

**2004 2nd International IEEE
Conference “Intelligent Systems”**

Proceedings

VOLUME I

EDITORS:

Ronald R. Yager and Vassil S. Sgurev

CONTENTS

Plenary SESSION	
PART I	
Intelligent Decision Making and Information Fusion <i>R. Yager, USA</i>	A New Method for Reaching Equilibrium Points in Fuzzy Cognitive Maps <i>T.L. Kottas, Y. S. Boutalis, Greece, G. Devedsic, Serbia and Montenegro, B. G. Mertziros, Greece</i>
1	53
Some Issues of Linguistic Approximation <i>G. Klir, USA</i>	Short-Term Load Forecasting Based on a Rough Fuzzy-Neural Network <i>F. Li and Q. Jia-ju, China</i>
5	61
Diversity and Unity of Uncertainty Theories <i>G. Klir, USA</i>	Self-Organising Map Representations of Greyscale Images Reflect Human Similarity Judgements <i>T. M. Gale, N. Davey, K. R. Laws, R. J. Frank, UK</i>
6	66
Computational Intelligence Approach to Real-World Cooperative Vehicle Dispatching Problem <i>K. Hirota, K. Chen, and F. Dong, Japan</i>	SORT: A Fast and Compact Neural Classifier Based on a Sorting Preprocessor <i>R. Dogaru, Romania, M. Glesner, Germany</i>
7	71
PART II	
Uncertain Variables and Their Applications in Knowledge-Based Decision Systems <i>Z. Bubnicki, Poland</i>	Nonlinear System Identification using Takagi-Sugeno Type Neuro-fuzzy Model <i>P. C. Panchariya, India, A. K. Palit, D. Popovic, Germany, A.L. Sharma, India</i>
13	76
Bioinformatics: A Knowledge Engineering Approach <i>N. Kasabov, New Zealand</i>	Two-Stage Learning Algorithm for Fuzzy Cognitive Maps <i>E. I. Papageorgiou and P. P. Groumpos, Greece</i>
19	82
Perspectives of Fuzzy Control: Lights and Shadows <i>P. Albertos and A. Sala, Spain</i>	PART III
25	
SESSION: Neural Networks and Neuro-Fuzzy Systems	
PART I	
Local Minima Free Neural Network Learning <i>I. Jordanov, T. Rafik, UK</i>	Dimensionality Reduction of Face Images for Gender Classification <i>S. Buchala, N. Davey, R. J. Frank, T. M. Gale, UK</i>
34	88
Automatic Text Summarization With Neural Networks <i>K. Kaikhah, USA</i>	Structured Connectivity in an Associative Memory Model <i>S. Turvey, S. Hunt, R. Frank, R. Adams, N. Davey, UK</i>
40	94
Applied Adaptive Fuzzy-Neural Inference Models: Complexity and Integrity Problems <i>G. M. Dimirovski, Turkey, I. I. Lokvenec, and D. J. Tanevska, F.Y. Rep. of Macedonia</i>	Improving of Classification Abilities of Neural Network by Modification of Training Data Set <i>S. Velikov, L. Dakovski, Bulgaria</i>
45	100
Modeling of Non-Linear Dynamic Systems via Discrete-Time Recurrent Neural Networks and Variational Training Algorithm <i>S. V. Minchev, G. I. Venkov, Bulgaria</i>	105
45	105

