Expert Systems with Applications 37 (2010) 3177–3183



Contents lists available at [ScienceDirect](http://www.sciencedirect.com/science/journal/09574174)

Expert Systems with Applications

journal homepage: [www.elsevier.com/locate/eswa](http://www.elsevier.com/locate/eswa)

A comparison of feature selection models utilizing binary particle swarm optimization and genetic algorithm in determining coronary artery disease using support vector machine

I\_smail Babaoglu \*, Og˘uz Findik, Erkan Ülker

*Department of Computer Engineering, Selcuk University, Konya, Turkey*

## a r t i c l e i n f o

*Keywords:*

Binary particle swarm optimization Genetic algorithm

Support vector machine Exercise stress testing Coronary artery disease

## a b s t r a c t

The aim of this study is to search the efﬁciency of binary particle swarm optimization (BPSO) and genetic

algorithm (GA) techniques as feature selection models on determination of coronary artery disease (CAD) existence based upon exercise stress testing (EST) data. Also, increasing the classiﬁcation performance of

the classiﬁer is another aim. The dataset having 23 features was obtained from patients who had per-

formed EST and coronary angiography. Support vector machine (SVM) with *k*-fold cross-validation method is used as the classiﬁer system of CAD existence in both BPSO and GA feature selection tech- niques. Classiﬁcation results of feature selection technique using BPSO and GA are compared with each other and also with the results of the whole features using simple SVM model. The results show that fea- ture selection technique using BPSO is more successful than feature selection technique using GA on determining CAD. Also with the new dataset composed by feature selection technique using BPSO, this study reached more accurate values of success on CAD existence research with more little complexity of classiﬁer system and more little classiﬁcation time compared with whole features used SVM.

© 2009 Elsevier Ltd. All rights reserved.

1. Introduction

Exercise stress test (EST) is most commonly used method on diagnosis of angina. It is a noninvasive, inexpensive, easy operable, safe and reproducible method. The main cause of angina is coro- nary artery disease (CAD) and occurs due to atherosclerosis of the cardiac arteries ([Sprigings, 2002](#_bookmark13)). Therefore, EST is one of the ﬁrst choice noninvasive diagnostic tools in the diagnosis of sus- pected CAD. Nonetheless, the relatively low sensitivity and speci- ﬁcity of EST for diagnosing CAD has led to limit its clinical usage ([San Roman, Vilacosta, Castillo, et al., 1998; Thom et al., 2006](#_bookmark13)).

Artiﬁcial intelligence techniques are commonly used in medical

diagnosis with an amazing increment day by day, and also it could

be seen in the literature.

Fuzzy discrete hidden Markov model is used to classify trans- cranial Doppler signals to predict the patients whether they are brain diseased or not ([Ug˘uz, Öztürk, Saraçog˘lu, & Arslan, 2008](#_bookmark14)). Sepehri et al. proposed a method for automated screening of con- genital heart diseases in children by means of heart sound analysis techniques. The method relies on categorizing the pathological murmurs by examining the heart sound energy over speciﬁc fre- quency bands based on the heart sections initiating those ([Sepehri](#_bookmark13)

\* Corresponding author. Tel.: +90 332 223 20 63; fax: +90 332 241 06 35.

*E-mail address:* [ibabaoglu@selcuk.edu.tr](mailto:ibabaoglu@selcuk.edu.tr) (I\_. Babaoglu).

[et al., 2008](#_bookmark13)). Tantimongcolwata et al. proposed a method for the interpretation of ischemic heart disease pattern of magnetocar- diography recordings using backpropagation neural network and direct kernel self-organizing map machine learning approaches ([Tantimongcolwata, Naennab, Isarankura-Na-Ayudhy-](#_bookmark13) [aa, Embrechtsc, & Prachayasittikula, 2008](#_bookmark13)). Least squares support vector machine and backpropagation artiﬁcial neural network methods are employed to classify the extracted features obtained from Doppler signals of the heart valve ([Comak](#_bookmark17) et al., 2007). Zhi- dong proposed noninvasive diagnosis method of coronary artery disease based on the instantaneous frequency estimation of dia- stolic murmurs and support vector machine (SVM) classiﬁer ([Zhi-](#_bookmark18) [dong, 2005](#_bookmark18)). Kurt et al. compare performances of machine learning approaches which are 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 ([Kurt, Ture, & Kur-](#_bookmark13) [um, 2008](#_bookmark13)).

