A Comparison of Artificial Intelligence Methods on Determining Coronary Artery Disease

İsmail Babaoğlu[1], Ömer Kaan Baykan[1], Nazif Aygül[2], Kurtuluş Özdemir[3], and Mehmet Bayrak[4]

[1] Selçuk University, Department of Computer Engineering, 42075, Konya, Turkey
[2] Selçuk University, Department of Cardiology, Selçuklu Faculty of Medicine, 42075, Konya, Turkey
[3] Selçuk University, Department of Cardiology, Meram Faculty of Medicine, 42080, Konya, Turkey
[4] Mevlana University, Department of Electrical and Electronics Engineering, 42075, Konya, Turkey
{ibabaoglu,obaykan,naygul,kozdemir}@selcuk.edu.tr, mbayrak@mevlana.edu.tr

Abstract. The aim of this study is to show a comparison of multi-layered perceptron neural network (MLPNN) and support vector machine (SVM) on determination of coronary artery disease existence upon exercise stress testing (EST) data. EST and coronary angiography were performed on 480 patients with acquiring 23 verifying features from each. The robustness of the proposed methods is examined using classification accuracy, k-fold cross-validation method and Cohen's kappa coefficient. The obtained classification accuracies are approximately 78% and 79% for MLPNN and SVM respectively. Both MLPNN and SVM methods are rather satisfactory than human-based method looking to Cohen's kappa coefficients. Besides, SVM is slightly better than MLPNN when looking to the diagnostic accuracy, average of sensitivity and specificity, and also Cohen's kappa coefficient.

Keywords: Exercise stress testing, coronary artery disease, support vector machine, artificial neural networks.

1 Introduction

Coronary artery disease (CAD) is the primary cause of mortality and morbidity in both developed and developing countries and its incidence is increasing rapidly worldwide. More than a half of the deaths in the United States are being occurred depending on cardiovascular diseases, many of which are consist of CAD [1]. Coronary angiography (CAG) is the gold standard diagnostic tool in the diagnosis of suspected CAD. However; because it is invasive and expensive method, it is not suggested as the first choice method. Exercise stress testing (EST) is a non-invasive, relatively cheap, reproducible and safe method; therefore, it can be used as one of the first choice non-invasive diagnostic tools in the diagnosis of suspected CAD. Nonetheless, the relatively low sensitivity and specificity of EST for diagnosing CAD, has led to limit its clinical usage [2,3].

B. Papasratorn et al. (Eds.): IAIT 2010, CCIS 114, pp. 18–26, 2010.

Recently, the artificial intelligence (AI), out of than invasive and non-invasive diagnostic tools, becomes the promising method in the diagnosis of heart diseases. Least squares support vector machine and backpropagation artificial neural network are used to classify the extracted features which are obtained from Doppler signals of the heart valve [4]. Electrocardiography (ECG) signals are classified to 10 different arrhythmias using a new fuzzy clustering neural network architecture for early diagnosis [5]. Fuzzy weighted pre-processing and artificial immune recognition system are used to classify ECG arrhythmia as a new method for the medical diagnosis [6]. An expert diagnosis system is presented for interpretation of the Doppler signals of the heart valve diseases. To make the feature extraction from Doppler signals on the time-frequency domain, wavelet transforms and short times Fourier transform methods are used. Wavelet entropy method and back-propagation neural network are employed to classify the extracted features [7]. An adaptive neuro-fuzzy network is developed to classify heart abnormalities in 10 different cardiac states and classification accuracy is more than 94% [8]. For classification of carotid artery Doppler signals in the early phase of atherosclerosis, principle component analysis and fuzzy c-means clustering methods with complex-valued artificial neural network are used [9].

