ROBUST FEATURE SELECTION AND CLASSIFICATION USING HEURISTIC ALGORITHMS BASED ON CORRELATION FEATURE GROUPS

Canan BATUR
Yıldız Technical University, Department of Computer Engineering, Istanbul- TURKEY
canan@ce.yildiz.edu.tr

Banu DİRİ
Yıldız Technical University, Department of Computer Engineering, Istanbul- TURKEY
banu@ce.yildiz.edu.tr

Abstract: The complexity of multidimensionality is one of the frequently encountered problems in the high-dimensional data space. The fact that multidimensionality in the data space increases and reaches great numbers brings about the problem that the number of non-informative ones among the features associated with the target class increases along with the data set complexity. The fact that all features included in the high-dimensional data space are not distinctive or do not contain critical information generally leads to difficulties at the learning stage. At this point, the importance of feature selection emerges. Feature selection is a problem of minimum subset selection from the original feature set for the best accuracy estimation. The neglected subject in the feature selection is ensuring that the inconsistency problem of the selected feature sets is brought to a solution. Studies carried out in recent years have focused on obtaining feature groups with which the group to which each feature belongs is associated with a class label rather than standard feature selection methods in which a single feature set is obtained. Within the scope of this study, each feature group obtained by group-based learning was presented as a solution candidate to heuristic methods. This paper proposes a novel feature selection method to Artificial Immune Recognition System (AIRS) variations in order to find robust gene sets from high dimensional microarray data. The unique feature of this feature selection method is that utilizes correlation based feature groups in order to increase the reliable classifying accuracy from optimal feature groups. We test the performance of the proposed local feature selection method for AIRS variations on high dimensional microarray data sets. We compare proposed LFSAIRS variations which are LFSAIRS1, LFSAIRS2, Parallel LFSAIRS1, and Parallel LFSAIRS2 with the Standard Genetic Algorithm (SGA), Sequential Forward Selection (SFS) and Sequential Backward Elimination (BES) approaches. Results of the robustness were evaluated by the Jaccard test and classification accuracy was evaluated using k-NN, SVM, Naïve Bayes and Random Forest classifiers. Results show that the proposed methods capable of finding robust gene subsets with high classification accuracy.
Keywords: Machine Learning, Local Feature Selection, Heuristic Approaches, Classification

Introduction

Gene expression microarray data sets due to the have characteristics such as high-dimensionality and small sample dimension their classification becomes hard. The fact that multidimensionality in the data space increases and reaches great numbers brings about the problem that the number of non-informative ones among the features associated with the target class increases along with the data set complexity. Many feature selection algorithms were developed for the purpose of reducing the dimensionality of this kind of data and improving the accuracy of classifiers. Feature selection is a problem of minimum subset selection from the original feature set for the best accuracy estimation.

Feature selection algorithms encountered inconsistency problem in many cases. One reason for the encountered inconsistency situation is the selection of the minimum feature set composed of features that give the best classifier accuracy, which is the classical purpose of feature selection algorithms. Furthermore, features having a high correlation with each other in the feature set defined by the feature selection algorithm can select different features in cases where the parameters of the feature selection algorithm are set differently. For the same feature selection algorithm, minor variations within the train data can also result in the selection of different feature subsets every time. Another reason for inconsistency encountered in feature selection algorithms is the small number of samples in the high-dimensional data space. The main observation here is that the groups formed by the associated features are generally present in high-dimensional data and these groups are resistant to the variations of training samples. Furthermore, sub-sampling of training data and the determination of stable feature

groups also make convergence possible to the structure of the original feature groups. Another observation is that each feature group is improved by being optimized with heuristic algorithms and learning is performed at the group level. Within the scope of this study, the feature groups formed by the correlation-based strategy within the feature selection framework were taken as a basis. An attempt to develop the correlation feature groups defined using high-dimensional data with the meta-dynamics of the local feature selection method to Artificial Immune Recognition System (AIRS) variations (LFSAIRS1, LFSAIRS2, Parallel LFSAIRS1, and Parallel LFSAIRS2) and the Standard Genetic Algorithms (SGA) was made. Then robust feature selection framework compared with Sequential forward selection and backward elimination methods.

