International Journal of
Intelligent Systems and
Applications in Engineering
ISSN:2147-6799 www.atscience.org/IJISAE
Advanced Technology and Science
Original Research Paper

Feature Selection on MR Images Using Genetic Algorithm with SVM and Naive Bayes Classifiers

Nihat Adar[1], Savaş Okyay*[1], Kemal Özkan[1], Suzan Şaylısoy[2], Belgin Demet Özbabalık Adapınar[3], Baki Adapınar[2]

Accepted 3rd September 2016

Abstract:Dementias are termed as neuropsychiatric disorders. Brain images of dementia patients can be obtained through magnetic resonance imaging systems. The relevant disease can be diagnosed by examining critical regions of those images. Certain brain characteristics such as the cortical volume, the thickness, and the surface area may vary among dementia types. These attributes can be expressed as numerical values using image processing techniques. In this study, the dataset involves T1 medical image sets of 63 samples. Each particular sample is labeled with one of the three dementia types: Alzheimer's disease, frontotemporal dementia, and vascular dementia. The image sets are processed to create different feature groups. These are cortical volumes, gray volumes, surface areas, and thickness averages. The main objective is seeking brain sections more effective in establishing the clinical diagnosis. In other words, searching an optimal feature subset process is carried out for each feature group. To that end, a wrapper feature selection technique namely genetic algorithm is used with Naive Bayes classifier and support vector machines. The test phase is performed by using 10-fold cross validation. Consequently, accuracy results up to 93.7% with different classifiers and feature selection parameters are shown.

Keywords:*Genetic algorithm, Feature selection, Dementia, Classification, Magnetic resonance imaging.*

1. Introduction

Dementia diseases are termed as neuropsychiatric disorders, and encounter rate of them increases severely with age[1].Brains of patients with dementia denote several differences in some ways such as cortical volume, thickness, surface area, according to disease type. The most common three dementia types may be sorted as Alzheimer's disease (AD), vascular dementia (VaD) and frontotemporal dementia (FTD) respectively. In neuroimaging science, an unknown diagnosis is tried to be determined via medical images. Magnetic resonance imaging (MRI) is a technique interprets the anatomy. The T1 pulse sequence is one of the basic sequences in MRI that achieving remarkable tissue contrast and provides a good correlation when larger amounts of iron are present[2].2-dimensional brain images in slices can be acquired with MRI. These sliced output files have *digital imaging and communications in medicine* (dicom) file extension frequently, and project brain to a particular axis. Besides, aforementioned files contain patient information. It is possible to reach measurements of brain regions using those medical imaging files with image processing techniques. In theory, unknown diseases can be labeled with the aid of classification algorithms taking numerical expressions of brain sections as input parameters. Meanwhile, some features may be qualified as more precious.

In the beginning of this research, some classification tests were performed with taking all extracted brain features as input set. The accuracy results were not satisfying at all. Therefore, finding the lowest size feature set having high precision classification result constitutes the motivation of the study. Briefly in this study, samples with three dementia types are tested to seek a valuable feature subset with genetic algorithm (GA) based wrapper feature selection method over different classification algorithms.

1.1. Literature Review

When it comes to the literature research of interdisciplinary brain imaging studies, there are numerous indicative papers. The number of the studies using free-access datasets such as Alzheimer's Disease Neuroimaging Initiative (ADNI), AddNeuroMed, etc. is quite high. The number of the studies using their own dataset is also a great deal more. In addition, some software tools are utilized for operations such as the construction of the virtual brain, feature extraction, etc. in computer science studies.

In 2011, Freesurfer v4.5.0 brain analyzing software tool was used for analyzing the brains of 295 AD, 444 mild cognitive impairment (MCI), 335 control subjects from the ADNI database. 23 regional volume and 34 cortical thickness features for 1074 MRI scans were used for the orthogonal partial least squares (OPLS) classification[3]. In another research, the performance of different methods through several brain regions was compared upon the features obtained from T1 MRI scans after the Freesurfer process. 509 AD, MCI or elderly control individuals

[1] *Department of Computer Engineering, Eskişehir Osmangazi University, Eskişehir, Turkey*
[2] *Department of Radiology, Eskişehir Osmangazi University, Eskişehir, Turkey*
[3] *Department of Neurology,Eskişehir Osmangazi University, Eskişehir, Turkey*
** Corresponding Author:Email: osavas@ogu.edu.tr*
Note: This paper has been presented at the 3rd International Conference on Advanced Technology & Sciences (ICAT'16) held in Konya (Turkey), September 01-03, 2016.

