

**Original article:**

**DATA MINING OF MAGNETOCARDIOGRAMS FOR PREDICTION  
OF ISCHEMIC HEART DISEASE**

Yosawin Kangwanariyakul<sup>1</sup>, Chanin Nantasenamat<sup>2</sup>, Tanawut Tantimongcolwat<sup>2</sup>,  
Thanakorn Naenna<sup>3,\*</sup>

<sup>1</sup> Master of Science Program in Technology of Information System Management, Faculty of Engineering, Mahidol University, Nakhon Pathom 73170, Thailand

<sup>2</sup> Department of Clinical Microbiology, Faculty of Medical Technology, Mahidol University, Bangkok 10700, Thailand

<sup>3</sup> Department of Industrial Engineering, Faculty of Engineering, Mahidol University, Nakhon Pathom 73170, Thailand

\* Corresponding author. Tel.: +6628892138 Ext 6211; Fax.: +6628892138 ext 6221.  
E-mail address: egtnn@mahidol.ac.th (T. Naenna)

**ABSTRACT**

Ischemic Heart Disease (IHD) is a major cause of death. Early and accurate detection of IHD along with rapid diagnosis are important for reducing the mortality rate. Magnetocardiogram (MCG) is a tool for detecting electro-physiological activity of the myocardium. MCG is a fully non-contact method, which avoids the problems of skin-electrode contact in the Electrocardiogram (ECG) method. However, the interpretation of MCG recordings is time-consuming and requires analysis by an expert. Therefore, we propose the use of machine learning for identification of IHD patients. Back-propagation neural network (BPNN), the Bayesian neural network (BNN), the probabilistic neural network (PNN) and the support vector machine (SVM) were applied to develop classification models for identifying IHD patients. MCG data was acquired by sequential measurement, above the torso, of the magnetic field emitted by the myocardium using a J-T interval of 125 cases. The training and validation data of 74 cases employed 10-fold cross-validation methods to optimize support vector machine and neural network parameters. The predictive performance was assessed on the testing data of 51 cases using the following metrics: accuracy, sensitivity, and specificity and area under the receiver operating characteristic (ROC) curve. The results demonstrated that both BPNN and BNN displayed the highest and the same level of accuracy at 78.43 %. Furthermore, the decision threshold and the area under the ROC curve was -0.2774 and 0.9059, respectively, for BPNN and 0.0470 and 0.8495, respectively, for BNN. This indicated that BPNN was the best classification model, BNN was the best performing model with sensitivity of 96.65 %, and SVM employing the radial basis function kernel displayed the highest specificity of 86.36 %.

**Keywords:** Ischemia, magnetocardiography, data mining, Back-propagation neural network, Bayesian Neural network, Probabilistic Neural network, Support vector machine

**INTRODUCTION**

Heart diseases are major causes of death worldwide. The World Health Organization reported that the global numbers of death

caused by cardiovascular diseases are increasing every year. This is especially critical in developing countries which have increased from 14 to 25 million for 1990 to 2020, respectively. Ischemic heart disease (IHD) is considered to be a foremost disorder

der of the heart. It is a disease affected by long-term deficiency of oxygen and nutrient of the cardiac muscle due to inadequate supply of blood circulation. This may lead to cardiac tissue damage and cause sudden cardiac death as a consequence of heart attack.

Early diagnosis of IHD can help reduce the rate of mortality (Faculty of Science Thaksin University, 2006). Electrocardiography (ECG) is traditionally applied to monitor defective electrophysiological activity of the heart. Nonetheless, in some circumstances, notification of normal ECG may be obtained albeit in the presence of angina (Tantimongcolwat et al., 2008). Measurement of ECG signal by electrode-skin contact is disturbed by many type of noises (Ayari et al., 2009). Therefore, magnetocardiography (MCG), a highly sensitive and contactless monitoring method of the physiological activity of the heart, has been developed. MCG essentially measures the magnetic field emission by electrophysiological activity of the heart. The magnetic field is recorded using superconducting quantum interference device (SQUID) that is without any sensor contact to the body, which is an advantage of MCG over that of ECG (Koch, 2001).

MCG is more sensitive to tangential currents in the heart than ECG, and it is also sensitive to vortex current, which cannot be detected by ECG. In a normal heart, the main direction of the activation waveform is radial, from endocardium to epicardium. For these reasons, MCG may show ischemia-induced deviations from the normal direction of depolarization and repolarization with better accuracy than that of ECG (Tsukada et al., 2000; Yamada and Yamaguchi, 2005). MCG is affected less by conductivity variations in the body (lungs, muscles, and skin) than ECG. In addition, because MCG is a fully non-contact method, therefore problems in the skin-electrode contact as encountered in ECG are avoided (Kanzaki et al., 2003; Tsukada et al., 1999; Tavarozzi et al., 2002).

Nowadays, medical informatics has increasingly been used to analyze large quantities of data that are stored in large databases which is made possible by the availability of powerful hardware and software along with the use of automated tools. But the interpretation of MCG recordings remains a challenge since there are no databases available from which precise rules could be deduced. The analysis of MCG data by experts is time-consuming and this is in concomitant with the shortage of experts possessing knowledge on the analysis of MCG data. Thus, methods to automate the interpretation of MCG recordings by minimizing human efforts are important for diagnosis of IHD in patients.

In this study, we used three different types of neural network and three different types of support vector machine for identification of IHD patients. The neural network classification models that are employed is comprised of back-propagation neural network, bayesian neural network, and probabilistic neural network. The support vector machine classification models that are employed is comprised of linear, polynomial, and radial basis function kernel. All models were trained using 10-fold cross-validation technique and the optimal parameters for neural network and support vector machine were empirically determined and finally used for comparing the efficiency and accuracy of these classification models.

## MATERIAL AND METHODS

### *Data acquisition and preparation*

MCG data was obtained by making four sequential measurements of cardiac magnetic field at 36 locations ( $6 \times 6$  matrices) above the torso. Each of the 36 MCG signals was recorded for 90 seconds using nine sets of SQUID sensor at a sampling rate of 1000 Hz. For noise reduction purposes, the default filters setting at 0.05–100 Hz was applied and followed by an additional digital low pass filter at 20 Hz (Bick et al., 2001). The MCG signals at a time

window between J-point and T-peak (J-T interval) of the cardiac cycle was subdivided into 32 uniformly spaced points resulting in a set of 1,152 descriptors for all of 36 MCG signals (Froelicher et al., 2002). These were assigned as inputs to build up artificial neural network models for automatic classification of IHD. The MCG datasets exploited in this study were acquired from 55 confirmed IHD cases and 70 normal individuals. Seventy-four MCG signals of both IHD and normal control groups were randomized to the training sets and the remainders were assigned as the testing set.