Creating correlation based feature groups from high-dimensional data is mentioned in the second part of the paper, the methods used in the study are mentioned in the third part, the novel local feature selection method to Artificial Immune Recognition System (AIRS) is mentioned in the fourth part, the data set is mentioned in the fifth part, stability performance measurements is mentioned in the sixth part and the performance measurements of the optimal feature sub-groups obtained are mentioned in the seventh part.

Creating Correlation Based Feature Groups from High-Dimensional Data

Feature groups are an effective method to reduce the complexity of multidimensionality. At the same time, factors such as the use of feature groups in learning with high-dimensional data, reducing the complexity of the model, increasing the constancy of selected features, and the decrease of the variability of the estimator is also very effective. Studies carried out in recent years have focused on obtaining feature groups with which the group to which each feature belongs is associated with a class label rather than standard feature selection methods in which a single feature set is obtained. The fact that relational features have a very high correlation in high-dimensional data sets makes it possible to use feature groups by being taken as a basis. Thus, obtaining stable feature groups resistant to the sample number variation is ensured (Loscalzo et al., 2009). The fact that feature groups are obtained from data sets with characteristics such as small sample size and high dimensionality leads to obtaining unstable results or results that are not completely optimal. Therefore, the first step is to produce a set of feature groups by targeting the convergence to the original feature groups with the frames formed to obtain feature groups. The second step is to perform the feature selection process based on the set of feature groups produced. The idea of converging to original feature groups by creating a set of feature groups is based on the principles of group-based learning method.

In this study, relational feature groups were obtained by the CFG (Correlation Based Feature Group) algorithm. The CFG is a filter-based feature selection method that sorts the feature subset by the correlation-based intuitive function. The CFG algorithm examines the usefulness of subset of attributes based on a heuristic evaluation function. In choosing a correlation-based feature, each attribute is taken into account in the correlation between the attributes, as well as the predictive predicting of the class label. The value of the heuristic evaluation function used in the evaluation of the attributes is determined by equation 1. The intuitive usability of a subset of S attributes with k attributes is represented by $meritS$, the mean attribute-class correlation is presented by rcf for ($f \epsilon S$), and the correlation between the mean attributes is presented by rff parameters.

$$meritS = \frac{k*rcf}{\sqrt{k+(k-1)*rff}} \quad (1)$$

Each of the feature groups represents a solution candidate, and the presence of the related feature in a feature group was encoded with 1 while the absence of it was encoded with 0.

Methods

A. Artificial Immune Recognition Systems

Artificial Immune Recognition system is a novel immune inspired supervised learning algorithm and consist of biological immune systems metaphors. Artificial Immune Recognition Systems consist of the stages of initialization, memory cell recognition, resource competition and the selection of memory cells.
At the initialization stage, the data set is normalized to the range of [0,1]. After normalization, the affinity threshold is calculated by equation (2). At the next stage, antigens are presented to the storage pool with antigen training. At the memory cell recognition stage, a stimulation value is assigned to these cells by stimulating the recognition cells in the memory pool. Affinity is calculated by equation (3), the stimulation values are calculated by equation (4) and (5). The recognition cell with the highest stimulation value is calculated by equation (6) then M_{cmatch} cell is cloned and mutated. The number of clones is calculated by equation (7),

$$affinity\ \ threshold = \sum_{i=1}^{n} \sum_{j=j+1}^{n} \left(\frac{affinity(\text{agi,agj})}{n(n+1)/2} \right) \quad (2)$$

$$affinity(\text{agi, agj}) = 1 - \text{Euclidean distance}(\text{agi, agj}) \quad (3)$$

$$stimulation = 1 - \text{affinity} \quad (4)$$

$$stimulation(\text{mc, ag}) = \begin{cases} \text{affinity(mc, ag)} & \text{if mc. class} = \text{ag. class} \\ 1 - \text{affinity} & \text{otherwise} \end{cases} \quad (5)$$

$$\boldsymbol{Mcmatch} = \mathbf{argmax}(\mathbf{stimulation}(\mathbf{mc}, \mathbf{ag})) \quad (6)$$

$$\boldsymbol{numClones} = \mathbf{stimulation} * \mathbf{clonalRate} \quad (7)$$