Since the use of optimization and feature selection techniques, these literature studies could be more effective and less complex. Binary particle swarm optimization (BPSO) is used as a feature selection method by implementing to the data obtained by mutual information and rough set to increase the effectiveness of the SVM classiﬁer and the classiﬁcation accuracy ([Zhou, Zhou, Liu, & Zhu,](#_bookmark21) [2006](#_bookmark21)). Chuang et al. used improved BPSO in cancer-type classiﬁca- tion based on the gene expression proﬁles ([Chuang, Chang, Tu, &](#_bookmark19)

***acer***

*2018-05-21 11:40:55*

--------------------------------------------

objective

***acer***

*2018-05-21 11:42:53*

--------------------------------------------

intro

0957-4174/$ - see front matter © 2009 Elsevier Ltd. All rights reserved. doi:[10.1016/j.eswa.2009.09.064](http://dx.doi.org/10.1016/j.eswa.2009.09.064)

3178 *I*\_*. Babaoglu et al. / Expert Systems with Applications 37 (2010) 3177–3183*

[Yang, 2008](#_bookmark19)). Yang et al. proposed a feature selection and classiﬁca- tion method for hyperspectral images by combining the global optimization ability of particle swarm optimization algorithm and the superior classiﬁcation performance of a SVM ([Yang, Zhang,](#_bookmark20) [Deng, & Du, 2007](#_bookmark20)).

In this paper, SVM is used as the classiﬁer. The efﬁciency of BPSO and genetic algorithm (GA) techniques as feature selection models on determination of CAD existence based upon EST data is investigated. Also, increasing the classiﬁcation performance of the classiﬁer is aimed.

1. Materials and methods
   1. *Data acquisition*

Four hundred eighty patients who underwent EST and coronary angiography (CAG) were included in this study. A total of 23 fea- tures are obtained from EST data. Basal demographic characteris- tics, rest and peak exercise heart rate, blood pressure and

exercise time were recorded. The EST results were evaluated by two experienced cardiologists. ST segment depression and eleva- tion occurred 60 ms after the J point were recorded at each deriva- tion in peak exercise. EST was accepted as positive in case P1 mm ST depression or ST elevation in P2 contiguous leads seems. With- in the ﬁrst month following the EST, CAG was performed to all pa- tients, and the angiographic images were evaluated by two skilled cardiologists. Presence of P50% narrowing in left main coronary artery or P70% narrowing in other major epicardial coronary

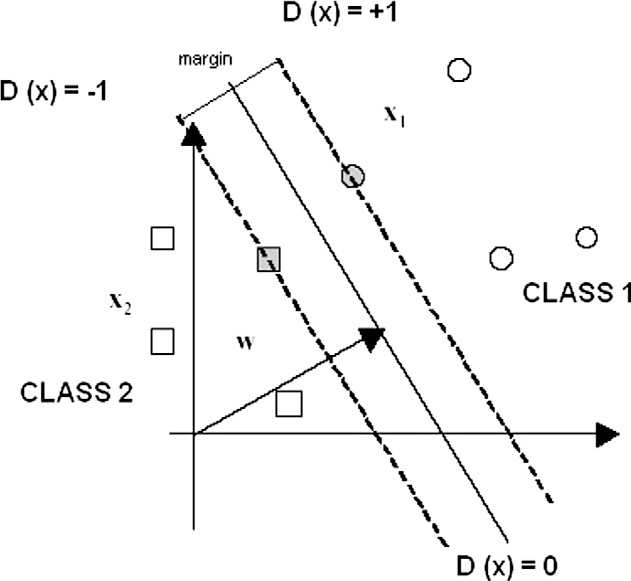


Fig. 1. SVM structure.

the optimal separating hyperplane maximizes the margin illus- trated in [Fig. 1](#_bookmark0). Problem is transformed into following dual form of quadratic optimization problem

*n n*

*i*¼1

*i*;*k*¼1

ð3Þ

arteries indicated severe CAD. Patients with bundle branch blocks

(right or, left bundle branch block), pre-excitation syndromes, at- rial ﬁbrillation, left ventricular hypertrophy and taking the digoxin

Maximize *w*ð*a*Þ¼ P *ai* — 1=2 P *aiyiakyk* ð*xi*; *xk*Þ;

*n*

were excluded from the study.

* 1. *Support vector machine*

Support vector machine has been invented by [Vapnik (1995)](#_bookmark15)

Subject to P *aiyi* ¼ 0; *ai* ≥ 8*i*:

According to *ai* Lagrange multipliers computed in [(3)](#_bookmark1), following decision function is built

*i*¼1

and proposed for classiﬁcation and regression tasks. SVM has been constructed on a strong statistical learning theory including Vap- nik–Chervonenkis dimension and structural risk minimization. Since SVM includes many reliable properties for learning and pre-

*sv*

*f* ð*x*Þ¼ *sign*

X

*i*¼1

*aiyi*ð*x*; *xi*Þþ *b*

### !: ð4Þ

sents good experimental results, it has been used in many applica- tion ﬁelds ([Kulkarni, Jayaraman, & Kulkarni, 2004; Takeuchi &](#_bookmark13) [Collier, 2003; Chen & Wang, 2007).](#_bookmark13)

* + 1. *Linear SVM classiﬁer*

Employed training data obey a form; (*x*1, *y*1), .. ., (*xn*, *yn*), *x* e *RN* and *y* e { + 1, 1}. Each data is formed with N dimensional vector and belonging only one of two classes (+1 or 1). Hyperplanes sep-

