International Journal of Engineering

Journal Homepage: www.ije.ir

# Algorithm of Predicting Heart Attack with using Sparse Coder

S. Mohamadzadeh*, M. Ghayedi, S. Pasban, A. K. Shafiei

*Department of Electrical and Computer Engineering, University of Birjand, Birjand, Iran*

*PAPER INFO*

*Paper history:*
Received 15 July 2023
Received in revised form 09 August 2023
Accepted 10 August 2023

*Keywords*:
*Heart Attack*
*Classification*
*Prediction*
*Machine Learning*
*Sparse Representation*

*ABSTRACT*

One of the most serious causes of disease in the world's population, which kills many people worldwide every year, is heart attack. Various factors are involved in this matter, such as high blood pressure, high cholesterol, abnormal pulse rate, diabetes, etc. Various methods have been proposed in this field, but in this article, by using sparse codes in the classification process, higher accuracy has been achieved in predicting heart attacks. The proposed method consists of two parts: preprocessing and sparse code processing. The proposed method is resistant to noise and data scattering because it uses a sparse representation for this purpose. The spars allow the signal to be displayed at its lowest value, which leads to improve computing speed and reduce storage requirements. To evaluate the proposed method, the Cleveland database has been used, which includes 303 samples and each sample has 76 features. Only 13 features are used in the proposed method. FISTA, AMP, DALM and PALM classifiers have been used for the classification process. The accuracy of the proposed method, especially with the PALM classifier, is the highest among other classifiers with 96.23%, and the other classifiers are 95.08%, 94.11% and 94.52% for DALM, AMP, FISTA, respectively.

***doi***: *10.5829/ije.2023.36.12c.08*

## 1. INTRODUCTION

Numerous studies indicate that hospitals and medical centers possess a significant amount of patient information. However, this data is rarely utilized for decision-making, treatment, and patient services. Extracting and utilizing this information can greatly contribute to decision-making and the quality of healthcare services. Heart disease is especially critical due to its high sensitivity and potential for saving lives through early diagnosis and treatment. It is a well-known cause of death worldwide, resulting in significant financial and human losses. Implementing preventive methods plays a vital role in reducing the incidence of heart disease, such as clogged arteries. Angiography, the established method for diagnosing clogged arteries, is associated with numerous side effects and is costly. Therefore, researchers are seeking alternative non-invasive methods. Figure 1 demonstrates an example of an ECG signal, which is crucial for diagnosing heart disease [1].

The importance of predicting heart disease can be assessed from two perspectives: heart disease itself and machine vision methods. Machine vision methods, in terms of modeling, are more accurate than traditional approaches. By 2030, the estimated global death toll from heart disease is expected to reach 23 million. Based on these statistics, early and accurate prognosis of heart disease is crucial for saving countless lives worldwide [2].

In this paper, we proposed using sparse coding to extract generative features from input data that are robust to noise, diversity, and other factors. Our goal is to reconstruct the signal using the fewest number of signals. One of the advantages of using sparse codes is that it is resistant to noise and data dispersion. In other words, according to the sampling environment, most processed signals are associated with noise or dispersion. One of the best ways to deal with these two problems is to use L1-norm that is used in this paper. In the following, various classifiers have been used in the learning process, which achieved higher accuracy by using the PALM. To

*Corresponding Author Email: *s.mohamadzadeh@birjand.ac.ir*
(S. Mohamadzadeh)

Please cite this article as: S. Mohamadzadeh, M. Ghayedi, S. Pasban, A. K. Shafiei, Algorithm of Predicting Heart Attack with using Sparse Coder, ***International Journal of Engineering, Transactions C: Aspects***, Vol. 36, No. 12, (2023), 2190-2197

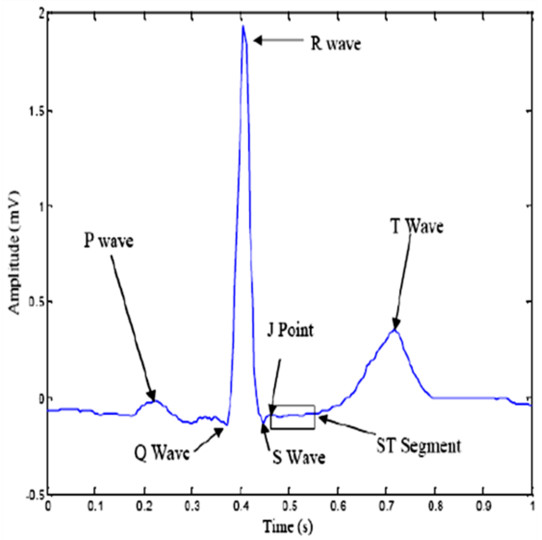


**Figure 1.** Electrical signal wave of the heart [1]

evaluate the proposed method, the Cleveland database has been used, which only 13 features are used in the proposed method. In order to select efficient features from the database, the influencing variables have an effect on the occurrence of heart disease. In this article, two methods are used to select the features, which are: using the opinions of experts and specialists, as well as existing articles in this field. After the investigations, 13 features were selected to evaluate the proposed method, the results of which are compared in section 4.