Mining Multi-Cross-Level Fuzzy Weighted Association Rules <i>M. Kaya, Turkey, and R. Alhaji, Canada</i>	225	Quality of Service Ensuring in Urban Solid Waste Management <i>N. V. Karadimas, V. G. Loumos and O. D. Mavrantza, Greece</i>	288
Fuzzy Regression Analysis by Entropy <i>Ch. Kao and P. - H. Lin, China</i>	231	Industrial Process Monitoring Using Nonlinear Principal Component Models <i>D. Antory, U. Kruger, G. W. Irwin, and G. McCullough, UK</i>	293
SESSION: Intelligent Agents, Ontology and Semantic Web		DEDS Control Synthesis Problem Solving <i>F. Čapkovič, Slovak Republic</i>	299
PART I			
Meta-Context Mediation to Attain Semantic Interoperability <i>Y. Biletskiy, C. Câmpeanu, Canada, Z. Dudar, and O. Vorochek, Ukraine</i>	238	Fuzzy Logic Based Intelligent Motion Control of Robot Swarm Simulated by Khepera Robots <i>Z. Minchev, O. Manolov, S. Noykov, Bulgaria, U. Witkowski, and U. Rückert, Germany</i>	305
The Semiotics Contribution on the Web <i>V. Damjanović, D. Gašević, V. B. Devedžić, and D. Djurić, Serbia and Montenegro</i>	244	SESSION: Applications	
PART I			
Soft Computing Agent Approach to Remote Learning of Disabled <i>D. V. Lakov, Bulgaria</i>	250	Ternary Grid as a Potentially New Technique for Knowledge Elicitation/Acquisition <i>Y. Erdani, A. Hunger, S. Werner, and S. Mertens, Germany</i>	312
Explore Agent Learning Process by Using Mechanical Features in Agent-Based Simulation <i>Zh. Bi, T. Takashina, K. Tanaka, Sh. Watanabe, Japan</i>	256	Reinforcement Learning for Process Identification, Control and Optimisation <i>J. J. Govindhasamy, UK, S. F. McLoone, Ireland, and G. W. Irwin, UK</i>	316
PART II			
A Hybrid Local-Global Approach for Handling Ontologies in a Multiagent System <i>R. F. Brena, H. Ceballos, Mexico</i>	261	Electric-Power Protection System Fuzzy Critical Analysis <i>M. Dumitrescu, T. Munteanu, A. P. Ulmeanu, Romania</i>	322
The Design Process of Intelligent Agents as Parallel Elaboration <i>S. Kojnov, V. Sgurev, V. Jotsov, Bulgaria</i>	267	Fuzzy Logic in Power System Performability <i>M. Dumitrescu, T. Munteanu, A. P. Ulmeanu, Romania</i>	326
ADMI: A Multi-Agent Architecture to Autonomously Generate Data Mining Services <i>S. Z. H. Zaidi, Malaysia, S. S. R. Abidi, Canada, S. Manikam, and Cheah Yu-N, Malaysia</i>	273	Neural Networks Modelling of Two Biotechnological Processes <i>I. Simeonov, E. Chorukova, Bulgaria</i>	331
PART II			
SESSION: Intelligent Control and Decision Support Systems		An Adaptive Integral Plus States Neural Control of Aerobic Continuous Stirred Tank Reactor <i>I. S. Baruch, Mexico, P. Georgieva, Portugal, L. A. Hernandez P., Mexico, B. Nenkova, Bulgaria</i>	337
ArgueNet: An Argument-Based Recommender System for Solving Web Search Queries <i>C. I. Chesñevar, Spain and A. G. Maguitman, USA</i>	282	A Cascading Support Vector Machines System for Gene Expression Data Classification	

<i>D. K. Iakovidis, I. N. Flaounas, S. A. Karkanis, D. E. Maroulis, Greece</i>	344	Applying a Reconfigurable Multi-Agent Scheduler to Product Distribution <i>S. P. Walsh and S. Nahavandi, Australia</i>	411
Fuzzy Two-Level Control for Anaerobic Wastewater Treatment <i>S. T. Yordanova, Bulgaria</i>	348	Collective Intelligence as a Framework for Chain Management <i>L. Sheremetov and L. Rocha-Mier, Mexico</i>	417
From Local Actions to Global Tasks: Simulation of Stigmergy Based Foraging Behavior <i>D. D. Tsankova, V. S. Georgieva, Bulgaria</i>	353	PART II	
Enhancement of DeLC for the Provision of Intelligent Mobile Services <i>I. Ganchev, Ireland, S. Stojanov, Bulgaria, M. O'Droma, Ireland, I. Popchev, Bulgaria</i>	359	Intuitionistic Fuzzy Estimation and Generalized Net Model of E-learning within a University Local Network <i>A. Shannon Australia, E. Kerre, Belgium, E. Szmidi, Poland, E. Sotirova, Bulgaria, I. Petrounias, UK, J. Kacprzyk, Poland, K. Atanassov, Bulgaria, M. Krawczak, Poland, P. Melo-Pinto, Portugal, P. Georgiev, Bulgaria, S. Melliani, Morocco, T. Kim, Korea</i>	423
Neural Network Adaptive Wavelets for Sizing of Stand-Alone Photovoltaic Systems <i>A. Mellit, M. Benghanem, A. H. Arab, A. Guessoum, K. Moulai, Algeria</i>	365	Intuitionistic Fuzzy-Valued Fuzzy Measures <i>A. I. Ban, Romania</i>	427
<hr/> Invited SESSION <hr/>			
PART I			
Multilingual On-Line Dictionary Breaking the Language Barriers in the Advent of Open Markets <i>P. Malo, R. Goncalves, R. Saraiva, S. Garcao, Portugal</i>	376	An Intuitionistic Fuzzy Component Based Approach for Identifying Web Usage Patterns <i>I. Petrounias and A. Tseng, UK, B. Kolev, Bulgaria, P. Chountas and V. Kodogiannis, UK</i>	430
Enhancing the Design of a Multi-Party Collaboration Framework with the Use of Ontologies <i>B. Roberts, UK, A. Koumpis, Greece</i>	382	Intuitionistic Fuzzy Graph Interpretations of Multi- Person Multi-Criteria Decision Making: Generalized Net Approach <i>G. Pasi, Italy, R. Yager, USA, and K. Atanassov, Bulgaria</i>	434
GAs and Simulation Techniques for Dynamic Resources Sharing and Reallocation across Workgroups <i>A. Orsoni, UK</i>	387	On Intuitionistic Fuzzy Multigraphs and Their Index Matrix Interpretations <i>A. Shannon, Australia and K. Atanassov, Bulgaria</i>	440
<hr/> Poster SESSION <hr/>			
Adding Context-Awareness to Knowledge Management in Modern Enterprises <i>W. Huang, and T. Tao, UK</i>	393	Decision Support System for Customs Examination <i>A. K. Singh and R. Sahu, India</i>	446
Generic Web-Based Platform for Virtual Organizations: the SICOV Case <i>F. Feltz, G. Simon, C. Lambert, and B. Otjacques, Luxembourg</i>	399	Question-Answer Processor for Cooperative Work in Human-Computer Environment <i>P. I. Sosnin, Russia</i>	452
MRP-Production Planning In Agile Manufacturing <i>V. T. Le, B. M. Gunn, and S. Nahavandi, Australia</i>	405	Aids to Bayesian Belief Network Construction <i>E. Rajabally, P. Sen, S. Whittle, and J. Dalton, UK</i>	457