aged between 55 and 90 from ADNI database were processed. Features were extracted and grouped as voxel-based (grey matter, white matter and cerebrospinal fluid in given voxel), vertex-based (the features are defined on the cortical surface) or regions of interest-based (includes only the hippocampus). The accuracy of these approaches was reported over 84%[4]. An alternative study classified MRI samples of 524 AD, MCI, and control subjects from ADNI database over various techniques some like logistic regression, support vector machine (SVM), radial basis function, and C4.5 tree learner. Freesurfer v4.3 handled feature extraction from T1 MRI scans. 328 numerical expressions per subject, the whole variable set computed from the tool, moved to the next stage. In detail, volumetric values were normalized. For feature selection, a filter method independent of any classifier was executed. 10-fold cross-validation testing technique over classifiers was realized. At that phase, above-stated classifiers took input from not only MRI features, but also age, gender, years of education, and the number of ApoEε4 alleles were performed with the algorithms. Average performance for control samples v AD classification with SVM was reported as 89.17% ±5.08. Besides, features that were common to all classification tasks were listed as age, the number of ApoE ε4 alleles, right and left hippocampal, left entorhinal cortex, left amygdale volumes, and average cortical thickness of the left middle temporal cortex in a comprehensive manner[5]. In 2015, Freesurfer v5.1 was used to extract cortical thickness, subcortical volume, and white matter integrity features from 27 subjective memory impairment (SMI), 18 MCI, and 27 AD patients of Pusan National University Hospital. SVM eliminated features recursively. Thus, subjects were classified using nonlinear SVM. The procedure was repeated 1000 times, and 2-class based classification accuracy was reported as various percentages between %84.4 and %96.3, in other 3-class based case, performance was %70.5 (±11.5)[6].In 2013, MR images of 345 people from the AddNeuroMed cohort were analyzed with FreeSurfer v4.5.0. The dataset contained 116 AD, 119 MCI, and 110 control samples. 34 cortical thickness and 23 volumetric normalized MRI measures were the inputs for multivariate analysis. Tests were performed with different techniques. Accuracies of the control samples v AD classification were reported between 81.4% and 88.1% for various additional input parameters like age, years of education, etc. Significant features that were chosen in all techniques were listed as hippocampus, amygdala, entorhinal cortex, inferior lateral ventricles, cerebrospinal fluid, inferior, superior and middle temporal gyri and temporal pole [7].

1.2. Paper Organization

In Section 1, an introduction to the research and literature review are presented. In Section 2, the dataset is described. In Section 3, the methodology of this research is explained, additionally in the subsections, feature extraction, classification, feature selection, and testing methods are noticed succinctly. In the following section, tests and results are reported. Finally in the last section, the conclusion is made and future plans are mentioned.

2. Dataset

The dataset used in this research belongs to the picture archiving communication systems (PACS) of Eskişehir Osmangazi University's Radiology Department.

The set consists of 63 dementia patients namely AD, FTD and VaD. Each sample is labeled only with one of these three clinical diagnoses. Samples are either male or female and ages of them vary between 50 and 90. Image sets were obtained from *Discovery MR750w* (GE, Milwaukee) and *Magnetom Vision plus* (Siemens, Erlangen) MRI systems last 2 years. Counts of dementia types over genders and MRI systems are shown in Table 1.

Table 1. Distribution of Dementia Types over Genders and MRI Systems

Dementia Type	***Male***	***Female***	***Discovery MR750w***	***Magnetom Vision plus***	***All***
AD	5	14	13	6	19
FTD	6	13	13	6	19
VaD	12	13	17	8	25
Total	23	40	43	20	**63**

In this work, T1 weighted image sets are used. Each sliced T1 weighted scan has MR modality and 2-dimensional MR acquisition type. Slice thicknesses of all image sets are between 4.5-5 millimeters. Also, all images are grayscale at least 256×256 image resolution and have dicom file extension.

3. Methodology

The basic idea of this study is achieving the optimal feature subset having the highest accuracy. The method incorporates several consecutive operations. T1 weighted image sets, obtained from brain imaging studies via MRI systems, are preprocessed with brain modeling software. Freesurfer brain analyzing tool is used for preprocessing the sliced medical images and feature extraction from the virtual brain. After then, feature matrices in different measurements are created. Feature selection algorithm is applied to each feature group separately to determine the subset having high classification accuracy result. For this purpose, GA based wrapper feature selection algorithm over different classifiers is performed. Tools and algorithms mentioned in the following subsections are put into practice to obtain results.

