An Effective Combination of Multiple Classifiers for Toxicity Prediction

Gongde Guo[1,2], Daniel Neagu[2], Xuming Huang[1,2], and Yaxin Bi[3]

[1] Dept. of Computer Science, Fujian Normal Univ., Fuzhou, 350007, China
[2] Dept. of Computing, Univ. of Bradford, Bradford, BD7 1DP, UK
{G.Guo, D.Neagu, X.Huang5}@bradford.ac.uk
[3] School of Computing and Mathematics, Univ. of Ulster, BT37 0QB, UK
Y.Bi@ulster.ac.uk

Abstract. The performance of individual classifiers applied to complex data sets has for predictive toxicology a significant importance. An investigation was conducted to improve classification performance of combinations of classifiers. For this purpose some representative classification methods for individual classifier development have been used to assure a good range for model diversity. The paper proposes a new effective multi-classifier system based on Dempster's rule of combination of individual classifiers. The performance of the new method has been evaluated on seven toxicity data sets. The classification accuracy of the proposed combination models achieved, according to our initial experiments, 2.97% better average than that of the best individual classifier among five classification methods (Instance-based Learning algorithm, Decision Tree, Repeated Incremental Pruning to Produce Error Reduction, Multi-Layer Perceptrons and Support Vector Machine) studied.

1 Introduction

Multiple Classifier System (MCS) has been widely applied to various fields of pattern recognition, including character recognition [1], speech recognition [2], text categorization [3],[4] and toxicity prediction [5], [6]. The idea of combination of classifiers is motivated by the observation of their complementary characteristics. It is desirable to take advantage of the strengths of individual classifiers and to avoid their weaknesses, resulting in the improvement of classification accuracy [4]. The work presented here is inspired by an idea from common sense reasoning and also from artificial intelligence research, i.e. a decision made on the basis of the multiple pieces of evidence should be more effective than one based on single piece of evidence. A classification problem is seen as a process of inferences about class concepts from concrete examples [7]. The inference process can be modeled as forward reasoning under uncertainty, as in production rule systems, which allows prior knowledge (prior performance assessments of classifiers) to be incorporated and multiple pieces of evidence from the classifiers to be combined to achieve precise classification decisions [4].

In the context of combining multiple classifiers for applications of toxicity prediction of chemical compounds, a number of researchers [5],[6] have shown that

L. Wang et al. (Eds.): FSKD 2006, LNAI 4223, pp. 481–490, 2006.

combining different classifiers can improve classification accuracy. Guo et al. [5] studied four similarity-based classifier combination methods which include Majority Voting-based combination (MV), Maximal Similarity-based Combination (MSC), Average Similarity-based Combination (ASC) and Weighted Similarity-based Combination (WSC). MV is the simplest approach, where the classification decision on each class is made on the basis of majority classifiers being in favor of that class for a given input [9]. MSC is based on the local highest similarity among a set of individual classifiers for combination. The classifier with highest local similarity will be dynamically selected for classifying the instances. ASC is a global combination method, where the similarities to each class are determined by individual classifiers and averaged together. The averaged similarities are then used for class label assignment to each test instance [5]. WSC is an intermediate approach between MSC and ASC, where instead of selecting the best classifier with the highest local similarity or considering all the classifiers' similarities to each class into account, WSC uses a control parameter α, where $0 < \alpha < 1$, to control the balance between the local optimization and global optimization [5].

In this paper, we propose to use Dempster's rule of combination to combine multiple classifiers for toxicity prediction of chemical compounds. Dempster's rule of combination provides a theoretical underpinning for achieving more accurate prediction through aggregating the majority voting principle and the belief degrees of decisions. The work presented in this paper mainly focuses on combining the outputs from different classifiers at the measurement level and incorporating the prior performance (prior knowledge) of each classifier into the definition of the mass functions, which is different from the work done by Xu et al. [1] and Bi et al. [4]. Xu et al aimed at combining the outputs from classifiers at the label level, and Bi et al. incorporate the prior performance of each classifier into the classification decision process.