A Cascading Support Vector Machines System for Gene Expression Data Classification

Dimitris K. Iakovidis, Ilias N. Flaounas, *Student Member, IEEE*,
Stavros A. Karkanis, *Member, IEEE*, and Dimitris E. Maroulis, *Member, IEEE*

Abstract—Microarray technology provides the ability of monitoring the gene expression levels of thousands of genes in parallel. Gene expression data classification applies for diseases' diagnosis or prediction. We propose a novel intelligent system for the classification of multiclass gene expression data. It is based on a cascading Support Vector Machines (SVM) scheme and utilizes Welch's *t*-test for the detection of differentially expressed genes. The system was applied for the discrimination of normal and lung cancer subtypes' specimens. The overall accuracy achieved was 98.5%. The results show that the proposed system can be efficiently used for microarray data analysis.

Index Terms—Classification, Gene Expression Data, Gene Selection, Microarrays, SVM

I. INTRODUCTION

A variety of techniques have been developed by molecular biologists in order to study gene expression changes associated with biological evolution mechanisms and diseases. Microarray technology first provided the advantage of monitoring the gene expression levels of thousands of genes in parallel. Microarrays consist of large numbers of individual DNA sequences printed as spots in a systematic order on a microscope's glass. Each spot produced by a DNA microarray hybridization experiment represents the expression levels' ratio of a particular gene under two different experimental conditions [1].

Microarray technology motivated computer scientists to focus on solving biological problems such as the identification of the functional roles of the genes, the way they are organized, the way they interact and the way their expression levels are changed by various diseases. The major related research areas include the detection of differential expression, pattern discovery, class prediction and inference of regulatory pathways and networks [2].

Class prediction methods involve supervised machine learning techniques for diseases' diagnosis or prediction.

This work was realized under the framework of the Operational Program for Education and Vocational Training Project "Pythagoras" cofunded by European Union and the Ministry of National Education and Religious Affairs of Greece.

D. K. Iakovidis, I. N. Flaounas, and D. E. Maroulis are with Realtime Systems and Image Analysis Group, Department of Informatics and Telecommunications, National and Kapodestrian University of Athens, 15784 Panepistimiopolis, Ilisia, Athens, Greece (e-mail: rtsimage@di.uoa.gr).

S. A. Karkanis is with the Department of Informatics and Computer Technology, Technological Educational Institute of Lamia, Lamia 35100, Greece (e-mail:sk@teilam.gr).

This is a challenging task mainly due to the following reasons:

1. Microarray data consist of a large number of features (gene expression measurements), while the number of samples involved is disproportionally small.
2. A significant percentage of genes is usually not associated with the problem under investigation.
3. The biochemical procedure used to produce microarrays, adds a lot of noise to the measurements.

The first two issues could lead to peaking phenomena associated with the "curse of dimensionality" [3], while the third introduces a large amount of uncertainty in our measurements, making the classification task harder. In order to remove irrelevant genes, identify the differentially expressed genes and reduce the feature space dimensions, gene selection algorithms are usually applied prior to the classification stage [2].

Several classification approaches have been proposed in the literature on microarray data including linear discriminant analysis, k-nearest neighbors (k-NN), parzen windows, decision trees, Neural Networks (NN) and Support Vector Machines (SVM) [4]-[8]. Comparative studies suggest that SVMs outperform other methods [5][9]. SVMs are remarkably robust machine learning algorithms that are based on statistical learning theory [10]. Their performance is not easily affected by sparse or noisy data, they resist overfitting and to the "curse of dimensionality".

The afore mentioned approaches have been applied to binary classification problems, such as the discrimination among normal and cancerous samples of colon, breast and ovarian cancer cases as well as the discrimination among two leukemia subtypes. The classification task becomes more complex as the number of classes increases. Multiclass classification approaches that have been proposed for microarray data classification include Multicategory SVMs for the classification of leukemia subtypes [11]; binary classifiers in conjunction with three combination scenarios, namely one-vs-one, one-vs-all and hierarchical partitioning for the discrimination of 14 common tumor types [12].

