

One-Year Survival Prediction of Myocardial Infarction

¹Abdulkader Helwan, ²Dilber Uzun Ozsahin

^{1,2}Department of Biomedical Engineering, Near East University, Near East Boulevard, TRNC, Nicosia, 99138 Cyprus

³Rahib Abiyev, ⁴John Bush

^{3,4}Department of Computer Engineering, Near East University, Near East Boulevard, TRNC, Nicosia, 99138 Cyprus

Abstract—Myocardial infarction is still one of the leading causes of death and morbidity. The early prediction of such disease can prevent or reduce the development of it. Machine learning can be an efficient tool for predicting such diseases. Many people have suffered myocardial infarction in the past. Some of those have survived and others were dead after a period of time. A machine learning system can learn from the past data of those patients to be capable of predicting the one-year survival or death of patients with myocardial infarction. The survival at one year, death at one year, survival period, in addition to some clinical data of patients who have suffered myocardial infarction can be used to train an intelligent system to predict the one-year survival or death of current myocardial infarction patients. This paper introduces the use of two neural networks: Feedforward neural network that uses backpropagation learning algorithm (BPNN) and radial basis function networks (RBFN) that were trained on past data of patients who suffered myocardial infarction to be capable of generalizing the one-year survival or death of new patients. Experimentally, both networks were tested on 64 instances and showed a good generalization capability in predicting the correct diagnosis of the patients. However, the radial basis function network outperformed the backpropagation network in performing this prediction task.

Keywords—Machine learning; myocardial infarction; backpropagation; radial basis function network; generalization; one-year survival prediction

I. INTRODUCTION

Myocardial is simply described as thick muscular wall of the heart whereas, infarction is simply referred to dead portion of tissue caused by loss of blood supply; a localized necrosis. Hence, myocardial infarction describes the dead portion of thick muscular wall of the heart induced by a loss of blood supply. In cardiovascular system, the heart is the main organ which also includes different types of blood vessels. Coronary arteries are some of the most important vessels in cardiovascular system. These arteries take oxygen-rich blood to the heart as well as all other organs in the body. Gradual buildup of plaque blocks or narrows the arteries as a result; the blood flowing to the heart decreases significantly or stops completely. This may lead to myocardial infarction.

Irreversible necrosis in acute myocardial infarction of heart muscle secondary to prolonged ischemia is the most deadly presentation of coronary arteries disease. Usually, imbalance between oxygen supply and demand leads to infarction which is most often caused by thrombus formation and plaque rupture in a coronary vessel, leading to an acute reduction in the blood

supply to a portion of the myocardium [1]. Myocardial infarction may lead to diastolic or systolic dysfunction and may increase the susceptibility to arrhythmias and other complications such as ischemic, mechanical, embolic and inflammatory disturbances [2]. Because of the high cost of care, effective drugs and treatments, the prevention of myocardial infarction is a desirable goal. To predict the likelihood of myocardial infarction many factors such as laboratory data, history and physical examination findings are used. Some of the results have been hopeful but none of these studies were successful in accurately predicting the likelihood of myocardial infarction [3], [4].

At some point in the past, the myocardial infarction patients suffered heart attacks. Some patients survived and are still alive but some died since they could not withstand the attack. Researchers that studied this problem addressed the prediction from the other variables whether or not the patient will survive or not for at least one-year.

The proposed research is targeted to investigating the use of backpropagation neural network (BPNN) and radial basis function network (RBFN) in learning the past clinical and historical data of patients who had myocardial infarction and use them to generalize or predict the one-year survival or death of new patients. Acknowledging the importance of the prediction of survivals after myocardial infarction as well as the lack of sufficient studies designed to test methods of prediction, the implementation of this research work prompts to compare the capability of two types of neural networks to perform this prediction task, i.e. predict the one-year survival of patients who have myocardial infarction.

Both networks are trained using data of some patients who suffered myocardial infarction. These data include some historical attributes that show if patients have survived or not for one year. Other attributes are correspondent to some medical variables that indicate some abnormalities in patient vital conditions. This helps the networks to learn historical and medical data of both types of patients, i.e., the ones who survived at one year and those who died before or at one year. Both networks are evaluated and showed a good capability in predicting the one-year survival or death of myocardial infarction patients when tested on unseen data.