—

—

arate two classes from each other to provide following forms for all training data. Thus,

ð*w* · *xi*Þþ *w*0 ≥ þ1 — *ni*; if *yi* ¼ þ1

* + 1. *Nonlinear SVM classiﬁer*

In a nonlinear input space (including all training data), SVM fails to build optimal separating hyperplane. In this case, nonlinear in- put space is transformed into higher dimensional linear feature space via several kernel functions. A kernel function can be deﬁned as following formula

*K*ð*x*; *x*0Þ¼ ð*U*ð*x*Þ · *U*ð*x*0ÞÞ ¼ *U*ð*x*Þ*U*ð*x*0Þ: ð5Þ

Kernel functions must satisfy the Mercer’s condition. Thus, aim of quadratic optimization problem and decision function of SVM in Section [2.2.1](#_bookmark2) are transformed into following formula

*w xi*Þþ *w*0 ≤ —1 — *ni*; if *yi* ¼ —1 *n n*

ð · ð1Þ Maximize *w*ð*a*Þ¼ X *ai* — 1=2 X *aiyiakykK*ð*xi*; *xk*Þ;

¼

### or

*yi*½ð*w* · *xi*Þþ *w*0]≥ 1; *i* ¼ 1; .. . ; *n*;

X

where *ni* P 0 are slack variables and used for providing a tolerance to some data with small error. If all data satisfy [(1)](#_bookmark3) correctly, *ni* vari-

*sv*

*f* ð*x*Þ¼ *sign*

*i*¼1

*i*¼1

*aiyiK*ð*x*; *xi*Þþ *b*

*i*;*k* 1

### !: ð6Þ

ables will not be used. Optimal hyperplane among all hyperplanes is found by minimizing following formula

Commonly used kernel functions are as follows:

X 2 ● Dot product kernels: *K*ð*x*; *x*0Þ¼ *x* · *x*0. *d*

*n*

* ð Þ¼ ð · þ Þ

*C*

*ni* þ 1=2jj*w*jj ; ð2Þ

Polynomial kernels: *K x*; *x*0 *x x*0 1 ; where *d* is the degree of kernel and positive integer number.

*i*¼1

where *C* is a regularization parameter and providing a trade-off be- tween complexity and classiﬁcation performance. In other words,

* RBF kernels: *K*ð*x*; *x*0Þ¼ expð—jj*x* — *x*0jj2 =*r*2 Þ; where *r* is a positive

real number.

*I*\_*. Babaoglu et al. / Expert Systems with Applications 37 (2010) 3177–3183* 3179

* 1. *Binary particle swarm optimization*

Particle swarm optimization is an optimization technique based on swarm intelligence such as ﬁsh schooling and bird ﬂocking. The aim of this technique is to ﬁnd optimum solution in the solution

set.

lized in binary optimization problems ([Kennedy & Eberhart,](#_bookmark13) [1997](#_bookmark13)). In the BPSO technique, the probability of the particle being as 0 or 1 is speciﬁed by the velocity value using sigmoid function ([Kennedy & Eberhart, 1997](#_bookmark13)). This determination of the position is performed using the following formula

The PSO is ﬁrstly developed by Kennedy in 1995 ([Kennedy &](#_bookmark13) [Eberhart, 1995](#_bookmark13)). This technique is developed to optimize the prob-

*Xi*;*j*

*t* 1 0 *if rand*ðÞ >¼ *S*ð*mi*;*j*ð*t* þ 1ÞÞ 9

1 *if rand*ðÞ < *S*ð*mi*;*j*ð*t* þ 1ÞÞ

ð þ Þ¼ ð Þ

lems that could be solved using real numbers. Later, it is success- fully implemented in many research areas.

In PSO technique, particles are composed of cells called posi-

where *rand*() is the random numbers uniformly distributed be- tween 0 and 1. *S*(·) is the sigmoid function and it is given as follows

tion. The swarm composed from these particles separates in the solution space randomly. Every particle in the swarm is a part of the solution set. Best values of each particle (local best value –

*pbesti* , global best value – *gbest* ) in the swarm and the swarm itself

*S*ð*v*

*i*;*j*

### 1

ð*t* þ 1ÞÞ ¼ 1 þ *e*—*vi*;*j* ð*t*þ1Þ : ð10Þ

;*j i*;*j*

are accumulated to be used in the next step and also to obtain opti- mum values. The velocity and the position of the particle are calcu- lated as follows

*vi*;*j* ð*t* þ 1Þ¼ *wvi*;*j* ð*t*Þþ *c*1*R*1ð*pbesti*;*j* — *xi*;*j*ð*t*ÞÞ þ *c*2*R*2ð*gbesti*;*j* — *xi*;*j*ð*t*ÞÞ; ð7Þ

*xi*;*j*ð*t* þ 1Þ¼ *xi*;*j*ð*t*Þþ *vi*;*j* ð*t* þ 1Þ; ð8Þ

where *i* is the index of particle, *j* is the index of position in particle, *t* shows the iteration number, *mi*,*j*(*t*) is the velocity of the *i*th particle in swarm on *j* th index of position in particle *v*min 6 *vi*,*j*(*t*) 6 *v*max

and *xi*,*j*(*t*) is the position. *R*1 and *R*2 are the random numbers uni- formly distributed between 0 and 1. *c*1 and *c*2 are the acceleration numbers and default values 2 and *w* is the inertia weight and is usu- ally used less than 1. PSO is brieﬂy illustrated in [Fig. 2](#_bookmark5).