3.1. Feature Extraction

In neuroimaging science studies, software tools used often as exemplified in the literature section. In this work, FreeSurfer v5.3.0 was used for medical image processing and analysis. Freesurfer is a functional, connectional and structural human brain analyzing tool that comprises a set of image processing, numeric, etc. algorithms[8]. In a few words, the tool associates sliced medical imaging files structurally with the help of header information. Later on, the virtual 3-dimensional brain is modeled by following image processing techniques iteratively. For each individual sample, the same procedure is repeated. Working principle of Freesurfer can be summarized as three major steps basically. Firstly, correction and verification operations for input files are processed. Secondly, volumetric registrations, removing neck actions, white matter segmentations, and also some visual smoothing transactions occur at this step. At the end of the last step, containing spherical instructions followed by cortical parcellation, brain modeling is completed[9], [10].During the procedure, data is transferred the way that the output of each sub-step will be the input of another step, analysis process continues gradually. When the procedure is completed with success, visual and statistical files become accessible. Depending on computer hardware, the entire Freesurfer process described may take quite some time.(Cuignet et al; 2011) reported that a single Freesurfer modeling could take roughly a day [4]. Likewise (Gronenschild et

al; 2012)uttered that this whole process found time to be completed in 30 hours and what is more, analysis may be affected by tool versions or operating system differences [11]. In this research, Freesurfer operations over T1 medical images were executed with a computer having Intel® Core™ i7-4700 2.40 GHz CPU, 1600 MHz 16 GB RAM hardware, and Ubuntu 14.04 x64 operating system. Multiple Freesurfer processes run parallelly for groups of samples. In detail, 3-dimensional modeling of brain structure for each sample was accomplished at the end of approximately 15 hours time period. No manual editing was used after analysis.

In the *Literature Review* section, some of the significant features are listed. Accordingly, from the Freesurfer statistics files, *cortical volume* features for whole brain structure, moreover, *gray volume*, *surface area* and *thickness average* features for left and right parts of the brain are taken into consideration in this study. Values are exported via bash scripts. Lastly, each feature is normalized to a length of 1 while creating feature matrices.

3.2. Classifiers

3.2.1. Naive Bayes Classifier

Characteristics for each class are thought independently. Theoretically, features are supposed to be uncorrelated by implementing the Bayes' rule with naive independence assumptions. Conditional probabilities are estimated for all classes. The sample, whose class label is unknown, is labeled with the name of the class having maximum probability[12].Naive Bayes (NB) method is also known as a conditional classifier.

3.2.2. Support Vector Machines

SVM is based on the methodology of *perceptron algorithm*. The principle of SVM is searching for a line that separates the plane into two classes. With using kernel functions, the algorithm tries to maximize the distance between the optimal hyperplane and nearest support vectors[13]. Furthermore, *one to all* strategy may be preferred for multi-class problems in like manner. This algorithm is also known as a maximum margin classifier.

3.3. Feature Selection

In classification studies, when all features are used as inputs, the input set may contain insignificant features and consistent results may not be achieved. Classification results can be enhanced using the worthwhile feature subset. In this work, GA was used for eliminating insignificant features from whole feature set. The principle of the algorithm is based on the disappearance of weak genes by natural selection and survival of the best ones evolving from one generation to another during biological adaptation[14], [15].In mathematical approach, GA searches an optimal solution for a problem from one set to another. In the methodology, the solution comes from a set of numerical values namely chromosomes. Populations are configured to any constant size and contain a set of various chromosomes. A fitness function is needed to compute the cost when a particular chromosome is selected for a possible solution to a problem. This function calculates the fitness cost of each chromosome in the population. Finding an optimal solution is performed in generation cycles. In each generation, the search is directed toward to find the best solution until then. The algorithm transfers better chromosomes to build new generation population iteratively. Sometimes, elites having the best fitness are preferred to be placed in new generation directly. Likewise, reproduction functions such as *crossover*, *selection*, *mutation*, are used for creating next generations. Crossover function produces two child chromosomes that are synthesized by parents. Selection function describes which individuals are chosen for the next level. Mutation function changes chromosome parts randomly[14], [16]. In this study, the main objective of using GA is seeking an optimal feature subset by wrapper feature selection approach via classification accuracy for selected chromosome. Chromosomes specify which features are paid attention to in bit strings. In other words, chromosomes act like selection masks. Related bit indices in the living chromosome, that are equal to 1, are entitled to be in the sub-feature set. After then, feature subset takes place as the input of fitness cost function. From this point of view, the fitness cost function is defined as the accuracy result of the selected features in the classification process.