At the resource competition stage, when mutated clones are added to the ARB (artificial recognition spheres, antibody) pool, competition begins for the time source. According to the stimulation value, limited resource assignment to the ARB pool is made according to the stimulation value. ARBs without enough resources are removed from the system. When the stop criterion is achieved, the process ends, and the ARB with the highest stimulation value is selected as the candidate memory cell. At the selection of memory cells stage dynamically and evolving developed Memory cell pool in the algorithm is used for the classification process.

The basic steps of the AIRS1 algorithm, the first version of artificial immune recognition systems, and the AIRS2 algorithm, the second version, are same. The main difference between them is that the ARB pool is used as a permanent resource in the AIRS1 algorithm, it is used as a temporary resource in the AIRS2 algorithm. In the case of being used as a permanent resource, ARBs remaining from previous steps cause the algorithm to spend more time by being involved in the competition for limited resources. Therefore, the complexity of the AIRS2 algorithm is less. While AIRS1 uses the mutation parameter that can be defined by the user, AIRS2 uses the concept of somatic hyper mutation where the mutation ratio of a clone is proportional to the affinity (Torres et al, 2016). While the classes of clones may change after the mutation process in the AIRS1 algorithm, classes are not allowed to change in the AIRS2 algorithm.

Parallel-AIRS1 and Parallel-AIRS2 versions demonstrate the distributed nature of the immune systems and their parallel processing qualities. At first, each part of the training data set is assigned to np number of processes. Thus, it is ensured that np number of the memory pool is created by running the AIRS algorithm on each process. As a result, the memory pools obtained are merged (Vijendra & Laxman , 2013).

In this study, the affinity threshold value, clonal ratio, mutation ratio, np, total source, stimulating value, hypermutation ratio, run number and iteration number parameters of the Artificial Immune Recognition Algorithms took the values of 0.1, 10, 0.15,2, 150, 0.9, 2.0, 30 and 50, respectively. The fitness function of each solution candidate was calculated according to accuracy of the KNN classifier.

B. Standard Genetic Algorithms

Standard Genetic Algorithms (SGAs) are adaptive heuristic search algorithms based on natural selection and evolutionary ideas. It is based on Darwin's principle of "survival of the fittest" and firstly proposed by John Holland, his colleagues and their students at the University of Michigan.

In genetic algorithm the initial population is generated first. Every individual in the population represents a solution candidate. Each representing a possible solution to a given problem. The evaluation of each candidate solution is performed according to the fitness function determined for the problem. Each individual goodness is represented by its fitness. Individuals in the population should be selected to form a new population. Individuals in a population compete for resources and mates. To ensure the formation of a new generation following the crossover and mutation operators of the two selected individuals. Genes from "survival" individuals propagate throughout the population.

In this study, the regular crossing of the crossing methods and tournament selection of the selection methods were used. While the population size of the genetic algorithm parameter takes the value of 100, chromosome length varies according to the dynamic size of the feature group. The mutation ratio, number of tournaments, number of runs and number of iterations of the algorithm took the values of 0.8, 2, 30 and 50, respectively. The fitness function of each solution candidate was calculated according to accuracy of the KNN classifier.

C. Sequential Forward Selection and Backward Elimination

Sequential forward and backward selection techniques are simple and effective methods for selecting an attribute. While these methods create a set of attributes, they perform an extraction or addition of an attribute from the attribute subset according to the method selected at each step. The selection criterion here is the performance ratio of the classifier algorithm. According to the performance status of the specified classifier algorithm, the discriminative attributes are determined at each step. For both approaches, KNN classifier was used. In this study, run number and iteration number parameters of the Sequential forward and backward selection techniques took the values of 30 and 50, respectively.