2 Background Knowledge

Consider a number of exhaustive and mutually exclusive propositions $h_i, i = 1, \ldots, m$, which form a universal set Θ, called the frame of discernment. For any subset $H_i = \{h_{i1}, \ldots, h_{ik}\} \subseteq \Theta$, h_{ij} $(0 < j \leq k)$ represents a proposition, called a focal element, and when H_i is one element subset, i.e. $H_i = \{h_i\}$, it is called a singleton. All the subsets of Θ constitute a powerset 2^{Θ}, i.e. for any subset $H \subseteq \Theta$, $H \in 2^{\Theta}$. The D-S theory uses a numeric value in a range [0, 1] to represent the strength of some evidence supporting a proposition $H \subseteq \Theta$ based on a given evidence, denoted by $m(H)$, called the mass function, and uses a sum of strength for all subsets of H to indicate a belief degree to the proposition H on the basis of the same evidence, denoted by *bel(H)*, often called belief function. The formal definitions for these functions are given below [10]:

Definition 1. *Let Θ be a frame of discernment, given a subset* H $\subseteq \Theta$*, a mass function is defined as a mapping* $m : 2^{\Theta} \rightarrow [0, 1]$*, and satisfies the following conditions:*

$m(\phi) = 0$
$\sum_{H \subseteq \Theta} m(H) = 1$

Definition 2. *Let Θ be a frame of discernment and* m *be a mass function on Θ, the belief of a subset $H \subseteq \Theta$ is defined as*

$$bel(H) = \sum_{B \subseteq H} m(B) \tag{1}$$

and satisfies the following conditions:
$bel(\phi) = 0$
$bel(\Theta) = 1$

When H is a singleton, $m(H) = bel(H)$. It can be seen that a belief function gathers all of the support that a subset H gets from all of the mass functions of its subsets.

Definition 3. *Let m_1 and m_2 be two mass functions on the frame of discernment Θ, and for any subset $H \subseteq \Theta$, the orthogonal sum of two mass functions on H is defined as:*

$$m(H) = m_1 \oplus m_2(H) = \frac{\sum_{X,Y \subseteq \Theta, X \cap Y = H} m_1(X) \times m_2(Y)}{1 - \sum_{X,Y \subseteq \Theta, X \cap Y = \phi} m_1(X) \times m_2(Y)} \tag{2}$$

This formula is also called Dempster's rule of combination. It allows two mass functions to be combined into a third mass function, pooling pieces of evidence to support propositions of interest.

3 Proposed Combination Technique

3.1 Definition of Mass Function

Let φ be a classifier, $C = \{c_1, c_2, \ldots, c_{|C|}\}$ be a list of class labels, and d be any test instance, an assignment of class labels to d is denoted by $\varphi(d) = \{s_1, s_2, \ldots, s_{|C|}\}$, where $s_i \geq 0, i = 1, 2, ..., |C|$ represents the relevance of the instance d to the class label c_i. The greater the score assigned to a class, the greater the possibility of the instance being under this class. For convenience of discussion, we define a function ϖ, $\varpi(c_i) = s_i + \delta$ for all $c_i \in C$, where $1 > \delta > 0$ represents the prior knowledge of classifier φ. It is clear that $\varpi(c_i) > 0, i = 1, 2, ..., |C|$. Alternatively, $\varphi(d)$ is written as $\varphi(d) = \{\varpi(c_1), \varpi(c_2), \ldots, \varpi(c_{|C|})\}$ which is treated as a general form of the output information at the measurement level.

A formal definition of mass function in this context is described as follows:

Definition 4. *Let C be a frame of discernment, where each class label $c_i \in C$ is a proposition that the instance d is of class label c_i, and $\varphi(d)$ be a piece of evidence that indicates a possibility that the instance comes from each class label*

$c_i \in C$, then a mass function is defined as a mapping, m: $2^C \rightarrow [0,1]$, i.e. mapping a basic probability assignment (bpa) to $c_i \in C$ for $1 \leq i \leq |C|$ as follows:

$$m(\{c_i\}) = \frac{\varpi(c_i)}{\sum_{j=0}^{|C|} \varpi(c_j)} \quad where \quad 1 \leq i \leq |C| \tag{3}$$

This expresses the degrees of belief in propositions of each class label to which a given instance should belong. The mass function defined in this way satisfies the conditions given in Definition 1.

With formula (3), the expression of the output information $\varphi(d)$ can be rewritten as $\varphi(d) = \{m(\{c_1\}), m(\{c_2\}), \ldots, m(\{c_{|C|}\})\}$. Therefore two or more outputs derived from different classifiers as pieces of evidence can be combined by using formula (2) to obtain a combined output as a new piece of evidence, forming a combined classifier for classification tasks.