Binary particle swarm optimization is introduced in 1997 ﬁrstly by Kennedy and Eberhart. Like GA, BPSO could be effectively uti-

* 1. *Genetic algorithm*

Based on long-term observation, Darwin asserted his theory of natural evolution. According to this theory, living beings compete with each other to survive. At the end of this competition, the suc- cessful beings transfer their genes to the beings in the next generation.

Inspiring by Darwin’s evolution theory, Genetic algorithm was ﬁrst introduced by Holland as a powerful computational model in 1975 ([Holland, 1975](#_bookmark23)). It is commonly used for optimization problems which could take discontiguous or continuous values. Prime aim of GA is to ﬁnd optimum solution within the potential solution set. Each solution set is called as population. Populations are composed of vectors, namely, chromosome or individual. Each item in the vector is called as gene. The structure of the GA is given in [Fig. 3](#_bookmark6) and described as follows.

# 

Fig. 2. The PSO algorithm. Fig. 3. The structure of the GA.

3180

*I*\_*. Babaoglu et al. / Expert Systems with Applications 37 (2010) 3177–3183*



Fig. 4. A structure of the particle.

* + 1. *Initialization*

At the beginning of the process, each chromosome is randomly composed within the values 0 and 1. These chromosomes are indi- vidual solutions in the potential solution set.

* + 1. *Evaluation*

In the optimization process with GA, the proper degree of each chromosome in the potential solution set is calculated using ﬁtness function. One of the most important factors that effect the achieve- ment of GA is the selection of the ﬁtness function. Thus, ﬁtness function must be adopted effectively focused to the solution.

* + 1. *Selection*

This is the process of the selection of the chromosomes to trans- fer their genes to the next generation considering ﬁtness values. Selection process is implemented using roulette wheel selection technique. The probability of the selection of each chromosome using roulette wheel selection technique is calculated as follows

—*f*—*i*

Population includes 10 particles, and the values of these parti- cles are taken as 0 at the beginning of the process. In the optimiza- tion process, training and test sets are composed considering the features deﬁned by the particles, and SVM is trained and tested using mentioned datasets. As a result, classiﬁcation accuracy rate, training error rate and the sum of the times elapsed for training and test processes are obtained for each particle. The success rate of each particle is calculated using the following ﬁtness function formula

### ð12Þ

*f* ð*i*Þ¼ *A*ð*i*Þ— *E*ð*i*Þ;

where *f*(*i*) is the success rate, *A*(*i*) is the classiﬁcation accuracy rate and *E*(*i*) is the training error rate of *i*th particle. Velocity of the par-

ticles is calculated using [(7)](#_bookmark4) and *v*min and *v*max are used as 6 and 6,

—

respectively. The value of the cells within each particle is updated using (8). For each iteration, *pbesti* and *gbest* is updated if necessary.

;*j i*;*j*

At the end of the optimization process, *gbesti* is found as the opti-

;*j*

mum solution.

*3.2. Feature selection technique using GA (GA–FST) architecture*

### 11Þ

*P*ð*ci*Þ ¼ P*N f i* ; ð

*j*¼1 *i*

where *ci* is the chromosome in question, *fi* is the ﬁtness value of the chromosome, *N* is the number of the chromosomes in the population.

* + 1. *Crossover*

The chromosome pairs are crossed over to generate the chro- mosomes in the next generation using a predetermined crossover rate. Crossover is the process of replacement of one or more seg- ments which are selected randomly.

* + 1. *Mutation*

Mutation process is utilized to enhance the variation of the pop- ulation. Value of the each gene in the chromosomes is changed considering mutation rate.

* + 1. *Termination*

The selection, crossover and mutation processes are repeated to the end of the iteration. The algorithm is also terminated in the sit- uation of obtaining desired ﬁtness value.

1. Proposed feature selection methods
   1. *Feature selection technique using BPSO (BPSO-FST) architecture*

In the BPSO–FST, each particle is composed of 23 binary cells, which refer to the whole features in the dataset. The value of these cells shows whether the feature that the cell refers to would be se- lected. A cell value of 1 shows the feature that the cell refers to is selected, a cell value of 0 shows the feature that the cell refers to is not selected into the dataset. The structure of the particle is given below, where *n* is the number of the features, *x* is the value of the

cell and *x* e {0, 1}, and *v* denotes the velocity of the cell.