II. THE PROPOSED METHODOLOGY

Artificial neural networks have opened new horizons in learning about the natural history of diseases and predicting

cardiac disease. In this work, we propose the use of two types of artificial neural networks to predict the one-year survival of patients who suffer myocardial infarction. All the patients suffered heart attacks at some point in the past. Some are still alive and some are not. Therefore, this work is to develop an intelligent system that will be trained using a database of many patients who have had myocardial infarction [5]. The database consists of 11 parameters as inputs such as survival period after infarction, a measure of contractility around the heart, etc. The parameters used in the dataset are shown in Table 1. As shown in Table 1 some parameters are historical data such as the survival and still-alive variables, when taken together, indicate whether a patient survived for at least one year following the myocardial infarction. The other attributes are clinical data correspondent to some medical variables which help the network to find differences in the two classes.

For output, one attribute is used to show if the patient has survived for one year or not. Upon training, the system will be capable of predicting whether the patient is going to survive for one year or will die before as shown in Fig. 1.

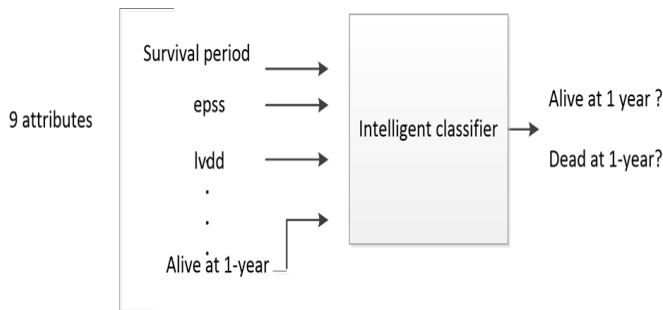


Fig. 1. Proposed prediction system.

A. Database Description

The database consists of 131 cases of different patients obtained for the public online database [5]. The data consist of 13 attributes that indicate the patient’s conditions after having myocardial infarction; however, three of them were discarded since they make no sense and don’t contribute to the network learning because they have no relation or indication to the myocardial disease. The three attributes are the patient name, group and other derivative variable that has no usage or benefits. Note that it was recommended to discard these three parameters by the database developers. Those parameters are described in Table 1. As seen in Table 1, some attributes are medical variables such as the occurrence of myocardial infarction pericardial-effusion which represents the fluid found around the myocardium, in addition to some other clinical parameters such as the wall-motion-score and wall-motion-index, etc.

Some other attributes such as survival and still-alive can be considered as historical data that show if the patients have survived or not at one year. These two parameters in addition to the other clinical attributes can be enough for the prediction of the one-year survival of myocardial infarction patients.

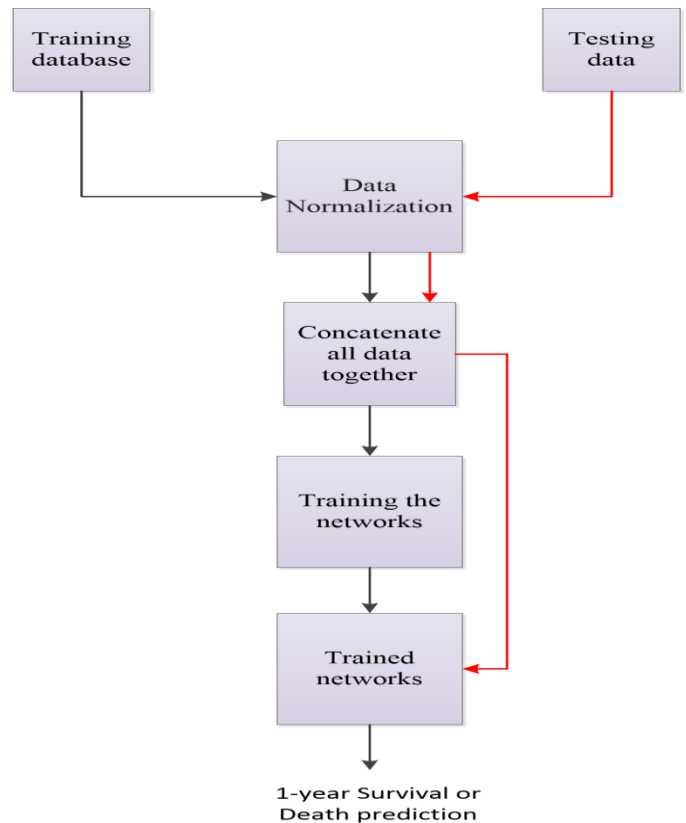


Fig. 2. Flowchart of the developed network system.