More specific usage of AI methods for the determination of CAD diagnosis could also be found in literature. Lapuerta et al. investigated artificial neural network (ANN) performance to predict the occurrence of CAD based on information from serum lipid profile [10]. Süt et al. examined the diagnostic performances of multilayer perceptron neural networks (MLPNNs) for predicting coronary artery disease and compared them with different types of artificial neural network methods, namely recurrent neural networks as well as two statistical methods (quadratic discriminant analysis and logistic regression) [11]. Scott et al. searched that ANN's can determine the presence or absence of CAD with predictive accuracy equal to that of standard (expert reader clinical interpretation using imaging) clinical and stress test data [12]. Kurt et al compare performances of logistic regression, classification and regression tree, multi-layer perceptron, radial basis function and self-organizing feature maps in order to predict the presence of CAD by using demographic and medical data [13]. Zhidong proposed noninvasive diagnosis method of coronary artery disease based on the instantaneous frequency estimation of diastolic murmurs and support vector machine (SVM) classifier [14]. And also many applications carried out for diagnosing coronary artery stenosis [15-17].

In this study, it is aimed to explore the MLPNN and SVM significance on determination of CAD existence upon EST data. A proper comparison is also performed for both the MLPNN and SVM methods.

2 Materials and Method

2.1 Data Acquisition

Four hundred and eighty patients who underwent EST and CAG were included to the study. Baseline demographic characteristics, rest and peak exercise heart rate, blood pressure, exercise time were recorded. The EST results were evaluated by 2 experienced cardiologists (human-based method). ST segment depression and elevation

occurred 60 ms after the J point were recorded at each derivation in peak exercise. According to human-based method, an exercise test result was considered positive if there was ≥1 mm horizontal or downsloping ST depression or ST elevation in two contiguous leads. Within the first month following the EST, CAG was performed to all patients, and the angiographic images were evaluated by 2 skilled cardiologists. Presence of ≥50% narrowing in left main coronary artery, or ≥70% narrowing in other major epicardial coronary arteries indicated significant CAD. Patients with bundle branch blocks (right or, left bundle branch block), pre-excitation syndromes, atrial fibrillation, left ventricular hypertrophy and taking the digoxin were excluded from the study.

2.2 Multi Layered Perceptrons

A neural network generally consist of a set of neurons, a pattern of connectivity, a propagation rule, an activation rule, a transfer function and a learning rule [18]. Artificial neural networks can be design many architectures and structures. Multi-layered Perceptrons (MLPs) are simple and most frequently used ANN architectures [19]. MLP model is a feed-forward network and as shown Fig. 1 which consists of one input layer, one or more hidden layers and one output layer. Layers can have different number of neurons. The input signals x_i are dispatched to neurons in the hidden layer by using input layer neurons. Each neuron at hidden layers or output layer receives a weighted sum from all neurons in the previous layer [20]. Outputs of the hidden layer neurons together with the output layer neurons are calculated with defined transfer functions (f). Neuron outputs are calculated as:

$$y_j = f\left(\sum w_{ji} x_i\right) \tag{1}$$

here f can be a sigmoidal or a hyperbolic tangent function, w_{ji} is the weight. For setting the weights of ANN, many learning algorithms can be adopted. In this study, MLP neural networks are trained with backpropagation (BP) learning algorithm. In BP algorithm, error (E) calculated as the sum of squared differences between the desired and actual values of the output neurons which propagate through the layers of neurons to update the weights. E is defined in Eq.(2) below as;

$$E = \frac{1}{2}\sum_j \left(y_{dj} - y_j\right)^2 \tag{2}$$

where, y_j is the actual value of output neuron and y_{dj} is the desired value of that neuron [20,21].

2.3 Support Vector Machines

Support vector machine is proposed by Vapnik (1995) based on structural risk minimization (SRM) principle. SVM as a new machine learning technique is used for many purpose such as classification, recognition, regression [22]. SVMs can classify