Under this framework we developed a novel system of cascading SVMs, for multiclass classification of gene expression data, which utilizes Welch's *t*-test for the detection of differentially expressed genes. The system was applied for the classification of normal and lung cancer subtypes samples [13].

The rest of this paper is organized in 3 sections. In section 2 the proposed system is described. In section 3 the

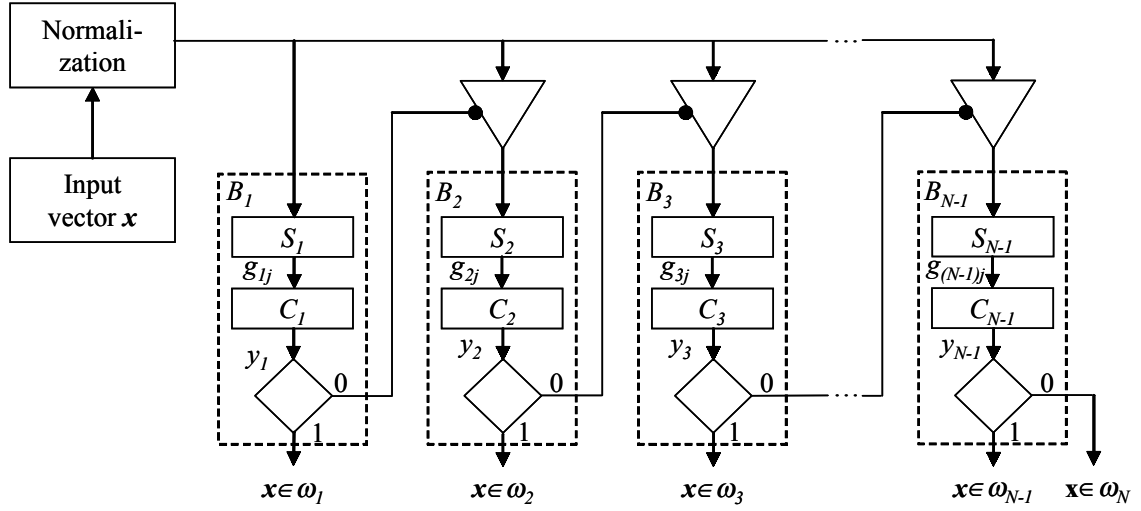


Fig. 1. Cascading SVMs system for microarray data classification.

results of the system's experimental evaluation on lung cancer data are apposed. The last section summarizes the conclusions of this study.

II. SYSTEM DESCRIPTION

The proposed system aims to the classification of a gene expression vector \mathbf{x} to its appropriate class ω_i , $i=1,2,\dots,N$. The gene expression levels are normalized to conform to zero mean and unitary variance in order to obtain directly comparable sample measurements. The system implements a cascading scheme of SVM classifiers as illustrated in Fig.1. It consists of $N-1$ blocks. Each block B_i consists of two modules. The first module noted as S_i , realizes gene selection and the second noted as C_i , implements classification. System's free parameters are tuned during training phase. Each block B_i is trained separately with a samples' subset X_i of the available training set X , where

$$X_i = \{x \in (\omega_i \cup \omega_h)\}_i, \omega_h = \bigcup_{k \neq i} \omega_k \quad (1)$$

Module S_i selects a subset of v genes g_{ij} , $j=1,2,\dots,v$ which best discriminates class ω_i from class ω_h , via Welch's t -test. The number of selected genes is determined by maximizing the performance of the classification module C_i .

Presenting a vector \mathbf{x} of unknown class to the system, module C_i is fed with the selected subset of genes, g_{ij} and outputs $y_i=1$ if $x \in \omega_i$ or $y_i=0$ if $x \notin \omega_i$. If $y_i=0$, the sample enters to the next block B_{i+1} . If $y_i=1$, the classification task terminates and \mathbf{x} is assigned to class ω_i . The last block B_{N-1} decides whether $x \in \omega_{N-1}$ or $x \in \omega_N$.

A. Welch's t -test

Welch's t -test is a statistical test that assumes unequal variances among classes and it can be applied in problems involving a small number of samples [14]. The genes are ranked based on how well they lead to a large between-class distance and a small within-class variance in the feature's space. Genes' ranking is achieved by calculating the absolute value of the t -statistic $Z(j)$ for each gene j :

$$Z(j) = \frac{m_j^i - m_j^h}{\sqrt{\frac{\sigma_j^i}{N_i} + \frac{\sigma_j^h}{N_h}}} \quad (2)$$

where (m_j^i, σ_j^i) and (m_j^h, σ_j^h) correspond to the mean and standard deviation of gene's j expression levels of the training samples that belong to ω_i and ω_h classes respectively. The number of samples belonging to each of the above classes is denoted by N_i and N_h . The larger the absolute value of $Z(j)$ the higher the expression of gene j .