Details about the parameters related to our proposed method are provided in Section 2. In section 3, we explained the sparse coding-based method we have developed. We compared our proposed method with other techniques using evaluation metrics, which are presented in section 4 along with the results. Finally, we conclude our study in section 5.

## 2. RELATED WORKS

The section presents a review of previous methods for predicting heart attacks. Gheitasi et al. [3] used the C-Means fuzzy clustering method to predict heart disease. The study evaluated the proposed method with 270 samples and found that it was more accurate than the K-Means clustering method, with an accuracy of 92%. Mohan et al. [4] proposed a method based on the backpropagation neural network. All 13 features of the Cleveland dataset were used to train the neural network, which had a specific structure. The accuracy of this method was 92% for the learning data and 86% for the test data. Verma et al. [5] presented a combined model of heart disease prediction which uses the genetic algorithm to update weights in an artificial neural network. This model has a higher speed in updating weights compared to the error back-propagation learning method. Finally, Mohamadzadeh et al. [6] utilized a hybrid model of the fuzzy-particle swarm system and decision tree to predict heart disease. The decision tree algorithm is used to select disease rules, the fuzzy system implements the selected rules, and the particle swarm algorithm optimizes the fuzzy system membership functions. The accuracy and sensitivity obtained from this model are 94.4% and 62%, respectively. Numerous methods have been proposed for predicting heart disease. Nazir wt al. [7] proposed an adaptive neural-fuzzy inference system to predict heart disease in the Cleveland Database. All 13 features are used for this purpose. Initially, a fuzzy inference system is formed. Then, the parameters of the fuzzy system are updated using an adaptive neural-fuzzy inference system and a combined training method to achieve the lowest possible error. The system is designed to have 9 rules and achieves an accuracy of 93.88%. Bahtiar et al. [8] presented a hybrid model based on the majority vote. This model uses three types of artificial neural networks: two multilayer perceptron neural networks and one radial base function network. The model uses a total of eight learning algorithms. The multilayer perceptron neural network achieves an accuracy of 93%, while the radial base neural network achieves 79.6%. Fooladi et al. [9] proposed the use of a hybrid model based on three simple Bayesian classification algorithms, decision trees, and support vector machines. The model uses the Cleveland Database and makes a final prediction of heart disease based on the majority vote of the results from the three categories. To predict heart attacks, previous studies have suggested using a combination of neural networks and genetic algorithms [10]. The genetic algorithm is used in this method to reduce the dimensions of the features, thereby increasing the accuracy of the prediction. Wadhawan and Maini [11] used the particle swarm optimization method based on fuzzy logic in order to classify patients in the prediction of cardiac patients. First, they fuzzified the rules in the patient data set by fuzzy logic and then optimized these rules by the particle swarm method. Many methods have used supervised data mining techniques to detect heart disease.Ahmed et al. [12] have used two algorithms for this purpose, which are: genetic algorithm and particle swarm algorithm. Patients are divided into two classes: disease and non-disease. The use of Bagging method to classify cardiac patients is suggested by Yuan et al. [13]. The authors of this article have used the Begin method, which is based on the decision tree base algorithm. The accuracy of the Bagging-based method has been higher than the conventional decision tree method.

The use of deep learning techniques based on neural networks has had a remarkable growth in the last several years, which have been used in many medical applications, including the risk of heart failure. Choudhury and Akbar [14] suggested the use of Convolutional Neural Networks (CNN), which deals with the early identification of people at risk of heart failure. Adaptive multilayer networks are also introduced by Banu and Vanjerkhede [15] to predict the risk of heart failure. This method works better than classical neural networks.

Various perspectives can be considered in this regard, which can achieve higher efficiency in the field of heart attack prediction by using new technologies such as artificial intelligence. An example is an artificial intelligence (AI) model that can predict the risk of death from a heart attack or stroke over a 10-year period with just a chest X-ray. This risk is calculated using a score based on variables such as age, sex, race, blood pressure, high blood pressure treatment, smoking, type 2 diabetes and blood tests. In this method artificial intelligence researchers trained deep learning to search X-ray images for patterns associated with atherosclerosis, a leading cause of cardiovascular disease. Because chest X-ray imaging is more readily available than other imaging modalities, it helps to identify individuals at risk [16].

## 3. THEORETICAL BACKGROUND

Figure 2 shows a block diagram of the heart attack prediction algorithm using a sparse classifier. In this part, the features used are from the Cleveland Dataset, and each sample includes 13 features. Various sparse classifier methods have been employed for the learning process, such as Approximate Message Passing (AMP), Fast Iterative Soft-Thresholding Algorithm (FISTA), Primal Augmented Lagrangian Method (PALM), and Dual Augmented Lagrangian Method (DALM). The proposed method has two stopping conditions: reducing the error rate to the default threshold value and reaching the specified number of repetitions.

