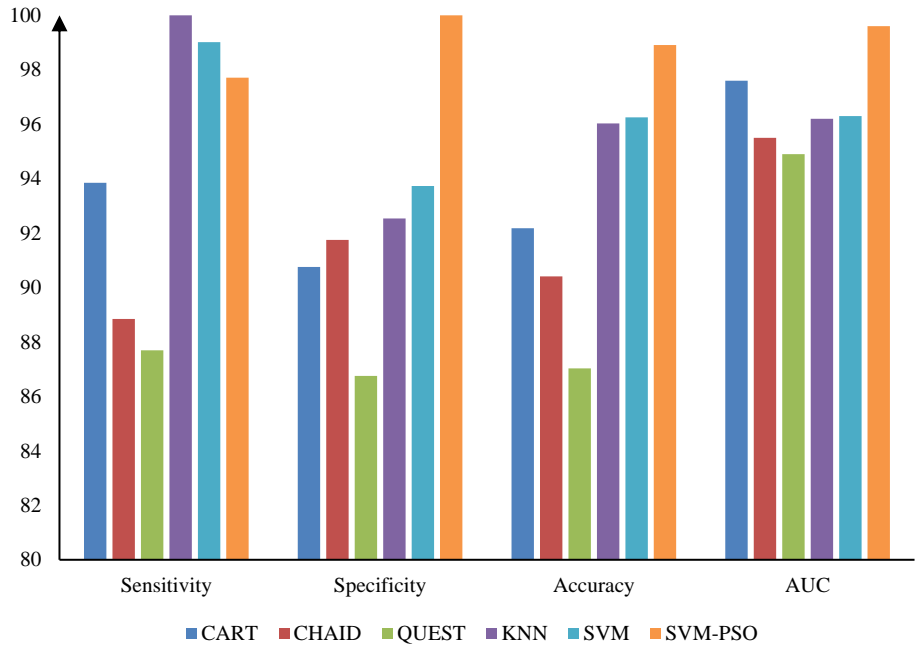


Figure 2: Comparison of different methods for males

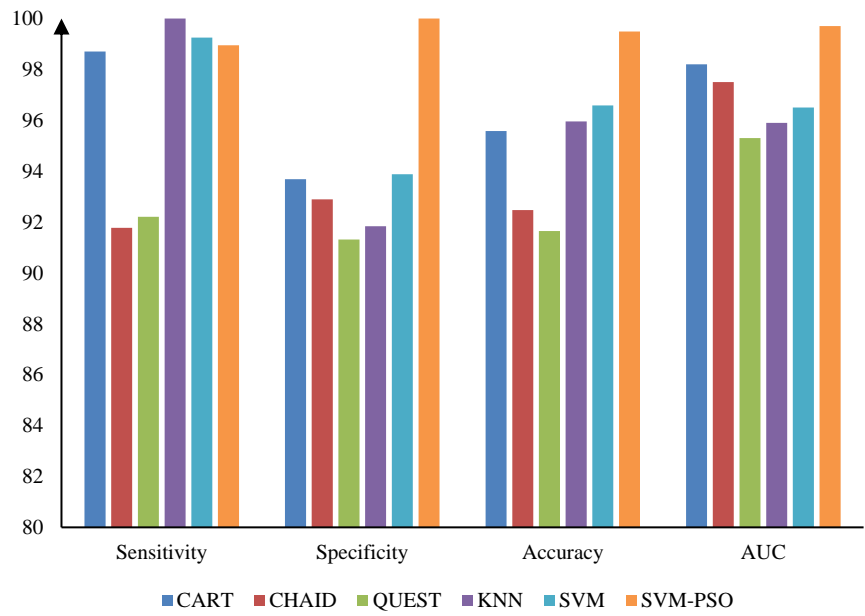


Figure 3: Comparison of different methods for females

As the results show, the proposed method performs better in most parameters for individuals of both male and female genders. However, the sensitivity parameter reports a lower value compared to the KNN method.

In the following, the results of the proposed method’s testing are presented. Table 3 displays the comparison results of different methods, broken down by age. To better understand the results, Figure 4 shows the comparison results for the age group of 6 to 11 years, while Figure 5 illustrates the comparison results for the age group of 12 to 18 years. It's important to note that in the presented graphs in Figures 4 and 5, the AUC metric has been scaled by a factor of 100 for better visualization.