Fig. 2 represents a flowchart that illustrates the proposed system for the prediction of one-year survival myocardial infarction.

Input coding

Among the 10 attributes used, 9 are considered as inputs and one as output which is the variable that shows if the patient has survived or not at one year. Basically, the pre-processing stage of the data includes the normalization of input features or attributes into the range of 0 to 1 so that they be suitable to be fed into the networks. Equation (1) shows how the data were normalized. It should be noted that this equation is used in case where the data are all positive; which is our case here.

$$ND = \frac{\text{Parameter value} - \text{min value}}{\text{Range of attribute value}} \quad (1)$$

Where, ND represents the normalized data.

Output coding

One of the data attributes is used as output since it shows if the patient has survived or not after having myocardial infarction. Therefore, the output was coded such that two neurons are used. Thus, one of these neurons switches on to one of the two classes; survive or not at one year (Table 2).

TABLE. I. DATABASE DESCRIPTION

Attribute	Attribute Description
Survival	The number of months patient survived (has survived, if patient is still alive). Because all the patients had their heart attacks at different times, it is possible that some patients have survived less than one year but they are still alive.
still-alive	A binary variable. 0=dead at end of survival period, 1 means still alive
age-at-heart-attack	Age in years when heart attack occurred
Pericardial effusion	Binary. Pericardial effusion is fluid around the heart. 0=no fluid, 1=fluid
fractional-shortening	A measure of contractility around the heart lower numbers are increasingly abnormal
Epsc	E-point septal separation, another measure of contractility. Larger numbers are increasingly abnormal.
Ivdd	Left ventricular end-diastolic dimension. This is a measure of the size of the heart at end-diastole. Large hearts tend to be sick hearts.
wall-motion-score	A measure of how the segments of the left ventricle are moving
wall-motion-index	Equals wall-motion-score divided by number of segments seen. Usually 12-13 segments are seen in an echocardiogram.
alive-at-1	Boolean-valued. Derived from the first two attributes. 0 means patient was either dead after 1 year or had been followed for less than 1 year. 1 means patient was alive at 1 year.

TABLE. II. OUTPUT CODING AND CLASSES

Output classes	Coding
Dead at 1-year	[1 0]
Alive at 1-year	[0 1]

III. BPNN TRAINING

The back propagation algorithm is a sort of supervised learning scheme. The neural network that uses such a learning algorithm is referred to as a back propagation neural network. BP algorithm is one of the most popular ANN algorithms. Rojas, (1996) in [6] claimed that BP algorithm could be packed up to four major steps. Once the weights chosen randomly, compute of necessary corrections are done by back propagation algorithm. The algorithm can be expressed in the following four steps [7]:

- 1) Computation of feed-forward.
- 2) Back propagation to the output layer.
- 3) Propagation to the hidden layer.
- 4) Weight updates.

While the function error value may become small enough, the algorithm is stopped. It considers being the basic formula for BP algorithm. With the variations proposed by other scientists, Rojas definition seems to be fairly accurate and simple to follow. The last step, weight updates is happening throughout the algorithm [8]. Equations (2) & (3) are used to

update the output-hidden layer weights and input-hidden layer weights, respectively.

$$W_{jh} \text{ (new)} = W_{jh} \text{ (old)} + \eta \Delta_j O_h + \alpha [\delta W_{jh} \text{ (old)}] \quad (2)$$

$$W_{hi} \text{ (new)} = W_{hi} \text{ (old)} + \eta \Delta h O_i + \alpha [\delta W_{hi} \text{ (old)}] \quad (3)$$

Where, $\delta W_{jh}(\text{old})$ represents the previous weight change, and η is the learning rate. $\Delta_j O_h$ stands for the error signal for output layer neurons, W_{jh} represent the weights that feed the output layer, and W_{hi} are weights that feed the hidden layer [9].