The algorithm presented in this section uses sparse representation and weighted elements to predict heart attacks. The flowchart for Algorithm 1, which outlines the steps of the proposed algorithm, is depicted below. Sparse coding, a powerful tool for analyzing various types of signals, is utilized in this method. The term "sparse" refers to a small number and is employed to represent non-zero values in vectors [17]. These methods require a learning process. A detailed explanation of sparse coding can be found in Algorithm 1.

In specific orthogonal transformations, the number of base signals is equal to the dimension of the processing signal. These transformations are suitable for representing a small number of signals. Specifically, it is not feasible to use a small number of signals to represent a signal, due to the limitations of unique orthogonal transformations. As a result, the use of sparse codes has

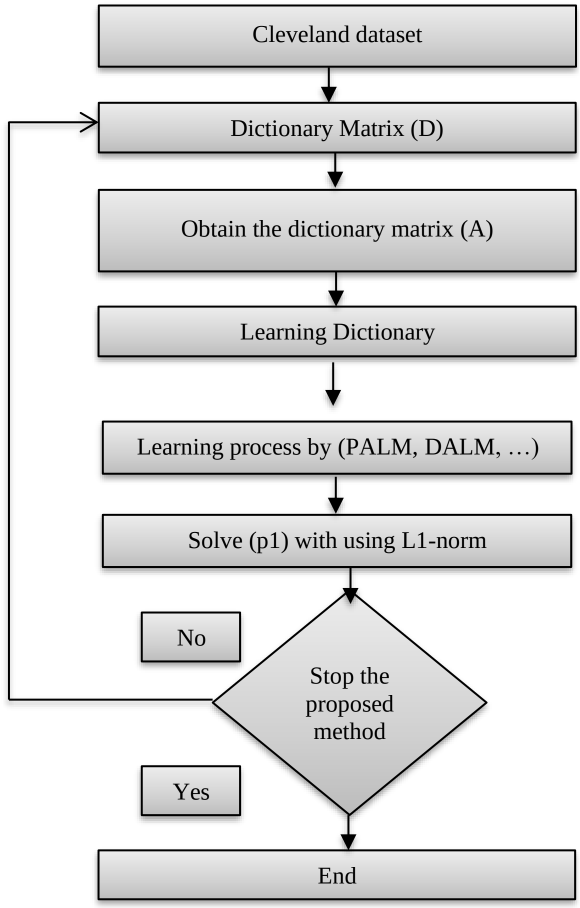


**Figure 2.** Schematic of the proposed method

**Algorithm 1.** Pseudo code of predicting heart attack algorithm method

1: Get data from new database;
2: Learning process by (PALM, Homotopy, …) algorithm;
3: solve $(p1)$: $\min \|x\|_1$ $subject\ to\ y = Dx$ and find x0:

$$E = \| b - Ax \|_F^2 = \| Y - \sum_{j=1}^{m} a_j x_j^T \|_F = \| (b - \sum_{j \neq j_0}^{m} a_j x_j^T) - a_{j0} x_{j0}^T \|_F^2$$

**While** stop condition is not satisfied, **do**:
4: Construct dictionary A:
- $A_i = [v_{i,1}, v_{i,2}, \dots, v_{i,m}] \in R^{m \times 1}$, $v_{i,j}$ denotes the j[th] feature of the i[th] image;
- $A = [A_1, A_2, \dots, A_n] \in R^{m \times n}$

5: Construct $b \in R^{m \times 1}$;
6: Seek sparse representation, $x_0 \epsilon R^n$ by solving Equation (1) and using SL0, DALM, PALM, homotopy, FISTA nd AMP techniques. Thus, some elements of x0 are zero except those associated with the $k^{th}$ column.

7: Separate elements of A and x0;

$$x_0 = [\underbrace{a_1, a_2}_{1}, \underbrace{\dots, a_j}_{2}, \underbrace{\dots, \dots}_{k}, a_n] = [x_{0,1}, x_{0,2}, \dots, x_{0,k}]$$

$A = [A_1, A_2, ..., A_j, ..., ..., A_n]$

$$\underset{1}{\longleftrightarrow} \quad \underset{2}{\longleftrightarrow}\underset{k}{\longleftrightarrow} = [D_1, D_2, ..., D_k]$$

8: Define $C_k = D_k x_{o,k} \in R^m$;

9: Compute Euclidean Distance (ED);

$$ED_k = \sqrt{\sum_{i=1}^{m} (b_i - C_i)^2}$$

10: Compute the weighting elements;

$$x_{weighted} = [\frac{|x_{o,1}|}{ED_1}, \frac{|x_{o,2}|}{ED_2}, ..., \frac{|x_{o,k}|}{ED_k}] \in R^m$$

-

**End**

11: computed the accuracy query sample by using element of $X_{weighted}$.

been recommended. Sparse coding, which utilizes base signals instead of signal dimensions, provides a more straightforward representation of the signal by setting specific criteria for it.