Table 3: Comparison of various methods by age group

		CART	CHAID	QUEST	KNN	SVM	SVM-PSO
age-group [6-11]	Sensitivity	96.914	95.370	95.988	99.679	99.358	<b>99.660</b>
	Specificity	95.380	91.419	87.789	92.356	93.949	<b>100</b>
	Accuracy	96.170	93.460	92.030	96.006	96.645	<b>99.840</b>
	AUC	0.975	0.974	0.929	0.960	0.966	<b>1</b>
age-group [12-18]	Sensitivity	97.396	97.396	94.531	<b>100</b>	99.234	99.230
	Specificity	97.354	91.534	92.593	92.753	94.444	<b>100</b>
	Accuracy	97.380	94.490	93.570	96.277	96.774	<b>99.630</b>
	AUC	0.987	0.982	0.957	0.963	0.968	<b>1</b>

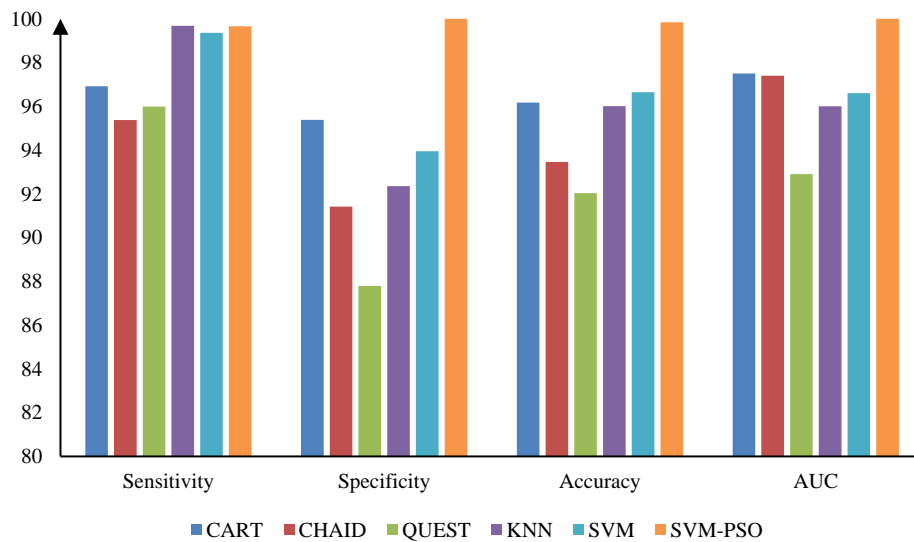


Figure 4: Comparison of different methods for individuals aged 6 to 11 years



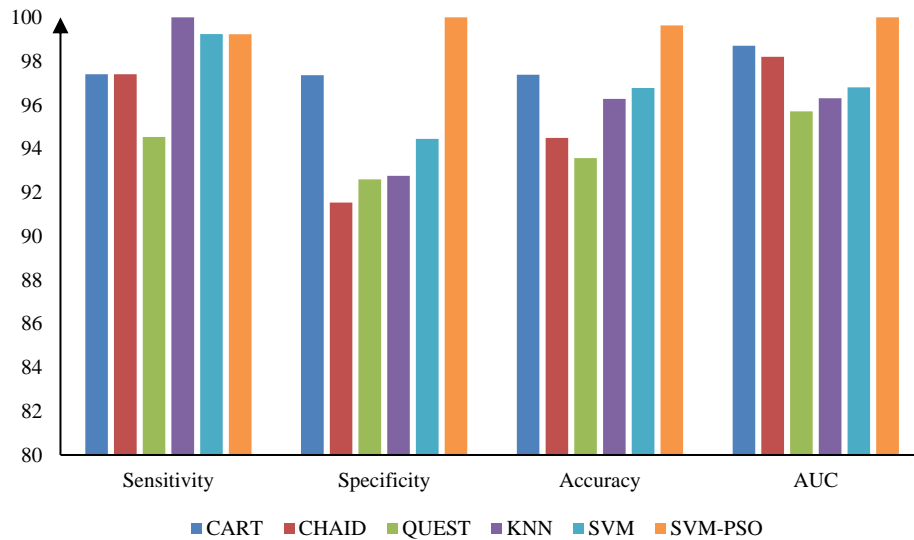


Figure 5: Comparison of different methods for individuals aged 12 to 18 years

Upon analyzing the results, it becomes evident that the proposed method outperforms other methods for all individuals in the age range of 6 to 11 years. In the age range of 12 to 18 years, except for the sensitivity parameter, where the KNN method performs better, the proposed method reports better results.