Furthermore, η stands for the learning rate of the network which should has a value of range between 0 and 1. The α stands for the momentum rate which is added to increase the convergence speed of the network [10]. The backpropagation neural network is trained on 67 of 131 instances of patients who have had myocardial infarction. The number of parameters used in the database is 9; therefore the number of inputs neurons in the input layer is 9 where each one represents a different attribute as shown in Fig. 3. The number of neurons in the input layer is 2 since the proposed system is to classify two classes: death at 1 year, or alive at one year. The number of hidden neurons was taken as 100 by experience.

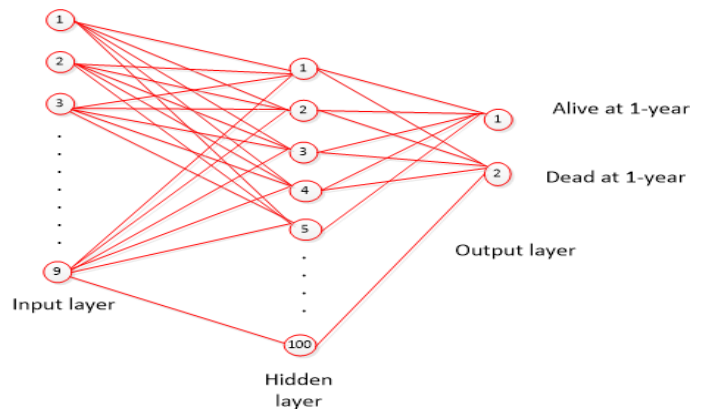


Fig. 3. The BPNN architecture.

Table 3 represents the training and testing set which consists of patients from two classes. It also shows the total number of database instances used for each set.

TABLE. III. TRAINING AND TESTING DATA

Data	Dead	Alive	Total
Training	45	22	67
Testing	44	20	64
Total	89	42	131

The pre-processing of the inputs data take place first in the system so that the data are normalized to values between 0 and 1 before feeding into network. Once the data are normalized, they are fed into a backpropagation neural network, respectively with their targets. Table 4 shows the input parameters values, as well as the training time of the network.

TABLE. IV. INPUT NETWORK PARAMETERS

Network data	Values
Number of training images	67
Number of hidden neurons	100
Activation function	Sigmoid
Learning rate (η)	0.3
Momentum rate (α)	0.7
Epochs	1000
Training time (secs)	20 sec
Mean square error (MSE) reached	0.012

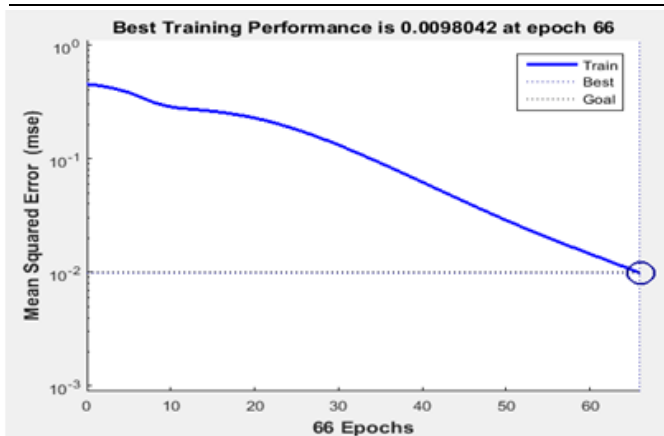


Fig. 4. Error variation with the iteration number (BPNN).

Fig. 4 shows the learning curve of the trained network. It can be seen that the network learned well since the error is decreasing after each epoch or iteration. The network has reached an error of 0.012 at epoch 66 which is good enough for this phase. It should be noted that the time taken for the network to learn and achieve the minimum square error is 20 seconds shown in Table 4.

IV. RBFN TRAINING

A radial basis function network is somehow different from the back propagation neural networks especially, in the way the weights in the hidden layer are updated. The output layer of a RBFN can be seen as that of a BPNN with linear activation functions [10]. Radial functions are simply a class of functions. In principle they could be employed in any sort of model linear or nonlinear and any sort of network single layer or multilayer. The output layer consists of neurons which combine linearly the bases computed in the hidden layer.

The motivation behind RBFN and some other neural classifiers is based on the knowledge that pattern transformed to a higher-dimensional space which is nonlinear is more probable to be linearly separable than in the low-dimensional vector representations of the same patterns (cover's separability theorem on patterns) [11].

