Journal of Theoretical and Applied Information Technology
15th August 2016. Vol.90. No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

AUTOMATIC CLUSTERING OF MICROARRAY DATA USING ART2 NEURAL NETWORK

[1]FADOUA RAFII, [2]BADR DINE ROSSI HASSANI, [3]M'HAMED AIT KBIR
[1,3]LIST Laboratory, University of Abdelmalek Essaadi (UAE), Tangier, Morocco
[2]LABIPHABE Laboratory, University of Abdelmalek Essaadi (UAE), Tangier, Morocco
E-mail: [1]fadoua.rafii1@gmail.com, [2]bd.rossi@fstt.ac.ma, [3]m.aitkbir@fstt.ac.ma

ABSTRACT

The Microarray technology has offered an overall view on the activity levels of numerous genes simultaneously. A typical Microarray dataset is characterized by a large number of genes; which is usually high with respect to the number of experiments. Due to the huge number of genes and the intricacy of biological data, clustering represents a gainful exploratory technique for analyzing Microarray data. The aim of clustering is to allocate objects to the adequate groups or clusters. On one hand, the objects belonging to the same cluster are more identical to each others. On the other hand, objects from dissimilar clusters are as unalike as possible. In the present work, the major goal is to extract automatically the number of clusters in the lung cancer Microarray dataset. In fact, a clustering technique based on Adaptive Resonance Theory is used; without any prior information about the number or the form of the clusters behind the processed data. The chosen form of Adaptive Resonance Theory architectures is ART2, which quickly self-organizes categories of pattern recognition in response to arbitrary presented sequences of both analog or binary input patterns [1]. The proposed approach was simulated on raw data and reduced data, using the PCA technique. The efficiency of the applied method was evaluated by measuring the performance of the classifiers against lung cancer Microarray data.

Keywords: *Microarray technology; Clustering; Microarray Data; ART2*

1. INTRODUCTION

The Microarray technology has been involved in tremendous applications such as:

- Gene discovery;
- Diagnosis and prognosis of diseases;
- Drug discovery;
- Toxicological research.

Microarray data constitute worthwhile information obtained by gathering results of specific experiments. Microarray experiments are following determined steps to get the desired data. The major goals of these experiments include the studies of gene expression in different types of cells; which could be exposed to distinct conditions (physical, chemical or biological). However, the generated Microarray data are affected by various factors; such as the instruments imperfections, the procedures and the materials involved. By preference, a numerical value in the Microarray matrix should reflect the right level of transcript abundance in the measured combination of gene sample. Thereby, it is essential to appeal the pretreatments techniques to extract the significant data. Thus, it has been concentrating on the application of these techniques on Microarray data [2].

The present work highlights the application of unsupervised learning technique; that is crucial in the context of Microarray data. It is the groundwork of large exploratory analysis for realizing the data structure. The algorithms of Adaptive Resonance Theory (ART) family are neural networks, and conceived in order to imitate the manner the human brain acknowledges the patterns [3]. The motivation of using Adaptive Resonance Theory for Microarray data is to handle the dilemma of stability-plasticity. It is related to the learning manner of a neural network that maintains the new information found in new patterns without discarding the information stored before. It is a problem that impacts lot of artificial and biological neural learning systems [4]. ART methods are able to dynamically update and recognize new prototypes; while data structure varies over time.

Journal of Theoretical and Applied Information Technology
15th August 2016. Vol.90. No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Which provides this method an advantage beyond static methods like K-means [5].

The aim of the proposed approach is to present iteratively a new input and update parameters related to the closest cluster. If this input is not closer enough to the existing clusters, a new cluster will be created, by the identification of the new structure. Besides, the algorithm controls continually whether the current prototypes are relevant choices. The plasticity of the algorithm is extremely appropriate for situations in which the data structure is unknown or may vary over time.

The selected dataset concerns lung cancer gene expression data. It has been used for testing performances of many learning techniques, in supervised and unsupervised cases. We have already used MLP network to classify this data [6], using both reduced and raw data versions. Also we are interested to evaluate the efficiency of the classical version of ART2, by comparing it with K-means. The obtained results from implementing the classical ART2 have shown its forcefulness.