BPSO is iterated 200 times to ﬁnd the optimum solution set

for each value of the parameters *c* and *c* that effect SVM’s

performance.

In the GA–FST, each chromosome is composed of 23 genes,

which refers to the whole features in the dataset similarly particles used in BPSO–FST. The value of these genes shows whether the fea- ture that the gene refers to would be selected. A gene value of 1 shows the feature that the gene refers to is selected, a gene value of 0 shows the feature that the gene refers to is not selected into the dataset.

The same SVM kernel parameters ranges are used in the GA–FST to compare with BPSO–FST. Also, GA–FST is implemented for 200 iterations.

The size of the population, the mutation rate and the crossover

rate is used as 10, 0.05 and 0.25, respectively, in this study. At the beginning of the process, the values of each chromosome are used as 0. In the optimization process, training and test sets are com- posed considering the features deﬁned by the chromosomes, and

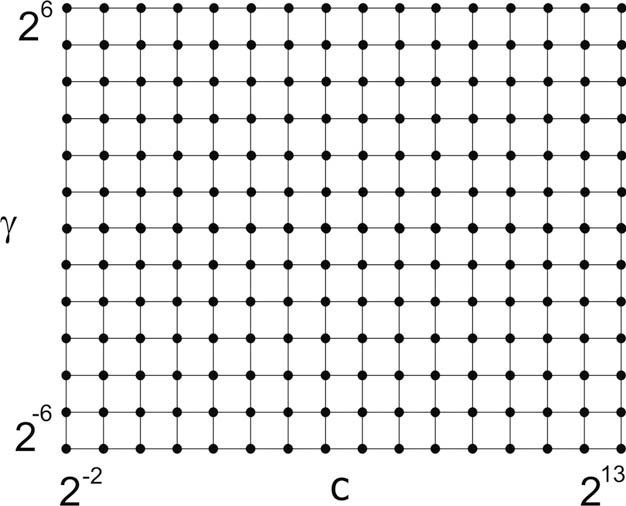


Fig. 5. SVM parameter search grid.

*I*\_*. Babaoglu et al. / Expert Systems with Applications 37 (2010) 3177–3183* 3181

SVM is trained and tested using mentioned datasets. As a result, classiﬁcation accuracy rate, training error rate and the sum of the times elapsed for training and test processes are obtained for each chromosome. The success rate of each chromosome is calculated

Table 1

Test results.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Method | *c* | *c* | ERR | (%) | ACC (%) | TIME | (s) | NoF |
| BPSO–FST | 4 | 8 | 0.26 | 81.46 | | 0.52 | 11 | |
| GA–FST | 512 | 4 | 0.00 | 79.17 | | 0.53 | 12 | |
| SVM | 8 | 1 | 0.31 | 76.67 | | 0.67 | 23 | |

Method, Method used in classiﬁcation; *c*, RBF kernels parameter; *c*, RBF kernels parameter; ERR, Training error; ACC, Classiﬁcation accuracy; TIME, Sum of the time elapsed in training and test processes of SVM classiﬁcation; NoF, Number of fea- tures selected by the feature selection method; BPSO–FST, SVM classiﬁcation uti- lizing BPSO–FST; GA–FST, SVM classiﬁcation utilizing GA–FST; SVM, Simple SVM classiﬁcation.

using [(12)](#_bookmark7). Chromosomes are selected using roulette wheel com- posed according to the success rates mentioned. Chromosomes are crossed over and mutated according to the given crossover and mutation rates. At the end of the process, the chromosome that has the optimum ﬁtness value is found as the optimum solution.

1. Results and discussion

[Fig. 4](#_bookmark8) Datasets are implemented by normalizing into the range [ 1, 1]. Radial basis function (RBF) kernel, which is commonly used

—

for SVM, is employed for the default kernel. The values of the RBF

kernel parameters *c* and *c* are found using grid search algorithm ([Cormen, Leiserson, Rivest, & Stein, 2001](#_bookmark22)) and both of the values of these parameters are used as 2*n*. The search grid used in this study is diagrammatized in [Fig. 5](#_bookmark9). In this technique, *n* is used in the range [—2, 13] and [—6, 6] for *c* and *c*, respectively.

