Advances in Computer Science and its Applications
Vol. 1, No. 1, March 2012

www.worldsciencepublisher.org

An Improved Algorithm with Gene Selection and Decision Rules for Ovarian Cancer

Shih-Wei Lin[1], Chou-Yuan Lee[2], So-Tsung Chou[3], and Zne-Jung Lee[4*]

[1]Department of Information Management, Chang Gung University, No. 259, Wen-Hwa 1st Road, Tao-Yuan, Taiwan
[2]Department of Information Management, Lan Yang Institute of Technology
[3]Department of Information Management, Kang-Ning Junior College of Medical Care and Management, Taiwan
[4*]Department of Information Management, Huafan University, Taiwan

E-mail: johnlee@hfu.edu.tw

Abstract: The microarray data of ovarian cancer consists of tens of thousands of genes on a genomic scale. To avoid higher computational complexity, it needs gene selection to find the gene subsets that are able to classify ovarian cancer. Most of gene selections use traditional statistics or data mining techniques to build the model. However, traditional statistics have no consideration about the variable gene selection and block effect. Data mining techniques may suffer the problem of parameter settings. Therefore, this paper applies scatter search to obtain suitable parameter settings for support vector machine and decision tree. Additionally, it selects a subset of beneficial genes without reducing the classification accuracy, and provides the decision rules for medical experts and biologists to evaluate the block effect of selected genes. In order to evaluate the proposed algorithm, the microarray data of ovarian cancer collected from China Medical University are used as the source datasets. From experimental results, it shows that the proposed algorithm can reduce unnecessary genes, and significantly improve the classification accuracy for ovarian cancer.

Keywords: Ovarian Cancer; Support Vector Machine; Decision Tree; Scatter Search; Gene Selection

1. Introduction

Ovarian cancer is one of the gynecological cancer deaths for women in the United States [1-3]. It is important to early detect ovarian cancer, because 70% of women with the epithelial ovarian cancer are not diagnosed until the disease is spread to upper abdomen [4]. Recently, microarray technologies have been developed that can be used to simultaneously assess the level of expression of genes [5-10]. Then, there is increasing interest in the emphasis of gene selection and cancer classification for microarray data [11-14]. The microarray data of ovarian cancer contains genes with tens of thousands of dimension and could harm the performance of the classification accuracy for the formation of decision rules. Thus, the gene selection is an important issue for evaluating the classification accuracy of microarray data. Several studies have been used to select genes from microarray data such as correlation methods, nonparametric scoring approach, Bayesian variable selection approach, and machine learning methods [15-30]. For above literature, they have not simultaneously considered the gene selection and the optimal parameter setting for the microarray data of ovarian cancer. Furthermore, there is no systematic approach for achieving a better insight into global gene expression analysis. In previous study, author has proposed an integrated algorithm for gene selection and classification applied to microarray data of ovarian cancer [31]. Its result can interpret the variable gene selection and block effect in microarray data, but there is no decision rule for biological interpretations. In this paper, an improved algorithm with gene selection and decision rules for ovarian cancer is proposed. The real microarray data of ovarian cancer is obtained from China Medical University Hospital. The purpose of this study is to apply scatter search (SS) to perform gene selection and the parameter determinations for support vector machine (SVM) and decision tree (DT), respectively. Furthermore, the decision rules of this study can be obtained as guidance for biologists.

The remainder of this paper is organized as follows. Section 2 reviews SVM, SS and DT. Section 3 then introduces the proposed algorithm. Simulation results are compared with other existing approaches in Section 4. Conclusions are finally drawn in Section 5.

2. Introduction of Support Vector Machine, Scatter Search and Decision Tree

The proposed algorithm is based on support vector machine, scatter search and decision tree. In this section, we briefly describe the basic concepts of support vector machine, scatter search and decision tree.