### Overview of artificial neural network

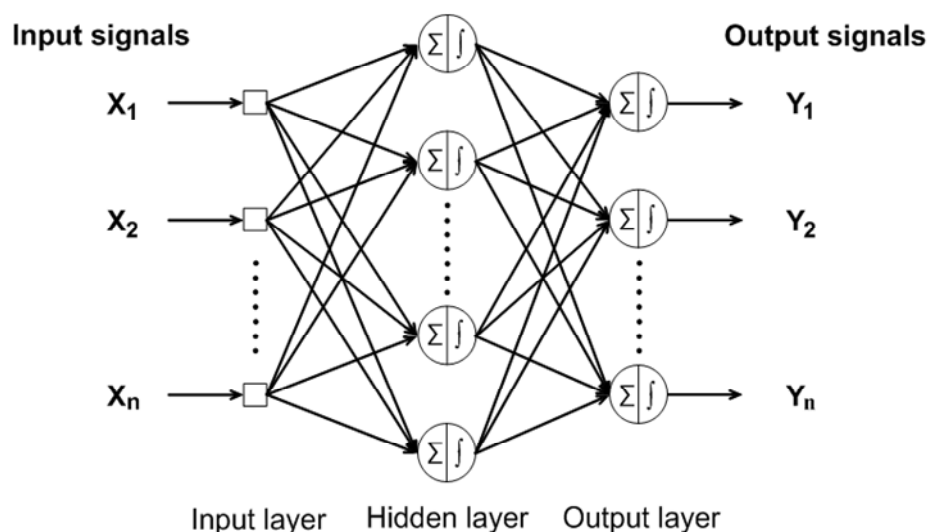
Neural networks are the most widely used learning approach in a variety of disciplines such as business, industries and academia. A majority of the applications of neural networks has been applied for biomedical prediction and classification tasks. Mobley (Mobley et al., 2000) developed neural network models for automatic detection of coronary artery stenosis using 14 risk factors for coronary artery disease. Chou (Chou et al., 2004) integrated the use of artificial neural networks with multivariate adaptive regression splines approach for the classification of breast cancer. Liew compared the predictive accuracy of artificial neural networks and conventional logistic regression for the prediction of gall-

bladder disease in obese patients. In addition, neural network was applied for the prediction of acute coronary syndrome and was compared with multiple logistic regression (Green et al., 2006).

From the above listed researches, the predictive accuracy of neural networks illustrates good performance for prediction and classification tasks. Therefore, the aim of this study was to compare the predictive accuracy of three algorithms of neural networks for prediction of ischemic heart disease from magnetocardiograms.

### Overview of back-propagation neural network (BPNN)

BPNN is one of the most accepted supervised feed-forward neural networks implementing the use of back-propagation learning algorithm. Input vectors and corresponding target vectors are used to train a neural network. A BPNN is composed of at least three or more layers. In this study, BPNN is consisted of three layers, where the first layer is the input layer which receives data from outside the neural network, second layer is the hidden layer which performs processing and transformation of the input data, and third layer is the output layer which sends data out of the neural network (Haykin, 1999). Fig. 1 shows the general BPNN architecture.



**Figure 1:** Architecture of a three-layer feed-forward back-propagation neural network

The back-propagation algorithm is comprised of a forward and backward path. Forward path involves creation of a feed-forward network and weight initialization. The input vectors are propagated through the network layer by layer. Finally, a set of outputs is produced as the actual response of the network. The network weight and bias are updated in a backward path from the error, which compares the actual response with the target output. The weight updating rule can be calculated using Eq. 1:

$$\Delta w_{ji}(n) = \eta \delta_j x_{ji} + \alpha \Delta w_{ji}(n-1) \quad (1)$$

where  $\Delta w_{ji}(n)$  is the weight update performed during the  $n^{\text{th}}$  iteration,  $\alpha$  is the momentum constant, and  $\delta_j$  is the error for any node (Mitchell, 1997).

### **Overview of Bayesian neural network (BNN)**

BNN uses a logistic regression model and back-propagation algorithm based on a Bayesian framework (Gao et al., 2005). The class label is binary-valued  $y = (1,0)$  corresponding to normal and abnormal (IHD), respectively. A logistic regression on a Bayesian method estimates the class probability for a given input by:

$$P(y = k | X) \quad k = 1,0 \quad (2)$$

The outputs from hidden and output neurons were denoted by  $S_j$  and  $S_k$ , respectively.

Hidden layer:

$$S_j = \tanh\left(\sum_i w_{ij} X_i + w_{j0}\right) \quad (3)$$

Output layer:

$$S_k = \sum_k w_{kj} S_j + w_{k0} \quad (4)$$

To allow the outputs to be interpreted as probabilities, the logistic regression is used to model the risk of occurrence of IHD. The logistic regression model is a S-

shaped distribution which is similar condition to the estimated probabilities to lie between 0 and 1.

Let  $P(y = k|X)$  be the probability of the event  $y = 1$  (normal case), given the input vector  $X$ . The logistic regression model as a function of network output  $y$  by:

$$P(y = k | X) = \frac{1}{1 + \exp(-S_k)} \quad (5)$$

### **Overview of probabilistic neural network (PNN)**

PNN is a supervised neural network (Fig. 2) that is derived from the Radial Basis Function (RBF) Network. RBF is a Gaussian function that scales the variable nonlinearly. The advantage of PNN is that it requires less training time (Comes and Kelemen, 2003). In this paper, PNN has three layers: input, radial basis, and competitive layer. Radial basis layer consists of a set of RBF for evaluating the euclidean distances between input vector and row weight vectors in the weight matrix. These distances are nonlinearly scaled by Gaussian function in the radial basis layer, which is fully interconnected to a competitive layer where it passes the results to the competitive layer. Therefore, the competitive layer finds the shortest distance among them, and thus discovers the training pattern closest to the input based on their distance (Specht, 1990).

### **Overview of support vector machine (SVM)**

SVMs are a kind of supervised learning based on the statistical learning theory that was introduced by Vapnik and co-workers (Vapnik, 1998; Boser et al., 1992). The techniques have been further adapted by a number of other researchers (Cristianini and Shawe-Taylor, 2000; Abe, 2005) and they are also showing high performances in biological, chemical and medical applications (Yao et al., 2001; Mehta and Lingayat, 2008; Nantasenamat et al., 2008).