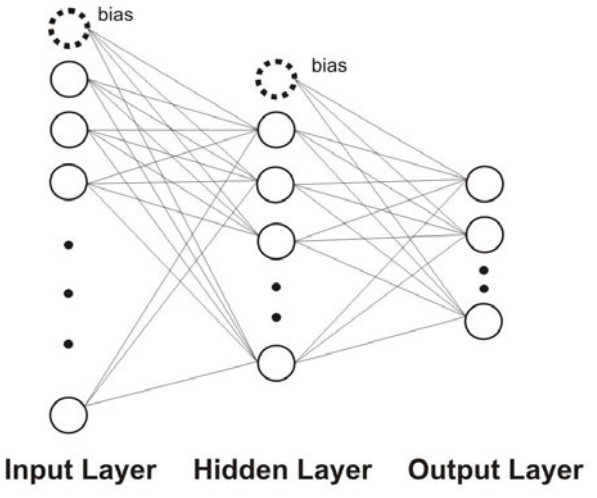

Fig. 1. A Multi Layered Perceptron

the input samples when the classes are linearly separable. If input space is nonlinear, SVMs need mapping N dimensional input space to a high-dimensional feature space using several kernel functions such as polynomial kernel, dot product, radial basis function kernel [23]. SVMs search an optimal separating hyper-plane that maximizes the margin between itself and the nearest training examples in the new high-dimensional space. A separating hyper-plane is a linear function and it can divide the training data into two classes. The training examples that are closest to the hyper-plane are called as support vectors. SVMs can be used both binary classifications and multi-classes problems [24-27].

Commonly used kernel functions can be briefly given as:

Dot product kernels : $K\left(x, x^{'}\right) = x.x^{'}$

Polynomial kernels : $K\left(x, x^{'}\right) = \left(x.x^{'} + 1\right)^{d}$; where d is the degree of kernel.

RBF kernels : $K\left(x, x^{'}\right) = \exp\left(-\left\|x - x^{'}\right\|^{2} / \sigma^{2}\right)$; where σ is positive real number.

In this study, prediction of CAD problem is considered a binary classification problem and SVMs are also applied to solve it. Four kernel functions namely linear, polynomial, radial basis, and sigmoid are tested and the best one is adopted.

2.4 K-Fold Cross Validation

To make the test results more meaningful and benefitable, *k*-fold cross validation method which minimizes the bias association with the random sampling of the training can be used [28,29]. Whole data is randomly divided to *k* mutually exclusive and approximately equal size subsets. Training and test processes are performed *k* times.

In each case, one of the folds is taken as test data and the remaining folds are added to form training data. So *k* times different test results are obtained. The average of these results gives the test accuracy of the algorithm.

2.5 Screening Test

Performance evaluations of proposed methods are implemented using screening test in point of sensitivity, specificity, positive and negative predictivity and accuracy. In this test TP, FP, TN and FN as described as follow [30];

True Positives (TP) : Those who test positive for a condition and are positive
False Positives (FP) : Those who test positive, but are negative.
True Negatives (TN) : Those who test negative and are negative.
False Negatives (FN) : Those who test negative but are positive.

Positive Predictive Value (PPV): Percent of patients with positive test having disease. PPV is calculated as;

$$PPV = \frac{TP}{TP + FP} \tag{3}$$

Negative Predictive Value (NPV): Percent of patients with negative test that do not have disease. NPV is calculated as;

$$NPV = \frac{TN}{TN + FN} \tag{4}$$

Sensitivity (SEN) and Specificity (SPE): Independent of disease prevalence in the community. SEN, SPE and Diagnostic Accuracy (ACC) are calculated as [6, 30];

$$SEN = \frac{TP}{TP + FN} \tag{5}$$

$$SPE = \frac{TN}{FP + TN} \tag{6}$$

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \tag{7}$$

2.6 Cohen's Kappa Coefficient

Cohen's kappa coefficient (κ) is a statistical measure of inter-rater reliability and measures the agreement between the evaluations of two raters when both are rating the same object. A kappa coefficient of 1 indicates the perfect agreement. A kappa coefficient of 0 indicates the agreement is no better than chance. However by no means universally accepted the ranges of κ mentioned on the previous studies are $\kappa < 0$ no agreement, $0 \leq \kappa \leq 0.20$ very low agreement, $0.21 \leq \kappa \leq 0.40$ low agreement, $0.41 \leq \kappa \leq 0.60$ moderate agreement, $0.61 \leq \kappa \leq 0.80$ full agreement, $0.81 \leq \kappa \leq 1.00$ almost perfect agreement [31,32].

In our study, we use Cohen's kappa coefficient to compare specifically the three methods (human-based, SVM and MLPNN) in order to outline the rate of agreement with CAG results. Cohen's kappa coefficients together with the test results are shown in Table 3.