B. Support Vector Machines

Let Φ be a non-linear mapping from the input space $I \subseteq \mathfrak{R}^n$ to the feature space $F \subseteq \mathfrak{R}^m$. The SVM algorithm is capable of finding a hyperplane defined by the equation

$$w\Phi(x) + b = 0 \quad (3)$$

so that the *margin of separation* is maximized. It is easy to prove [10][15] that for the *maximal margin* hyperplane,

$$w = \sum_{i=1}^N \lambda_i y_i \Phi^T(x_i) \quad (4)$$

where the variables λ_i are Lagrange multipliers that can be estimated by maximizing the quantity

$$L_D = \sum_{i=1}^N \lambda_i - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \lambda_i \lambda_j y_i y_j K(x_i, x_j) \quad (5)$$

with respect to λ_i , where the following constraints should be satisfied: $\sum_{i=1}^N \lambda_i y_i = 0$ and $0 \leq \lambda_i \leq c$, for $i = 1, 2, \dots, N$, and a given cost value c . Increasing c corresponds to a higher penalty for errors.

$K(x_i, x_j)$ is called kernel function and it is defined as the inner product

$$K(x_i, x_j) = \Phi^T(x_i)\Phi(x_j) \quad (6)$$

Linear, polynomial, Radial Basis (RBF) and sigmoid are the most common functions used as SVM kernels. We used the RBF kernel:

$$K(x_i, x_j) = e^{-\|x_i - x_j\|^2 / \gamma} \quad (7)$$

where γ is a strictly positive constant. This kernel, usually has better boundary response as it allows for extrapolation,

and most high-dimensional data sets can be approximated by Gaussian-like distributions similar to that used by RBF networks [15].

III. RESULTS

The experimentation presented in this study aims to the evaluation of the proposed system's performance. The dataset used has been first studied by Bhattacharjee *et al.* [13], who applied hierarchical unsupervised classification to reveal unknown adenocarcinoma subclasses. It consists of 203 samples spanning 6 different classes which correspond to normal lung specimens, Small-Cell Lung Carcinomas (SCLC), Adenocarcinomas (AC), Large-Cell Lung Carcinomas (LCLC), Squamous Carcinomas (SC) and ACs which are suspected to be extrapulmonary metastases (MAC). The number of samples per class is 17, 6, 127, 21, 20 and 12 respectively. Each sample is represented by a 12600 dimensional vector formed by the expression levels of the measured genes.

A 5-block cascading SVMs architecture was used for the 6-class classification problem. The block sequence used for the discrimination of the corresponding classes is presented in Table I.

TABLE I
SYSTEM'S BLOCK SEQUENCE FOR LUNG CANCER DATA CLASSIFICATION

Block	ω_i	ω_h
B_1	Normal	{SCLC, LCLC, SC, MAC, AC}
B_2	SCLC	{LCLC, SC, MAC, AC}
B_3	LCLC	{SC, MAC, AC}
B_4	SC	{MAC, AC}
B_5	MAC	AC

In each block all genes were ranked in descending significance using Welch's *t*-test. System's parameters were selected by grid search. The search parameters were the number of genes and SVM's cost *c*. Among the available genes only the 50 top-ranked were considered. Preliminary tests showed that a further increase of this number did not result in any significant increase of the classification performance. The classification performance was evaluated by adopting a Leave-One-Out (LOO) cross validation approach. LOO is commonly used when the available dataset is small providing an almost unbiased estimate of the generalization ability of a classifier [16].

Under this experimental framework the minimum number of differentially expressed genes which maximizes the classification performance of each block was determined. The classification accuracy vs. the number of genes used in blocks B_1 , B_2 , B_4 and B_5 is illustrated in Fig. 2, 3, 4 and 5 respectively. The diagram corresponding to the third block's performance was omitted because it reached 100% accuracy by using only the first ranked gene. Maximum accuracies are designated with vertical dashed lines within figures. The classification performances achieved as well as the number of selected genes per block are summarized in Table II. The overall accuracy of the proposed system reaches 98.5% (3 out of 203 samples were misclassified). It manages to accurately discriminate among normal specimens and different lung cancer types utilizing a rather small number of genes ranging from 1 to 40.

The results achieved are comparable with the results reported in [8]. In that study two gene selection methods

namely Recursive Feature Elimination (RFE) and Univariate Association Filtering (UAF) were combined with linear and polynomial SVM, NN and k-NN classifiers for the discrimination of (i) normal - cancerous, (ii) SC - {MAC, AC} and (iii) MAC - AC specimens from the same dataset.

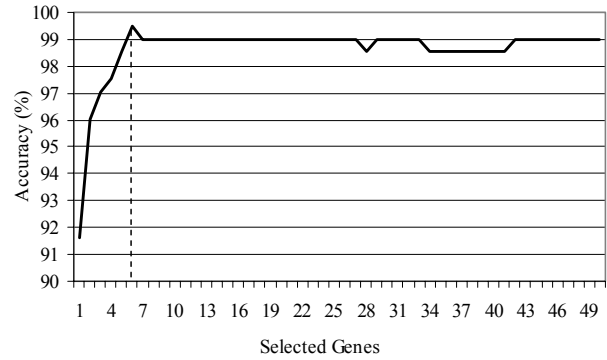


Fig. 2. Normal samples classification vs. number of genes (B_1).