2. MICROARRAY TECHNOLOGY

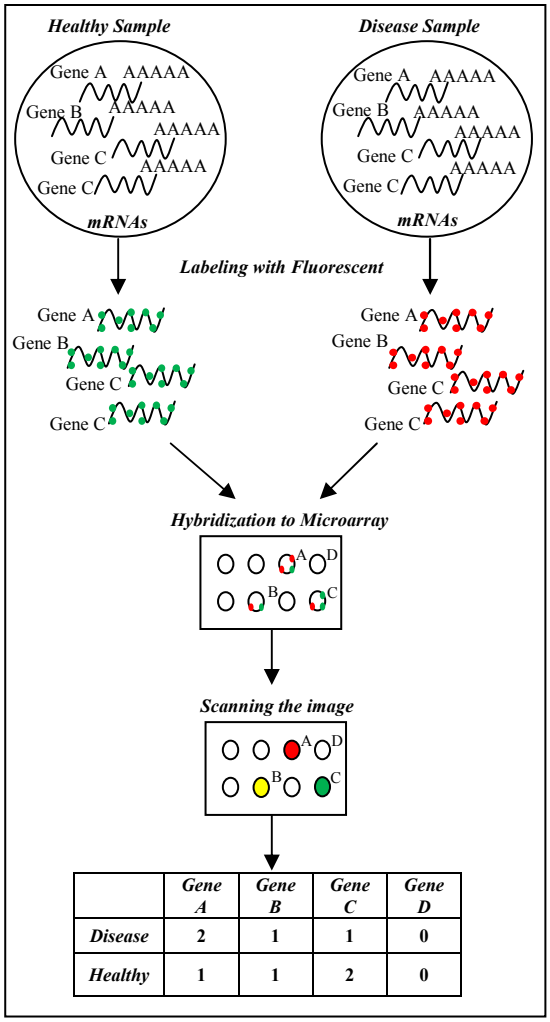
Microarray technology has become one of the essential tools that lot of biologists use to control the gene expression levels in a defined organism [7]. At the beginning, Microarray methods were emerged to study differential gene expression by utilizing complex populations of RNA [8]. This advanced technology is used for profiling the gene expression in cells and tissues. It allows monitoring thousands of genes concurrently.

2.1 Microarray Experiment

Microarray experiments are implemented in order to answer enormous biological questions; where the answers are established by investigating hundreds or thousands of genes.

The Microarray experiment workflow depicted on the Figure 1 contains different stages:

- Obtaining the samples for the labeling step; that includes the isolation of RNA, the labeling by using a reverse procedure of transcription with fluorescent markers (the most commonly used are Cy3 and Cy5). Then the labeled products should be purified.
- Hybridizing the labeled fluorescent DNAs onto Microarrays.



	Gene A	Gene B	Gene C	Gene D
Disease	2	1	1	0
Healthy	1	1	2	0

Figure 1: Microarray Experiment Workflow

A crucial part of each Microarray experiment is the hybridization of the labeled targets by dyes to the probes of surface-immobilized [9]. The hybridization of cDNA strands on the glass supports is relatively well ascertained in molecular biology [10][11][12] and is interesting to the Microarrays quality where both the affinity and specificity of probe–target interaction are defining mainly the Microarray quality [13].

- Scanning the image such that after the hybridization and the application of appropriate washing steps, the Microarrays should be scanned by using Microarray scanners. These scanners are characterized by lasers that are excited at wavelengths especially for Cy3 and Cy5 [14]. At each spot on the Microarray, the present

Journal of Theoretical and Applied Information Technology
15th August 2016. Vol.90. No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

fluorescent dyes are excited by the scanner. Then the dye emits a signal; such that the amount of signal emitted is proportional to the amount of dye at the spot. The final results are represented by the obtained values that were quantitated on the scanner.

2.2 Microarray Data Matrix

The term of gene expression profile describes the expression values for one gene through many experimental conditions or samples [15]. On the other side, there is the term of array profile which determines the expression values for many genes under one sample or condition.

The gene expression profile and the array profile are depicted on the Figure 2. The Microarray matrix format shown on the Figure 2 is composed by two dimensions:

- *M* represents the samples dimension;
- *N* represents the genes dimension.

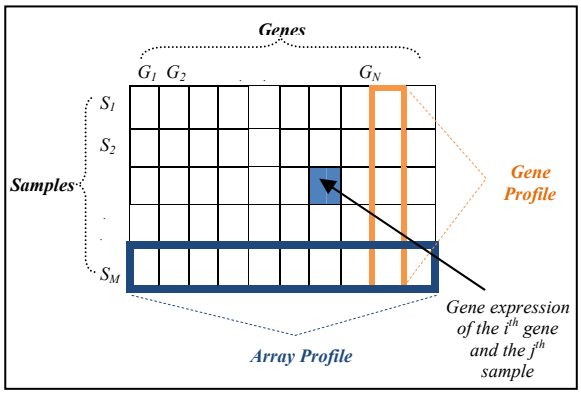


Figure 2: Microarray Data Matrix

3. MICROARRAY DATA ANALYSIS

Microarray data analysis is defined as a complex process. Exploring methods and tools of data analysis methods is the helpful way for investigators.

Regarding the schema of Microarray data analysis process in Figure 3, it shows the main steps followed in order to obtain Microarray Data Matrixes and to explore the generated data. The starting step consists in specifying the scientific aims by formulating clear hypothesis.