3.2 Combination Method

Given a group of learning algorithms and a training data set, each of learning algorithms can build one or more classifiers (models) based on different subsets, e.g., feature subsets, of training data set. Moreover, different classification algorithms can build different classifiers on the same subsets. The combination task of multiple classifiers, in this context, is to summarize the classification results by the classifiers derived from diverse learning algorithms on different feature subsets.

Let ψ be a group of L learning algorithms, $\varphi_1^k, \varphi_2^k, \ldots, \varphi_n^k$ be a group of classifiers associated with learning algorithm L_k, where $1 \leq k \leq L$ and n is a parameter that is related to the number of feature subsets, then each of the classifiers, φ_i^k assigns an input instance d to Y_i^k, i.e. $\varphi_i^k(d) = Y_i^k$ and $1 \leq i \leq n$. The results output by multiclassifiers are represented as a matrix:

$$\begin{bmatrix} Y_1^1 & Y_2^1 & \ldots & Y_n^1 \\ Y_1^2 & Y_2^2 & \ldots & Y_n^2 \\ \ldots & \ldots & \ldots & \ldots \\ Y_1^L & Y_2^L & \ldots & Y_n^L \end{bmatrix} \tag{4}$$

where Y_i^k is a vector denoted as $(m_i^k(c_1), m_i^k(c_2), ..., m_i^k(c_{|C|})$. Each row in the matrix corresponds to one of learning algorithms, and each column corresponds to one of the feature subsets, i.e. Y_i^k is the result yielded by the classifier φ_1^k - a classifier built by L_k learning algorithm on i feature subset. If the number of classification algorithms $L = 5$, and the number of feature subsets is 5, 5 classifiers will be generated by each of the classification algorithms, denoted by $\{\varphi_1^k, \varphi_2^k, \ldots, \varphi_5^k\}_{k=1}^5$. Thus the combination task based on this matrix is made both on the columns and rows, i.e. for each column, all the rows will be combined using formula (5), and the combined results in each column will be combined again using formula (6), thereby producing a new mass distribution over all the class labels that represents the consensus of the assignments of the multiple

classifiers to test class labels. The final classification decision will be made by using the decision rule of formula (7).

$$m_i^{'}(c_i) = m_i^1 \oplus m_i^2 \oplus \ldots \oplus m_i^L = [\ldots[[m_i^1 \oplus m_i^2] \oplus m_i^3] \oplus \ldots \oplus \ m_i^L](c_i) \quad (5)$$

$$bel(c_i) = m_1^{'} \oplus m_2^{'} \oplus \ldots \oplus m_K^{'} = [\ldots[[m_1^{'} \oplus m_2^{'}] \oplus m_3^{'}] \oplus \ldots \oplus \ m_K^{'}](c_i) \quad (6)$$

With all belief values of class labels to which class labels could belong obtained by using Equation (5) and (6), we can define a decision rule for determining a final class label in general cases below:

$$\varphi_{DRC}(d) = c_i \ if \ bel(c_i) = argmax_{c_i \in C} \ \{bel(c_i) | i = 1, 2, \ldots, |C|\} \quad (7)$$

In Equation (7) the abbreviation *DRC* stands for Dempster's rule of combination.

4 Experiments and Evaluation

4.1 Data Sets

To evaluate the effectiveness of our proposed classifier combination method, seven toxicity data sets: Trout, Bee, Daphnia, Dietary_Quail, Oral_Quail, APC and Phenols from the real-world applications have been collected for evaluation. Among these data sets five of them, i.e. Trout, Bee, Daphnia, Dietary_Quail and Oral_Quail come from DEMETRA project [11], each of them contains all the descriptors from both 2D_MDL_ABLegend and 2D_Pallas subsets; APC data set is proposed by CSL [12]; Phenols data set comes from TETRATOX database [13]. Some general characteristics of the data sets are given in Table 1.

Table 1. General information about the data sets

Data set	NF	NFFS	NN	NO	NB	NC	NI	CD
Trout	248	22	0	22	0	3	282	129:89:64
Bee	252	11	0	11	0	5	105	13:23:13:42:14
Daphnia	182	20	0	20	0	4	264	122:65:52:25
Dietary_Quail	254	12	0	12	0	5	123	8:37:34:34:10
Oral_Quail	253	8	0	8	0	4	116	4:28:24:60
APC	248	6	0	6	0	4	60	17:16:16:11
Phenols	173	11	0	11	0	3	250	61:152:37

Titles of columns in Table 1 have the following meanings: NF - Number of Features; NFFS - Number of Features after Feature Selection; NN - Number of Nominal features; NO - Number of Ordinal features; NB - Number of Binary features; NC - Number of Classes; NI - Number of Instances; CD - Class Distribution.