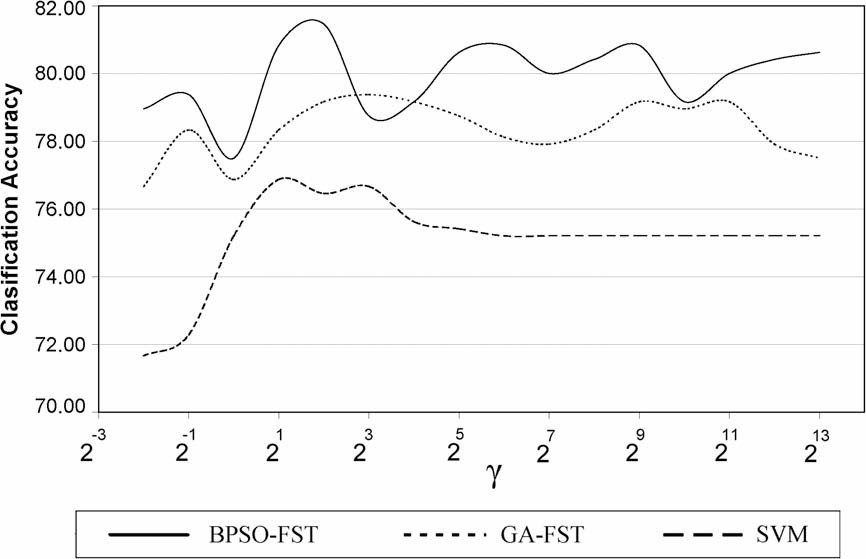


Fig. 6. Classiﬁcation accuracy graph depend on *c*. The values of the ﬁxed *c* parameters are 4, 512 and 8 for BPSO–FST, GA–FST and simple SVM classiﬁcation model, respectively. BPSO–FST, SVM classiﬁcation utilizing BPSO–FST; GA–FST, SVM classiﬁcation utilizing GA–FST; SVM, Simple SVM classiﬁcation.

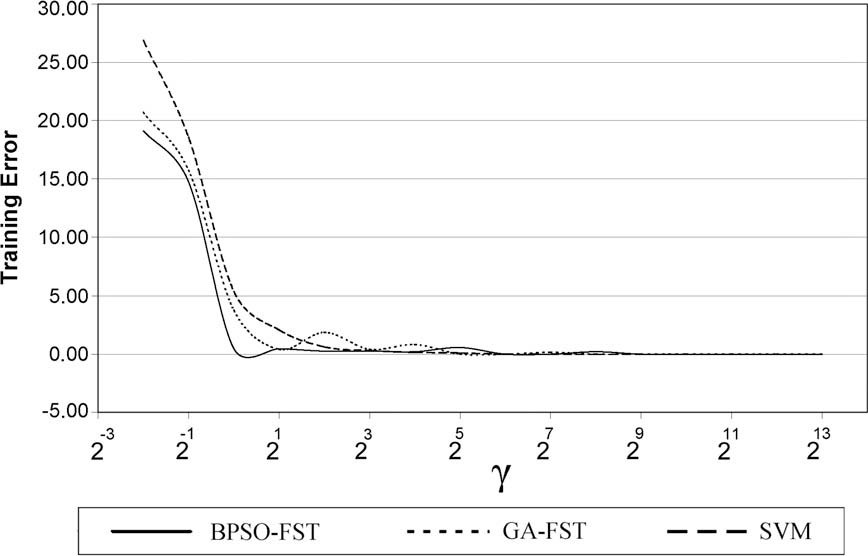


Fig. 7. Training error graph depend on *c*. The values of the ﬁxed *c* parameters are 4, 512 and 8 for BPSO–FST, GA–FST and simple SVM classiﬁcation model, respectively. BPSO– FST, SVM classiﬁcation utilizing BPSO–FST; GA–FST, SVM classiﬁcation utilizing GA–FST; SVM, Simple SVM classiﬁcation.

3182 *I*\_*. Babaoglu et al. / Expert Systems with Applications 37 (2010) 3177–3183*

The results of the BPSO-FST, the results of the GA–FST and the results of the simple SVM classiﬁcation are compared using the optimum *c* and *c* parameters obtained for each technique.

Classiﬁcation accuracy, training error and the sum of the train- ing and test time is considered ﬁnding the best classiﬁcation archi- tecture. k-Fold cross-validation method is employed using *k* = 5 to improve the reliability of the results ([Shao, 1993; An, Liu, & Venk-](#_bookmark13) [atesh, 2007](#_bookmark13)).

The implementations are employed using LIBSVM package ([Chang & Lin, 2001](#_bookmark16)) in the Matlab 7.0 application platform in a computer with Pentium IV with a 3.2 GHz CPU and 1 GByte memory.

The optimum parameters *c* and *c* are found as 4 and 8, respec-

tively, in the BPSO–FST. The optimum particle has 11 features se- lected, and the sum of the training and test times is 0.52 s.

The optimum parameters *c* and *c* are found as 512 and 4,

respectively, in the GA–FST. The optimum particle has 12 features selected, and the sum of the training and test times is 0.53 s.

The optimum parameters *c* and *c* are found as 8 and 1, respec- tively, in the simple SVM classiﬁcation technique. Whole 23 fea- tures are used in this technique, and the sum of the training and test times is 0.67 s.

The values of the optimum parameters and the result of the SVM classiﬁcation process for all techniques are given in [Ta-](#_bookmark10) [ble 1](#_bookmark10).