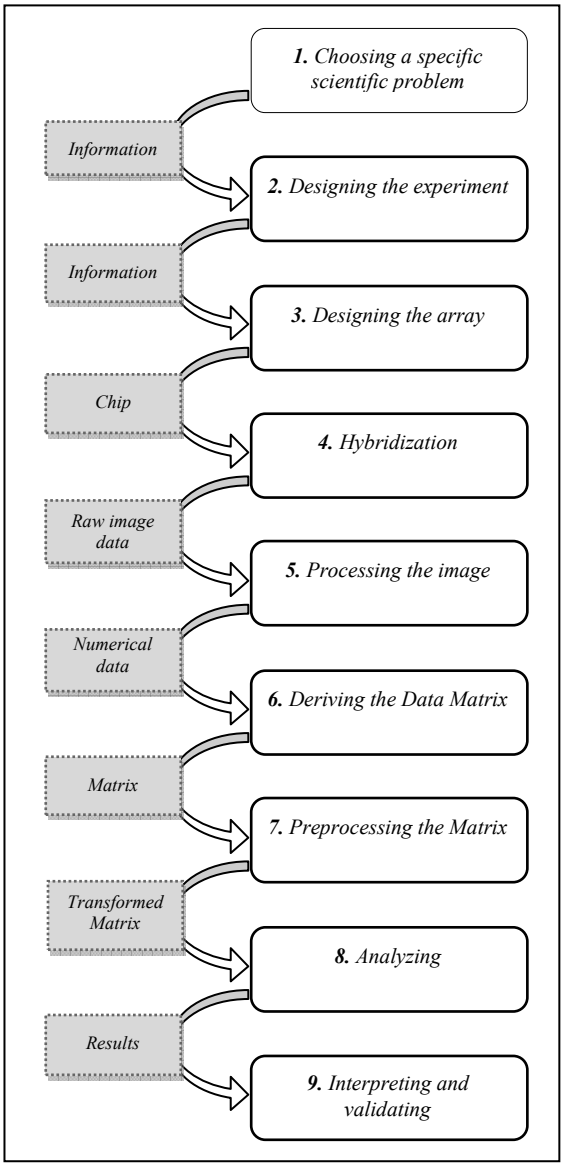


Figure 3: Microarray Data Analysis Process

The step 2 is concerning the choices made by the investigators; such as selecting the factors, defining the experimental conditions, deciding on replications and defining the decision criteria. For the step 3, it is about the decisions on which DNA sequences or genes to place on the chip and the manner of representing each gene. After that, the hybridization experiment produces digital image data that is analyzed on the step 5; where the numerical measurement estimates are derived. The step 6 is depicting the phase of collecting and storing the numerical values in the form of data sets. The main object of the step 7 is to get the significant Microarray data; by implementing the

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

normalization and transformation techniques in order to get rid of the noisy data. On the step 8, it is necessary to explore the generated Microarray data; by applying the classification or clustering techniques, the correlation analysis, regression or approximation. The results obtained on the step 8 are useful for the step 9 to get the visual inspection of results and the biological validation. The last phase of interpreting and validating the results represents a gainful phase; especially for deciding new hypotheses, insights and knowledge.

The present paper is interested to perform tasks included in the step 8, depicted on the Figure 3. In fact, clustering biological samples in Microarray experiments is a big challenge that receives a great attention of researchers in the field. A reliable and precise clustering is essential in the context of diagnosis of tumors. Some of the clustering algorithms that have been applied to gene expression data are:

- K-means [16];
- Fuzzy C-means [17];
- Model based clustering [18];
- Cluster Affinity Search Technique (CAST) [19].

4. PROBLEMATIC

Despite the advances known in Microarray technologies, analysis of the tremendous amounts data generated by this technology's researches remains as a considerable challenge [20]. The major problem faced by most of researchers in Microarray data analysis is known as the dimensionality curse. It is referring to the fact that many Microarray experiments contain a large number of genes than the number of samples. The present work focuses on using unsupervised learning approach; to be able to analyze and interpret the huge amount of Microarray data. The analysis and interpretation of gene expression data are considered as the challenging tasks for researchers. The elaborated studies in medical research have shown that Cancer research represents one of the principal research areas [21]. The ability of predicting diverse tumor types; especially with high accuracy; has great value in order to decrease harmfulness and grant suitable treatments for the patients. This work is concerned on the prognostic of Lung cancer gene expression data. There are researches that have pointed out the clustering techniques on Microarray data such as k-means, CAST, model based clustering. In the proposed approach, the object is to take advantage of ART2 and to develop a method that could determine the clusters automatically.

5. ARTIFICIAL NEURAL NETWORKS AS A POWERFUL CLUSTERING TOOL

Artificial Neural Networks (ANNs) have been implemented in many fields, especially in applications for many biological systems [22]. ANNs represent robust tools; that are recently used as either for clustering or classification of Microarray data [23].

Adaptive Resonance Theory (ART) neural network was induced as a processing theory of human cognitive information [24][25]. It is a large family of neural networks architectures; that allows learning any input pattern in a quick, steady and self-organizing manner.

