A Method for Estimating the Risk of a Heart Attack Using Imperialist Competitive Algorithm and Neural Networks

Hajar Vatani¹, Farsad Zamani², Mohamad Hosein Nadimi³
1-3-Faculty of Computer Engineering, Najafabad Branch, Islamic Azad University, Najafabad, Iran
Big Data Research Center, Najafabad Branch, Islamic Azad University, Najafabad, Iran
1-Email: hajar.vatani@gmail.com
2-Khorsgan Branch, Islamic Azad University, Isfahan, Iran

ABSTRACT:
The mortality from coronary heart disease is much higher than those from natural events. The World Health Organization (WHO) estimates that around 17 million deaths are due to heart and artery attacks. It should be noted that coronary heart disease is one of the main causes of mortality in advanced and developed countries such as Iran. Several methods have been proposed for the estimation and recognition of the risk of heart attacks, each of which has several advantages and disadvantages. Some disadvantages are as follows: low accuracy in the diagnosis of risk factors for coronary artery disease, time costs for selecting appropriate features, large number of diagnostic parameters, and the possibility of error. In this paper, we evaluate some of these methods and their advantages and disadvantages. The main scope is to review the methods and their advantages and disadvantages.

Keywords: risk estimation of heart attack, neural network, coronary artery attack,

Received: September 2016 Revised: Nov. 2016 Accepted: December 2016

1. INTRODUCTION
Today, with the advancement of science and the use of data mining and machine tools, it attempts to prevention of heart attack by estimating the risk and identifying the signs of cardiovascular attacks. Coronary artery disease is a condition in which the blood flow to the heart muscle is partially or completely blocked. This blockage is due to gradual accumulation of cholesterol and other fatty substances called Atheromas or atherosclerotic plaques (Zand, 2010). Considering the fundamental changes in lifestyle, these statistics are rapidly growing and the age of heart attacks and heart disease is decreasing. Therefore, new and reliable methods are required to provide statistics and determine the risk factors for coronary artery disease.

Recently, scientists have created a new approach to predict the risk of heart attacks so that people who are susceptible to heart attacks can be identified. By correcting the risk factors such as smoking, no physical activity, high blood pressure etc., coronary artery diseases are prevented. The main problem to estimate the risk of heart attack is the existence of high diagnostic parameters in this area. Most of the time, these parameters are not critical for risk estimation. For this reason, a reduction in the number of parameters of risk estimation is one of the major issues in this area. Many researches have been performed in this field. One of the solutions that is available to estimate the risk of heart attack is to employ the neural networks to extract useful information from the diastolic heart signals that show the blockage of the blood flow. Since the signals related to the sound of the heart are complex and very weak, the risk estimation in this method has low accuracy (Akay, 1992).

2. LITERATURE REVIEW
High diagnostic parameters are one of the common problems to diagnose coronary artery stenosis for specialists in this field. This information requires processing and analysis by relevant experts that is a time-consuming process. In all cases, the expert may not be sufficiently careful. Therefore, computer methods such as data mining techniques have been developed to help the experts to analyze the information in less time using the minimum number of related resources with more precision. Several articles are presented relevant researches correspond to the risk of the heart attack.

In 1989, Detreno et al. used a new probability algorithm to diagnose coronary artery disease obtained from the Cleveland algorithm and a Bayes algorithm. They concluded that the method is useful for the
patients with chest pain syndrome; otherwise the results are not acceptable (Detrano, 1989).

In 1997, Haddad et al. used the feasibility analysis method based on myocardial scan for the automatic diagnosis of coronary artery disease. Myocardial perfusion imaging is a non-invasive diagnostic method for the assessment of coronary artery disease. The performance of this method in the risk assessment is 48% (Haddad, 1997).

In 2008, Yan et al. used a coded genetic algorithm (binary) to detect features in heart diseases. They focused solely on the selection of features and there was no specific method to estimate the selected features to evaluate them. The selected features were related to five different diseases which this assessment method does not provide any assurance that the selected features are appropriate for the effective diagnosis of coronary artery stenosis. But the genetic algorithm can be useful by choosing the most effective features in this field. This is due to the variety of selection of a subset of the features and the use of the fit function to satisfy the selection of more appropriate features and higher diagnostic power (Van, 2008).

In 2010, Srinivas et al. used the methods based on data mining rules of Bayesian separators, neural networks, and decision tree for diagnosis. The odanb and ncc2 were used for pre-processing and making decision. The results of the two methods were tested on three data sets and the highest accuracy of 84.4% was obtained (Srinivas, 2010).

In 2010, Anbarasi et al. used a genetic algorithm to isolate the subset of the superior set of all features to diagnose heart disease. There are 13 characteristics in all used data and only 6 ones are needed for the optimal state. Three classifications were used to predict the disease diagnosis include Bayesian separator, cluster and decision tree. The highest accuracy corresponds to Bayesian separator method (Anbarasi, 2010).

In 2011, Atkov et al. investigated the development of a neural network for coronary artery stenosis using traditional complexity and genetic factors. The data set in the neural network includes medical, laboratory, functional, vascular and genetic data of 487 people. A model was developed that increased the accuracy from 64% to 94% by changing the type of neural network and the number of input factors.

In 2012, Bhatla et al. investigated the need to reduce the characteristics required to diagnose the disease, reduce the patient's necessary tests and also increase the accuracy of the system diagnosis. They used Bayesian separator, decision trees, and fuzzy logic techniques (Bhatla, 2012).