Local Feature Selection Method to Artificial Immune Recognition System

A. LFSAIRS steps:

1. The initial set of feature group sets are created based on CFG algorithm.
2. Do for each Antigen (Ag) until training process is completed:
    - 2.1. Calculating fitness value of the each feature set is calculated by taking into account only best matching cell
    - 2.2. Until termination do :
    - 2.3. The highest fitness value of the feature set is selected as a best feature set
    - 2.4. Generation of l clones of the best feature set
    - 2.5. Mutation of the each clone
    - 2.6. Calculating fitness of the each clone by taking into account only best candidate cell
    - 2.7. Set the highest fitness value of the feature set as a candidate optimum feature subset
    - 2.8. If best candidate cell is sufficient calculate the optimum subset then go step 3. else go 2.3
3. After memory cell replacement stage, set the optimum subset of attributes as the subset of the new attribute. If training process is completed go to Step 4, else go to step 2.
4. Selection of the best optimized feature set
5. Classification of the best optimized feature sets based on test set

B. LFSAIRS Flowchart:

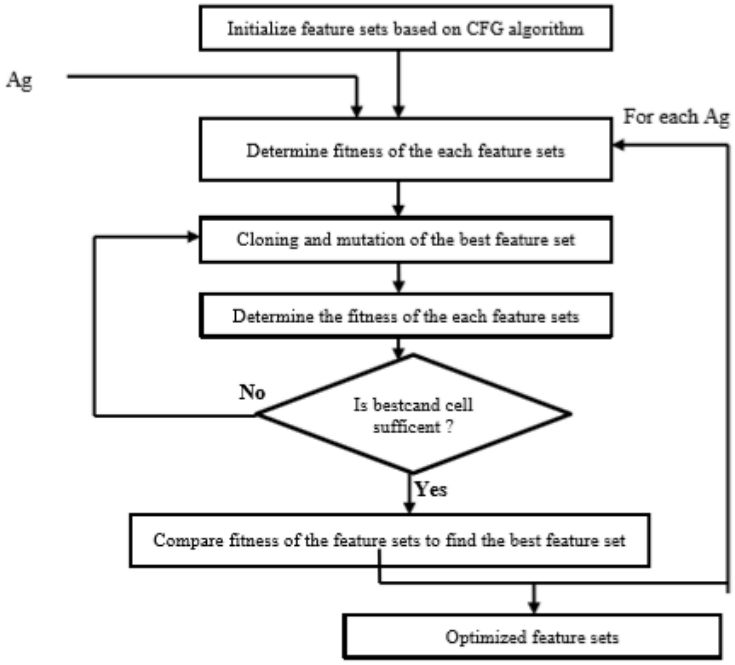

Figure 1. LFSAIRS local feature selection

Data Sets

The most common six microarray data sets were used in this study. Table 1 includes information on the genes, samples and class numbers contained in the data sets used in this study (Loscalzo et al, 2008).

Table 1: Microarray Data set

Dataset	Gene	Sample	Class
Colon	2000	62	2
Lungstd	5000	181	2
Prostate	6034	102	2
SRBCT	2308	63	4
Lymphoma	4026	62	3
Leukemia	7129	72	2

Experimentally obtained performance values were obtained by dividing the data sets as 70% training and 30% test set. A number of bootstrap data sets were obtained from the training data set in order to ensure the resistance of training samples against variations. Then, *n* number of feature groups was selected by separately running the CFG algorithm on *t* number of bootstrap data sets. We set the *t* and *n* parameters respectively to 10 and 10 for all algorithms within the scope of this paper. The number of features contained in the feature groups obtained as a result of the CFG algorithm varies dynamically specific to each data set. Learning was performed at the group level by improving the feature groups presented to the LFSAIRS1, LFSAIRS2, Parallel-LFSAIRS1, Parallel-LFSAIRS2 and Standard Genetic Algorithm like a single cell. While stability results were evaluated by the Jaccard test, their classifying accuracy was evaluated using k-NN, SVM, Naïve Bayes and Random Forest classifiers. WEKA was used to obtain classifying accuracies. For all algorithms the classifying accuracy of the most optimal solution candidate obtained at the end of each run was obtained using the test data set with 10 cross validations. The performance values added to the results were calculated by taking the average of the number of runs.