The output of neuron units are calculated using k-means

clustering similar algorithms, after which Gaussian function is applied to provide the unit final output. During training, the hidden layer neurons are centered usually randomly in space on subsets or all of the training patterns space (dimensionality is of the training pattern) [10]; after which the Euclidean distance between each neuron and training pattern vectors are calculated, then the radial basis function (also or referred to as a kernel) applied to calculated distances. The radial basis function is so named because the radius distance is the argument to the function [12].

$$Weight = RBFN(distance) \tag{4}$$

It is to be noted that while other functions such as logistic and thin-plate spline can be used in RBFN networks, the Gaussian functions is the most common. During training, the radius of Gaussian function is usually chosen; and this affects the extent to which neurons have influence considering distance.

The best predicted value for the new point is found by summing the output values of the RBF functions multiplied by weights computed for each neuron [12]. The equation relating Gaussian function output to the distance from data points ($r>0$) to neurons center is given below.

$$\Phi(r) = e^{\frac{-r^2}{2\sigma^2}} \tag{5}$$

Where, σ is used to control the smoothness of the interpolating function [11]; and r is the Euclidean distance from a neuron center to a training data point.

Similarly, same data are used for the RBFN where 67 and 36 instances of data are used for training and testing, respectively. Table 5 shows the parameters values set during the training phase of this network. As seen in Table 5; the network is trained with 50 hidden neurons and spread constant of 0.5.

It is observed that RBFN with 50 hidden neurons and spread constant of 0.5 reached the lowest mean square error (MSE) (0.0330) in a very short time of 10 seconds. Moreover, this network was capable of reaching that low MSE with only 50 maximum epochs which is smaller than that of BPNN. Moreover, it is observed that this network was able to learn and converge in a shorter time than that of BPNN. The learning curve for RBFN is shown in Fig. 5.

TABLE. V. RBFN TRAINING PARAMETERS

Network parameter	RBFN
# of training samples	67
# hidden neurons	50
Spread constant	0.5
Maximum epochs	50
Training time (secs)	10
Mean Square Error reached	0.0330

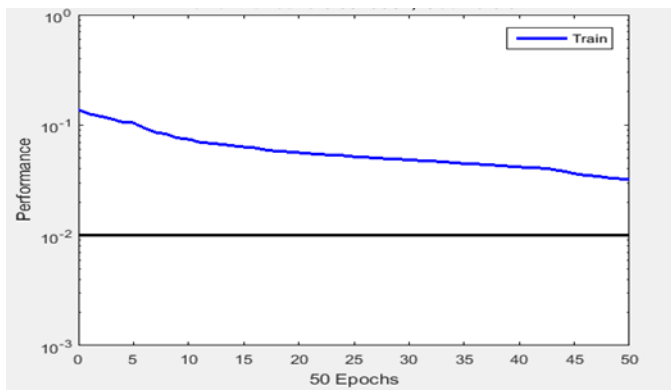


Fig. 5. Learning curve of the RBF network.

V. RBFN TRAINING

The networks were tested on a dataset of 64 records or patients; 44 for dead at one year, and 20 for alive at one year. Table 5 represents the total prediction rate of the designed one-year survival myocardial infarction prediction system. Note that the prediction rate shows the capability of the trained network to generalize, i.e., to predict the correct diagnosis while tested with unseen data. This prediction rate is defined as the total number of correctly classified instances of patient of the two classes divided by the total number of patients or instances.

Table 6 shows the backpropagation network's training and testing recognition rate of each set and class of data. It represents the number of cases that were accurately classified by the network in the training and the testing phases. In addition, it shows the percentage of instances that were not correctly classified by the network during the testing phase.

TABLE VI. TOTAL PREDICTION RATE

Networks		Total number of data (patients)	Number of correctly classified cases	Prediction rate
	Training	67	66	66/67 98.5%
BPNN	Testing	64	61	95.3%
	Total	131	127	96.9%
	Training	67	67	67/67 100%
RBFN	Testing	64	62	96.8%
	Total	131	127	98.5%

The experimental results of the developed one-year survival myocardial infarction prediction system were as follows: 98.5% using the training data set (67), and 95.3% using the testing data set (64) for the BPNN. While, for the radial basis function network the training and testing classification rate results were 100% and 96.8%, respectively. The overall prediction rate for BPNN and RBFN was eventually calculated to be 96.9% and 98.5%, respectively.