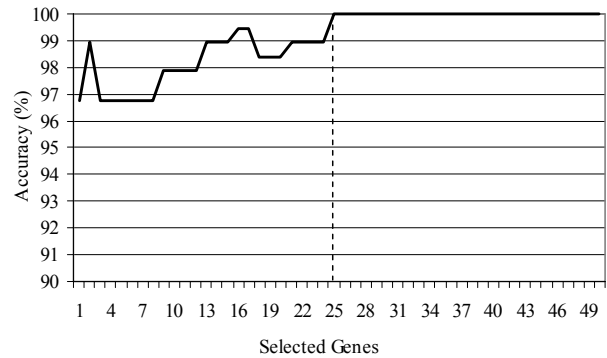


Fig. 3. SCLC samples classification vs. number of genes (B_2).

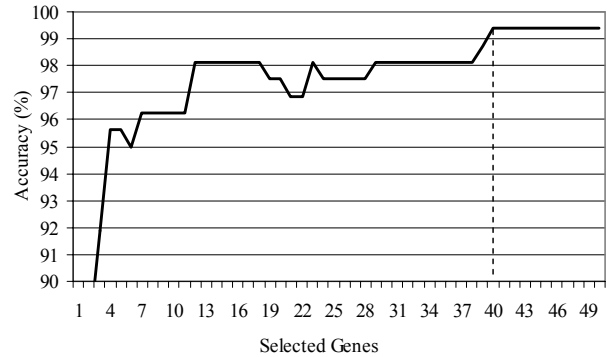


Fig. 4. SC samples classification vs. number of genes (B_4).

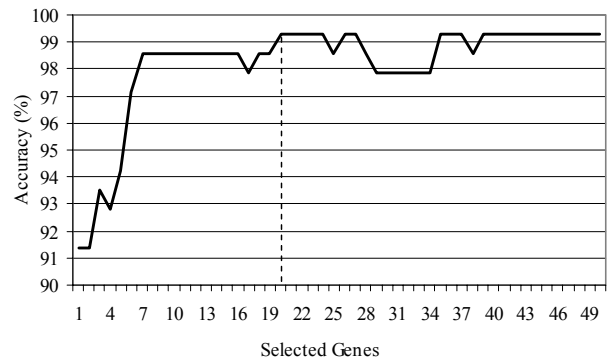


Fig. 5. MAC samples classification vs. number of genes (B_5).

These pairs of classes correspond to $\omega_i - \omega_h$ pairs handled by the B_1 , B_4 and B_5 blocks of the cascading SVMs system. The best results reported in [8] as well as the results of our approach are compared in Table III. In cases (i) and (ii) we achieved a comparable accuracy by using a significantly

smaller number of genes. In case (iii) the accuracy we achieved was higher by using only 14 genes more.

TABLE II
MAXIMUM CLASSIFICATION ACCURACY AND PARAMETERS USED PER BLOCK

Block	Selected Genes	Accuracy (%)
B_1	6	99.5
B_2	25	100
B_3	1	100
B_4	40	99.4
B_5	20	99.3

TABLE III
COMPARATIVE RESULTS

$\omega_i - \omega_h$	Cascading SVMs System		Results reported in [8]	
	Accuracy (%)	Selected Genes	Accuracy (%)	Selected Genes
Normal-Cancerous	99.5	6	99.8	100
SC- $\{\text{MAC}, \text{AC}\}$	99.4	40	99.6	500
MAC-AC	99.3	20	97.6	6

IV. CONCLUSIONS

In this paper we presented a novel system for the classification of multiclass gene expression data. It implements a cascading scheme of SVMs combined with gene selection modules. The proposed system was applied for the classification of lung cancer data. A 5-block cascading architecture was used for the discrimination of the six classes comprising the dataset. The results showed that the lung cancer classes could be characterized by a very small number of genes compared to the total 12600 genes involved in the experiment. The overall system's accuracy for this dataset was estimated 98.5%.

This study shows that the proposed system can be successfully used for the classification of gene expression data. A straightforward application of this system is disease diagnosis or even prediction under a medical decision support framework.