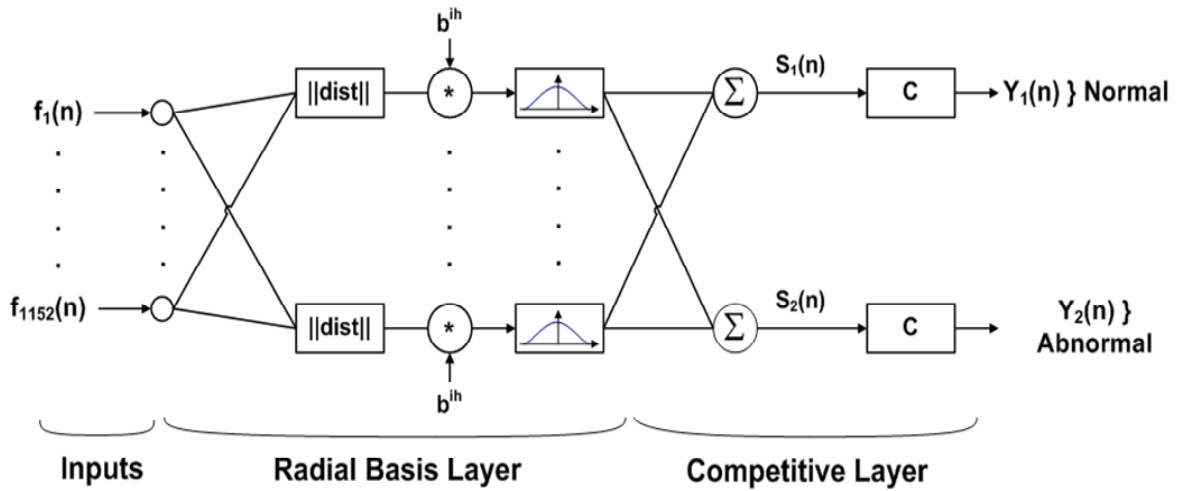


Figure 2: Network structure of PNN

SVM is a robust technique for data classification and regression. SVM models search for a hyperplane that can linearly separate classes of objects (Fig. 3a). Let  $x_i \in \mathbf{R}^m$ , ( $i = 1, 2, \dots, m$ ) represents the vectors and  $y_i \in \{-1, 1\}$  represents the class labels. The term  $f(x_i)$  can be represented by a linear function of the form by  $y_i = f(x_i)$

$$f(x_i) = \langle w_i, x_i \rangle + b \quad (6)$$

where  $\langle w_i, x_i \rangle$  represents the inner product of  $w$  and  $x$ ,  $w$  is weight vector and  $b$  is bias. SVM approximates the set of data with linear function as follows

$$y = \sum_{i=1}^m w_i \Phi(x_i) + b \quad (7)$$

where  $\{\Phi(x_i)\}_{i=1}^m$  represents the feature of input variables subjected to kernel transformation while  $\{w_i\}_{i=1}^m$  and  $b$  are coefficients.

SVM can be applied to non-linear classification using non-linear kernel functions to map the input data onto a higher dimensional feature space in which the input data can be separated with a linear classifier (Fig. 3b). Kernel function  $K(\mathbf{x}, \mathbf{y})$  represents the inner product  $\langle \phi(\mathbf{x}), \phi(\mathbf{y}) \rangle$  in feature space. This study uses the following kernel functions:

Linear kernel

$$K(x, y) = x \cdot y \quad (8)$$

Polynomial kernel

$$K(x, y) = (x \cdot y)^d \quad (9)$$

where  $d$  is the degree of the polynomial ( $d = 0$  for linear kernel).

Radial basis function kernel

$$K(x, y) = \exp(-\gamma \|x - y\|^2) \quad (10)$$

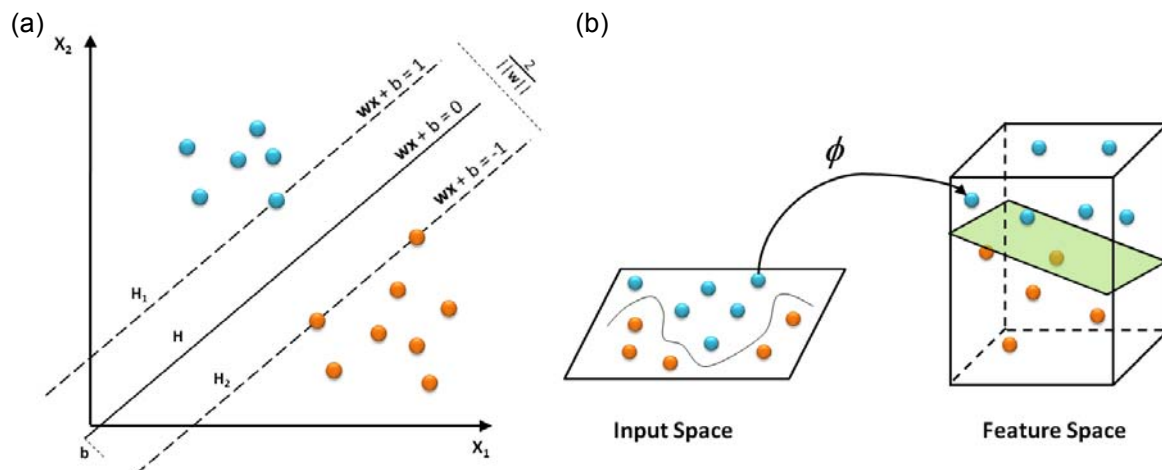
where  $\gamma > 0$ .

#### Parameter optimization of neural network

The neural network architecture was selected by tuning various parameters by trial-and-error using root mean square error (RMSE) according to eq. 11.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (p_i - a_i)^2}{n}} \quad (11)$$

where  $p_i$  is the predicted output,  $a_i$  is the actual output, and  $n$  is the number of compounds in the dataset. The neural network parameters include the number of nodes in



**Figure 3:** Schematics of SVM depicting objects that are linearly separable (a) and non-linear mapping of input space onto the feature space (b)

the hidden layer, the number of learning epochs, learning rate and momentum. For each of the parameter calculations, 10-fold cross-validation was performed and the average RMSE of each parameter was calculated. The lowest RMSE value of each parameter was chosen as the optimal value. In this study, to avoid overtraining 10-fold cross-validation was used for model validation (Kohavi, 1995), which essentially involved the random separation of the data set into 10 fold where each fold was used once as the test set and the remaining nine-tenth was used as the training set.

The PNN training set comprising of 74 cases was trained using stratified 10-fold cross-validation. The PNN models were trained and validated for 10 times using parameters from the range of 0.1 to 5 at an interval of 0.1 (Comes and Kelemen, 2003). For each case, one fold was used as the testing set while the remaining folds were used as the training set.

### Confusion matrix

A confusion matrix contains the results about the actual and predicted classifications made by the neural network on the test set. Such information is often displayed as a two dimensional matrix. The confusion matrix is easy to understand as the correct and incorrect classification are displayed in

the table (Witten and Frank, 2005). The confusion matrix for a two class classifier is shown in Table 1.