Stability Performance Measurements

The stability of feature selection approaches is obtained by measuring the similarity between feature sets. In this study, the Jaccard index was calculated using the formula given in equation (8). Parameter *m* was used to specify the number of feature sets while expressing two feature sets used for *Si* and *Sj* similarity measurement.

$$Jaccard_{Index}=IJ(Si,Sj)=| Si \cap Sj | / | Si \cup Sj | \quad (8)$$

The stability estimation was calculated using the Jaccard test formula specified in equation (9). The fact that the obtained result was high means that the stability of the relevant feature set was also high.

$$\sum(S)= \frac{2}{m(m-1)} \sum_{i=1}^{m-1} \sum_{j=i+1}^{m} IJ(Si,Sj) \quad (9)$$

Performance Results and Discussion

Within the scope of this study, it was focused on the problem of stability encountered in feature selection algorithms. As a solution to this problem, stable feature groups were obtained by combining group-level learning with the meta-dynamics of heuristic approaches.

Table 2: Average Classification accuracy of Algorithms with CFG based on k-NN and SVM

	k-NN					*SVM*				
Data set	***LFSAIRS2 %***	***LFSPAIRS2 %***	***LFSAIRS1 %***	***LFSPAIRS1 %***	***SGA %***	***LFSAIRS2 %***	***LFSPAIRS2 %***	***LFSAIRS1 %***	***LFSPAIRS1 %***	***SGA %***
Colon	80	81.1	75.8	78.8	80	82.1	85.6	83.7	83.4	81.4
Lungstd	**91.6**	**86.8**	**87.2**	**86.9**	**89.1**	**91.6**	**93**	**94.1**	**92.3**	**93.3**
Prostate	53	58.5	59	59.7	54.2	48.2	45.4	48	47.8	50
SRBCT	**86.5**	**85.3**	**86.1**	**83**	**84**	**88.4**	**89.7**	**87.6**	**87.7**	**88.1**
Lymphoma	76.9	76.5	76.1	77.6	73	88.4	86	83.4	82.8	80
Leukemia	**84.6**	**84.6**	**82.2**	**83**	**82**	**83.3**	**86.1**	**85.3**	**86**	**82.6**

Table 3: Average Classification accuracy of Algorithms with CFG based on NB and RF

	Naïve Bayes					*Random Forest*				
Data set	***LFSAIRS2 %***	***LFSPAIRS2 %***	***LFSAIRS1%***	***LFSPAIRS1 %***	***SGA %***	***LFSAIRS2 %***	***LFSPAIRS2 %***	***LFSAIRS1%***	***LFSPAIRS1 %***	***SGA %***
Colon	74.6	75	74.7	70.3	70.7	69.2	71.1	72	68.4	68.4
Lungstd	**97.2**	**95.9**	**95.8**	**95.8**	**95**	**88.8**	**89.7**	**90.2**	**89.4**	**90**
Prostate	52.3	51.1	51.4	54.7	55.2	65.3	65	63.3	62.8	61.9
SRBCT	37.6	36.1	38.4	38.8	36.1	38.4	41.9	41.5	42.3	38.4
Lymphoma	**76.9**	**76.9**	**76.9**	**78**	**76.9**	**78.4**	**77.6**	**78.4**	**78**	**76.1**
Leukemia	**86.6**	**86.6**	**86.6**	**86.6**	**86.6**	**84.6**	**86**	**84**	**84.6**	**84.6**

Table 4: Average Classification accuracy of Sequential Forward Feature Selection (SFS), Backward Elimination Feature Selection (BES) based on SVM, NB, RF and k-NN