In each technique, the classiﬁcation accuracy and training er- ror graphs depend on the *c* parameter of SVM for the results based on ﬁxed *c* parameters in the optimum solutions are given in [Figs. 6 and 7](#_bookmark11). The classiﬁcation accuracy and training error graphs depend on the *c* parameter of SVM for the results based on ﬁxed *c* parameters in the optimum solutions are given in [Figs.](#_bookmark12) [8 and 9](#_bookmark12).

As seen in [Figs. 6–9](#_bookmark11), BPSO–FST has highest classiﬁcation accu- racy rate considered to GA–FST and simple SVM classiﬁcation tech- nique for each *c* and *c* parameters within the search grid. Sum of the training and test times depends on number of the features used in SVM classiﬁcation process. Hence, BPSO–FST has the minimum

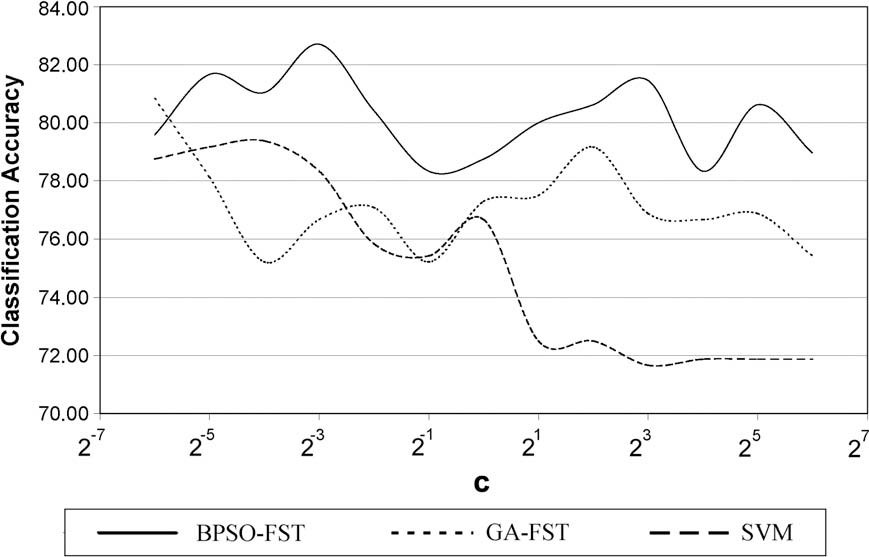


Fig. 8. Classiﬁcation accuracy graph depend on *c*. The values of the ﬁxed *c* parameters are 8, 4 and 1 for BPSO–FST, GA–FST and simple SVM classiﬁcation model, respectively. BPSO–FST, SVM classiﬁcation utilizing BPSO–FST; GA–FST, SVM classiﬁcation utilizing GA–FST; SVM, Simple SVM classiﬁcation.

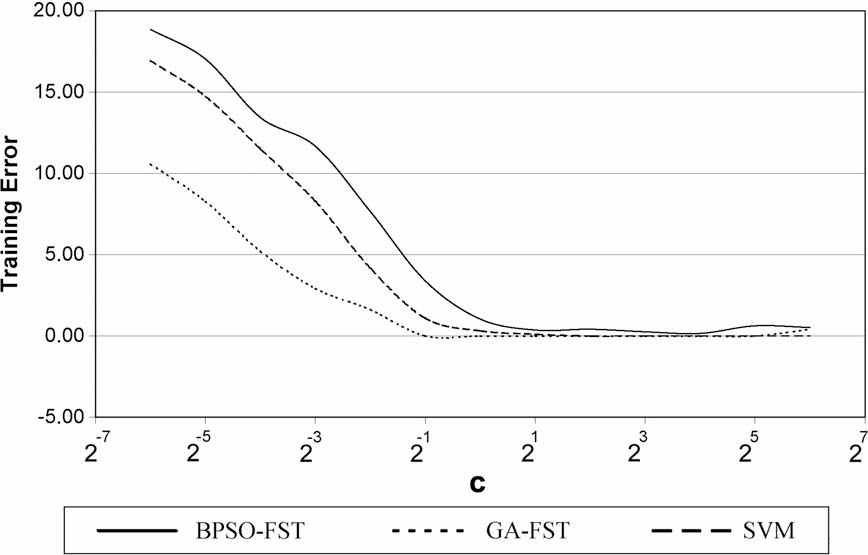


Fig. 9. Training error graph depend on *c*. The values of the ﬁxed *c* parameters are 8, 4 and 1 for BPSO–FST, GA–FST and simple SVM classiﬁcation model, respectively. BPSO– FST, SVM classiﬁcation utilizing BPSO–FST; GA–FST, SVM classiﬁcation utilizing GA–FST; SVM, Simple SVM classiﬁcation.

*I*\_*. Babaoglu et al. / Expert Systems with Applications 37 (2010) 3177–3183* 3183

feature size, so that the sum of the training and test times is also minimum compared to the other classiﬁcation techniques in addi- tion to providing the best classiﬁcation accuracy.