The information in the input data has many repetitive structures that they are known as sparsity. Also, in many real-world applications, the input data is accompanied by noise. In this regard, sparse display has been used to solve these two problems. There are various reasons for the existence of these two problems, one of which is the sampling environment.

It is possible that most of the natural signals, taking into account personal bases, have a display sparse. Natural signals often do not cover the entire space and are placed on a manifold subspace. For this purpose, various methods such as Matching Pursuit and Orthogonal matching, can be used. In this paper, another criterion is used to measure thinness and stability against noise, which is confirmed as $l_1$-norm in the following.

In the sparse representation method, which is considered one of the best methods, a large signal can be represented using few non-zero coefficients. In the condition that k<< n, a 'k' non-zero coefficient presents a signal with the length of 'n'. This enables signal compression and reduction in the number of required measurements. Researchers in various fields, such as speech recognition and image processing, have therefore employed the sparse representation method:

$$b = Ax = \sum_{i=1}^{n} a_i x_i \tag{1}$$

If the main signal is $b \in R^m$, and the dictionary is defined as A in which $\alpha_i \in R^m$ $(1 \leq i \leq n)$ are named atoms. Clearly, Equation (1) solution is dependent on the relation between 'm' and 'n'. On the condition that 'm' is equal to 'n', the equation has one unique solution; and on the condition that 'm' is smaller than 'n', no unique solution is possible for the equation. Therefore, in order to achieve the specific solution, the following condition is considered for the equation [18]:

$$P_j = \min J(x) \quad subject.to \quad b = Ax \tag{2}$$

The problem can be classified into numerous forms based on J(x) function. Sezavar et al. [19] showed that the $l_0$-norm of x is found as follows:

$$\|x\|_0 = \lim_{q \to 0} \sum_{i=1}^{n} |x_i|^q \tag{3}$$

where

$$P_0: \min \|x\|_0 \quad subject.to \quad b = Ax \tag{4}$$

The output of this equation indicates the sum of non-zero components of the vector x. However solving this problem with the $l_0$-norm function can make results sparse; this is not a convex problem and is difficult to solve. Therefore, the closest solution to $l_0$ called $l_1$-norm is used instead. Therefore, the problem becomes a convex optimization problem [18-21]:

$$P_1: min\|x\|_1 \quad subject.to \quad b = Ax \tag{5}$$

Because, in fact, the signals are accompanied by noise, Equation (5) is rewritten as Equation (6) as follow:

$$p_{1,2}: min\|x\|_1 \quad subject.to \quad \|b - Ax\|_2 \ll \varepsilon \tag{6}$$

Finally, it can be displayed as the famous Equation (7) as follows [18, 21]:

$$QP_\lambda: \min \frac{1}{2} \|b - Ax\|_2^2 + \lambda \|x\|_1 \tag{7}$$

Sezavar et al. [19] showed that PALM and DALM methods are more efficient than other methods. So PALM and DALM methods are used in this paper to solve the optimization problem (Equation (7)).

## 4. ANALYSIS OF EXPERIMENTAL RESULTS

In the sparse representation method, which is considered one of the best approaches, a large signal can be represented using only a few non-zero coefficients. This allows for efficient signal compression and reduces the number of required measurements. As a result, the sparse representation method has found applications in various fields, including speech recognition and image processing.

**4. 1. Evaluation Metrics** In order to evaluate the proposed method with other existing methods, three evaluation criteria of sensitivity, specificity, and accuracy have been used [22]. In other words, the following terms are used for evaluation: TP (True Positive), TN (True Negative), FP (False Positive), FN (False Negative), PC= TP + FN, and NC= FP +TN. Also, the sensitivity, specificity, and accuracy are defined as follows:

$$Sensitivity = \frac{TP}{TP + FN} \tag{8}$$

$$Specificity = \frac{TN}{TN+FP} \quad (9)$$

$$Accuracy = \frac{TP+TN}{Pc+Nc} \quad (10)$$

**4. 2. Dataset** The Cleveland database has been used to evaluate methods for predicting heart disease [16]. It is a collection of standardized and reliable data that is provided to researchers, who can use this data to compare their methods with others. Heart disease has many different symptoms, and analyzing patterns in heart data is important for diagnosing the condition. The Cleveland Database was compiled by the Cleveland Medical Center in 1998. It contains 303 samples, including 297 complete samples and 6 samples with missing values. The database includes 76 raw attributes, but experiments are only performed on 13 of these features. The data set is categorized, and Table 1 provides a description of the data [16].