**Table 1:** Confusion matrix for a two class classifier

Actual \ Predicted	Positive	Negative
	Positive	TP
Negative	FP	TN

TP = True positive: Diseased individuals correctly diagnosed as sick

FP = False positive: Healthy individuals wrongly identified as sick

TN = True negative: Healthy individuals correctly identified as healthy

FN = False negative: Diseased individuals wrongly identified as healthy

### Accuracy, sensitivity and specificity

Clinical research often investigates the statistical relationship between the test results and the presence of disease. In this study, we used three components that determine the performance of neural networks. Sensitivity is the probability that a test is positive (a symptom is present), given that the person has the disease.

$$\text{sensitivity} = \frac{TP}{TP + FN} \quad (12)$$

Specificity is the probability that a test is negative (a symptom is not present), given that the person does not have the disease.

$$\text{specificity} = \frac{TN}{TN + FP} \quad (13)$$

Accuracy is the proportion of correct classifications in the set of predictions.

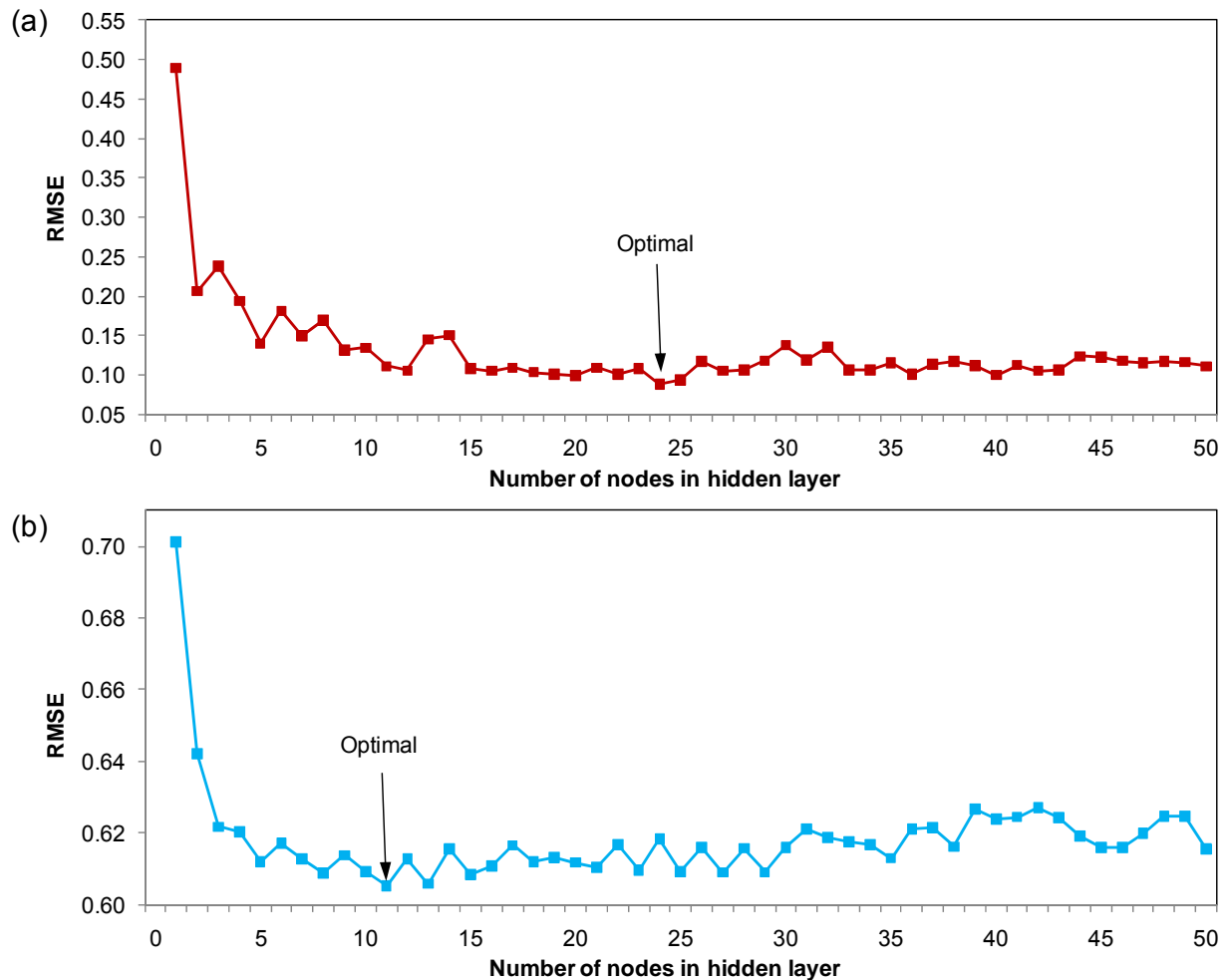
$$\text{accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (14)$$

### Optimization of artificial neural network

To achieve maximal performance fine-tuning of the parameters were performed. Particularly, number of nodes in the hidden layer, learning epoch size, learning rate ( $\eta$ ) and momentum ( $\mu$ ) was adjusted when using BPNN and BNN while the spread parameter was optimized when using PNN.

The optimal number of nodes in the hidden layer was determined by varying the number of nodes from 1 to 50 while fixing the settings for the remaining parameters. The results were plotted as a function of RMSE. The optimal value was found to be 24 and 11 for BPNN (Fig. 4a) and BNN (Fig. 4b), respectively.