2.1 Support Vector Machine (SVM)

SVM is proposed by Vapnik and successively applied to many applications [32-35]. Let (x_1, y_1), …, $(x_m, y_m) \in X \times \{-1, +1\}$ be a set of training data, where X represents some nonempty set from which the pattern x_i, and y_i are called the target $y_i \in \{-1, +1\}$. The objective of SVM is to find an optimal separating hyper-plane with the maximum margin (w) and a real value b for

classification of data. Consider the class of hyper-planes in the dot product space *H*. The parameters *w* and *b* are described as follows:

$$\begin{aligned} &\left(\left\langle w \cdot x_i \right\rangle + b\right) = 0 \\ &\text{where } w \in H,\ b \in R,\ i = 1, \cdots, m. \end{aligned} \tag{1}$$

A linear separation is obtained if a hyper-plane that satisfies Eq. (1) exists. The separating hyper-planes include one optimal separating hyper-plane (OSH) which has the largest distance between two support vector points on its two sides. The minimal distance to OSH can be derived from Eq. (2).

$$\underset{w,b}{Min}\ \frac{1}{2} w^T w \tag{2}$$

Subject to:

$$y_i \left(\left\langle w \cdot x_i \right\rangle + b\right) \geq 1,\ i = 1,\ ..., m. \tag{3}$$

The margin of a separating hyper-plane can be regarded as the hyper-plane's generalization ability, and the OSH has the maximal margin among separating hyper-planes. Let α denote $(\alpha_1, \ldots, \alpha_m)$. Combining Lagrange's polynomial (in the order of *m*) with Eq. (3) produces the following equation:

$$\underset{\alpha}{Max}\ L(\alpha) = \sum_{i=1}^{m} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{m} \alpha_i \alpha_j y_i y_j \left\langle x_i \cdot x_j \right\rangle \tag{4}$$

Subject to:

$$\begin{aligned} &0 \leq \alpha_i \leq C,\ i = 1,\ ..., m \\ &\text{and } \sum_{i=1}^{m} \alpha_i y_i = 0 \end{aligned} \tag{5}$$

Where $\alpha_i \geq 0$ denotes the Lagrange multiplier and *C* is the penalty parameter. Given a vector which satisfies Eq. (4) in maximization, the OSH can be written as follows:

$$w = \sum_{i=1}^{m} \alpha_i y_i x_i \tag{6}$$

Where the support vector points must satisfy $\alpha_i \geq 0$ and Eq. (1).

In practice, the data may not be linearly separable and could be mapped to a higher dimensional feature space. It means that SVM will map the data into a higher dimensional space for classification if the data cannot be classified explicitly in the current dimensional space. The input data are mapped to a higher dimensional feature space by plotting a nonlinear curve. The OSH is built into the feature space. The feature space vectors x_i, x_j are constructed in terms of the kernel *k*, evaluated on input patterns x_i, x_j where $k\left(x_i, x_j\right) = \left\langle x_i, x_j \right\rangle$. The nonlinear SVM can map the input data into a high-dimension feature space via a mapping function $\varphi\left(x\right)$. By virtue of constructing the feature space, we can substitute $\varphi\left(x\right)$ into Eq. (4) and have the following:

$$\begin{aligned} &\underset{\alpha}{Max}\ L(\alpha) = \sum_{i=1}^{N} \alpha_i - \\ &\frac{1}{2} \sum_{i,j=1}^{N} \alpha_i \alpha_j y_i y_j \left\langle \phi\left(x_i\right), \phi\left(x_j\right) \right\rangle \end{aligned} \tag{7}$$

Given a kernel function $k\left(x, y\right) = \left\langle \phi(x), \phi(y) \right\rangle$, Eq (7) is showed as follow:

$$\begin{aligned} &\underset{\alpha}{Max}\ L(\alpha) = \sum_{i=1}^{N} \alpha_i - \\ &\frac{1}{2} \sum_{i,j=1}^{N} \alpha_i \alpha_j y_i y_j k(x_i, x_j) \end{aligned} \tag{8}$$

Several kernel functions help the SVM in finding the optimal solution. The most frequently used such functions are the polynomial kernel, sigmoid kernel and radial basis kernel function (RBF) [34]. The RBF $k\left(x_i, x_j\right) = \exp(-\gamma \left\| x_i - x_j \right\|^2)$ is the most often used in general cases, since it can classify multi-dimensional data unlike a linear kernel function. Furthermore, the RBF needs fewer parameters than the polynomial kernel. The two major parameters of the RBF (*C* and γ) applied in SVM have to be set appropriately. The classification accuracy is very high in the training stage and very low in the testing stage if *C* is set too large. The partitioning outcome in the feature space is dominated by the parameter γ. An excessive value for parameter γ may lead to over-fitting, while a disproportionately small value may result in under-fitting [35]. In this paper, scatter search will be conducted to find the best values for these two parameters, *C* and γ.