	SFS				*BES*			
Dataset	***k-NN***	***SVM***	***NB***	***RF***	***k-NN***	***SVM***	***NB***	***RF***
Colon	71.5	70.7	59.2	76.9	76.9	78.8	76.9	71.1
Lungstd	98	93.3	95.8	95	**98.1**	94	94.5	95
Prostate	72.3	75.2	75.7	64.7	74.5	78.2	76.1	74.5
SRBCT	86.9	60	61.5	57.6	84.6	82	**41**	46.1
Lymphoma	92.3	76.9	76.9	81.5	84.6	80.7	76.9	76.9
Leukemia	86.6	86.6	80.6	80.6	83.7	82.8	80.6	83.1

Table 5: Stability Results of Algorithms Based CFG and SFS, BES

Dataset	*LFSAIRS2*	*LFSPAIRS2*	*LFSAIRS1*	*LFSPAIRS1*	*SGA*	*SFS*	*BES*
Colon	0.8	**0.9**	0.85	**0.9**	0.8	0.51	0.5
Lungstd	0.8	**0.9**	0.81	**0.9**	0.8	0.54	0.55
Prostate	0.71	**0.9**	0.83	**0.92**	0.8	0.6	0.51
SRBCT	0.8	**0.9**	0.8	**0.9**	0.8	0.67	0.63
Lymphoma	0.89	**0.9**	0.8	**0.91**	0.83	**0.72**	0.69
Leukemia	0.81	0.89	0.79	**0.9**	0.85	0.71	**0.7**

Table 6: Average Selected Feature size Results of Algorithms Based CFG and SFS, BES

Dataset	*μ LFSAIRS2*	*μLFSPAIRS2*	*μ LFSAIRS1*	*μ LFSPAIRS1*	*μ SGA*	*μ SFS*	*μ BES*
Colon	48.5	49.9	47.4	48.9	40.4	4.8	43.75
Lungstd	54.7	61.8	57.5	60	45	6.0	48
Prostate	52	62	62.3	61.1	41.8	8.4	44.2
SRBCT	69.3	68	68.6	66.7	44.9	9.2	47
Lymphoma	99.6	99.1	95.9	95.3	62	4.4	63
Leukemia	123.6	124.8	123.1	**125.1**	61.6	**4.0**	64.1

According to the average classification accuracy results shown in Table 2 and Table 3, it was observed that the algorithms achieved highest classifying accuracy result respectively in the Lungstd, SRBCT and Leukemia data sets for KNN and SVM classifiers and It was observed that the algorithms achieved highest classifying accuracy result respectively in the Lungstd, Leukemia, Lymphoma data sets for Naïve Bayes and Random Forest classifiers. It was observed that the highest classifier performance achieved by the LFSAIRS2 algorithm on Lungstd data set by 97.2% based on Naïve Bayes classifier and the lowest classifier performance showed by LFSPAIRS2 and SGA algorithms on SRBCT data set by 36.1 % based on Naïve Bayes classifier.
In Table 4 including the average classification accuracy results of Sequential Forward Selection and Sequential Backward Elimination approaches applied on six microarray data sets. It was observed that the algorithms achieved highest classifier performance by the Sequential Backward Elimination approach on Lungstd data set by 98.1% based on K-NN classifier and the lowest classifier performance showed by the Sequential Backward Elimination approach on SRBCT data set by 41 % based on Naïve Bayes classifier.
Table 5 shows the stability results of the algorithms. The results showed respectively LFSPAIRS1 and LFSPAIRS2 algorithms gave the highest stability results and BES and SFS algorithms gave the lowest stability results. While it was observed that the LFSPAIRS2 algorithm generally gave close results by 90% stability on data sets except Leukemia. LFSPAIRS1 algorithm generally gave close stability results by 90% on data sets. LFSPAIRS1 algorithm respectively gave highest stability results on Prostate and Lymphoma. It was observed that the SFS algorithm achieved the highest stability performance by 0.72 on Lymphoma data set and the lowest stability performance by 0.51 on Colon data set. It was observed that the BES algorithm achieved the highest stability performance by 0.7 on Leukemia data set and the lowest stability performance by 0.5 on Colon data set.
In Table 6, we present the average selected of feature size of the algorithms. The comparison results showed that the feature reduction capacity of the SFS algorithm was better compared to the other algorithms. It was observed that the feature reduction capacity of the SGA algorithm was better compared to the artificial immune recognition algorithm versions and BES approach. It was observed that the highest average feature size showed by the LFSPAIRS1 algorithm on Leukemia by 125.1 and the lowest average feature size achieved by the SFS algorithm on Leukemia by 4.0.