**4. 3. Simulation Details** To go to the details of implementation, it is noticed that codes are written using Matlab. The simulation was done on 8 GB of RAM and core i-5 Intel CPU. For training, data is randomly divided to training (70%), validation set (15%), and test (15%). Because the model performs the same in training and validation data, it is understood that it can be used generally for new data with high performance. The size of feature vector is 13 and the dataset has 297 signals. Regularization parameter is 0.5 and maximum iteration is 150 epochs.

**4. 4. Experimental Results** In this section, the obtained results of the proposed method and state-of-the-art methods can be shown by computing the accuracy, the sensitivity, and the specificity measure. To solve Equation (1), the PALM algorithm is used. The results of the proposed method, ANNGA [23], TSB [24], FISGA [25], FISBFS [26], ANFIS1 [27], ANFIS2 [28] on Cleveland databases are presented. The index results on the Cleveland database are shown in Table 2.

**TABLE 1.** Cleveland Dataset

| No | Features | Descriptions |
|---|---|---|
| 1 | Age | Age of patient (years) |
| 2 | Sex | 1: male, 0: female |
| 3 | Chest pain (CP) | 1 = typical angina 2 = atypical angina<br>3 = nonangina pain 4 = asymptomatic |
| 4 | Rest BP | Resting blood pressure |
| 5 | Chol | Serum cholesterol in mg/dl |
| 6 | FBS | Fasting blood sugar larger 120 mg/dl (1 true) |
| 7 | RestECG | Resting electrocardiographic result |
| 8 | Thalach | Maximum heart rate accomplished |
| 9 | Exang | Exercise-induce angina (1 yes) |
| 10 | Oldpeak | ST depression induce: exercise relative to rest |
| 11 | CA | Number of major vessels (0–3) |
| 12 | Slope | Slope of peak exercise ST |
| 13 | Thal | No explanation provided, but probably thalassemia |
| 14 | Num | Diagnosis of cardiac disease: 1: yes 0: no |

**TABLE 1.** the evaluated metrics on Cleveland dataset

| Method | The Accuracy % | The Sensitivity % | The Specificity % |
|---|---|---|---|
| ANNGA | 89 | -- | -- |
| TSB | 82 | 74 | 93 |
| FIGSA | 86 | 80 | 90 |
| FISBFS | 81 | 81 | 82 |
| ANFIS1 | 92.3 | -- | -- |
| ANFIS2 | 83.8 | -- | -- |
| The proposed method (PALM) | **96.23** | 88.65 | 91.21 |

As observed in Table1, the best score has achieved Accuracy= 96.23%, Sensitivity= 88.65% and Specificity= 91.21% by the proposed method with using PALM classifier.

In the Cleveland database, we have compared the proposed model against state-of-the-art methods [23]. Saifudin et al. [23] proposed 6 classifiers, Logistic regression, K Neighbors, SVM, Random forest, Decision tree, DL. As an observer in Table 3, the accuracy rate by the proposed method, Logistic regression, K Neighbors, SVM, Random forest, Decision tree, DL classifier defined by Saifudin et al. [23] are 96.23%, 83.3%, 84.8%, 83.2%, 80.3%, 82.3% and 94.2%, respectively.

**TABLE 3.** Evaluated metrics for Cleveland dataset

| Method | Accuracy % | Sensitivity % | Specificity % |
|---|---|---|---|
| Logistic regression | 83.3 | 86.3 | 82.3 |
| K Neighbors | 84.8 | 85.0 | 77.7 |
| SVM | 83.2 | 78.2 | 78.7 |
| Random forest | 80.3 | 78.2 | 78.7 |
| Decision tree | 82.3 | 78.5 | 78.9 |
| DL | 94.2 | 82.3 | 83.1 |
| The proposed method (PALM) | **96.23** | **88.65** | **91.21** |

The performance of the proposed method on the Cleveland database is shown in Table 4, in which the proposed method has been achieved for FISTA, AMP, DALM, and PALM classifiers. The accuracy of the proposed approach using the PALM classifier is the best metric among the other classifiers.

The accuracy rate measures resulting from the proposed method and the methods reported in literture [29-34] on the Cleveland database are presented in Figure 3.

As an observer in Figure 3, the accuracy rate by the proposed method, and the data reported in literature [29-34] are 96.23%, 85.48%, 87%, 82.75%, 93%, 89.71%, and 93.44%, respectively. The proposed method provides the best performance in the Cleveland database.

As mentioned earlier, the sampling environment causes noise in the patients' data. In this regard, we have also evaluated the proposed method for the noise in the data, and the results of the accuracy evaluation criteria are shown in Figure 4. As an observer in Figure 4, the accuracy rate by PALM method is rather than other methods that is 91.87%.