## 5 CONCLUSIONS

One of the main concerns in medical societies today is the prevalence of metabolic syndrome (MS). Urbanization, lack of physical activity, and overall lifestyle changes have directly contributed to the increasing number of individuals affected by this condition. Preventive measures play an effective role in early detection and, consequently, better treatment of individuals. In this paper, a solution for more accurate diagnosis of this condition is presented, and the results obtained from it are tested and analyzed on the data of children and adolescents in the city of Birjand. The results, in comparison to various methods such as decision trees, KNN, and support vector machines, demonstrate the reliable performance of this method. An important point to note about the results is that, unlike the compared methods which exhibit fluctuations in the parameters being compared, the proposed method shows consistent and reliable performance with minimal variance.

## REFERENCES

- [1] O'Neill, S. and O'Driscoll, L., "Metabolic syndrome: a closer look at the growing epidemic and its associated pathologies", *Obesity Reviews*, vol. 16, pp. 1-12, 2015. <https://doi.org/10.1111/obr.12229>.
- [2] Sookoian, S. and Pirola, C.J., "Metabolic Syndrome: From the Genetics to the Pathophysiology", *Curr Hypertens Rep* vol. 13, pp. 149–157, 2011. <https://doi.org/10.1007/s11906-010-0164-9>.
- [3] Shin, Y., Shim, S., and Oh, S., "Machine learning-based predictive model for prevention of metabolic syndrome", *PLoS One*, vol. 18, 2023. <https://doi.org/10.1371/journal.pone.0286635>.
- [4] Yki-Jarvinen H., "Non-alcoholic fatty liver disease as a cause and a consequence of metabolic syndrome", *Lancet Diabetes Endocrinol*, vol. 2, pp. 901–910, 2014. [https://doi.org/10.1016/S2213-8587\(14\)70032-4](https://doi.org/10.1016/S2213-8587(14)70032-4).
- [5] Argolo, D.F., Hudis, C.A., and Iyengar, N.M., "The impact of obesity on breast cancer diagnosis and treatment", *Curr Oncol Rep*, vol. 20:47, 2018. <https://doi.org/10.1007/s11912-019-0787-1>.