REFERENCES

- [1] M. K. Deyholos, and D. W. Galbraith, "High-Density Microarrays for Gene Expression Analysis," *Cytometry*, vol. 43, no. 3, pp. 229-238, 2001.
- [2] D. K. Slonim, "From patterns to pathways: gene expression data analysis comes of age," *Nature Genetics*, vol. 32, no. 12, pp. 502-508, 2002.
- [3] A. K. Jain, R. P. W. Duin, and J. Mao, "Statistical Pattern Recognition: A Review," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, pp. 4-37, Jan. 2000.
- [4] S. Dudoit, J. Fridlyand, and T. P. Speed, "Comparison of Discrimination Methods for the Classification of Tumors Using Gene Expression Data," *Journal of the American Statistical Association*, vol. 97, no. 457, pp. 77-87, 2002.
- [5] M. P. S. Brown, W. N. Grundy, D. Lin, N. Cristianini, C. W. Sugnet, T. S. Furey *et al.*, "Knowledge-based analysis of microarray gene expression data by using support vector machines," *Proceedings of the National Academy of Sciences USA*, vol. 97, no. 1, pp. 262-267, 2000.
- [6] J. Ryu, and S. Cho, "Gene expression classification using optimal feature/classifier ensemble with negative correlation," in *Proc. International Joint Conference on Neural Networks (IJCNN'02)*, 2000, pp. 198-203.
- [7] Y. Lu, and J. Han, "Cancer classification using gene expression data," *Information Systems*, vol. 28, no. 4, pp.243-268, 2003.
- [8] C. F. Aliferis, I. Tsamardinos, P. P. Massion, A. Statnikov, N. Pananapazir, and D. Hardin, "Machine learning models for classification of lung cancer and selection of genomic markers using array gene expression data," in *Proc. 16th International FLAIRS Conference*, 2003, pp. 67-71.
- [9] T. S. Furey, N. Christianini, N. Duffy, D. W. Bednarski, M. Schummer and D. Haussler, "Support vector machine classification and validation of cancer tissue samples using microarray expression data," *Bioinformatics*, vol. 16, no. 10, pp. 906-914, 2000.
- [10] V. Vapnik, "The Nature of Statistical Learning Theory," *Springer-Verlag*, 1995.
- [11] Y. Lee, and C. K. Lee, "Classification of multiple cancer types by multiclass support vector machines using gene expression data," *Bioinformatics*, vol. 19, no. 13, pp. 1132-1139, 2003.
- [12] C. Yeang, S. Ramaswamy, P. Tamayo, S. Mukherjee, R. M. Rifkin, M. Angelo *et al.*, "Molecular classification of multiple tumor types," *Bioinformatics*, vol. 17 Suppl., pp. S316-S322, 2001.
- [13] A. Bhattacharjee, W. G. Richards, J. Staunton, C. Li, S. Monti, P. Vasa *et al.*, "Classification of human lung carcinomas by mRNA expression profiling reveals distinct adenocarcinoma subclasses," *Proceedings of the National Academy of Sciences USA*, vol. 98, no. 24, pp.13790-13795, 2001.
- [14] W. Pan, "A comparative review of statistical methods for discovering differentially expressed genes in replicated microarray experiments," *Bioinformatics*, vol. 18, no. 4, pp. 546-554, 2002.
- [15] C. Burges, "A Tutorial on Support Vector Machines for Pattern Recognition," *Kluwer Academic Publishers*, Boston, 1998.
- [16] G. C. Cawley, and N. L. C. Talbot, "Efficient leave-one-out cross validation of kernel Fisher discriminant classifiers," *Pattern Recognition*, vol. 36, no. 11, pp. 2585-2592, 2003.

Author Index

Abidi, S. S. R.	273	De Cecco, M.	600
Adams, Rod	94, 150	Debei, S.	600
Aguirre, José L.	200	Debnath, Rameswar	520
Albertos, Pedro	25	Devaraj, D.	612
Alexandropoulos, Theodoros	169	Devedzic, Goran	53
Alexiev, Kiril M.	213	Devedžić, Vladan B.	244, 462, 485
Alhajj, Reda	225	Dimirovski, Georgi M.	45
Angrilli, F.	600	Dimitrova, Maya	552
Antory, David	293	Ding, Yongsheng	178
Arai, Masahiko	138	Djurić, Dragan	244, 485
Atanassov, Krassimir	423, 434, 440	Dogaru, Radu	71
Bakir, Goekhan	174	Domínguez-Brito, Antonio C.	560
Ban, Adrian I.	427	Dong, Fangyan	7
Barakova, Emilia I.	116	Dudar, Zoya	238
Baruch, Ieroham S.	337, 514	Dumitrescu, Mariana	322, 326
Bede, Barnabas	208	Erdani, Yuliadi	312
Beltran L., Rafael	514	Feltz, Fernand	399
Benghanem, M.	365	Flaounas, Ilias N.	344
Berber, Stevan M.	530	Frank, Ray J.	66, 88, 94
Bharadwaj, Vijayanand	134	Gale, Tim M.	66, 88
Bi, Zhenbo	256	Ganchev, I.	359
Bianchini, G.	600	Garcao, Steiger	376
Biletskiy, Yevgen	238	Gašević, Dragan	244, 462, 485
Boeva, Veselka	110	Gegov, Emil	557
Boutalis, Yiannis S.	53, 144	Georgiev, Peter	423
Bramer, Max	190	Georgieva, Olga I.	213
Brena, Ramon F.	261	Georgieva, Petia	337
Bubnicki, Zdzislaw	13	Georgieva, Velichka S.	353
Buchala, Samarasena	88	Georgios, Skemperis	557
Cabrera-Gámez, Jorge	560	Glesner, Manfred	71
Câmpeanu, Cezar	238	Goncalves, Ricardo	376
Cao, Aize	491	Govindhasamy, James J.	316
Čapkovič, František	299	Groumpos, Peter P.	82
Ceballos, Hector	261	Guardiola, Carlos	184
Chen, An Pin	503	Guessoum, A.	365
Chen, Kewei	7	Gunn, Bruce M.	405
Cheng, Tung Wan	503	Hadj Arab, A.	365
Chesñevar, Carlos Iván	282	Halvey, Martin	586
Chorukova, Elena	331	Hammad, Sherif A.	547
Chountas, Panagiotis	430	Hassan, Ahmed	547
Dakovski, Ljudmil	100	Hernandes P., Luis Alberto	337
Dalton, John	457	Hernández-Sosa, Daniel	560
Damjanović, Violeta	244, 485	<i>Hirota, Kaoru</i>	7, 208
Danilowicz, Czeslaw	130	Huang, Weihong	393
Darvish, Alireza	595	Hunger, Axel	312
Davey, Neil	66, 88, 94, 150	Hunt, Steve	94
Davidsson, Paul	508	Iakovidis, Dimitris K.	344
De Baets, Bernard	110	Irwin, George W.	293, 316