Both networks successfully classified the two classes consistent with the clinical data. The RBF network was capable to achieve a better performance in predicting the correct

diagnosis of the unknown data. Moreover, it was found that the RBF network's performance was the highest in the testing phase, as well as during training. On the other hand, this network reached a mean square error of 0.0330 which is higher than that of the BPNN "0.012". However, the BPNN required a longer training time "20 seconds" to reach that error than that of RBFN which achieved its lowest mean square error in "10 seconds". Note that the difference of the training time was not that high between both types of networks; however, the maximum number of iterations needed for the BPNN "1000 epochs" to converge was roughly higher than that of the RBFN "50 epochs".

VI. CONCLUSION

The deficiency in the myocardial infarction survival prediction systems was the motivation behind this work. In spite of the unreliability of survival prediction systems of myocardial infarction patients, our work showed that an intelligent system can learn from the past data of patients who suffered this disease to be capable of generalizing the correct diagnosis (survival or death at one year) of new myocardial infarction patients. This study compared the capability of two neural networks: BPNN and RBFN, to perform this task. As a result, both networks learned accurately to predict the two classes: dead at one year and survived at one year. However, one network "RBFN" outperformed the other "BPNN" when generalizing or predicting the diagnosis of the unseen instances. This outperformance was in terms of accuracy, training time, and number of maximum iterations needed for the network to converge.

REFERENCES

- [1] G. Davi and C. Patrono, "Platelet activation and atherothrombosis", *New England Journal of Medicine*, 357, 24, 2482-2494, 2007.
- [2] AS. Mullasari, P. Balaji, and T. Khando, "Managing complications in acute myocardial infarction", *Journal of the Association of Physicians of India* 59, pp. 43-8, 2011.
- [3] BW. Karlson, J. Herlitz, O. Wiklund, A. Richter, and A. Hjalmarson, "Early prediction of acute myocardial infarction from clinical history, examination and electrocardiogram in the emergency room". *The American journal of cardiology* 68 (2), pp. 171-175, 1991.
- [4] T. Mocan, L. Açoşton-Coldea, M. Gatfossé, S. Rosenstingl, LC. Mocan, and DL. Dumitraşcu, "A new prediction score for myocardial infarction: MINF SCORE", *Romanian Journal of Internal Medicine* 46, pp. 145-151, 2008.
- [5] M. Lichman, UCI Machine Learning Repository [http://archive.ics.uci.edu/ml]. Irvine, CA: University of California, School of Information and Computer Science. 2013.
- [6] R. Rojas. *Neural Networks-A Systematic Introduction* Springer-Verlag. New York. (1996).
- [7] A. Helwan, and DP. Tantua, IKRAI: Intelligent Knee Rheumatoid Arthritis Identification. *International Journal of Intelligent Systems and Applications* 8, pp. 18-24, 2016.
- [8] A. Helwan, and D. Ozsahin, "Sliding Window Based Machine Learning System for the Left Ventricle Localization in MR Cardiac Images", *Applied Computational Intelligence and Soft Computing*, 2017, doi:10.1155/2017/3048181.N.
- [9] A., Helwan, & R. Abiyev, "Shape and Texture Features for the Identification of Breast Cancer", In *Proceedings of the World Congress on Engineering and Computer Science*, 2016.
- [10] P. Strumillo and W. Kamiński, "Radial basis function neural networks: theory and applications", *Neural Networks and Soft Computing - Advances in Soft Computing*, 19, pp. 107-119, 2003.

- [11] A. Helwan, A. Khashman, E. O. Olaniyi, O. K. Oyedotun, & , O. A. Oyedotun, "Seminal Quality Evaluation with RBF Neural Network", *Bulletin of the Transilvania University of Brasov, Series III: Mathematics, Informatics, Physics*, 9, 2, 2016.
- [12] S. Garg, K. Patra, and SK. Pal, D. Chakraborty, "Effect of different basis functions on a radial basis function network in prediction of drill flank wear from motor current signals", *Soft Computing* 12, pp. 777-787, 2008.