1. Conclusion

Instead of using the whole features in the dataset, SVM classiﬁ- cation process is implemented employing the reduced dataset on

the determination of CAD using EST data. The dataset’s dimension is reduced by utilizing BPSO–FST or GA–FST. The classiﬁcation pro- cess implemented by feature selection techniques achieves more successful classiﬁcation accuracy. Besides, they decrease the com- plexity of the system by reducing the dimensions of the dataset. Classiﬁcation processes implemented by using mentioned tech- niques are compared to each other. The classiﬁcation process implemented by utilizing BPSO–FST has the best classiﬁcation

accuracy and minimal process time compared to others.

References

An, S., Liu, W., & Venkatesh, S. (2007). Fast cross-validation algorithms for least squares support vector machine and kernel ridge regression. *Pattern Recognition, 40*, 2154–2162.

Chang, C. C., & Lin, C. J. (2001). LIBSVM: A library for support vector machines, available: [<http://www.csie.ntu.edu.tw/~cjlin/libsvm>](http://www.csie.ntu.edu.tw/~cjlin/libsvm).

Chen, K. Y., & Wang, C. H. (2007). 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.

Chuang, L. Y., Chang, H. W., Tu, C. J., & Yang, C. H. (2008). Improved binary PSO for feature selection using gene expression data. *Computational Biology and Chemistry, 32*, 29–38.

Comak, E., Arslan, A., & Turkoglu, I. (2007). A decision support system based on support vector machines for diagnosis of the heart valve diseases. *Computers in Biology and Medicine, 37*, 21–27.

Cormen, T. H., Leiserson, C. E., Rivest, R. L., & Stein, C. (2001). *Introduction to algorithms* (2nd ed.). MIT Press.

Holland, J. (1975). *Adaptation in natural and artiﬁcial systems*. The Michigan University Press.

Kennedy, J., & Eberhart, R. C. (1995). Particle swarm optimization. In *Proceedings of IEEE international conference on neural networks* (vol. 4; pp. 1942–1948).

Kennedy, J., & Eberhart R. C. (1997). A discrete binary version of the particle swarm algorithm. In *Proceedings of 1997 conference systems man and cybernetics* (pp. 4104–4108).

Kulkarni, A., Jayaraman, V. K., & Kulkarni, B. D. (2004). Support vector classiﬁcation with parameter tuning assisted by agent-based technique. *Computers and Chemical Engineering, 28*, 311–318.

Kurt, I., Ture, M., & Kurum, A. T. (2008). Comparing performances of logistic

regression, classiﬁcation and regression tree, and neural networks for

predicting coronary artery disease. *Expert Systems with Applications, 34*, 366–374.

San Roman, J. A., Vilacosta, I., Castillo, J. A., Rollan, M. J., Hernandez, M., Peral, V., et al. (1998). Selection of the optimal stress test for the diagnosis of coronary artery disease. *Heart, 80*(4), 370–376.

Sepehri, A. A., Hancq, J., Dutoit, T., Gharehbaghi, A., Kocharian, A., & Kiani, A. (2008). Computerized screening of children congenital heart diseases. *Computer methods and programs in biomedicine, 92*(2), 186–192.

Shao, J. (1993). Linear model selection by cross-validation. *Journal of American Statistical Association, 88*, 486–494.

Sprigings, D. (2002). Exercise stress testing. *Medicine, 30*(3), 31–33.

Takeuchi, K., & Collier, N. (2003). Bio-medical entity extraction using support vector machines. *Artiﬁcial Intelligence in Medicine, 33*(2), 125–137.

Tantimongcolwata, T., Naennab, T., Isarankura-Na-Ayudhyaa, C., Embrechtsc, M. J., & Prachayasittikula, V. (2008). Identiﬁcation of ischemic heart disease via machine learning analysis on magnetocardiograms. *Computers in Biology and Medicine, 38*, 817–825.

Thom, T., Haase, N., Rosamond, W., Howard, V. J., Rumsfeld, J., Manolio, T., et al. (2006). Heart disease and stroke statistics – 2006 update: A report from the American Heart Association Statistics Committee and Stroke Statistics Subcommittee. *Circulation, 113*, e85–e151.

Ug˘uz, H., Öztürk, A., Saraçog˘lu, R., & Arslan, A. (2008). A biomedical system based on fuzzy discrete hidden Markov model for the diagnosis of the brain diseases. *Expert Systems with Applications, 35*, 1104–1114.

Vapnik, V. (1995). *The nature of statistical learning theory*. New York: Springer.

Yang, H. C., Zhang, S. B., Deng, K. Z., & Du, P. J. (2007). Research into a feature selection method for hyperspectral imagery using PSO and SVM. *Journal of China University of Mining and Technology, 17*(4), 473–478.

Zhidong, Z. (2005). 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).

Zhou, W., Zhou, C., Liu, G., & Zhu, H. (2006). *Feature selection for microarray data analysis using mutual information and rough set theory*. Boston: Springer.