2.2 Scatter Search (SS)

Scatter search (SS), proposed by Glover in 1977, is an evolutionary approach that starts with a collection of reference solutions obtained by applying preliminary heuristic processes [36]. SS uses strategies for search diversification and intensification based on formulations for combining decision rules and problem constraints. Glover presented a simplification of the description for SS method known as SS template in 1998. It is considered as a milestone in SS literature and has regarded as the main reference for most of the SS implementations up to date [37]. It has proved effective in a variety of optimization problems and has shown potential for solving various applications [38-39].

Generally, there are five principal components in SS. (1) *The diversification generation method* (DGM) generates trial solutions that satisfy a critical level of diversity by using an arbitrary trial solution as an input. (2) *The improvement method* (IM), a local optimizer, transforms trial solutions obtained from DGM into enhanced feasible trial solutions. (3) *The reference set update method* (RSUM) builds and maintains a reference set consisting of high-quality and diverse solutions. The reference set is the basis for creating new combined solutions. (4) *The subset generation method* (SGM) is conducted to the reference set and then produces a subset of solutions as a basis for creating combined solutions. (5) *The solution combination method* (SCM) transforms a given subset of solutions produced by the SGM into one or more combined new solutions. Since above components can be implemented in a variety of ways, SS is very adaptable to solve different problems.

2.3 Decision Tree (DT)

Decision tree is an important area of artificial intelligence. It is a rule induction approach that utilizes a divide-and-conquer strategy to recursively partition the data set into smaller subdivisions by generating a tree-like structure [40-45]. This tree-like structure is composed of a root node (formed from all of the data), a set of internal nodes (splits), and a set of terminal nodes (leaves). The C4.5 rule is one of the primary approaches in DT [46]. There are two major phases for C4.5. One is growth phase and the other is pruning phase [43]. For the growth phase, C4.5 uses an information entropy evaluation function as the selection criteria [44]. The entropy evaluation function is calculated as follows.

Step 1: Calculate *Info*(*S*) to identify the class in the training set *S*.

$$Info_x(S) = -\sum_{i=1}^{k}\left\{\left[freq(C_i,S/|S|)\right]\log_2\left[freq(C_i,S/|S|)\right]\right\} \quad (9)$$

where $|S|$ is the number of cases in the training set. C_i is a class, i=1,2,...,k. k is the number of classes and $freq(C_i,S)$ is the number of cases included in C_i.

Step 2: Calculate the expected information value $Info_x(S)$ for feature X to the partition S.

$$Info_x(S) = -\sum_{i=1}^{L}\left[\ (|S_i|/|S|)Info(S_i)\ \right] \quad (10)$$

where L is the number of outputs for feature X, S_i is a subset of S corresponding to the i^{th} output and $|S_i|$ is the number of cases of the subset S_i.

Step 3: Calculate the information gained after partitioning according to feature X.

$$Gain(X) = Info(S) - Info_x(S) \quad (11)$$

Step 4: Calculate the partition information value $SplitInfo(X)$ acquired for S partitioned into L subsets.

$$SplitInfo(X) = -\sum_{i=1}^{L}\left[\frac{|S_i|}{|S|}\log_2\frac{|S_i|}{|S|}\right] \quad (12)$$

Step 5: Calculate the gain ratio of *Gain*(*X*) over *SplitInfo*(*X*).

$$GainRatio(X) = Gain(X)\ /\ SplitInfo(X) \quad (13)$$

Where the *GainRatio*(*X*) compensates the weak point of *Gain*(*X*), which represents the quantity of information provided by *X* in the training set. Therefore, the feature with the highest *GainRatio*(*X*) is taken as the root of the decision tree. The pruning phase aims to avoid over-fitting the training data for the generated DT. In the pruning phase, the most difficult task is to get the balance between accuracy and simplicity. Unfortunately, the minimum cases for the leaf (*M*) and the pruning confidence factor (*CF*) are varied with different cases. The decision of these two parameters becomes an optimization problem. In this paper, we will use scatter search to find the best values for these two parameters, *M* and *CF*.