## RESULTS



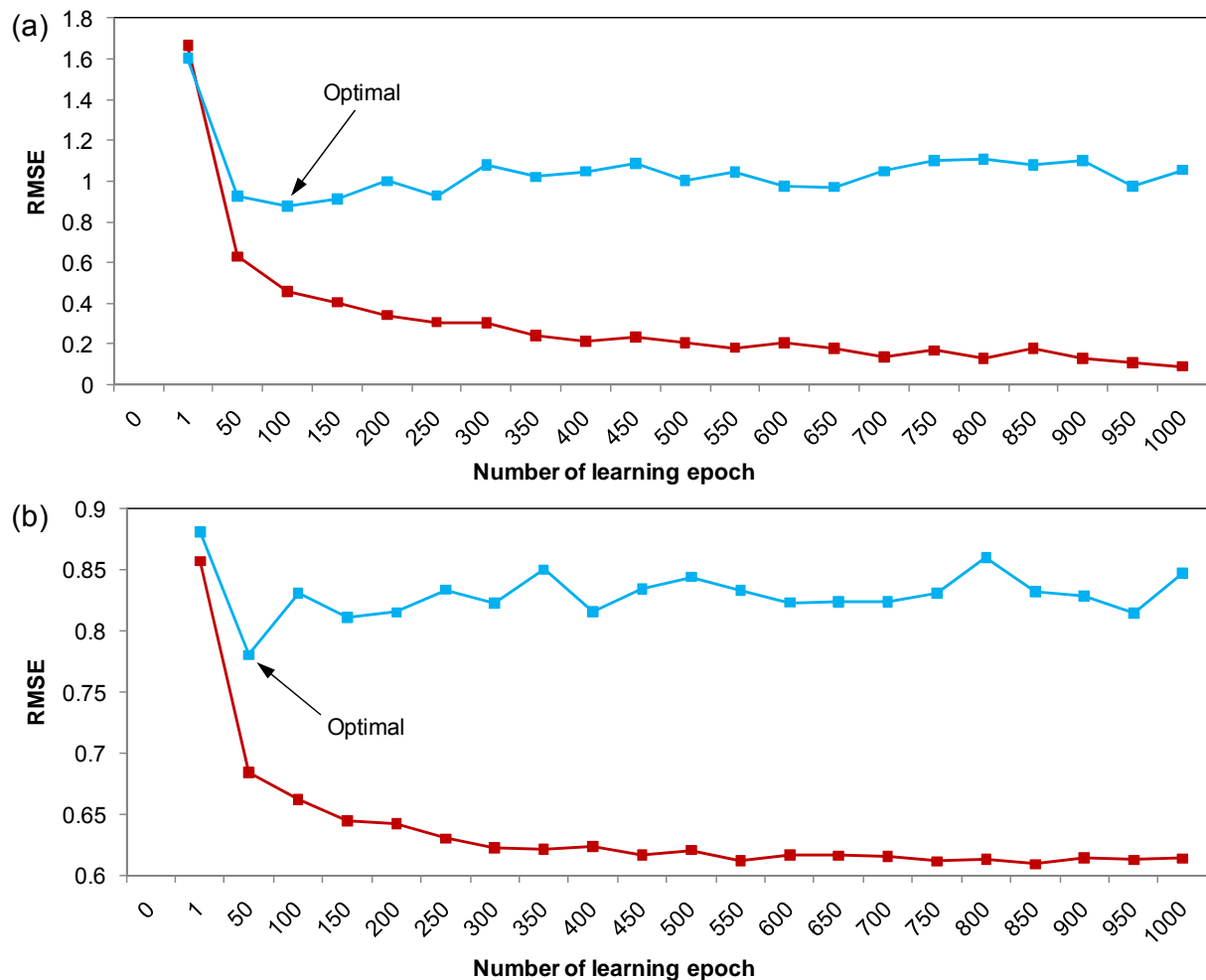
**Figure 4:** Plot of the number of nodes in hidden layer as a function of RMSE (■) using BPNN (a) and plot of the number of nodes in hidden layer as a function of RMSE for BNN (■) using BNN (b)

In order to avoid overtraining of the predictive model, the learning epoch size was subsequently optimized from 50 to 1000 in increments of 50 and learning was stopped once a detectable rise in RMSE for the leave-one-out cross-validated testing set was observed. By making a plot of the RMSE as a function of the learning epoch size, the best learning time was found to be 100 and 50 for BPNN (Fig. 5a) and BNN (Fig. 5b), respectively.

In the same manner, the optimal settings for the learning rate and momentum was determined by simultaneously varying the number of learning rate and momentum in the range of 0 to 1 and subsequently making a surface plot of RMSE against the relevant learning rates and momentums (Fig. 6). The momentum and learning rate

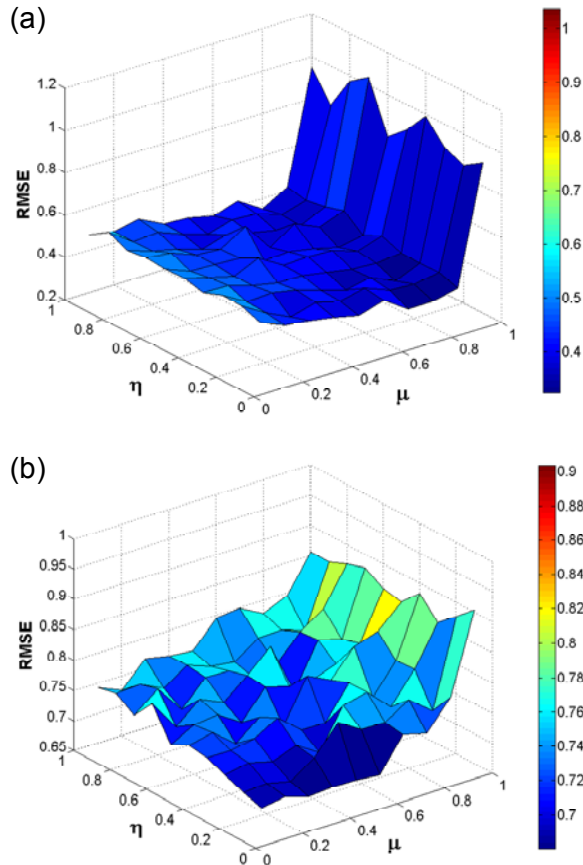
providing a minimum global surface of error was then chosen as the optimum. It had been found that the optimal learning rate and momentum of BPNN was 0.1 and 0.7, which was nearly similar to those observed in BNN (0.1 and 0.6).

The optimization of PNN calculation was carried out in a trial-and-error manner by optimizing the spread parameter in the range of 0.1 to 5 with intervals of 0.1. The resulting prediction accuracy was then plotted as a function of the spread parameter as shown in Fig. 7. It was observed that increasing the magnitude of the spread parameter resulted in an enhancement of the accuracy of prediction that was kept stable once spread parameter was in excess of 0.9. Therefore, the spread parameter of 0.9 was chosen as the optimum.