In most of researches exploring clustering techniques on Microarray data, the typical and classical ART2 architecture wasn't applied enough. Microarray data are represented by a set of samples with numerical values. ART2 is very suitable to deal with gene expression data, especially to find automatically the number of clusters and the prototypes of each one.

5.1 ART Family

The ART family is formed by versions of ART; that have been developed depending to specific requirements:

- *ART1* self-organizes the categories of recognition for arbitrary sequences of binary input patterns [26].
- *ART2* architectures are conceived to process the analog, as well as binary, input patterns [27].
- *ART2-A* is a derivative of ART2 that assimilates the dynamic nature of the original, while ameliorating the computational efficiency by a magnitude order [5].
- *ART3* is based on ART2 by utilizing the neurotransmitter regulation simulation of synaptic activity [28]. It is a rugged mechanism for parallel searching of recognition code of a learned pattern [29].

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

- The *Fuzzy ART* model [30] is developed to generalize ART1 in order to be able to learn categories of stable recognition, in responding to both analog and binary input patterns.
- *ARTMAP* was originally involved in the object of learning mappings between the binary input and the vectors of binary output [31].
- *Fuzzy ARTMAP (FAM)* is one of the typical structures that is frequently used; where the code and the representation of training data are performed by utilizing hyper rectangle shape node [32].

5.2 ART2

The typical ART2 architecture [1] is depicted on the Figure 4.

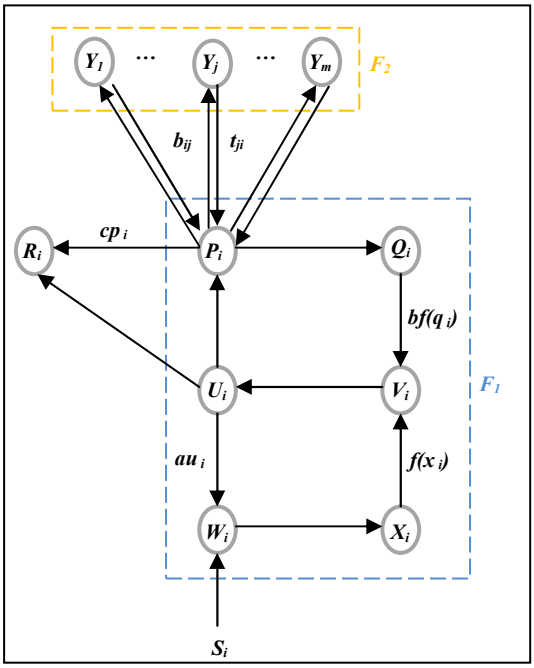


Figure 4: Typical ART2 Architecture

The two ART2 layers are:

- **F_1** Layer which consists of six units types that are *W*, *X*, *U*, *V*, *P* and *Q*. There are *n* units of each type; where *n* is the dimension of an input pattern. The Figure 4 illustrates just one unit of each type.
- **F_2** Layer which contains the units Y_j that compete in the mode of winner-take-all; in order to learn each input pattern.

There is a supplemental unit between the *P* and *Q* units, another one between the *V* and *U* units and the third one is between the *W* and *X* units. For example, for the *W* and *X* units as shows the Figure 5, the supplemental unit existing between them receives signals from all the *W* units. Then, it computes the norm of the vector w, and sends the obtained signal to each of *X* units.

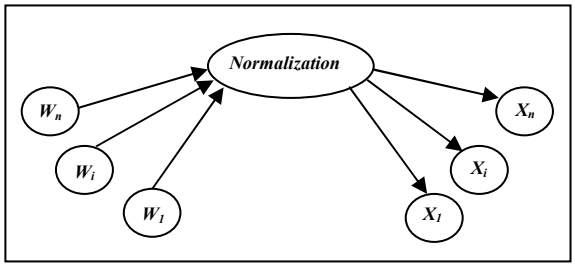


Figure 5: The supplemental unit performing normalization between two types of units

In the F_1 layer, there are connections between each two type units; that indicate the signal transformation when the signal passes from one unit type to the next one.

ART2 is characterized by specific parameters that are:

- ***n***: number of input units
- ***m***: generated number of clusters; which is initialized by zero and updated depending to the presented inputs
- ***a***, ***b***: fixed weights in the F_1 layer
- ***c*** : fixed weight used in testing for reset
- ***d***: activation of winning F_2 unit
- ***e***: small parameter used for preventing the division by zero when the vector norm is zero
- ***θ***: parameter of noise suppression, where the sample value is $\theta = \frac{1}{\sqrt{n}}$
- ***α***: learning rate; such that small value slows the learning and ensures that the weights reach equilibrium
- ***ρ***: vigilance parameter

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Excluding the parameter n, the other ART2 parameters are restricted on choosing the appropriate values [33].