3 Results

In this study, on determination of coronary artery disease existence upon EST data, 3 different methods are used. Data set consist of 480 patient and 23 verifying features. 346 of these patients have CAD and rest of 134 is healthy.

Each feature used in the study is normalized into [-1 1] range. Training and test implementations are applied by using k-fold cross validation method being k as 5. In this situation, for each fold, size of training data set is 384*23 and size of test data set is 96*23.

Human-based method can be expressed as the evaluation of the EST results by 2 experienced cardiologists. On the other hand, different MLP and SVM architectures are trained and tested in the study. Grid search algorithm is used on these training and test processes for both MLPNN and SVM techniques [33]. In grid search algorithm, the values of each parameter across the specified search range is tried to find optimum ones using geometric steps. The value ranges of the parameters used by the grid search algorithm for both different MLPNN and SVM models and are given in Table 1 and Table 2 respectively. The highest value of the difference between the diagnostic accuracy and the training error is selected as the optimum classification model. In other words, MLPNN and SVM models are attempted to minimize the training error while maximizing the diagnostic accuracy.

According to the MLPNN evaluation, tangent sigmoid transfer function is used at the hidden layer and logarithmic sigmoid transfer function is used at the output layer in the best result. For this network, hidden layer has got 75 neurons. Average test accuracy and training errors are found to be 78.13% and 1.86%.

In SVM evaluation, the best results are obtained by using radial basis kernel. In this best SVM model, γ value is 0.4 and C value is 10. Average test accuracy and training errors are found to be 79.17% and 1.30%.

Results obtained with all 3 methods are given in Table 3. The results obtained from different MLP and SVM methods are examined to select the optimum models for each method. While selecting the optimum models, the primary aim is to select higher

Table 1. MLPNN parameters search grid

Prm	SVal	EVal	Int
lr	0.1	0.5	0.1
mc	0.5	0.9	0.1
neuron	5	200	5
iteration	1000	10000	1000

Prm, Name of the parameter; SVal, Start value of the parameter; EVal, End value of the parameter; Int, Interval of the parameter between start value and end value; lr, Learning rate; mc, Momentum coefficient; neuron, Number of the neurons in the hidden layer; iteration, Training iterations size.

Table 2. SVM parameters search grid

Kernel	Prm	SVal	EVal	Int
Linear	c	1	10	1
Polynomial	d	1	5	1
Polynomial	γ	0	1	0.1
Polynomial	r	0	5	1
Polynomial	c	1	100	1
RBF	γ	0	5	0.1
RBF	c	1	100	1

Kernel, Kernel type used in SVM; Prm, Name of the parameter; SVal, Start value of the parameter; EVal, End value of the parameter; Int, Interval of the parameter between start value and end value; Linear, Linear type kernel; Polynomial, Polynomial type kernel; RBF, Radial basis type kernel; c, Parameter of linear kernel; d, γ, r and c, Parameters of polynomial kernel; γ and c, Parameters of RBF kernel.

diagnostic accuracy and sensitivity, the secondary aim is to select the lower training error. After optimum model selection for both methods, Cohen's kappa coefficients are calculated. Obtained values by human-based method were consistent with that of obtained in the literature. Nonetheless, sensitivity (mean 72%, range 45–92%) and specificity (mean 77%, range 17–92%) values are relatively variable in the diagnosis of CAD using EST by human-based evaluation [1-3]. Disparity in positive criterion of EST, prevalence/intensity of CAD and interobserver variability could be effective factors in this variability. In our study, a human-based interpretation of EST showed relatively high sensitivity of 78% but a poor specificity of 43% and poor agreement with CAG results (κ 0.21). In contrast, both SVM and MLPNN methods provided a better sensitivity, specificity and diagnostic accuracy and as expected NPV and PPV, compared with human-based method. It is also observed that both the two methods agreement with coronary angiographic results was rather satisfactory. Besides, SVM is slightly better than MLPNN when looking to the diagnostic accuracy, average of sensitivity and specificity, and also kappa coefficient. In terms of explication, well education is needed to improve the usage of EST. Therefore, the supporting of human-based method with AI methods will provide standardization, and remove disparity of personal explication.