4.2 Classifiers

Five classification methods involved in generating classifiers for combination are chosen in terms of their representability and diversity which include the Instance-based Learning algorithm (IBL), Decision Tree learning algorithm (DT), Repeated Incremental Pruning to Produce Error Reduction (RIPPER), Multi-Layer Perceptrons (MLPs) and Support Vector Machine (SVM). The IBL, DT, RIPPER, MLPs, and SVM used in our experiments are from the Weka software package [14]. A brief introduction of the five classifiers applied in this study is given below:

Instance Based Learners: IBLs classify an instance by comparing it to a set of pre-classified instances and choose a dominant class of similar instances as the classification result.
Decision Tree: DT is a widely used classification method in machine learning and data mining. The decision tree is grown by recursively splitting the training set based on a locally optimal criterion until all or most of the records belonging to each of the leaf nodes bear the same class label.
Repeated Incremental Pruning to Produce Error Reduction: RIPPER is a propositional rule learning algorithm that performs efficiently on large noisy data sets. It induces classification (if-then) rules from a set of pre-labeled instances and looks at the instances to find a set of rules that predict the class of earlier instances. It also allows users to specify constraints on the learned if-then rules to add prior knowledge about the concepts, in order to get more accurate hypothesis.
Multi-Layer Perceptrons: MLPs are feedforward neural networks with one or two hidden layers, trained with the standard backpropagation algorithm. They can approximate virtually any input-output map and have been shown to approximate the performance of optimal statistical classifiers in difficult problems.
Support Vector Machine: SVM is based on the Structural Risk Minimization principle from statistical learning theory. Given a training set in a vector space, SVM finds the best decision hyperplane that separates the instances in two classes. The quality of a decision hyperplane is determined by the distance (referred as margin) between two hyperplanes that are parallel to the decision hyperplane and touch the closest instances from each class.

4.3 Combination Schemes

(1) Majority Voting-based Combination (MVC)
Given x a new instance to be classified with true class label t_x and k predefined classifiers $A_1, A_2, \cdots, A_k$ respectively, where classifier A_i approximates a discrete-valued function $f_{A_i} : \Re^n \longrightarrow C$, then the final class label of x is:

$$f(x) \leftarrow argmax_{c \in C} \sum_{i=1}^{k} \delta(c, f_{A_i}(x)) \tag{8}$$

where $\delta(a, b) = 1$ if a=b, and $\delta(a, b) = 0$ otherwise.

Based on the hypothesis above, the classification result of x classified by A_j is a vector of probabilities of x to each class $P =< P_{j1}, P_{j2}, \cdots, P_{jm} >$, where $j = 1, 2, \cdots, k$ and m is the number of predefined classes. The final class label of x can be obtained either as:

(2)Maximal Probability-based Combination (MPC)

$$f_1(x) \leftarrow argmax_{c_v \in C}\{max_u\{P_{uv}|u = 1, 2, \cdots, k\}|v = 1, 2, \cdots, m\} \quad (9)$$

(3) Average Probability-based Combination (APC)

$$f_2(x) \leftarrow argmax_{c_v \in C}\{\sum_{u=1}^{k}(P_{uv}/k)|v = 1, 2, \cdots, m\} \quad (10)$$

4.4 Statistical Tool for Comparison

There are many approximate statistical tests for determining whether one learning method outperforms another on a particular learning task. Among these the *Signed Test* [15] is commonly used. Here we give a brief description of this method which will be used to measure the statistical difference between the performances of two classification methods in the next section.