Table 1: Restrictions of ART2 parameters

Parameter	Restrictions
a, b	The sample values are a=10, b=10. But determining either a=0 or b=0 generates the instability in the network. Excepting the value of zero, the network is not espiecially sensitive to the chosen values.
c	The sample value is c=0.1, the small value of c is giving a larger efficient range of vigilance parameter.
d	The sample value is d=0.9, the choice of c and d must satisfy the following inequality: $\frac{cd}{1-d} \le 1$ The choice of the ratio should be close to 1, in order to attain a larger efficient range of the vigilance.
e	The value of e is preventing the normalization to unity from being exact. And the value of zero may be used if there is a skip of the normalization step when the vector is zero
θ	The sample value is: $\theta = \frac{1}{\sqrt{n}}$ In some applications, the sample value could be larger than wished. When the components of the normalized input vector are less than the predefined value, they are set to zero.
α	The smaller value of α is slowing the learning in both fast and slow learning mode. And it is ensuring also that the weights attain equilibrium for slow learning mode.
ρ	With the initial bottom-up weights, the parameter ρ has a significant role on determining the number of clusters that will be formed. The values that are between 0.7 and 1 perform useful role in verifying the clusters number. It the value is less than 0.7, it has the same impact as ρ=0

For ART2 algorithm, the following calculations are required in the object of updating the F1 activations:

$$u_i = \frac{v_i}{e + \|v\|} \quad (1)$$

$$w_i = s_i + a u_i \quad (2)$$

$$p_i = u_i + dt_{Ji} \quad (3)$$

$$x_i = \frac{w_i}{e + \|w\|} \quad (4)$$

$$q_i = \frac{p_i}{e + \|p\|} \quad (5)$$

$$v_i = f(x_i) + bf(q_i) \quad (6)$$

The activation function is determined by:

$$f(x) = \begin{cases} x & if \quad x \ge \theta \\ 0 & if \quad x < \theta \end{cases} \quad (7)$$

The ART2 algorithm is depicted by the following steps:

---

a. *Initializing the network parameters:*
$a, b, \theta, c, d, e, \alpha, \rho$

b. *Implementing the steps from* ***c*** *to* ***n*** *for a determined number of epochs*

c. *For each input vector s, executing the steps from d to* ***m***

d. *Updating the F_1 unit activations:*

$$w_i = s_i \qquad q_i = 0$$

$$u_i = 0 \qquad x_i = \frac{s_i}{e + \|s\|}$$

$$p_i = 0 \qquad v_i = f(x_i)$$

e. *Updating the F_1 unit activations again:*

$$u_i = \frac{v_i}{e + \|v\|} \qquad w_i = s_i + a u_i$$

$$p_i = u_i \qquad x_i = \frac{w_i}{e + \|w\|}$$

$$q_i = \frac{p_i}{e + \|p\|} \qquad v_i = f(x_i) + bf(q_i)$$

f. *Computing the F2 units signals:*

$$y_i = \sum_i b_{ij} p_i$$

g. *Executing the two steps* ***h*** *and* ***i****, while the reset is true*

h. *Finding Y_J of F2 unit that has largest signal; by determining J as $y_{sJ} \ge y_j$ with j=1, …, m*

i. *Verifying for reset:*

$$u_i = \frac{v_i}{e + \|v\|} \qquad p_i = u_i + dt_{Ii}$$

$$r_i = \frac{u_i + cp_i}{e + \|u\| + c\|p\|}$$

If $\|r\| < \rho - e$, *then*

$$y_J = -1$$

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

If $\|r\| \geq \rho - e$, *then*

$$w_i = s_i + au_i$$

$$x_i = \frac{w_i}{e + \|w\|}$$

$$q_i = \frac{p_i}{e + \|p\|}$$

$$v_i = f(x_i) + bf(q_i)$$

j. *Executing the steps from* ***k*** *to* ***m*** *for the specified number of learning iterations.*

k. *Updating the weights for the winning unit J:*

$$t_{Ji} = \alpha d u_i + \{1 + \alpha d(d-1)\} t_{Ji}$$

$$b_{iJ} = \alpha d u_i + \{1 + \alpha d(d-1)\} b_{iJ}$$

l. *Updating the activations of* F_1 *:*

$$u_i = \frac{v_i}{e + \|v\|}$$

$$w_i = s_i + au_i$$

$$p_i = u_i + dt_{Ji}$$

$$x_i = \frac{w_i}{e + \|w\|}$$

$$q_i = \frac{p_i}{e + \|p\|}$$

$$v_i = f(x_i) + bf(q_i)$$

m. *Examining the stop condition for the updates of weight*

n. *Examining the stop condition for the epochs number*

6. EXPERIMENTAL RESULTS

6.1 Microarray Dataset

The chosen dataset interests Lung cancer of the human gene expression [37]. Its issue is to classify two types of lung cancer that are:

- Malignant pleural mesothelioma (MPM)
- Adenocarcinoma (ADCA).