Table 3. Test Results

Method	ACC (%)	SEN (%)	SPE (%)	PPV (%)	NPV (%)	κ
Human-based	68.33	78.03	43.28	78.03	43.28	0.213
SVM	79.17	84.76	63.98	86.70	59.69	0.473
MLPNN	78.13	86.35	59.64	83.42	60.23	0.411

ACC, Diagnostic accuracy; SEN, Sensitivity; SPE, Specificity; PPV, Positive predictive value; NPV, Negative predictive value; SVM, Support vector machine; MLPNN, Multi layered perceptron neural network.; κ, Cohen's kappa coefficient.

4 Discussion

The findings of our study suggest that assessment of EST using AI methods increase sensitivity, specificity and diagnostic accuracy compared to the assessment of EST by

human-based method. Improvement in the sensitivity and specificity of EST for diagnosis of CAD may improve the PPV and NPV of the test in population with suspected CAD. This situation enhances the usage of EST as a more reliable test which is still commonly used in clinical practice.

Acknowledgement

This study has been supported by Scientific Research Project of Selçuk University.

References

1. Gibbons, R., Balady, G., Bricker, J.T., et al.: ACC/AHA Guideline update for exercise testing: summary article. A report of the American college of cardiology/American heart association task force on practice guidelines (Committee to update the 1997 exercise testing guidelines). J. Am. Coll. Cardiol. 40(8), 1531–1540 (2002)
2. San Roman, J.A., Vilacosta, I., Castillo, J.A., Rollan, M.J., Hernandez, M., Peral, V., Garcimartin, I., del Mar de la Torre, M., Fernandez-Aviles, F.: Selection of the optimal stress test for the diagnosis of coronary artery disease. Heart 80(4), 370–376 (1998)
3. Thom, T., Haase, N., Rosamond, W., Howard, V.J., Rumsfeld, J., Manolio, T., Zheng, Z.J., Flegal, K., O'Donnell, C., Kittner, S., Lloyd-Jones, D., Goff Jr., D.C., Hong, Y., Adams, R., Friday, G., Furie, K., Gorelick, P., Kissela, B., Marler, J., Meigs, J., Roger, V., Sidney, S., Sorlie, P., Steinberger, J., Wasserthiel-Smoller, S., Wilson, M., Wolf, P.: Heart disease and stroke statistics–2006 update: A report from the American heart association statistics committee and Stroke statistics subcommittee. Circulation 113, e85–e151 (2006)
4. Comak, E., Arslan, A., Turkoglu, I.: A decision support system based on support vector machines for diagnosis of the heart valve diseases. Computers in Biology and Medicine 37, 21–27 (2007)
5. Ozbay, Y., Ceylan, R., Karlik, B.: A fuzzy clustering neural network architecture for classification of ECG arrhythmias. Computers in Biology and Medicine 36(4), 376–388 (2006)
6. Polat, K., Sahan, S., Gunes, S.: A new method to medical diagnosis: artificial immune recognition system (AIRS) with fuzzy weighted pre-processing and application to ECG arrhythmia. Expert Systems with Applications 31(2), 264–269 (2006)
7. Turkoglu, I., Arslan, A., Ilkay, E.: An expert system for diagnosis of the heart valve diseases. Expert Systems with Applications 23(3), 229–236 (2002)
8. Kannathal, N., Lim, C.M., Rajendra Acharya, U., Sadasivan, P.K.: Cardiac state diagnosis using adaptive neuro-fuzzy technique. Medical Engineering & Physics 28(8), 809–815 (2006)
9. Ceylan, M., Ceylan, R., Dirgenali, F., Kara, S., Ozbay, Y.: Classification of carotid artery doppler signals in the early phase of atherosclerosis using complex-valued artificial neural network. Computers in Biology and Medicine 37, 28–36 (2007)
10. Lapuerta, P., Azen, S.P., Labree, L.: Use of neural networks in predicting the risk of coronary artery disease. Computers and Biomedical Research 28(1), 38–52 (1995)
11. Süt, N., Şenocak, M.: Assessment of the performances of multilayer perceptron neural networks in comparison with recurrent neural networks and two statistical methods for diagnosing coronary artery disease. Expert Systems 24(3), 131–142 (2007)
12. Scott, J.A., Aziz, K., Yasuda, T., Gewirtz, H.: Integration of clinical and imaging data to predict the presence of coronary artery disease using neural networks. Journal of Nuclear Cardiology 11(4), 26 (2004)