Isern-González, Josep	560	Loutas, Evangelos	541
Ivanova, Elena	572	Lughofer, Edwin	184
Ivanova, Z.	580	Luján, José Manuel	184
Iyoda, Eduardo Masato	208	Maguitman, Ana Gabriela	282
Jia-ju, Qiu	61	Majeed, Basim	164
Jordanov, Ivan	34	Malo, Pedro	376
Jotsov, Vladimir S.	121, 267	Manikam, S.	273
Jovanović, Jelena	462	Manolov, Ognyan	305
Kacprzyk, Janusz	423	Marco, Santiago	174
Kaikhah, Khosrow	40	Maroulis, Dimitris E.	344
Kalpakam, N. V.	474	Martin, Trevor	164
Kankanahalli, Srinivas	134	Mavrantza, Ourania D.	288
Kao, Chiang	231	McCullough, Geoffrey	293
Karadimas, Nikolaos V.	288	McLoone, Seán F.	316, 326
Karkanis, Stavros A.	344	Melliani, Said	423
Karras, Dimitrios A.	144	Mellit, A.	365
Kasabov, Nikola	19	Melo-Pinto, Pedro	423
Kaya, Mehmet	225	Mertens, Sascha	312
Kayafas, Eleftherios	169	Mertzios, Basil G.	53, 144
Kaye, Paul	150	Michalas, Angelos	541
Keane, Mark T	586	Mikhailov, Ludmil	497
Kedar, Amol	595	Minchev, Stefan V.	105
Kerre, Etienne	423	Minchev, Zlatogor	305
Kim, Taekyun	423	Montaño, Omar	200
Klement, Erich Peter	184	Moulai, K.	365
<i>Klir, George</i>	5, 6	Munteanu, Toader	322, 326
Kodogiannis, Vassilis	430	Nahavandi, Saeid	405, 411
Koehl, Ludovic	178	Najarian, Kayvan	595
Kojnov, Stefan L.	267	Nauck, Detlef	164
Kolev, Boyan	430	Nenkova, Boyka	337, 514
Koprinkova-Hristova, Petya D.	219	Nikolov, Dimitar	552
Kottas, Theodore L.	53	Nobuhara, Hajime	208
Koumpis, Adamantios	382	Noykov, Svetoslav	305
Kraounakis, Stylianos	541	O'Droma, M.	359
Kräußling, Andreas	566	Orsoni, Alessandra	387
Krawczak, Maciej	423	Otjacques, Benoît	399
Kruger, Uwe	293	Paleru, Radhakrishna	595
Lakov, Dimitar V.	250	Palit, Ajoy K.	76
Lambert, Cécile	399	Panchariya, P. C.	76
Lavallée, Marc	468	Papageorgiou, Elpiniki I.	82
Lavesson, Niklas	508	Papavasileiou, Athanasios	557
Laws, Keith R.	66	Paquin, Louis-Claude	468
Le, Vu T.	405	Pasi, Gabriella	434
Lee, Beum-Seuk	164	Pechlivanos, Lambros	535
Lee, Keum-Chang	497	Peneva, Vania	606
Li, Feng	61	Pensuwon, Wanida	150
Lin, Pei-Huang	231	Perera, Alexandre	174
Lokvenec, Irena I.	45	Petrounias, Ilias	423, 430
Loriette, S.	196	Popchev, Ivan	359, 606
Loumos, Vassili G.	169, 288	Popovic, D.	76
Lourens, Tino	116	Qin, A. K.	524
Louta, Malamati	535, 541	Radeva, Petia	552