3.4. Cross Validation

k-fold cross validation technique carries out to compute overall classification performance. *k*-fold cross validation term implies that 1/*k* of the whole dataset is chosen as a test set, and the rest is chosen as a train set automatically for all *k* rounds. It is ensured that each sample becomes a test sample just once. In this study, before GA was implemented, the dataset had been partitioned into 10-folds randomly. Each test run started with the same folds. Classification performances were based on 10-fold cross-validation for NB & SVM classifiers. Eventually, test performances are reported as the percentage of correctly classified cells in confusion matrices.

4. Tests and Results

The testing phase proceeded in two different ways. According to the number of input classes, 3-Class and 2-Class GA based wrapper feature selection tests over NB and SVM classifiers were performed.

Table 2. 3-Class Genetic Algorithm Results with Naive Bayes Classifier

Feature Group	*Total Feat. Count*	*Best Feat Subset Count*	*Acc %*	*Confusion Matrix*
Cortical Volumes	41	17	84.1	$\begin{bmatrix}17 & 0 & 2\\ 1 & 16 & 2\\ 2 & 3 & 20\end{bmatrix}$
Left Gray Volumes	34	17	69.8	$\begin{bmatrix}10 & 5 & 4\\ 0 & 18 & 1\\ 4 & 5 & 16\end{bmatrix}$
Left Surface Area	34	16	74.6	$\begin{bmatrix}11 & 5 & 3\\ 1 & 14 & 4\\ 2 & 1 & 22\end{bmatrix}$
Left Thickness Average	34	11	69.8	$\begin{bmatrix}11 & 2 & 6\\ 4 & 11 & 4\\ 3 & 0 & 22\end{bmatrix}$
Right Gray Volumes	34	16	71.4	$\begin{bmatrix}7 & 4 & 8\\ 0 & 16 & 3\\ 0 & 3 & 22\end{bmatrix}$
Right Surface Area	34	17	63.5	$\begin{bmatrix}10 & 8 & 1\\ 4 & 12 & 3\\ 4 & 3 & 18\end{bmatrix}$
Right Thickness Average	34	9	73	$\begin{bmatrix}11 & 1 & 7\\ 1 & 14 & 4\\ 2 & 2 & 21\end{bmatrix}$

In 3-Class tests, for each feature group, best feature subsets that

classify all dementia types were found. To put it another way, the whole dataset was analyzed. Also in 2-Class tests, features that separate particular dementia type from other types with high classification accuracy results were examined. In other words, these tests were performed as one versus the others. For both tests, some GA parameters were set distinctly in order to arrange operation time considering the input class count.

GA parameters of 3-Class tests were chosen as 10% for the*elite count*, 90% for *crossover fraction*. *Roulette* selection function, *single point* crossover function and *uniform* mutation function with 13% rate created new generations. 3-Class (*AD*, *FTD*, *VaD*) classification tests performed over NB classifier with 800 for *population size* and the initial population was created randomly in *bit string* chromosome type. Algorithm worked with maximum 1000 *generations*. With given parameters, genetic algorithm results of 3-Class tests are shown in Table 2.

Before the 2-Class tests began, samples, that did not belong to the examined class, were labeled as *others* (Oth). For each dementia type, this procedure was repeated. Afterwards, GA parameters of 2-Class tests were chosen as 10% for the *elite count*, 90% for*crossover fraction*. *Roulette* selection function, *single point* crossover function and *uniform* mutation function with 13% rate created new generations. 2-Class (*AD v Others*, *FTD v Others*, *VaD v Others*) classification tests performed over NB and SVM classifiers with 400 for *population size* and theinitial population was created randomly in *bit string* chromosome type. Algorithm worked with maximum 500 *generations*. With given parameters, genetic algorithm results of 2-Class tests with NB and SVM are shown in Table 3 and Table 4 respectively.