The Signed Test [15] is a general statistical tool for comparing the performance of different classification methods. Given n data sets, let n_A (n_B, respectively) be the number of data sets in which classification method A does better (worse respectively) than classification method B in terms of the classification accuracy. Then we have:

$$z = \frac{\frac{n_A}{n_A+n_B} - p}{\sqrt{\frac{p \times q}{n_A+n_B}}} \approx N(0, 1) \quad (11)$$

where p is the probability that classification method A does better than classification method B; and q=1-p. Under the null hypothesis, p=0.5, so

$$z = \frac{\frac{n_A}{n_A+n_B} - 0.5}{\sqrt{\frac{0.5 \times 0.5}{n_A+n_B}}} \approx N(0, 1) \quad (12)$$

which has (approximately) a standard normal distribution $N(0, 1)$. We can reject the null hypothesis that two classification methods are the same in terms of performance if $|Z| > Z_{\infty,0.975} = 1.96$.

4.5 Evaluation

[Experiment 1]. In this experiment, we test both five classification methods, i.e. IBL, DT, RIPPER, MLPs and SVM, and four combination methods, i.e. MVC, MPC, APC and DRC (the abbreviation DRC here stands for the proposed combination method which is based on Dempster's rule of Combination), over seven toxicity data sets using a ten-fold cross validation. The class distribution

Table 2. Performance of individual classifiers evaluated on seven data sets

Data set	IBL	k	DT	RIPPER	MLPs	LR	SVM
TROUT	59.93	5	55.32	56.74	58.16	0.9	62.06
ORAL_QUAIL	57.76	5	62.93	60.34	51.72	0.3	65.52
DAPHNIA	54.17	5	50.38	50.00	53.41	0.3	54.55
DIETARY_QUAIL	48.78	10	45.53	39.84	55.28	0.3	48.78
BEE	58.09	5	45.71	46.67	51.43	0.3	53.33
PHENOLS	74.80	10	74.40	76.40	78.40	0.3	80.00
APC	43.33	5	43.33	40.00	40.00	0.3	43.33
Average	56.69	/	53.94	52.86	55.49	/	58.22

Table 3. Performance of different combination methods evaluated on seven data sets

Data set	MVC	MPC	APC	DRC
TROUT	63.12	56.38	59.22	64.93
ORAL_QUAIL	62.93	56.03	60.34	63.34
DAPHNIA	54.17	53.78	53.78	54.92
DIETARY_QUAIL	53.66	43.90	52.03	53.78
BEE	58.10	42.86	55.24	60.29
PHENOLS	80.40	79.20	82.40	82.40
APC	38.33	40.00	36.67	40.00
Average	58.67	53.16	57.10	59.95

of each data set is presented in Table 1. The experimental results are presented in Table 2 and 3.

In Table 2, each row recorded the best performances of different classification methods evaluated on a feature subset of the leftmost data set by CfsSubsetEval method which is implemented in the Weka software package [14]. Parameter k stands for the number of nearest neighbors chosen for IBL, which is tuned from 1 to 10 with step 1; LR represents the learning rate set for MLPs, which is tuned from 0.1 to 0.9 with step 0.1.

Table 3 reported the experimental results of different classifier combination methods carried out on the seven aforementioned data sets. The performances of MVC, MPC and APC in Table 3 are based on the results reported in Table 2. The performance of DRC is calculated on a $L \times n$ performance matrix by using Dempster's rule of combination where L stands of the number of classifiers and n stands for the number of feature subsets for each toxicity data set.

Eight feature selection methods are involved in extracting different subsets for each original toxicity data set, which are: Correlation-based Feature Selection; Chi-Chi squared ranking filter; Consistency Subset evaluator; Gain Ratio feature evaluator; Information Gain ranking filter; kNNMFS Feature Selection [8]; ReliefF ranking filter; SVM feature evaluator. All the feature selection methods except kNNMFS are implemented in the Weka software package [14], where kNNMFS is implemented in our own prototype system.

From Table 2 and 3 it is clear that the average classification accuracy of DRC based combination method over seven data sets is better than that of any other classification methods. Moreover, DRC based combination method performs best compared to other classifier combination methods.

[Experiment 2]. The goal of this experiment is to measure the statistical difference between the performances of any two methods studied. We compare the performance of any two classification methods based on the results obtained in Table 2 and 3. The statistical difference between the performances of any two methods is calculated using the signed test and is given in Table 4.