13. Kurt, I., Ture, M., Kurum, A.T.: Comparing performances of logistic regression, classification and regression tree, and neural networks for predicting coronary artery disease. Expert Systems with Applications 34, 366–374 (2008)
14. Zhidong, Z.: Noninvasive diagnosis of coronary artery disease based on instantaneous frequency of diastolic murmurs and SVM. In: Proceedings of the 2005 IEEE Engineering in Medicine and Biology 27th Annual Conference, Shanghai, China, pp. 5651–5654 (2005)
15. Cios, K.J., Goodenday, L.S.: Hybrid intelligence system for diagnosing coronary stenosis. Combining fuzzy generalized operators with decision rules generated by machine learning algorithms. IEEE Engineering in Medicine and Biology Magazine 13(5), 723–729 (1994)
16. Mobley, B.A., Moore, W.E., Schechter, E., Eichner, J.E., McKee, P.A.: Neural network predictions of significant coronary artery stenosis in women. Computational Intelligence Processing in Medical Diagnosis, 261–290 (2002)
17. Mobley, B.A., Moore, W.E., Schechter, E., Eichner, J.E., McKee, P.A.: Neural network predictions of significant coronary artery stenosis in men. Artificial Intelligence in Medicine 34, 151–161 (2005)
18. Rumelhart, D.E., McClelland, J.L.: PDP Research Group: Paralel distributed processing: exploration in the microstructure of cognition, vol. 1, p. 950. MIT Press, MA (1989)
19. Sagiroglu, S., Colak, I., Bayindir, R.: Power factor correction technique based on artificial neural networks. Energy Conversion and Management 47(18-19), 3204–3215 (2006)
20. Serhatlioglu, S., Hardalac, F., Guler, I.: Classification of transcranial doppler signals using artificial neural network. Journal of Medical Systems 27, 205–214 (2003)
21. Pao, Y.H.: Adaptive pattern recognition and neural networks. Addison-Wesley, Reading (1989)
22. Vapnik, V.: The nature of statistical learning theory. Springer, New York (1995)
23. Gunn, S.R.: Support vector machines for classification and regression. ISIS, Technical Report, University of Southampton, Department of Electrical and Computer Science (1998)
24. Chen, K.Y., Wang, C.H.: A hybrid SARIMA and support vector machines in forecasting the production values of the machinery industry in Taiwan. Expert Systems with Applications 32(1), 254–264 (2007)
25. Kulkarni, A., Jayaraman, V.K., Kulkarni, B.D.: Support vector classification with parameter tuning assisted by agent-based technique. Computers and Chemical Engineering 28, 311–318 (2004)
26. Seo, K.K.: An application of one-class support vector machines in content-based image retrieval. Expert Systems with Applications 33(2), 491–498 (2007)
27. Takeuchi, K., Collier, N.: Bio-medical entity extraction using support vector machines. Artificial Intelligence in Medicine 33(2), 125–137 (2003)
28. An, S., Liu, W., Venkatesh, S.: Fast cross-validation algorithms for least squares support vector machine and kernel ridge regression. Pattern Recognition 40, 2154–2162 (2007)
29. Shao, J.: Linear model selection by cross-validation. Journal of American Statistical Association 88, 486–494 (1993)
30. Nielsen, C., Lang, R.S.: Principles of screening. The Medical Clinics of North America 83(6), 1323–1337 (1999)
31. Agresti, A.: An introduction to categorical data analysis, 2nd edn. Wiley, Chichester (2007)
32. Sprent, P., Smeeton, N.C.: Applied non-parametric statistical methods, 3rd edn. Chapman & Hall, Boca Raton (1999)
33. Cormen, T.H., Leiserson, C.E., Rivest, R.L., Stein, C.: Introduction to algorithms, 2nd edn. MIT Press, Cambridge (2001)