3. The Proposed Algorithm

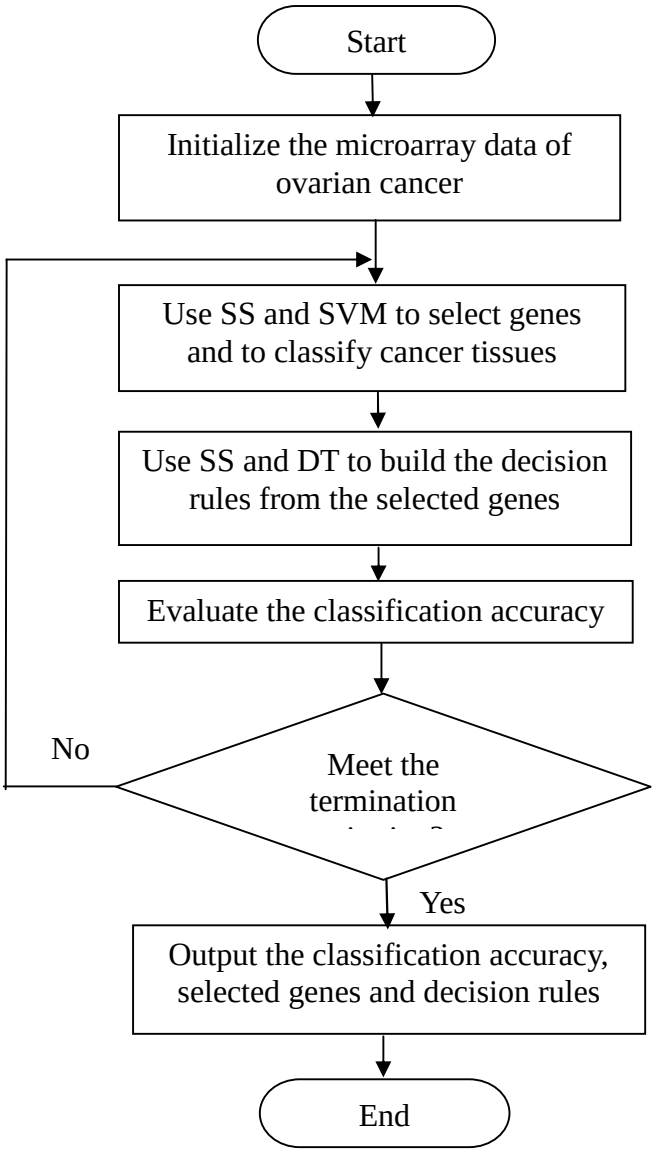

Figure 1: The flow chart of the proposed algorithm

In this paper, the datasets of microarray data of ovarian cancer were collected from China Medical University Hospital. The ovarian tissues, vaginal tissues, cervical tissues and myometrium of patients include 6 benign ovarian tumors (BOT), 10 ovarian tumors (OVT) and 25 ovarian cancers (OVCA). The 9,600 human cDNA clones in a sequence-verified human cDNA library were a kind gift from the National Health Research Institute of Taiwan. They were originally obtained from the Minimum Information About a Microarray Experiment (MIAME) consortium libraries through its distributor (Research Genetics, Huntsville, AL) [14]. To avoid higher computational complexity, it needs to select the most likely differentially expressed genes to explain the effects of ovarian cancer. Additionally, the microarray data with highly correlated genes will significantly increase the classification accuracy by selecting genes.

In this paper, an improved algorithm with gene selection and decision rules for ovarian cancer is proposed. It presents a novel algorithm based on SS that

provides the best parameter settings for the DT and SVM, and finds the beneficial subset of genes to maximize the classification accuracy for the microarray data of ovarian cancer. The flow chart of the proposed algorithm is shown in Figure 1. In the proposed algorithm, SVM is first applied to select genes for increasing the classification accuracy for the microarray data of ovarian cancer. Thereafter, DT is used to learn rules from the training data set. In SVM and DT, SS is conducted to parameter determination. To implement SS, the population *P* with P_{size} solutions is uniformly generated in the diversification generation method. In reference set update method, the size of the reference set is $b = b_1 + b_2 = |RefSet|$. Construction of the initial reference set starts with selecting b_1 best solutions (solutions with the highest classification accuracy) from *P*, and these solutions are added to *RefSet*. For each solution in the *P-RefSet*, the minimum Euclidean distance to the solutions in *RefSet* is calculated. The solution with the maximum of the minimum distances is selected, and this solution is then added to *RefSet*. Thus the minimum distances are updated accordingly. The resulting reference set has b_1 high-quality solutions and b_2 diverse solutions. In subset generation method, the size of subsets is set to 2; that is, only subsets consisting of all pair-wise combinations of solutions in *RefSet* are considered. In solution combination method, the method employed consists of finding linear combinations of reference solutions. Each combination of two reference solutions, denoted as X' and X'', are employed to create three trial solutions. These three trial solutions are (1) $X = X' - d$, (2) $X = X' + d$, and (3) $X = X'' + d$, where $d = u(X'' - X')/2$ and *u* is a random number with values within [0,1]. In the hybrid process of SS and SVM, *n* variable followed with two variables, *C* and γ, must be established if *n* number of genes are selected, The value of *n* variables ranges between 0 and 1. The corresponding gene is not selected if its value is less than or equal to 0.5. Conversely, the corresponding gene is selected if its value is greater than 0.5. For the hybrid process of SS and DT, two decision variables, designated *M* and *CF*, are necessary. SS will set the proper parameters of *M* and *CF* to increase the classification accuracy. The proposed algorithm is repeated until the stop criterion has met. Thereafter, the classification accuracy, selected genes, and extracted decision rules are reported.