For implementing ART2 algorithm, another dataset was formed from the predefined one. The implemented gene expression dataset contains 62 tissue samples; such that 31 samples represent the type of MPM and 31 samples for the type of ADCA. And each sample is depicted by 12533 genes.

Table 2: The selected Microarray dataset

Genes number	Number of the samples	
	MPM	ADCA
12533	31	31

6.2 Useful parameters

In the present work, the ART2 algorithm has been applied without determining the number of the clusters that will be formed; as a parameter of the algorithm. The object is to use the typical ART2 architecture, such that the clusters could be dynamically determined as a result of the clustering.

The ART 2 parameters have been fixed by choosing carefully the appropriate values; by taking into consideration the restrictions of the parameters choices.

Table 3: Values of ART2 parameters

ART2 parameter	Value
a, b	10
c	0.1
d	0.9
e	2.2204 e-16
θ	$\frac{1}{\sqrt{n}}$
α	0.6
ρ	0.93

The methodology was simulated by utilizing the Microarray dataset of lung cancer. In order to assess the clustering performance, the following measure was calculated:

$$Accuracy = (N_{correct} / N_{total}) * 100 \qquad (8)$$

- *Accuracy* represents the performance measure of the classifiers
- $N_{correct}$ represents the number of instances that were rightly classified
- N_{total} represents the total number of instances

6.3 Simulation on raw and reduced data

Journal of Theoretical and Applied Information Technology
15th August 2016. Vol.90. No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

In order to reduce the lung cancer data, the PCA technique was used.

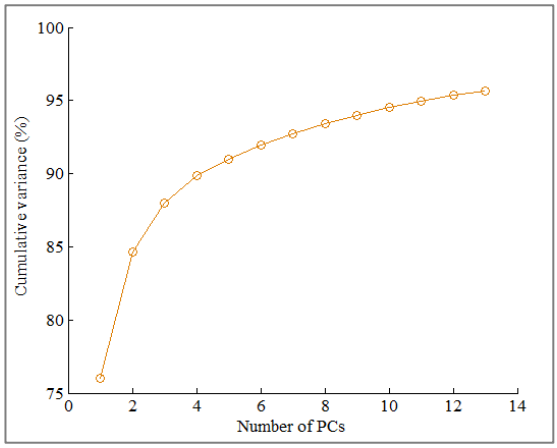


Figure 6: Variation of the cumulative variance with PCs number for Lung cancer Microarray dataset

The Figure 6 shows the number of principle components with the variation of the cumulative variance. The generated variation has indicated that by choosing 13 as a number of principal components, there is more than 95% of the cumulative variance. Thus, thirteen PCs were conserved to perform the genes of the selected Microarray dataset.

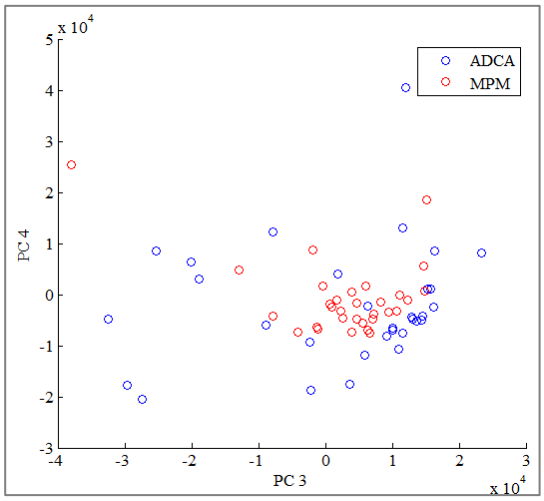


Figure 7: Distribution of the samples in PC3 and PC4 subspaces

To show the difference of the distribution of the samples on the obtained principal components, Figure 7 and Figure 8 were illustrated.

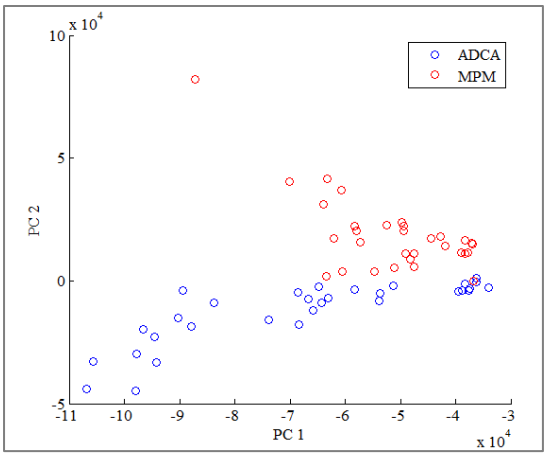


Figure 8: Distribution of the samples in PC1 and PC2 subspaces

On the Figure 7, the two principal components PC 3 and PC 4 were chosen arbitrary. On the other side, the selected PC 1 and PC 2 on the Figure 8 are the most dominated principal components. And it proved that the samples identifying ADCA are not distributed on the same way as those of MPM.