Table 3. 2-Class Genetic Algorithm Results with Naive Bayes Classifier

Feature Group	*Total Feat. Count*	*Classes*	*Best Feat Subset Count*	*Acc %*	*Confusion Matrix*
Cortical Volumes	41	ADvOth	18	92.1	[18 1;4 40]
		FTDvOth	17	90.5	[16 3;3 41]
		VaDvOth	24	87.3	[21 4;4 34]
Left Gray Volumes	34	ADvOth	18	88.9	[16 3;4 40]
		FTDvOth	17	90.5	[19 0;6 38]
		VaDvOth	11	88.9	[20 5;2 36]
Left Surface Area	34	ADvOth	17	79.4	[12 7;6 38]
		FTDvOth	17	88.9	[14 5;2 42]
		VaDvOth	10	88.9	[19 6;1 37]
Left Thickness Average	34	ADvOth	18	88.9	[16 3;4 40]
		FTDvOth	17	85.7	[14 5;4 40]
		VaDvOth	23	82.5	[16 9;2 36]
Right Gray Volumes	34	ADvOth	17	81	[16 3;9 35]
		FTDvOth	16	81	[9 10;2 42]
		VaDvOth	23	81	[16 9;3 35]
Right Surface Area	34	ADvOth	18	93.7	[17 2;2 42]
		FTDvOth	16	82.5	[10 9;2 42]
		VaDvOth	13	79.4	[13 12;1 37]
Right Thickness Average	34	ADvOth	18	88.9	[14 5;2 42]
		FTDvOth	16	87.3	[13 6;2 42]
		VaDvOth	23	82.5	[18 7;4 34]

When all test results are analyzed, feature subsets having significant classification accuracies are found. Also, subset findings are in smaller sizes than the input matrices. In classifier comparison, NB performs slightly better than SVM for 2-Class tests.

Table 4. 2-Class Genetic Algorithm Results with Support Vector Machines

Feature Group	*Total Feat. Count*	*Classes*	*Best Feat Subset Count*	*Acc %*	*Confusion Matrix*
Cortical Volumes	41	ADvOth	10	82.5	[8 11;0 44]
		FTDvOth	9	82.5	[8 11;0 44]
		VaDvOth	9	90.5	[19 6;0 38]
Left Gray Volumes	34	ADvOth	7	88.9	[13 6;1 43]
		FTDvOth	11	82.5	[8 11;0 44]
		VaDvOth	10	85.7	[17 8;1 37]
Left Surface Area	34	ADvOth	8	82.5	[12 7;4 40]
		FTDvOth	9	82.5	[8 11;0 44]
		VaDvOth	10	82.5	[14 11;0 38]
Left Thickness Average	34	ADvOth	11	85.7	[10 9;0 44]
		FTDvOth	8	82.5	[9 10;1 43]
		VaDvOth	12	77.8	[15 10;4 34]
Right Gray Volumes	34	ADvOth	9	82.5	[10 9;2 42]
		FTDvOth	12	81	[7 12;0 44]
		VaDvOth	11	87.3	[17 8;0 38]
Right Surface Area	34	ADvOth	10	85.7	[10 9;0 44]
		FTDvOth	12	79.4	[6 13;0 44]
		VaDvOth	11	81	[14 11;1 37]
Right Thickness Average	34	ADvOth	8	90.5	[14 5;1 43]
		FTDvOth	10	85.7	[11 8;1 43]
		VaDvOth	10	79.4	[15 10;3 35]

Cortical volume features are striking to the eyes as the most significant feature group. In addition to the groups, behind the confusion matrices and best feature subset counts, valuable features are observed as *left lateral ventricle*, *left putamen*, *brain-stem*, *left vessel*, *right cerebellum white matter*, *right putamen* and *5th ventricle*. Moreover, *left thalamus proper*, *left pallidum*, *3rd ventricle*, *4th ventricle*, *left choroid plexus*, *right lateral ventricle*, *right caudate*, *right amygdala*, *right ventraldc*, *white matter hypointensities* and *optic chiasm*may also be included to the list as the secondary valuables.

5. Conclusion

In neuroimaging science, comparing classification results among different researches may not be straightforward due to attributes of each dataset like sample size, disease types, etc. Additionally, image preprocessing algorithms for feature extraction step, and also testing phase parameters may affect the classification performance. These issues have been reported in some researches previously[4], [5]. Therefore, the consistency of the work within itself may become much more considerable. In this work, a wrapper feature selection approach through certain classifiers with specific feature groups was implemented successfully. Best feature subset findings with significant classification accuracies were listed. For future plans, the first target is ensuring the data reliability using increased sample size. Standard dementia datasets in literature may also be processed. The ultimate goal is analyzing whole brain structure via combined feature groups together with different classifiers and feature selection algorithms.