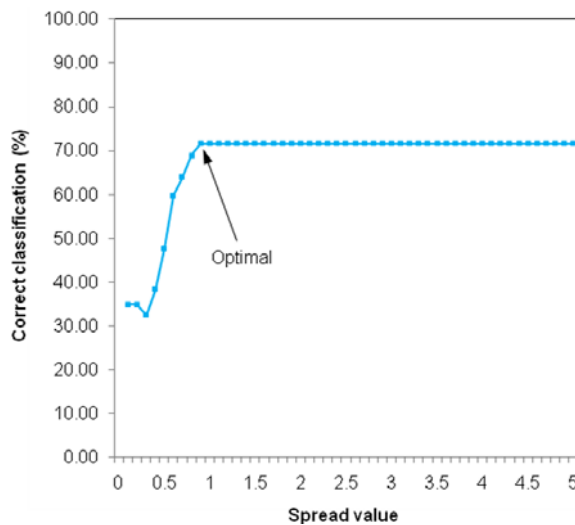


**Figure 5:** Plot of the number of learning epoch as a function of RMSE for the training set (■) and validation set (■) using BPNN (a) and plot of the number of learning epoch as a function of RMSE for the training set (■) and validation set (■) using BNN (b)





**Figure 6:** Surface plot of the learning rate and momentum as a function of RMSE for BPNN (a) and surface plot of the learning rate and momentum as a function of RMSE for BNN (b)



**Figure 7:** Plot of the spread value as a function of correct classification

### Optimization of support vector machine

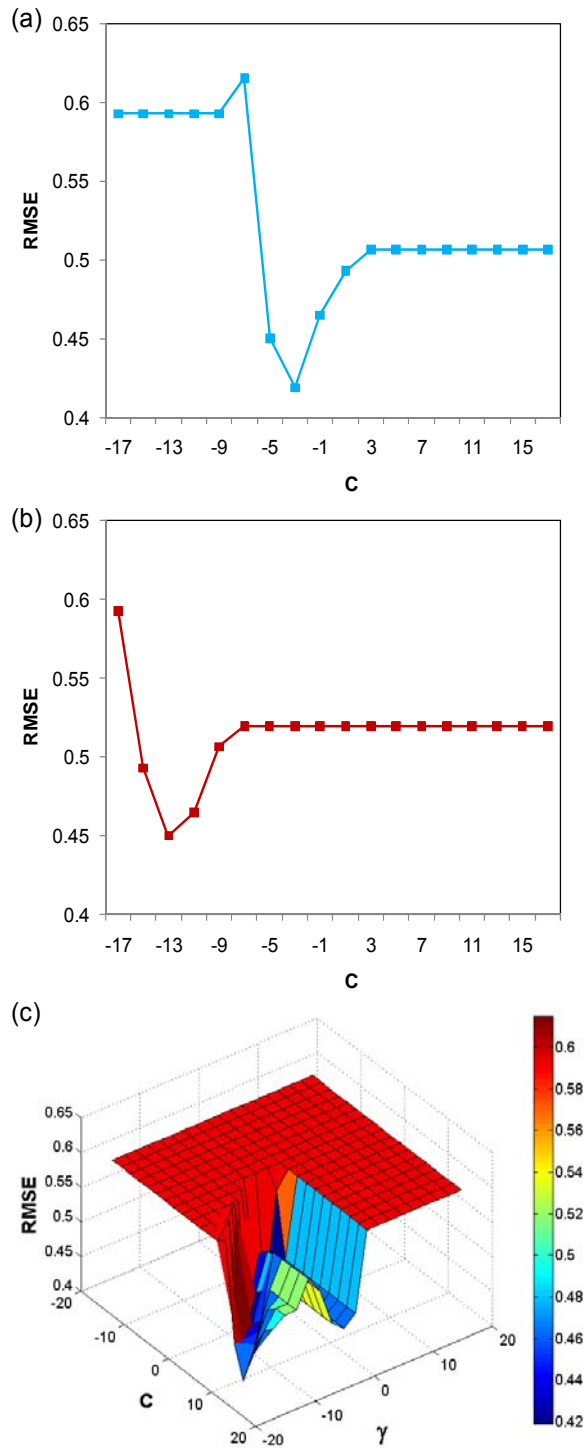
For support vector machine calculations, the C parameter was optimized when using linear kernel SVM in the range of  $2^{-17}$  to  $2^{17}$  at an interval of 2. Fig. 8a shows the plotting of RMSE versus C parameter. The minimum value of RMSE is optimal at C parameter of  $2^{-3}$  for linear kernel SVM.

Polynomial kernel SVM has two parameters (exponent parameter and C parameter) that must be optimized. The exponent parameter was adjusted from 2 until 10 and the C parameter was adjusted from  $2^{-17}$  until  $2^{17}$  by an interval of 2. We found that the optimal exponent parameter is  $2^2$  and C parameter is  $2^{-13}$  as shown in Fig. 8b.

Finally, the RBF kernel SVM was optimized by adjusting the C parameter and gamma ( $\gamma$ ) parameter by an interval of 2. The optimal value of parameter was determined by making a surface plot of RMSE as a function of C parameter and  $\gamma$  parameter. The optimal C parameter and  $\gamma$  parameter are  $2^{13}$  and  $2^{-17}$  (Fig. 8c), respectively.

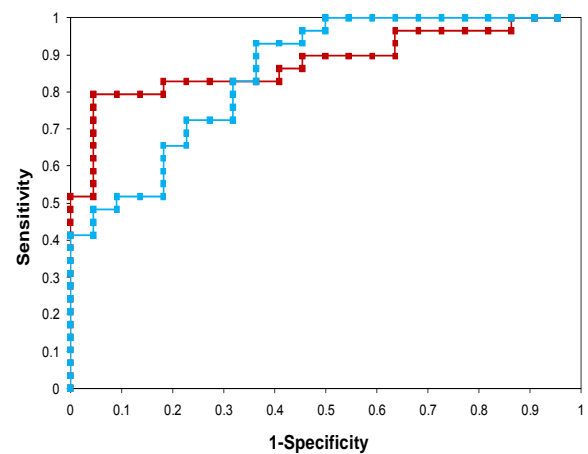
### Prediction of ischemic heart disease

Machine learning models for identification of IHD were based on three learning algorithms that included BPNN, BNN and PNN. Each of the predictive models was constructed by training with 74 MCG patterns (26 IHD and 48 normal cases) under optimal parameter settings. The performance of IHD classification was then tested on 51 MCG patterns consisting of 29 IHD cases and 22 normal controls. To investigate the prediction performance and to find the optimum decision threshold, ROC curve of each model was calculated and shown in Fig. 9. The decision threshold of BPNN and BNN was varied in the range of -1 to 1 and 0 to 1, respectively. The resulting prediction sensitivities were plotted as a function of one minus relevant specificities. Models giving greater area under the ROC curve provided better prediction performance than models offering the lesser area. It was observed that the area under the ROC curve given by the BPNN model was 0.9059 which



**Figure 8:** Plot of the C parameter of linear kernel SVM as a function of RMSE (a), plot of the C parameter of polynomial kernel SVM as a function of RMSE (b) and surface plot of C parameter and gamma ( $\gamma$ ) as a function of RMSE (c)

was slightly higher than the BNN model which was 0.8495. The optimal decision thresholds were 0.1 and 0.5 for BPNN and BNN models, respectively. As represented in Table 2, 25 out of 29 IHD cases could be correctly identified by BPNN and PNN, while only one IHD failed to be classified by BNN (see Table 2). In the classification of normal individuals, BPNN outperformed both BNN and PNN. Fifteen normal MCGs were correctly classified by BPNN, while 10 and 11 normal patterns were made by BNN and PNN.