4. Simulation Results

In simulation, we need to identify the range of parameter for SVM and DT. The searching range of parameter *C* of the SVM is between 0.01 and 50,000, while the searching range of parameter γ of the SVM is between 0.0001 and 50. Meanwhile, the searching range of the parameter *M* of DT is between 2 and 10, and the searching range of parameter *CF* of DT is between 0.01 and 0.5 [47].

Table 1: The simulation results for various approaches

	The hybrid process of SVM and SS	The hybrid process of SVM and DT	SVM	The proposed algorithm
Classification accuracy%	92.8571%	85.714%	78.571%	96.4376%
The number of selected genes	12	14	15	5

Table 2: The obtained 5 rules of DT from the proposed algorithm

Rule number	Rule	Classification
1	"peptidylprolyl isomerase D (cyclophilin D)" <= -0.7378 and "matrix metalloproteinase 2 (gelatinase A, 72kDa gelatinase, 72kDa type IV collagenase)" <= -0.67	BOT
2	"cadherin 4, type 1, R-cadherin (retinal)" <= 1.59 and "matrix metalloproteinase 2 (gelatinase A, 72kDa gelatinase, 72kDa type IV collagenase)" > -0.67	OVT
3	"fibronectin 1" > -0.38 and "cadherin 4, type 1, R-cadherin (retinal)" <= 1.59	OVT
4	"fibronectin" <= -0.38 and "cadherin 4, type 1, R-cadherin (retinal)" <= 1.59 and "fibronectin" <= -0.38	OVCA
5	"peptidylprolyl isomerase D (cyclophilin D)" > -0.7378 and "matrix metalloproteinase 2 (gelatinase A, 72kDa gelatinase, 72kDa type IV collagenase)" <= -0.67 and"matrix metalloproteinase 2 (gelatinase A, 72kDa gelatinase, 72kDa type IV collagenase)" > -0.73	OVCA

To verify the performance of the proposed algorithm, various approaches include the proposed algorithm, the hybrid process of SVM and SS, the hybrid process of SVM and DT, and SVM are used to compare the simulation results. For fair comparisons, the same values of parameters are used for these approaches. Simulations are performed to see which approaches can find the best classification accuracy with selected genes. The *k*-fold approach is used to evaluate the classification accuracy for the microarray data of ovarian cancer obtained form China Medical University [48]. This study set *k* as 3; that is, the data was divided into three portions. Two portions of data are retrieved as training data and the other one is used for testing data. The simulation results are shown in Tables 1. As shown in Table 1, the classification accuracy for the proposed algorithm is 96.4376%, and it outperforms other approaches. The proposed algorithm has the minimal selected genes among these compared approaches. Five selected genes are extracted from 9600 genes for the proposed algorithm. These genes are "peptidylprolyl isomerase D (cyclophilin D)", "fibronectin 1 (4390)", "cadherin 4, type 1, R-cadherin (retinal)", "matrix metalloproteinase 2 (gelatinase A, 72kDa gelatinase, 72kDa type IV collagenase)", "protein tyrosine phosphatase, receptor type, K". For the hybrid process of SVM and SS, the hybrid process of SVM and DT, and SVM, each approach also has good classification accuracy, but the process among them is kind of black box. Biologists cannot take these results into their future judgments. In the proposed algorithm, five rules provided by DT are shown in Table 2. From Table 2, it can be

observed the relation between the selected genes for ovarian cancer.

5. Conclusions

This study gives evidence for the improvement of gene selection and decision rules for ovarian cancer. In this paper, the process of SVM, SS and DT are hybridized to select genes, and the parameters for SVM and DT are automatically achieved. The proposed algorithm can classify the cancer tissues for microarray data of ovarian cancer by the selected genes. Additionally, the obtained gene markers are performed to classify cancer tissues by the improved fuzzy model. It has demonstrated that five selected genes can be obtained through the proposed algorithm. Additionally, the structure of tree which is obtained from the proposed DT architecture, medical experts and biologists can analyze the decision rules and thus make a further evaluation. Results of the experiments show that the proposed algorithm is effective in searching for the beneficial subset of genes and decision rules.