In 2012, Anooj used weighted fuzzy rules to support medical systems for the diagnosis of heart disease. This method includes two phases for prediction of the disease. In the first phase, weighted fuzzy rules are created and developed and the detection system was developed in the second phase. In the first phase, data mining techniques and feature selection and weighted rules were used and also fuzzy rules were created. For comparison, the data set of the site uci was used. The proposed method was compared to the neural network method (Anooj, 2012).

In 2012, Dangare et al. used hdps predictive system to predict the disease with a neural network using 15 parameters. They used the multilayer perceptron neural network with a backward learning algorithm to detect a pattern and predict a disease (Dangare, 2012).

In 2013, Alizadehsani et al. studied the subject by adding three characteristics and using a data set called Z with 303 samples and 54 characteristics and reached a precision of 94.88% (Alizadehsani, 2013).

In 2014, Myers et al. used the neural network to diagnose coronary artery disease in patients. A neural network involves multidimensional weighted data that is mathematically motivated to detect heart disease. The information of 2635 people was used in this research. A number of samples were randomly divided into two parts: training and test. The neural network improved the rate of diagnosis of arterial stenosis resulting in mortality (Myers, 2014).

In 2014, Florance et al. used the neural network and the Id3 decision tree to predict heart attacks and datasets with 6 characteristics for diagnostic process. The dataset used, is the heart attack data set provided in the UCI machine learning storage (Florence, 2014).

In 2015, El Bialy et al. pruned the characteristics tree of coronary artery disease using the rapid decision tree and C4.5 tree. The characteristics of the patients were selected and analyzed. The accuracy of this method was 48.75% (El-bialy, 2015).

In 2016, Alizadehsani et al. used SVM to predict coronary artery stenosis. The number of used features was 24. The accuracy of this method to predict was 83.50% (Alizadehsani, 2016).

In 2017, Torrecilha and colleagues used a radial basis neural network to predict the status of parasite of lymph glands in dogs and achieved an accuracy of 86.9% leads to an improvement of almost 12% compared to the previous work (Torrecilha, 2017).

In 2017, Sbarufatti et al. used a radial basis neural network to predict and estimate the state of lithium batteries and reached useful results (Sbarufatti, 2017).
<table>
<thead>
<tr>
<th>year</th>
<th>author</th>
<th>method</th>
<th>results</th>
<th>comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>1989</td>
<td>Detrano et al.</td>
<td>Cleveland algorithm and a Bayesian algorithm</td>
<td></td>
<td>Disadvantages: assessment is not well done. The exact method is not described.</td>
</tr>
<tr>
<td>1997</td>
<td>Haddad et al.</td>
<td>The feasibility analysis method of a reasoning system based on myocardial scan</td>
<td>48%</td>
<td>The need for high processing and scanning, as well as low accuracy, are the disadvantages of this method.</td>
</tr>
<tr>
<td>2008</td>
<td>Yan et al.</td>
<td>Genetic Algorithm</td>
<td></td>
<td>This article focuses solely on the selection of features and does not provide a specific method for evaluating the selected features for their evaluation.</td>
</tr>
<tr>
<td>2010</td>
<td>Srinivas et al.</td>
<td>Bayesian separators, Neural Networks, and Decision trees</td>
<td>84.14%</td>
<td>it has been used on three data sets.</td>
</tr>
<tr>
<td>2010</td>
<td>Anbarasi et al.</td>
<td>Genetic Algorithm + Bayesian separator, Cluster Classifier and Decision tree</td>
<td></td>
<td>Six features have been used. The best result has been the Bayes separator.</td>
</tr>
<tr>
<td>2011</td>
<td>Atkov et al.</td>
<td>neural network</td>
<td>94%</td>
<td>Neural network used is not stated.</td>
</tr>
<tr>
<td>2012</td>
<td>Batla et al.</td>
<td>Bayesian separator, Decision tree and fuzzy logic</td>
<td></td>
<td>Reduced feature has been done, but it is not exactly stated which method has been used.</td>
</tr>
<tr>
<td>2012</td>
<td>Anooj</td>
<td>Weighted Fuzzy Rules</td>
<td></td>
<td>The data taken from UCI is used.</td>
</tr>
<tr>
<td>2012</td>
<td>Dangare et al.</td>
<td>Multilayer Perceptron Neural Network with Backward Learning Algorithm</td>
<td></td>
<td>Has used 15 parameters.</td>
</tr>
<tr>
<td>2013</td>
<td>Alizadeh et al.</td>
<td>Adding three attributes and a neural network</td>
<td>94.08</td>
<td>The data set, called “Z-Alizadeh Sani”, has been used with 303 samples (individual) with 54 attributes</td>
</tr>
<tr>
<td>2014</td>
<td>Myers et al.</td>
<td>neural network</td>
<td></td>
<td>It is not precisely stated which algorithm has been used for the neural network.</td>
</tr>
<tr>
<td>2014</td>
<td>Florence et al.</td>
<td>Decision tree (ID3)</td>
<td></td>
<td>Six features have been used.</td>
</tr>
<tr>
<td>2015</td>
<td>El-bialy et al.</td>
<td>C4.5 Tree</td>
<td>75.48</td>
<td>Accuracy is relatively low.</td>
</tr>
<tr>
<td>2016</td>
<td>Alizadeh Sani et al.</td>
<td>Svm</td>
<td>83.50</td>
<td>24 features have been used.</td>
</tr>
</tbody>
</table>

**REFERENCES**