The results produced by the PCA technique using the same dataset were performed for evaluating the proposed methodology; named Dynamical ART2. The accuracy obtained was compared to the accuracies of classical ART2 and K-means. The results are summarized in the Table 4.

Table 4: Clustering accuracies on reduced data

Clustering technique	Lung dataset	
	Number of genes	Accuracy value (%)
Classical ART2	13	93.55
Dynamical ART2	13	98.38
K-means	13	77.42

The proposed methodology was also simulated on the raw data, characterized by 62 of samples and each sample is described by 12533 genes. The samples of the Lung cancer dataset; representing the two classes MPM and ADCA; were presented randomly to the network. In order to show the success of the proposed methodology, the results obtained by implementing the proposed method were compared to those generated by the classical ART2 and K-means. To point the difference, the ART2 parameters have been fixed

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

for the same values in both classical ART2 and dynamical ART2. It is remarkable that ART2 has reached a high accuracy comparing with K-means. A general comparison is summarized in the Figure 9. It shows the obtained accuracy from each method.

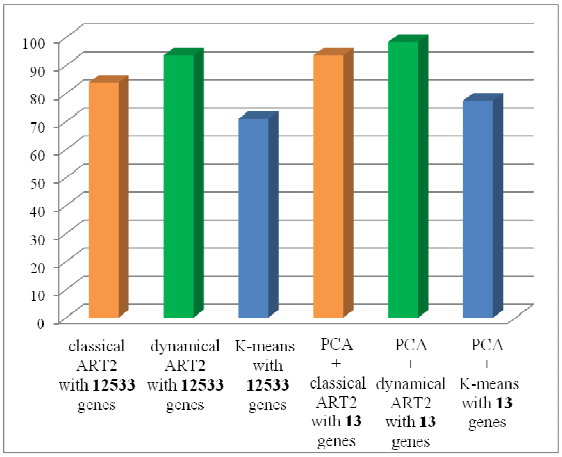


Figure 9: General comparison

It's remarkable that the adopted approach has shown its high power on both the raw and reduced data.

7. CONCLUSION

The main goal of this work was to find exactly the clusters behind the Microarray data (their number and their structure). For this reason, the investigations have been started by exploring the Adaptive Resonance Theory neural network (ART). The variant ART2 has been chosen to process the analog inputs; represented by gene expression data related to lung cancer. The effectiveness of the technique is proved by taking measurement of the clustering accuracy. Dynamical ART2 was able to extract the two clusters MPM and ADCA; representing the two types of lung cancer of the selected Microarray dataset. And it could derive the patterns of the presented samples. The created network has learned the distribution of the two clusters MPM and ADCA, and has generated the decision about the adequate cluster when applied to data samples out of the training set.

REFRENCES:

[1] G. A. Carpenter and S. Grossberg, "ART 2: Self-organization of stable category recognition codes for analog input patterns", *Applied optics*, vol. 26, no. 23, 1987, pp. 4919–4930.

[2] Fadoua Rafii, M'hamed Aït Kbir and Badr Dine Rossi Hassani, "Data Preprocessing and Reducing for Microarray Data Exploration and Analysis", *International Journal of Computer Applications*, Vol. 132, No. 16, 2015, pp.20-26 .

[3] S. Grossberg, "Adaptive Resonance Theory: how a brain learns to consciously attend, learn, and recognize a changing world", *Neural networks: the official journal of the International Neural Network Society*, Vol. 37, Jan. 2013, pp. 1–47.

[4] M. Mermillod, A. Bugaiska, and P. Bonin, "The stability plasticity dilemma: investigating the continuum from catastrophic forgetting to age-limited learning effects", *Frontiers in psychology*, Vol. 4, August 2013, p. 504.

[5] S. J. Chambers, I. H. Jarman, and P. J. Lisboa, "A framework for initialising a dynamic clustering algorithm: ART2-A", *Computational Intelligence and Data Mining (CIDM), 2014 IEEE Symposium*, 2014, pp. 273–280.

[6] Fadoua Rafii, M'hamed Aït Kbir and Badr Dine Rossi Hassani, "MLP Network for Lung Cancer Presence Prediction Based on Microarray Data", *Third World Conference on Complex Systems IEEE, Marrakech, Morocco*, 23 - 25 November 2015.

[7] M. M. Babu, "Introduction to microarray data analysis", *Computational genomics: Theory and application*, 2004, pp. 225–249.

[8] Michael F Ochs and Andrew K Godwin, "Microarrays in Cancer: Research and Applications", *BioTechniques, Vol*34 March 2003, pp. S4-S15.

[9] T. K. Karakach, R. M. Flight, S. E. Douglas, and P. D. Wentzell, "An introduction to DNA microarrays for gene expression analysis", *Chemometrics and Intelligent Laboratory Systems*, Vol. 104, No. 1, Nov. 2010, pp. 28–52.