Acknowledgments

Authors appreciate G. Steven Huang, Yao-Ching Hung, Meng-Hsiun Tsai, Yen-Po Huang, and Shih-Chieh Chen for providing the microarray data and discussions.

References

[1] M. Schena, D. Shalon, R.W. Davis and P.O. Brown, Quantitative monitoring of gene expression patterns with a complementary DNA microarray, *Science* 270 (1995) 467-470.
[2] A. Jemal, A. Thomas, T. Murray and M. Thun, Cancer statistics, *CA: A Cancer Journal for Clinicians* 52 (2002) 23-47.
[3] G. Steven Huang, Yao-Ching Hung, Ann Chen and Meng-Yen Hong, Microarray Analysis of Ovarian cancer, in: *IEEE International Conference on Systems, Man and Cybernetics*, IEEE Systems, Man and Cybernetics Society, New York (2005) 1036-1042.
[4] A. Jermal, T. Murray, A. Samuels, A. Ghafoor, E. Ward and M. J. Thun, Cancer statistic 2003, *CA: A Cancer Journal for Clinicians* 13 (2003) 5-26.
[5] Steen Knudsen, *A biologist's analysis of DNA microarray data* (John Wiley and Sons, Inc, Publication, New York, 2002).
[6] Jin-Tsong Jeng, Tsu-Tian Lee and Yung-Cheng Lee, Classification of ovarian cancer based on intelligent systems with microarray data, in: *IEEE International Conference on Systems, Man and Cybernetics*, IEEE Systems, Man and Cybernetics Society, New York (2005) 1053-1058.
[7] Chen-Chia Chuang, Jin-Tsong Jeng and Shun-Feng Su, Dimension reduction with support vector regression for ovarian cancer microarray data, in: *IEEE International Conference on Systems, Man and Cybernetics*, IEEE Systems, Man and Cybernetics Society, New York (2005) 1048-1052.
[8] S. Chao and C. Lihui, Feature dimension reduction for microarray data analysis using locally linear embedding, in: Yi-Ping Phoebe Chen and Limsoon Wong (eds.) in: *Proceeding of 3rd Asia-Pacific Bioinformatics Conference*, Imperial College Press, London (2005) 211-217.
[9] A. A. Alizadeh, M. B. Eisen, R. E. Davis, C. Ma, I. S. Lossos, A. Rosenwald, J. C. Boldrick, H. Sabet, T. Tran, X. Yu, J. I. Powell, L. Yang, G. E. Marti, T. Moore, J. Hudson, L. Lu, D. B. Lewis, R. Tibshirani, G. Sherlock, W. C. Chan, T. C. Greiner, D. D. Weisenburger, J. O. Armitage, R. Warnke, R. Levy, W. Wilson, M. R. Grever, J. C. Byrd, D. Botstein, P. O. Brown and L. M. Staudt, Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling, *Nature* 403 (2000) 503-511.
[10] J. DeRisi, L. Penland, P. O. Brown, M. L. Bittner, P. S. Meltzer, M. Ray, Y. Chen, Y. A. Su and J. M. Trent, Use of a cDNA microarray to analyse gene expression patterns in human cancer, *Nat Genet* 14 (1996) 457-460.
[11] D. Chen, D. Hua, J. Reifman and X. Cheng, Gene selection for multi-class prediction of microarray data, in: *Proceedings of the IEEE Computer Society Conference on Bioinformatics*, IEEE Computer Society, Washington (2003) 492-495.
[12] K. E. Lee, N. Sha, E. Dougherty, M. Vannucci and B.K. Mallick, Gene selection: a bayesian variable selection approach, *Bioinformatics* 19 (2003) 90-97.
[13] P. J. Park, M. Pagano and M. Bonetti, A nonparametric scoring algorithm for identifying informative genes from microarray data, in: Russ B. Altman, A. Keith Dunker, Lawrence Hunter, Kevin Lauderdale, and Teri E. Klein (eds.) *Pacific Symposium on Biocomputing 2001*, World Scientific, New Jersey (2001) 52-63.
[14] A. Brazma, P. Hingamp, J. Quackenbush, G. Sherlock, P. Spellman, C. Stoeckert, J. Aach, W. Ansorge, C. A. Ball, H. C. Causton, T. Gaasterland, P. Glenisson, F. C. Holstege, I. F. Kim, V. Markowitz, J. C. Matese, H. Parkinson, A. Robinson, U. Sarkans, S. Schulze-Kremer, J. Stewart, R. Taylor, J. Vilo and M. Vingron, Minimum information about a microarray experiment (MIAME)-toward standards for microarray data, *Nat Genet* 29 (2001) 365-71.
[15] T. Z. Tan, C. Quek and G. S. Ng, Ovarian cancer diagnosis using complementary learning fuzzy neural network, in: *Proceedings of the 2005 IEEE International Joint Conference on Neural Networks (IJCNN 2005)*, IEEE Service Center, Piscataway, New Jersey (2005) 3034-3039.
[16] J. D. Schaffer, A. Janevski and M. R. Simpson, A Genetic Algorithm approach for discovering diagnostic patterns in molecular measurement data, in: *Proceedings of the 2005 IEEE Symposium on Computational Intelligence in Bioinformatics and Computational Biology*, IEEE Computer Society, La Jolla, California (2005) 1-8.