Conclusion

In high-dimensional data space, the feature subsets obtained by the feature selection algorithms cannot be stable despite having good classifier performances. The lack of stability means that the feature subsets that field experts will use in their studies will decrease the reliability of the experiment. In this study we proposed a robust feature selection framework with a novel local feature selection technique. Within the scope of this study, each feature group obtained by group-based learning was presented as a solution candidate to heuristic methods except sequential forward selection and backward elimination algorithms. When the performance results obtained were examined, it was concluded that feature groups obtained at the correlation base increased their robustness by being improved with the meta-dynamics of heuristic approaches like a single cell. The classifier and stability results obtained were compared with six commonly used microarray data sets.

References

Loscalzo S , YU L., Ding C. (2009), *Consensus group stable feature selection* , June28–July1, Paris, France

M.G., S , Vela F.G., Batista B.M, . (2016), *High dimensional feature selection via feature grouping: a variable neighborhood search approach* , July 29, Las Vegas, Spain

Loscalzo S , YU L., Ding C. (2008), *Stable feature selection via feature groups*, August 24–27, Las Vegas, Nevada, USA.Torres

Vijendra S., Laxman S., (2013), *Subspace clustering of high-dimensional data: an evolutionary approach*, 11 November

http://www.cs.waikato.ac.nz/ml/weka/

Mazel J. (2011), *Unsupervised network anomaly detection*, Phd Thesis.

Polat K., Güneş S., Sahan S., Kodaz H., (2005), *A New classification method for breast cancer diagnosis: Feature selection artificial immune recognition system (FS-AIRS*) , Berlin

Brownlee J. (2005). *Artificial immune recognition system (AIRS): A review and analysis* ,Technical Report

Menon A. edited by. (2004). *Frontiers of evolutionary computation*, Kluwer academic Publishers, Pittsburgh,USA

Oh S. , Lee J.S., Moon B.R. (2004). *Hybrid genetic algorithms for feature selection* .Vol. 26, No.11, November

Brownlee J. (2005). *Clonal selection theory & clonalg the clonal selection classification algorithm (CSCA),* Technical Report.

Dudek G. (2012). *An artificial system for classification with local feature selection* ,IEEE Transaction on Evolutionary Computation. December 2012, Czestochowa.

Wang K., Chen K.,Adrian A., (2014), *An improved artificial immune recognition system with the opposite sign test for feature selection*, Knowledge-Based Systems ,Taiwan.

Daga M., Lakhwani K., (2013), *A novel content based image retrieval implemented by nsa of ais* , International Journal of Scientific & Technology Research.

Watkins A. and Timmis J.,(2002), *Artificial immune recognition system (AIRS): Revisions and refinements*, 1st International Conference on Artificial Immune Systems (ICARIS2002), University of Kent at Canterbury, pp. 173-181.

Kalousis A., Prados J. and Hilario M., (2007), *Stability of feature selection algorithms: A study on high-dimensional spaces*, Knowladge and Information Systems.

Saeys Y., Abeel T., and Peer Y.V. (2008), *Robust feature selection using ensemble feature selection techniques*, In Proceedings of the ECML Conference, pages 313-325.

Mishra P.K, Bhusry M.,(2015), *Artificial immune system: State of the art approach*, International Journal of Computer Applications (0975-8887) Volume 120 – No.20.

Farhan M., Mohsin M., Hamdan A.,Bakar A.A., (2014), *An evaluation of feature selection technique for dendrite cell algorithm*,IEEE.

Baykasoğlu A.,Saltabaş A., Taşan S., Subulan K., (2012), *Yapay bağışıklık sisteminin çoklu etmen benzetim ortamında realize edilmesi ve gezgin satıcı problemine uygulanması*, Journal of the Faculty of Engineering and Architecture of Gazi University.