Table 4. The signed test of different classifiers

Signed Test	IBL	DT	RIPPER	MLPs	SVM	MVC	MPC	APC
$n_A : n_B$	6:1	6:1	6:0	5:1	5:2	7:0	6:0	6:0
DRC	1.89(+)	1.89(+)	2.45(+)	1.63(+)	1.13(-)	2.65(+)	2.45(+)	2.45(+)

In Table 4, the item 1.63(+) in cell (3, 5), for example, means DRC is better than MLPs in terms of performance over the seven data sets. That is, the corresponding $|Z| > Z_{0.90} = 1.415$. The item 1.13(-) in cell (3, 5) means there is no significant difference in terms of performance between DRC and APC over seven data sets as the corresponding $|Z| < Z_{0.90} = 1.415$. From the statistical point of view the proposed DRC classifier combination algorithm outperforms individual classification algorithms and other combination systems with an exception of SVM. Although there is no significant difference in terms of performance between DRC and SVM, the average classification accuracy of DRC is still 2.97% better than that of SVM.

5 Conclusions

In this work, we proposed an approach for combining multiple classifiers using Dempster's rule of combination. Various experiments have been carried out on seven collected toxicity data sets from real-world applications to evaluate the performance of classification algorithms individually and in combination. Based on our experimental results, it is fairly to draw a conclusion: the performance of the combination method based on Dempster's rule of combination is better than that of any other combination method studied, i.e. MVC, MPC and APC, and is 2.97% on average, better than the best individual classification method SVM. The experimental results have shown the promise of the proposed approach. However more experiments both on toxicity data sets and also benchmark data are necessary for a full evaluation of the approach proposed.

Acknowledgments

This work is partially funded by the EPSRC project PYTHIA: GR/T02508/01. The authors acknowledge also the support of the EU EP5 project DEMETRA (www.demetra-tox.net).

References

1. L.Xu, A.Krzyzak, C.Y.Suen. Several Methods for Combining Multiple Classifiers and Their Applications in Handwritten Character Recognition, IEEE Trans. on System, Man and Cybernetics, Vol. 22 (3), 1992, pp.418-435.
2. T.Denoeux. A neural Network Classifier based on Dempster-Shafer Theory. IEEE transactions on Systems, Man and Cybernetics A, Vol. 30(2), 2000, pp.131-150.
3. Y.Yang, T.Ault, T.Pierce. Combining Multiple Learning Strategies for Effective Cross Validation. Proc. of The Seventeenth International Conference on Machine Learning (ICML'00), 2000, pp.1167-1182.
4. Y.Bi, D.Bell, H.Wang, G.Guo et al. Combining Multiple Classifiers for Text Categorization using Dempster-Shafer Theory of Evidence. In Proc. of the 1^{st} International Conference on Modeling Decisions for Artificial Intelligence (MDAI'2004), 2004, pp.127-138.
5. G.Guo, D.Neagu. Similarity-based Classifier Combination for Decision Making. Proc. of IEEE International Conference on Systems, Man and Cybernetics, 2005, pp.176-181.
6. E.Benfenati, P.Mazzatorta, D.Neagu, G.Gini. Combining Classifiers of Pesticides Toxicity through a Neuro-fuzzy Approach, Lecture Notes in Computer Science, Volume 2364, 2002, pp. 293.
7. T.K.Ho. Multiple Classifier Combination: Lessons and Next Steps, Tin Kam Ho, in A. Kandel, H. Bunke, (eds.), Hybrid Methods in Pattern Recognition, World Scientific, 2002, pp.171-198.
8. G.Guo, D.Neagu, M.T.D.Cronin. Using kNN Model for Automatic Feature Selection. In Proc. of ICAPR 2005, LNCS 3686, pp. 410-419, 2005.
9. Y.H.Li and A.K.Jain. Classification of Text Documents. The Computer Journal, Vol. 41(8), 1998, pp.537-546.
10. J.W.Guan, and D.Bell. Evidence Theory and its applications. 1991, Vol.1 & 2.
11. EU FP5 Quality of Life DEMETRA QLRT-2001-00691: Development of Environmental Modules for Evaluation of Toxicity of pesticide Residues in Agriculture.
12. CSL: Development of artificial intelligence-based in-silico toxicity models for use in pesticide risk assessment, 2004-2007.
13. T.W.Schultz. TETRATOX: Tetrahymena pyriformis population growth impairment endpoint-A surrogate for fish lethality. Toxicol. Methods 7: 289-309, 1997.
14. I.H.Witten and G.Frank. Data Mining: Practical Machine Learning Tools with Java Implementations, Morgan Kaufmann, San Francisco, 2000.
15. G.W.Snedecor and W.G.Cochran. Statistical Methods. Iowa State University Press, 1989.