References

[1] C. Güngen, T. Ertan, E. Eker, R. Yaşar and F. Engin (2002).Validity and reliability study on standardized mini mental test for the diagnosis ofmild dementia in the Turkish society.*Turkish Journal of Psychiatry.* Vol.13. Pages.273–281.

[2] H.G. Kreeftenberg, E.L. Mooyaart, J.R. Huizenga and W.J. Sluiter (2000). Quantification of liver iron concentration with magnetic resonance imaging by combining T1-, T2-weighted spin echo sequences and a gradient echo sequence.*Neth. J. Med.*Vol.56. Pages. 133–137.

[3] E. Westman, A. Simmons, J.S. Muehlboeck, P. Mecocci, B. Vellas, M. Tsolaki, I. Kloszewska, H. Soininen, M.W. Weiner, S. Lovestone, C. Spenger and L.O. Wahlund (2011). AddNeuroMed and ADNI: similar patterns of Alzheimer's atrophy and automated MRI classification accuracy in Europe and North America.*Neuroimage*.Vol.58.Pages. 818–828.

[4] R. Cuingnet, E. Gerardin, J. Tessieras, G. Auzias, S. Lehéricy, M.O. Habert, M. Chupin, H. Benali, O. Colliot, ADNI and others (2011). Automatic classification of patients with Alzheimer's disease from structural MRI: a comparison of ten methods using the ADNI database.*Neuroimage*.Vol.56. Pages. 766–781.

[5] J. Escudero, J.P. Zajicek and E. Ifeachor (2011). Machine Learning classification of MRI features of Alzheimer's disease and mild cognitive impairment subjects to reduce the sample size in clinical trials. in *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*, Pages. 7957–7960.

[6] W.B. Jung, Y.M. Lee, Y.H. Kimand C.W.Mun (2015). Automated classification to predict the progression of Alzheimer's disease using whole-brain volumetry and DTI.*Psychiatry Investig.* Vol.12.Pages. 92–102.

[7] C. Aguilar, E. Westman, J.S. Muehlboeck, P. Mecocci, B.Vellas, M. Tsolaki, L. Kloszewska, H. Soininen, S. Lovestone, C.Spenger and others (2013). Different multivariate techniques for automated classification of MRI data in Alzheimer's disease and mild cognitive impairment.*Psychiatry Res. Neuroimaging*. Vol.212. Pages. 89–98.

[8] B. Fischl (2012). FreeSurfer.*Neuroimage*.Vol.62.Pages. 774–781.

[9] A.M. Dale, B.Fischl and M.I. Sereno (1999). Cortical surface-based analysis: I. Segmentation and surface reconstruction.*Neuroimage*.Vol.9. Pages. 179–194.

[10] B. Fischl, A. Liu and A.M. Dale (2001). Automated manifold surgery: constructing geometrically accurate and topologically correct models of the human cerebral cortex.*Med. Imaging, IEEE Trans.*Vol.20. Pages. 70–80.

[11] E.H.B.M. Gronenschild, P. Habets, H.I.L. Jacobs, R. Mengelers, N. Rozendaal, J. Van Os and M. Marcelis (2012). The effects of FreeSurfer version, workstation type, and Macintosh operating system version on anatomical volume and cortical thickness measurements.*PLoS One*.Vol.7. Pages. e38234.

[12] I. Rish (2001). An empirical study of the naive Bayes classifier. in *IJCAI 2001 workshop on empirical methods in artificial intelligence*. Vol.3. Pages. 41–46.

[13] C.J.C. Burges (1998). A tutorial on support vector machines for pattern recognition.*Data Min. Knowl. Discov.* Vol.2. Pages. 121–167.

[14] M. Pei, E.D. Goodman, W.F.Punchand Y. Ding (1995). Genetic algorithms for classification and feature extraction. in *Classification Society Conference*.

[15] A. Konak, D.W. Coit and A.E. Smith (2006). Multi-objective optimization using genetic algorithms: A tutorial.*Reliab. Eng. Syst. Saf.*Vol.91. Pages. 992–1007.

[16] R.S. Sexton, R.E. Dorsey and J.D. Johnson (1999). Optimization of neural networks: A comparative analysis of the genetic algorithm and simulated annealing. *Eur. J. Oper. Res.* Vol.114. Pages. 589–601.