**TABLE 4.** The proposed method on Cleveland dataset

| Method | Accuracy % | Sensitivity % | Specificity % |
|---|---|---|---|
| The proposed method (FISTA) | 94.52 | 87.62 | 84.76 |
| The proposed method (AMP) | 94.11 | 86.41 | 87.28 |
| The proposed method (DALM) | 95.08 | 87.19 | 91.04 |
| The proposed method (PALM) | **96.23** | **88.65** | **91.21** |

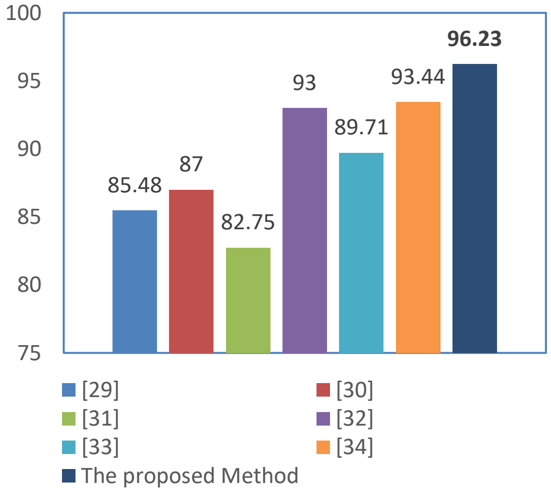


**Figure 3.** The Accuracy evaluated metrics for Cleveland dataset

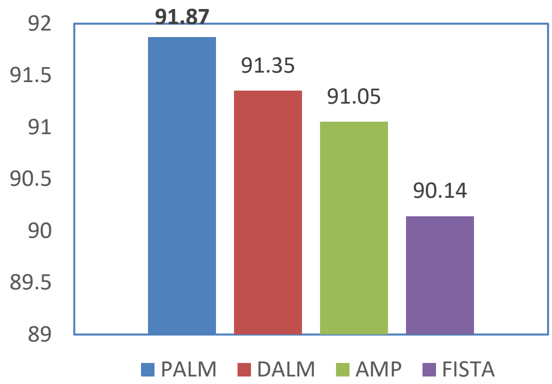


**Figure 4.** The Accuracy evaluated metrics of the proposed method with noisy data

## 5. CONCLUSION

In this paper, we proposed a novel method for predicting heart attacks based on sparse coding. Despite the extensive research conducted in this field, an efficient method has yet to be achieved. Most existing methods rely on extracting low-level features, which leads to low accuracy. The main difference between our proposed method and existing methods lies in the use of sparse classifiers. We modeled our method using FISTA, AMP, DALM, and PALM classifiers, which utilized sparse representation - a powerful method in this domain. Our proposed method for predicting heart attacks achieved better results compared to state-of-the-art methods. We evaluated the performance of our method on the Cleveland dataset, which is commonly used to assess prediction systems. As shown in the results section, our proposed method outperforms other methods in predicting heart attacks.

## 6. REFERENCES

1. Wang, M., Yao, X. and Chen, Y., "An imbalanced-data processing algorithm for the prediction of heart attack in stroke patients", ***IEEE Access***, Vol. 9, (2021), 25394-25404. doi: 10.1109/ACCESS.2021.3057693.
2. Mehmood, A., Iqbal, M., Mehmood, Z., Irtaza, A., Nawaz, M., Nazir, T. and Masood, M., "Prediction of heart disease using deep convolutional neural networks", ***Arabian Journal for Science and Engineering***, Vol. 46, No. 4, (2021), 3409-3422. https://doi.org/10.1007/s13369-020-05105-1
3. Gheitasi, A., Farsi, H. and Mohamadzadeh, S., "Estimation of hand skeletal postures by using deep convolutional neural networks", ***International Journal of Engineering, Transactions A: Basics***, Vol. 33, No. 4, (2020), 552-559. https://doi.org/10.5829/IJE.2020.33.04A.06
4. Mohan, S., Thirumalai, C. and Srivastava, G., "Effective heart disease prediction using hybrid machine learning techniques", ***IEEE Access***, Vol. 7, (2019), 81542-81554. doi: 10.1109/ACCESS.2019.2923707.
5. Verma, L., Srivastava, S. and Negi, P., "A hybrid data mining model to predict coronary artery disease cases using non-invasive

clinical data", ***Journal of Medical Systems***, Vol. 40, (2016), 1-7. https://doi.org/10.1007/s10916-016-0536-z