**Figure 9:** ROC curve showing possible trade-offs between true positives (sensitivity) and false positives (1-specificity) for BPNN (■) and BNN (■)

For the SVM models, the linear kernel SVM gave the highest correctly classified accuracy of 18 out of 22 normal cases and 20 out of 29 IHD cases (see Table 2) but the highest sensitivity is obtained using polynomial kernel SVM and the highest specificity is obtained from the RBF kernel SVM (see Table 3).

In addition, the prediction performance of these six machine learning models could be statistically demonstrated in terms of the prediction accuracy, sensitivity and specificity as shown in Table 3. Comparable accuracy of prediction (78.43 %) was observed for BPNN and BNN learning models. However, BNN exhibited relatively higher sensitivity but lesser specificity of predictions than those made by BPNN.

**Table 2:** Confusion matrix shows the prediction results of the 51 cases in the testing set

Actual \ Predicted	BPNN		BNN		PNN		Linear SVM		Polynomial SVM		RBF SVM	
	IHD	Normal	IHD	Normal	IHD	Normal	IHD	Normal	IHD	Normal	IHD	Normal
IHD	25	4	28	1	25	4	20	9	26	3	12	17
Normal	7	15	10	12	11	11	4	18	12	10	3	19

**Table 3:** The prediction performance of classification models

Classifiers	Accuracy (%)	Sensitivity (%)	Specificity (%)
BPNN	78.43	86.21	68.18
BNN	78.43	96.65	54.55
PNN	70.59	86.21	50.00
Linear SVM	74.51	68.97	81.82
Polynomial SVM	70.59	89.66	45.45
RBF SVM	60.78	41.38	86.36

## DISCUSSION

Machine learning plays an important role in knowledge discovery of medical data. It is widely applied for identification or classification of many diseases. For example, predicting of cancer diseases and heart diseases (Fenici et al., 2005; Vibha et al., 2006; Polat and Gunes, 2007; Mobley et al., 2000; Temurtas et al., 2009). Our previous works had employed various machine algorithms in attempts to automatically distinguish between normal MCG and ischemic patterns. Under utilization of direct kernel-self organizing map (DK-SOM), direct kernel-partial least square (DK-PLS) and back propagation neural network (BPNN) on wavelet and time domains of MCG.

We employed six types of machine learning model for identification of ischemic heart disease using J-T interval of abnormal MCG pattern as inputs (Berul et al., 1994; On et al., 2007). On the other hand, MCG had higher sensitivity for detecting myocardial ischemia than conventional 12-lead ECG (Sato et al., 2001).

As seen from previous diagnosis of IHD heart patterns with direct kernel methods, Embrechts et al. reported classification accuracy between 71 % and 83 % using direct kernel based self-organizing maps (DK-SOM), direct kernel partial least square (DK-PLS) and least-squares support vector machines (LS-SVM), which is also known as kernel ridge regression (Embrechts et al., 2003). Moreover, the back-propagation neural network (BPNN) and direct kernel self-organizing map (DK-SOM) were applied for identification of IHD. The BPNN obtained 89.7 % sensitivity, 54.5 % specificity and 74.5 % accuracy, while the DK-SOM obtained higher sensitivity, specificity and accuracy than BPNN. The DK-SOM obtained 86.2 % sensitivity, 72.7 % specificity and 80.4 % accuracy (Tantimongcolwat et al., 2008).

In this study, we employed various algorithms of neural network for classification task. Results of BPNN and BNN shows that both classifiers had equal performance of 78.43 % accuracy which was higher than that of the PNN classifier, which had 70.59 % classification accuracy. Sensitivity of BPNN classifier (86.21 %) was lower than that of BNN classifier (96.55 %), while specificity of BPNN classifier (68.18 %) was higher than that of BNN classifier (54.55 %). Although PNN classifier had the lowest accuracy but it also exhibited the highest specificity (73.33 %). In addition, linear kernel SVM came out to be the best performing classifier with a classification accuracy of 74.51 % but such results is still less than that of BPNN and BNN. The polynomial kernel SVM had the best sensitivity (89.66 %) but was also less

than BNN. Finally, the RBF kernel had the highest specificity (86.36 %) as compared with the other classifiers.

### CONCLUSION

This paper presents a comparative analysis of prediction model for IHD identification by using three algorithms of neural network and three kernels of support vector machine. The 125 cases were randomly separated into 74 cases for training set and 51 cases for testing set. In order to optimize the neural network structure, we used a 10-fold cross-validation on training set. The optimal parameters of neural network was determined by averaging the values from 10 runs.

Comparison of the prediction performance of IHD identification was performed using three neural network algorithms and three support vector machine kernels as applied on the 51 cases of testing sets. The result shows that BPNN and BNN gave the highest classification accuracy of 78.43 %, while RBF kernel SVM gave the lowest classification accuracy of 60.78 %. BNN presented the best sensitivity of 96.55 % and RBF kernel SVM displayed the lowest sensitivity of 41.38 %. Both polynomial kernel SVM and RBF kernel SVM presented the minimum and maximum specificity of 45.45 % and 86.36 %, respectively.

### ACKNOWLEDGMENTS

This study was supported by Mr. Karsten Sternickel, Prof. Mark J. Embrechts, and Prof. Boleslaw Szymanski for preparation of the data set used in this study and some advises for this study. Special thanks are due to Mr. Graham Pogers for partial revision of this manuscript.

### REFERENCES

Abe S. Support vector machines for pattern classification. New York: Springer-Verlag, 2005.

Ayari EZ, Tielert R, When N. A noise tolerant method for ECG signals feature extraction and noise reduction. In: Proceedings of the 3<sup>rd</sup> International Conference on Bioinformatics and Biomedical Engineering (ICBBE 2009), 2009 June 11-13. Beijing, China.

Berul CI, Sweeten TL, Dubin AM, Sham MJ, Vetter VL. Use of the rate-corrected JT interval for prediction of repolarization abnormalities in children. *Am J Cardiol* 1994;74:1254-7.

Bick M, Sternickel K, Panaitov G, Effern A, Zhang Y, Krause HJ. SQUID gradiometry for magnetocardiography using different noise cancellation techniques. *IEEE Trans Appl Supercond* 2001;11:673-6.

Boser BE, Guyon IM, Vapnik VN. A training algorithm for optimal margin classifiers. In: Proceedings of the 5<sup>th</sup> Annual ACM Workshop on Computational Learning Theory AMC Press, 1992 (pp 144-152). Pittsburgh, USA.