[10] S. P. Fodor, "Light-directed, spatially addressable parallel", science, Vol. 1990438, No. 767, 1991, p. 251.

[11] U. Maskos, E.M. Southern, "Oligonucleotide hybridizations on glass supports: a novel linker for oligonucleotide synthesis and hybridization properties of oligonucleotides synthesised in

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

situ", *Nucleic Acids Res., Vol. 20*, No. 7, 1992, pp.1679-1684.

[12] K. R. Khrapko, A.A. Lysov YuP, V.V. Khorlyn, V.L. Shick, Florentiev, A.D. Mirzabekov, "An oligonucleotide hybridization approach to DNA sequencing", *FEBS Letters, Vol. 256*, 1989, pp.118-122.

[13] M. Schena, "Microarray Analysis", *Wiley-liss, Hoboken, NJ,* 2003.

[14] A. Natarajan and T. Ravi, "A Survey on Gene feature selection using microarray data for cancer classification", *IJCSC,* Volume 5, 2014, pp. 126-129.

[15] M. A. Branca and N. Goodman, "DNA microarray informatics: Key technological trends and commercial opportunities", *Cambridge Healthtech Institute, CHI Genomic Reports,* 2001.

[16] S. Tavazoie, J.D. Huges, M.J. Campbell, R.J. Cho and G.M. and Church, "Systematic determination of genetic network architecture", *Nature Genetics*, Vol. 22, 1999, pp.281-285.

[17] D. Dembele and P. Kastner, "Fuzzy C-means method for clustering microarray data", *Bioinformatics*, Vol. 19, No. 8, May 2003, pp. 973–980.

[18] C. Fraley and A. E. Raftery, "How many clusters? Which clustering method? Answers via model-based cluster analysis", *the computer journal*, Vol. 41, No. 8, 1998, pp. 578–588.

[19] A. Ben-Dor, R. Shamir, and Z. Yakhini, "Clustering gene expression patterns", *Journal of Computational Biology*, Vol. 6, 1999, pp.281-297.

[20] N. E. Olson, "The microarray data analysis process: from raw data to biological significance", *NeuroRx*, Vol. 3, No. 3, 2006, pp. 373–383.

[21] B. R. Jeetha and M. Malathi, "Diagnosis Of Ovarian Cancer Using Artificial Neural Network", *International Journal of Computer Trends and Technology (IJCTT),* Volume 4 Issue10, Oct 2013, pp. 3602 – 3606.

[22] J. S. Almeida, "Predictive non-linear modeling of complex data by artificial neural networks", *Current Opinion in Biotechnology*, Vol. 13, No. 1, Feb. 2002, pp. 72–76.

[23] S. Agrawal and J. Agrawal, "Neural Network Techniques for Cancer Prediction: A Survey," *Procedia Computer Science*, Vol. 60, 2015, pp. 769–774.

[24] S. Grossberg, "Adaptive pattern classification and universal rccoding, II: Feedback, expectation, ollaction, and illusions", *Biological Cybernetics*, Vol. 23, 1976, pp. 187-202.

[25] S. Grossberg, "How does a brain build a cognitive code?", *Psychological Review*, Vol. 1, 1980, pp. 1-51.

[26] G.A. Carpenter & S. Grossberg, "A massively parallel architecture for a self-organizing neural pattern recognition machine", *Computer Vision Graphic, and Image Processing*, Vol. 37, 1987, pp.54-115.

[27] G. A. Carpenter and S. Grossberg, "ART 2: Self-organization of stable category recognition codes for analog input patterns", *Applied optics*, Vol. 26, No. 23, 1987, pp. 4919–4930.

[28] G.A. Carpenter, S. Grossberg & D.B. Rosen, "ART 2-A: An adaptive resonance algorithm for rapid category learning and recognition", *Neural Networks*, 1991, Vol. 4, pp. 493–504.

[29] G. A. Carpenter and S. Grossberg, "Hierarchical pattern recognition system with variable selection weights", *Google Patents*, 1994.

[30] G.A. Carpenter, S. Grossberg, & D.H. Reyqoids, "ARTMAP: Supervised real-time learning and classification of nonstationary data by a self-organizing neural network", *Neural Networks*, Vol. 4, 1991, pp.565-588.

[31] G. A. Carpenter, S. Grossberg, and D. B. Rosen, "Fuzzy ART: Fast stable learning and categorization of analog patterns by an adaptive resonance system", *Neural networks*, Vol. 4, No. 6, 1991, pp. 759–771.

[32] C. Liu, G. F. Wang, and Z. M. Li, "Incremental learning for online tool condition monitoring using Ellipsoid ARTMAP network model", *Applied Soft Computing*, Vol. 35, Oct. 2015, pp. 186–198.

[33] Laurene Fausett, "Fundamentals of Neural Networks", *Prentice Hall*, 1994.