6. Mohamadzadeh, S., Pasban, S., Zeraatkar-Moghadam, J. and Shafiei, A.K., "Parkinson's disease detection by using feature selection and sparse representation", ***Journal of Medical and Biological Engineering***, Vol. 41, No. 4, (2021), 412-421. https://doi.org/10.1007/s40846-021-00626-y
7. Nazir, N., Rajab, S. and Sharma, V., "State-of-the-art artificial intelligence techniques in heart disease diagnosis", in Emerging Trends in Expert Applications and Security: Proceedings of ICETEAS 2018, Springer. (2019), 651-657.
8. Bahtiar, A., Dwilestari, G., Basysyar, F.M. and Suarna, N., "Data mining techniques with machine learning algorithm to predict patients of heart disease", in IOP Conference Series: Materials Science and Engineering, IOP Publishing. Vol. 1088, (2021), 012035.
9. Fooladi, S., Farsi, H. and Mohamadzadeh, S., "Segmenting the lesion area of brain tumor using convolutional neural networks and fuzzy k-means clustering", ***International Journal of Engineering, Transaction B: Applications***, Vol. 36, No. 8, (2023), 1556-1568. doi: 10.5829/IJE.2023.36.08B.15.
10. Moameri, S. and Samadinai, N., "Diagnosis of coronary artery disease via a novel fuzzy expert system optimized by cuckoo search", ***International Journal of Engineering, Transactions c: Aspects***, Vol. 31, No. 12, (2018), 2028-2036. doi: 10.5829/ije.2018.31.12c.06.
11. Wadhawan, S. and Maini, R., "Ebpso: Enhanced binary particle swarm optimization for cardiac disease classification with feature selection", ***Expert Systems***, Vol. 39, No. 8, (2022), e13002. https://doi.org/10.1111/exsy.13002
12. Ahmed, R., Bibi, M. and Syed, S., "Improving heart disease prediction accuracy using a hybrid machine learning approach: A comparative study of svm and knn algorithms", ***International Journal of Computations, Information and Manufacturing***, Vol. 3, No. 1, (2023), 49-54. https://doi.org/10.54489/ijcim.v3i1.223
13. Yuan, X., Wang, X., Han, J., Liu, J., Chen, H., Zhang, K. and Ye, Q., "A high accuracy integrated bagging-fuzzy-gbdt prediction algorithm for heart disease diagnosis", in 2019 IEEE/CIC International Conference on Communications in China (ICCC), IEEE. (2019), 467-471.
14. Choudhury, R.P. and Akbar, N., "Beyond diabetes: A relationship between cardiovascular outcomes and glycaemic index", ***Cardiovascular Research***, Vol. 117, No. 8, (2021), e97-e98. https://doi.org/10.1093/cvr/cvab162
15. Banu, U. and Vanjerkhede, K., "Hybrid feature extraction and infinite feature selection based diagnosis for cardiovascular disease related to smoking habit", ***International Journal on Advanced Science, Engineering & Information Technology***, Vol. 13, No. 2, (2023).
16. AL-Hawamleh, A.M., "Predictions of cybersecurity experts on future cyber-attacks and related cybersecurity measures", ***International Journal of Advanced Computer Science and Applications***, Vol. 14, No. 2, (2023). doi: 10.14569/IJACSA.2023.0140292.
17. Sharma, H. and Rizvi, M., "Prediction of heart disease using machine learning algorithms: A survey", ***International Journal on Recent and Innovation Trends in Computing and Communication***, Vol. 5, No. 8, (2017), 99-104.
18. Shahsavani, F., Nasiripour, R., Shakeri, R. and Gholamrezaee, A., "Arrhythmia detection based on the reduced features with k-svd sparse coding algorithm", ***Multimedia Tools and Applications***, Vol. 82, No. 8, (2023), 12337-12350. https://doi.org/10.1007/s11042-022-13894-w
19. Sezavar, A., Farsi, H., Mohamadzadeh, S. and Radeva, P., "A new person re-identification method by defining cnn-based feature extractor and sparse representation", ***Multimedia Tools and Applications***, (2023), 1-17. https://doi.org/10.1007/s11042-023-15718-x
20. Hamidi, H. and Daraee, A., "Analysis of pre-processing and post-processing methods and using data mining to diagnose heart diseases", ***International Journal of Engineering, Transactions A: Basics***, Vol. 29, No. 7, (2016), 921-930. doi: 10.5829/idosi.ije.2016.29.07a.06.
21. Barati, A., Farsi, H. and Mohamadzadeh, S., "Integration of the latent variable knowledge into deep image captioning with bayesian modeling", ***IET Image Processing***, Vol. 17, No. 7, (2023), 2256-2271. https://doi.org/10.1049/ipr2.12790
22. Nasiripour, R., Farsi, H. and Mohamadzadeh, S., "Saliency detection in eye gaze prediction by using deep learning", ***Journal of Soft Computing and Information Technology***, Vol. 9, No. 3, (2020), 93-108.
23. Saifudin, A., Nabillah, U. and Desyani, T., "Bagging technique to reduce misclassification in coronary heart disease prediction based on random forest", in Journal of Physics: Conference Series, IOP Publishing. Vol. 1477, (2020), 032009.
24. Bashir, S., Qamar, U. and Javed, M.Y., "An ensemble based decision support framework for intelligent heart disease diagnosis", in International conference on information society (i-Society 2014), IEEE. (2014), 259-264.
25. Santhanam, T. and Ephzibah, E., "Heart disease prediction using hybrid genetic fuzzy model", ***Indian Journal of Science and Technology***, Vol. 8, No. 9, (2015), 797. doi: 10.17485/ijst/2015/v8i9/52930.
26. Oad, K.K., DeZhi, X. and Butt, P.K., "A fuzzy rule based approach to predict risk level of heart", ***Global Journal of Computer Science and Technology (C) Software & Data Engineering***, Vol. 14, No. 3, (2014).
27. Ziasabounchi, N. and Askerzade, I., "Anfis based classification model for heart disease prediction", ***International Journal of Electrical & Computer Sciences IJECS-IJENS***, Vol. 14, No. 02, (2014), 7-12.
28. Kumar, A.S., "Diagnosis of heart disease using advanced fuzzy resolution mechanism", ***International Journal of Science and Applied Information Technology***, Vol. 2, No. 2, (2013), 22-30.
29. Latha, C.B.C. and Jeeva, S.C., "Improving the accuracy of prediction of heart disease risk based on ensemble classification techniques", ***Informatics in Medicine Unlocked***, Vol. 16, (2019), 100203. https://doi.org/10.1016/j.imu.2019.100203
30. Reddy, K.V.V., Elamvazuthi, I., Aziz, A.A., Paramasivam, S., Chua, H.N. and Pranavanand, S., "Heart disease risk prediction using machine learning classifiers with attribute evaluators", ***Applied Sciences***, Vol. 11, No. 18, (2021), 8352. https://doi.org/10.3390/app11188352
31. Ananey-Obiri, D. and Sarku, E., "Predicting the presence of heart diseases using comparative data mining and machine learning algorithms", ***International Journal of Computer Applications***, Vol. 176, No. 11, (2020), 17-21.
32. Rustam, F., Ishaq, A., Munir, K., Almutairi, M., Aslam, N. and Ashraf, I., "Incorporating cnn features for optimizing performance of ensemble classifier for cardiovascular disease prediction", ***Diagnostics***, Vol. 12, No. 6, (2022), 1474. https://doi.org/10.3390/diagnostics12061474
33. Ahamad, G.N., Shafiullah, Fatima, H., Imdadullah, Zakariya, S., Abbas, M., Alqahtani, M.S. and Usman, M., "Influence of optimal hyperparameters on the performance of machine learning algorithms for predicting heart disease", ***Processes***, Vol. 11, No. 3, (2023), 734. https://doi.org/10.3390/pr11030734
34. Chandrasekhar, N. and Peddakrishna, S., "Enhancing heart disease prediction accuracy through machine learning techniques and optimization", ***Processes***, Vol. 11, No. 4, (2023), 1210. https://doi.org/10.3390/pr11041210