Chou SM, Lee TS, Shao YE, Chen IF. Mining the breast cancer pattern using artificial neural networks and multivariate adaptive regression splines. *Expert Syst with Appl* 2004;27:133-42.

Comes B, Kelemen A. Probabilistic neural network classification for microarray data. In: Proceedings of the International Joint Conference on Neural Networks 2003 July; 3:1714-7.

Cristianini N, Shawe-Taylor J. An introduction to support vector machines and other Kernel-based learning methods. Cambridge: Cambridge Univ. Pr., 2000.

Embrechts MJ, Szymanski B, Sternickel K, Naenna T, Bragaspathi R. Use of machine learning for classification of magnetocardiograms. *IEEE Int Conf Syst Man Cybernet* 2003;2:1400-5.

Faculty of Science, Thaksin University [Online]. 2006 [cited 2010 Apr 19]; available from: URL:www.sci.tsu.ac.th/sciptl/radio/tape043/colesterol.doc.

Fenici R, Brisinda D, Meloni AM, Sternickel K, Fenici P. Clinical validation of machine learning for automatic analysis of multichannel magnetocardiography. In: Frangi AF, Radeva PI, Santos A, Hernandez M (Eds). Functional imaging and modeling of the heart (pp 143-152). Berlin: Springer-Verlag, 2005.

Froelicher V, Shetler K, Ashley E. Better decisions through science: Exercise testing scores. *Prog Cardiovasc Dis* 2002;44:395-414.

Gao D, Madden M, Chambers D, Lyons G. Bayesian ANN classifier ECG arrhythmia diagnostic system: a comparison study. *IEEE Int Joint Conf Neural Networks* 2005;4:2383-8.

Green M, Bjork J, Forberg J, Ekelund U, Edenbrandt L, Ohlsson M. Comparison between neural network and multiple logistic regression to predict acute coronary syndrome in the emergency room. *Artif Intell Med* 2006;38:305-18.

Haykin S. *Neural networks: a comprehensive foundation*. 2<sup>nd</sup> ed. Upper Saddle River, NJ: Prentice Hall, 1999.

Kanzaki H, Nakatami S, Kandori A, Tsukada K, Miyatake K. A new screening method to diagnose coronary artery disease using multichannel magnetocardiogram and simple exercise. *Basic Res Cardiol* 2003;98:124-32.

Koch H. SQUID magnetocardiography: status and perspectives. *IEEE Trans Appl Supercond* 2001;11:49-59.

Kohavi R. A study of study cross-validation and bootstrap for accuracy estimation and model selection. In: Mellish CS (Ed). *Proceedings of the 14<sup>th</sup> International Joint Conference on Artificial Intelligence, 1995, San Francisco* (pp 1137-1143). San Francisco, CA: Morgan Kaufman, 1995.

Mehta SS, Lingayat NS. Development of SVM based classification techniques for the delineation of wave components in 12-lead electrocardiogram. *Biomed Signal Process* 2008;3:341-9.

Mitchell TM. *Machine learning*. New York: McGraw Hill, 1997.

Mobley BA, Schechter E, Moore WE, McKee PA, Eichner JE. Predictions of coronary artery stenosis by artificial neural network. *Artif Intell Med* 2000;18(3):187-203.

Nantasenamat C, Isarankura-Na-Ayudhya C, Naenna T. Prediction of bond dissociation enthalpy of antioxidant phenols by support vector machine. *J Mol Graph Model* 2008;27:188-96.

On K, Watanabe S, Yamada S, Takeyasu N, Nakagawa Y, Nishina H, Morimoto T, Aihara H, Kimura T, Sato Y, Tsukada K, Kandori A, Miyashita T, Ogata K, Suzuki D, Yamaguchi I, Aonuma K. Integral value of JT interval in magnetocardiography is sensitive to coronary stenosis and improves soon after coronary revascularization. *Circ J* 2007;71:1586-92.

Polat K, Gunes S. Breast cancer diagnosis using least square support vector machine. *Digit Signal Process* 2007;17:694-701.

Sato M, Terada Y, Mitsui T, Miyashita T, Kandori A, Tsukada K. Detection of myocardial ischemia by magnetocardiogram using 64-channel SQUID system. In: Nenonen J, Imoniemi RJ, Katila T (Eds). Proceedings of the 12<sup>th</sup> International Conference on Biomagnetism, 2001 (pp 523-526). Espoo: Helsinki University Technology, 2001.

Specht DF. Probabilistic neural networks. *Neural Networks* 1990;3:109-18.

Tantimongcolwat T, Naenna T, Isarankura-Na-Ayudhya C, Embrechts MJ, Prachayasittikul V. Identification of ischemic heart disease via machine learning analysis on magnetocardiograms. *Comput Biol Med* 2008;38:817-25.

Tavarozzi I, Comani S, Gratta CD, Romani GL, Luzio SD, Brisinda D et al. Magnetocardiography: current status and perspectives. Part I: Physical principles and instrumentation. *Ital Heart J* 2002;3(2):75-85.

Temurtas H, Yumusak N, Temurtas F. A comparative study on diabetes disease diagnosis using neural networks. *Expert Syst Appl* 2009;36:8610-5.

Tsukada K, Sasabuchi H, Mitsui T. Measuring technology for cardiac magneto-field using ultra-sensitive magnetic sensor for high speed and noninvasive cardiac examination. *Hitachi Rev* 1999;48:116-9.

Tsukada K, Miyashita T, Kandori A, Yamada S, Sato M, Terada Y et al. Magnetocardiographic mapping characteristic for diagnosis of ischemic heart disease. *Comput Cardiol* 2000;27:505-8.

Vapnik VN. *Statistical learning theory*. New York: Wiley, 1998.

Vibha L, Harshavardhan GM, Pranaw K, P Deepa Shenoy, Venugopal KR, Patnaik LM. Classification of mammograms using decision trees. In: Proceedings of the 10<sup>th</sup> International Database Engineering and Applications Symposium, 2006, Dec 11-13. Delhi, India.

Witten IH, Frank E. *Data mining: practical machine learning tools and techniques*. 2<sup>nd</sup> ed. San Francisco, CA: Morgan Kaufman, 2005.

Yamada S, Yamaguchi I. Magnetocardiograms in clinical medicine: Unique information on cardiac ischemia, arrhythmias, and fetal diagnosis. *Intern Med* 2005;44:1-19.

Yao Y, Frasconi P, Pontil M. Fingerprint classification with combinations of support vector machine. In: 3<sup>rd</sup> International Conference of Audio and Video-Based Person Authentication, 2001, June (pp 253-259). Halmstad, Sweden.