[17] A. Bertoni and G. Valentini, Randomized maps for assessing the reliability of patients clusters in DNA microarray data analyses, *Artificial Intelligence in Medicine* 37 (2006) 85-109.
[18] J. Khan, J. S. Wei, M. Ringnér, L. H. Saal, M. Ladanyi, F. Westermann, F. Berthold, M. Schwab, C.R. Antonescu, C. Peterson and P. S. Meltzer, Classification and diagnostic prediction of cancers using gene expression profiling and artificial neural networks, *Nature Medicine* 7 (2001) 673-679.
[19] F. Chu, W. Xie and L. Wang, Gene selection and cancer classification using a fuzzy neural network, in: S. Dick, L. Kurgan, P. Musilek, W. Pedrycz, and M. Reformat (eds.) *Proceedings of NAFIPS 2004, Annual Meeting of the North American Fuzzy Information Processing Society*, IEEE, Banff, Alberta, Canada (2004) 555-559.
[20] B. Ni and J. Li, A hybrid filter/wrapper gene selection method for microarray classification, in: *Proceedings of 2004 International Conference on Machine Learning and Cybernetics* 4. IEEE Systems, Man and Cybernetics Society, New York (2004) 2537-2542.
[21] Jin-Hyuk Hong and Sung-Bae Cho, The classification of cancer based on DNA microarray data that uses diverse ensemble genetic programming, *Artificial Intelligence in Medicine* 36 (2006) 43-58.
[22] T. S. Furey, N. Cristianini, N. Duffy, David W. Bednarski, Michèl Schummer and David Haussler, Support vector machine classification and validation of cancer tissue samples using microarray expression data, *Bioinformatics* 16 (2000) 906-914.
[23] Te Ming Huang and Vojislav Kecman, Gene extraction for cancer diagnosis by support vector machines-An improvement, *Artificial Intelligence in Medicine* 35 (2005) 185-194.
[24] P. C. H. Ma, Keith C. C. Chan, X. Yao and D. K. Y. Chiu, An evolutionary clustering algorithm for gene expression microarray data analysis, *IEEE Transactions on Evolutionary Computation* 10 (2006) 296-314.
[25] M, Schummer, W. V. Ng, R. E. Bumgarner, P.S. Nelson, B. Schummer, D. W. Bednarski, L. Hassell, R. L. Baldwin, B. Y Karlan and L. Hood, Comparative hybridization of an array of 21,500 ovarian cDNAs for the discovery of genes overexpressed in ovarian carcinomas, *Gene* 85 (1999) 238-375.
[26] K. Wang, L. Gan, E. Jeffery, M. Gayle, A. M. Gown, M. Skelly, P. S. Nelson, W.V. Ng, M. Schummer, L. Hood and J. Mulligan, Monitoring gene expression profile changes in ovarian carcinomas using cDNA microarray, *Gene* 229 (1999) 101-109.
[27] R. S. Ismail, Rae Lynn Baldwin, Junguo Fang, Damaris Browning, Beth Y. Karlan, Judith C. Gasson and David D. Chang, Differential gene expression between normal and tumor-derived ovarian epithelial cells, *Cancer Research* 60 (2000) 6744-6753.
[28] A. M. Martoglio, B. D. Tom, M. Starkey, A. N. Corps, D. S. Charnock-Jones and S. K. Smith, Changes in, tumor-genesis and angiogenesis-related gene transcript, abundance profiles in ovarian cancer detected by tailored high density cDNA arrays *Molecular Medicine* 6 (2000) 750-65.
[29] Kenji Ono, Toshihiro Tanaka, Tatsuhiko Tsunoda, Osamu Kitahara, Chikashi Kihara, Aikou Okamoto, Kazunori Ochiai, Toshihisa Takagi and Yusuke Nakamura, Identification by cDNA microarray of genes involved in ovarian carcinogenesis, *Cancer Research* 60 (2000) 5007-5018.
[30] J. B. Welsh, P. P. Zarrinkar, L. M. Sapinoso, S. G. Kern, C. A. Behling, B. J. Monk, D. J. Lockhart, R. A. Burger and G. M. Hampton, Analysis of gene expression profiles in normal and neoplastic ovarian tissue samples identifies candidate molecular markers of epithelial ovarian cancer, in: *Proceedings of the National Academy of Sciences of the United States of America* 98, National Academy of Sciences, Washington (2001) 1176-1181.
[31] Zne-Jung Lee, An integrated algorithm for gene selection and classification applied to microarray data of ovarian cancer, *International Journal Artificial Intelligence in Medicine* 42 (2008) 81-93.
[32] C. J. C. Burgers, A tutorial on support vector machines for pattern