## Persian Abstract

**چکیده**

هر ساله بیماری های قلبی جان افراد بی شماری را در سراسر جهان می گیرد. میزان بروز این نوع بیماری در نقاط مختلف جهان متفاوت است. این امر منجر به عدم تعادل بین سوابق موجود افراد سالم و افراد مبتلا به بیماری قلبی شده است. به عبارت دیگر، عدم تعادل بین داده های موجود در تشخیص بیماری قلبی وجود دارد. استفاده از روش‌های پیشگیرانه نقش مهمی در پیشگیری از این بیماری ها دارد. مشکل پیش بینی بیماری قلبی یک موضوع تقسیم بندی است. در فرآیند بخش‌بندی، یک مجموعه داده جدید و ناشناخته با شناسایی دسته‌ها و مفاهیم بین آنها پیش‌بینی می‌شود. در این مقاله فرآیند طبقه بندی داده ها با استفاده از کدهای تنک انجام می شود. روش پیشنهادی از دو بخش تشکیل شده است: بخش اول پیش پردازش و بخش دوم پردازش کد تنک است. بخش پیش پردازش شامل قرار دادن داده‌های ناموجود و در بخش پردازش بیماری قلبی پیش بینی می‌شود. کدگذاری پراکنده به دنبال ساده ترین نمایش برای سیگنالی است که سیگنال را با استفاده از حداقل تعداد سیگنال بازسازی می‌کند. این امر باعث افزایش سرعت محاسبات و کاهش فضای مورد نیاز برای ذخیره داده‌ها می‌شود. در روش پیشنهادی از نمایش تنک استفاده می‌شود که توسط طبقه‌بندی‌کننده FISTA، AMP، DALM و PALM مدل‌سازی شده است که در مقایسه با روش‌های پیشرفته، روشی موفق بوده است. روش پیشنهادی با استفاده از طبقه‌بندی‌کننده PALM به عنوان بهترین متریک در بین طبقه‌بندی‌کننده‌های دیگر با دقت ۹۶/۲۳ درصد است.