- [6] Ni, J., Zhu, T., Zhao, L., Che, F., Chen, Y., Shou H., et al. “Metabolic syndrome is an independent prognostic factor for endometrial adenocarcinoma”, *Clin Transl Oncol*, vol. 17, pp. 835–839, 2015. <https://doi.org/10.1007/s12094-015-1309-8>.
- [7] Esposito, K., Chiodini, P., Capuano, A., Bellastella, G., Maiorino, M.I., and Giugliano, D. “Metabolic syndrome and endometrial cancer: a meta-analysis”, *Endocrine*, vol. 45, pp. 28–36, 2014. <https://doi.org/10.1007/s12020-013-9973-3>.
- [8] Hernández-Pérez, J. G., Torres-Sánchez, L., Hernández-Alcaráz, C., López-Carrillo, L., Rodríguez-Covarrubias, F., Vázquez-Salas, R. A., Galván-Portillo, M., “Metabolic Syndrome and Prostate Cancer Risk: A Population Case-control Study”, *Archives of Medical Research*, Vol. 53, Issue 6, pp. 594-602, 2022. <https://doi.org/10.1016/j.arcmed.2022.07.003>.
- [9] Despres, J. P. and Lemieux, I. “Abdominal obesity and metabolic syndrome”, *Nature*. Vol. 444, pp, 881–887, 2006. <https://doi.org/10.1038/nature05488>.
- [10] Guarner, V. and Rubio-Ruiz, M. E., “Low-grade systemic inflammation connects aging, metabolic syndrome and cardiovascular disease”, *Interdiscip Top Gerontol*, vol. 40, pp. 99–106, 2015. <https://doi.org/10.1159/isbn.978-3-318-02730-3>.
- [11] Rochlani, Y., Pothineni, N. V., Mehta, J. L., “Metabolic syndrome: does it differ between women and men?”, *Cardiovasc Drugs Ther*, vol. 29, pp. 329–338, 2015. <https://doi.org/10.1007/s10557-015-6593-6>.
- [12] Pitsavos, C., Panagiotakos D., Weinem, M., and Stefanadis C., “Diet, exercise and the metabolic syndrome”, *Rev Diabet Stud*, vol. 3, pp. 118–126, 2006. <https://doi.org/10.1900/RDS.2006.3.118>.
- [13] Sun, K., Ren, M., Liu, D., Wang, C., Yang, C., and Yan L., “Alcohol consumption and risk of metabolic syndrome: a meta-analysis of prospective studies”, *Clinical Nutrition*, vol, 33, pp. 596–602, 2014. <https://doi.org/10.1016/j.clnu.2013.10.003>.
- [14] Shin, H. S., Oh, J. E., Cho, Y. J., “The association between smoking cessation period and metabolic syndrome in Korean men”, *Asia Pac J Public Health*, vol. 30, pp. 415–424, 2018. <https://doi.org/10.1177/1010539518786517>.
- [15] Gutiérrez-Esparza, G.O., Ramírez-delReal, T.A., Martínez-García, M., Infante Vázquez, O., Vallejo, M., Hernández-Torruco, J., “Machine and Deep Learning Applied to Predict Metabolic Syndrome without a Blood Screening”, *Appl. Sci.* vol. 11, pp. 4334, 2021. <https://doi.org/10.3390/app11104334>.
- [16] Stefanescu, A., Revilla, L., Lopez, T., Sanchez, S. E., Williams, M. A., and Gelaye, B., “Using A body shape index (absi) and body roundness index (bri) to predict risk of metabolic syndrome in Peruvian adults,” *The Journal of International Medical Research*, vol. 48, no. 1, 2019. <https://doi.org/10.1177/0300060519848854>.
- [17] Alzeidan, R., Fayed, A., Hersi, A. S., and Elmorshedy, H., “Performance of neck circumference to predict obesity and metabolic syndrome among adult Saudis: a cross-sectional study”, *BMC Obesity*, vol. 6, p. 13, 2019. <https://doi.org/10.1186/s40608-019-0235-7>.
- [18] Park, J. E., Mun S., and Lee S., “Metabolic Syndrome Prediction Models Using Machine Learning and Sasang Constitution Type”, *Evid Based Complement Alternat Med*, 8315047, 2021, <https://doi.org/10.1155/2021/8315047>.
- [19] Yang, H., Yu, B., OUYang, P. et al. “Machine learning-aided risk prediction for metabolic syndrome based on 3 years study”, *Sci Rep*, vol. 12, pp. 2248, 2022. <https://doi.org/10.1038/s41598-022-06235-2>.
- [20] Kim, S. Y., Nam, G. H., Heo, B. M. “Identification of Metabolic Syndrome Based on Anthropometric, Blood and Spirometric Risk Factors Using Machine Learning”, *Applied Sciences*, vol. 10, no. 21, pp. 7741, 2020. <https://doi.org/10.3390/app10217741>.
- [21] Choe, E. K., Rhee, H., Lee, S., Shin, E., Oh, S. W., Lee, J. E., and Choi, S. H., “Metabolic Syndrome Prediction Using Machine Learning Models with Genetic and Clinical Information from a Nonobese Healthy Population”, *Genomics Inform*, vol. 16, no., 4(e31), 2018. <https://doi.org/10.5808/GI.2018.16.4.e31>.
- [22] Taheri, F., Babaiyan, V., Saffarian, M., Kazemi, S. M., Namakin, K., Kazemi, T., and Ramazankhani, A., “Assessment of the Metabolic Syndrome in Children and Adolescents in Birjand, Iran: A Data Mining Approach”, *Journal of Computer and Knowledge Engineering*, Vol. 3, No. 2, pp. 21-30, 2020. <https://doi.org/10.22067/CKE.2021.63297.0>.
- [23] Cortes, C. and Vapnik, V., “Support-vector networks,” *Mach Learn*, vol. 20, pp. 273–297, 1995. <https://doi.org/10.1007/BF00994018>.
- [24] Kennedy, J., and Eberhart, R., “Particle swarm optimization”, *International Conference on Neural Networks*, Perth, WA, Australia, vol. 4, pp. 1942-1948, 1995. <https://doi.org/10.1109/ICNN.1995.488968>.
- [25] Chawla, N. V., Bowyer, K. W., Hall, L. O., Kegelmeyer, W. P., “SMOTE: Synthetic minority over-sampling technique”, *J Artif Intell Res* vol. 16, pp.321–357, 2002. <https://doi.org/10.1613/jair.953>.
- [26] Ishaq, A., et al., “Improving the Prediction of Heart Failure Patients’ Survival Using SMOTE and Effective Data Mining Techniques”, *IEEE Access*, vol. 9, pp. 39707-39716, 2021, <https://doi.org/10.1109/ACCESS.2021.3064084>.