recognition, *Data Mining and Knowledge Discovery* 2 (1998) 121–167.
[33] N. Cristiamini, J. Shawe-Taylor, (2000). An introduction to Support Vector Machines, Cambridge university press, England.
[34] B. SchÖlkopf, and A. J. Smola, (2002). Learning with kernels support vector machines, regularization, optimization, and beyond, MIT, London.
[35] K., -R. Müller, S. Mike, G. Rätsch, K. Tsuda, and B. Schölkopf, An introduction to kernel-based learning algorithms, *IEEE Transactions on Neural Networks*, 12 (2) (2001) 181–201.
[36] F. Glover, Heuristics for integer programming using surrogate constraints. *Decision Science* 8 (1977) 156–166.
[37] M. Laguna and R. Martí, Scatter search: methodology and implementations in C. Kluwer Academic Publishers, Boston (2003).
[38] F. G.. López, G. M. Torres, and B. M. Batista, Solving feature subset selection problem by parallel scatter search, *European Journal of Operation Research* 169 (2006) 477-489.
[39] R. Martí, Scatter search - wellsprings and challenges. *European Journal of Operational Research* (2006) 169 351–358.
[40] J. Gehrke, V. Ganti, R. Ramakrishnan, and W. -Y. Loh, BOAT-optimistic decision tree construction, *Proceedings of ACM SIGMOD International Conference Management of Data*, Philadelphia, Pennsylvania (1999) 169–180.
[41] J. R. Quinlan, Decision trees as probabilistic classifiers, *Proceedings of 4th International Workshop Machine Learning*, Irvine, California (1987) 31–37.
[42] H. Kim, and G. J. Koehler, Theory and practice of decision tree induction, *Omega* 23 (1995) 637–652.
[43] J. R. Quinlan, Introduction of decision trees, *Machine learning* 1 (1986) 81–106.
[44] J. R. Quinlan, Simplifying decision trees, *International Journal of Man-Machine studies* 27 (1987) 221–234.
[45] K. -M. Osei-Bryson, Post-pruning in decision tree induction using multiple performance measures, *Computers and Operations Research* 34 (2007) 3331–3345.
[46] J. R. Quinlan, *C4.5 : Programs for machine learning*, Morgan Kaufmann (1993).
[47] Y. Liao, S. -C. Fang, and H. L. W. Nuttle, A neural network model with bounded-weights for pattern classification, *Computers and Operations Research* 31 (2004) 1411–1426.
[48] H. -T. Lin, and C. -J. Lin, A study on sigmoid kernels for SVM and the training of non-PSD kernels by SMO-type methods, Technical report, University of National Taiwan, Department of Computer Science and Information Engineering (2003) March, 1-32.
[49] A. Ahmadivand, Enhancement of L-junction Plasmon Waveguides Properties Using Various Shapes of Au Nanoparticles Arrays at λ≈1550nm, Advances in Digital Multimedia, 1 (2012) 1-3.
[50] Y. Zhang, L. Wu, Rigid Image Registration by PSOSQP Algorithm, Advances in Digital Multimedia, 1 (2012) 4-8