Artificial Immunity and Features Reduction for effective Breast Cancer Diagnosis and Prognosis

Mafaz Mohsin Al-Anezi [1]*, Marwah Jasim Mohammed [2*], Dhufr Sami Hammadi [2*]
[a] PhD of Computer Sciences, [b] Computer Sciences
*** Computer Sciences Department, College of Computer Sciences and Mathematics, Mosul University, Mosul, Iraq**

Abstract

The diagnosis and prognosis of breast cancer is an important, real-world medical problem. As an intrusion detection problem is one of the applications of artificial immune system, in this paper proposes a novel scheme that uses a robust immune system formed from clonal selection theory and principal component analysis for breast cancer diagnosis and Prognosis. Like the job done by Antigen Presenting Cells APCs in natural immune system, this work use PCA as an aided tool for immune cells in the selection for the most important features that can detect the cancer and forward them for the immune system in training phase which generates an artificial lymphocytes ALCs and save them as immune memory. It is important to note that the training phase was done on 20% of the dataset, whereas the testing phase was done on the remaining 80% of the data set which are considered as unknown cases for the ALCs. The study proved that the best results obtained when the PCA select minimum reasonable number of features, while in the training phase the diagnostic accuracy is 0.99 and the prognostic accuracy is 0.9, and the memories ALCs achieved in the testing phase a diagnostic accuracy 0.97 and prognostic accuracy 0.88.

Keywords: *Artificial Immune System (AIS); Clonal Selection Algorithm (CLONALG); Principal Component Analysis (PCA); Wisconsin Diagnosis Breast Cancer (WDBC); Wisconsin Prognosis Breast Cancer (WPBC).*

1. Introduction

The second leading cause of death among women is breast cancer, as it comes directly after lung cancer. Most clinicians are prone to misjudge patients' survival furthermore heightening the need for accurate prognostication tools. In addition to that, the planning of public health and cancer treatment can be vastly improved by learning the pattern of survival and the related prognostic factors. According to the American Cancer Society, breast cancer is the second leading cause of cancer deaths among women today. In developing countries where prognosis is much poorer, many die from this disease although breast cancer could be treated with early diagnosis and treatment [1][13].

Although there was a great deal of public education and scientific research, Breast cancer considered the most common invasive cancer in women, with more than one million cases and nearly 600,000 deaths occurring worldwide annually [13]. Early diagnosis helps to save thousands of disease victims. This shows that a good prognostication tool could significantly influence survival prediction and death count.

Data mining approaches in medical domains is increasing rapidly due to the improvement effectiveness of these approaches to classification and prediction systems, especially in helping medical practitioners in their decision making. In addition to its importance in finding ways to improve patient outcomes, reduce the cost of medicine, and help in enhancing clinical studies [13].

The natural immune system NIS constitutes a weapon against intruders in a given body, for this goal several cells contribute to eliminate this intruder named antigen, these cells participate for a 'biologic immune response'. Artificial Immune System AIS simulate the most important functions of the natural immune system for pattern recognition. The main factors entering in the artificial immune system are antigens, antibodies, B and T memory cells.

The term peptide in the NIS refers to a short chain of amino acids, usually obtained by the fragmentation of an antigen, and presented to other cells of the immune system by antigen presenting cells (APC). Antigen presentation refers to processing a suspicious foreign particle. Such a particle is broken up into peptides, and then such peptides are held on the surface of APC, where T cells can recognize them. Several types of cells may serve as APC, including macrophages, dendritic cells, and B cells [5].

From APC behavior the inspiration come to use PCA to extract the most significant features and introduce them to Artificial LymphoCytes ALCs. Where PCA is a useful statistical technique that has found application in fields such as face recognition and image compression, and is a common technique for finding patterns in data of high dimension. PCA is used for dimensionality reduction. The goal of PCA is to reduce the dimensionality of the data while retaining as much as possible of the variation present in the original dataset.

2. Problem Background and Previous Works

Presently many studies have been done by researchers in predicting survival among cancer patients particularly from the statistical field. Nevertheless the statistical approaches face many challenges in handling the nature of the survival analysis datasets which often are censored data, and the difficulties in managing the complex, non-linear relationships between the prognostic factors and the patient's tumor progression. Also, many have argued that the statistical approach omits the need

in prediction of the patients prognosis since it does not take into account that all patients are individual and unique cases. This leads to the alternative technique that is the artificial intelligence method which is still largely uninvestigated. Although several reported works come from the application of ANN, genetic algorithm, decision trees, and Fuzzy classifier [1].

Anagnostopoulos, et al. [2] presented two neural network architectures are proposed for the breast cancer detection/prognosis problems. The first is a probabilistic classifier, which can detect malignancy while the second architecture consist of a probabilistic neural network that employs a generalised regression algorithm [2].

Balakrishnan, et al. [3] propose a feature selection embedded Hybrid Prediction model that combines two different functionalities of data mining; the clustering and the classification. The F-score feature selection method and k-means clustering selects the optimal feature subsets of the medical datasets that enhances the performance of the Support Vector Machine classifier.

Gupta, et al. [6] presented an overview of the current research being carried out using the data mining techniques to enhance the breast cancer diagnosis and prognosis.

Karabatak and Ince [7] presented an automatic diagnosis system for detecting breast cancer based on association rules (AR) used for reducing the dimension of breast cancer database and neural network (NN) is used for intelligent classification.

Khelil and Benyettou [8] developed four versions of Artificial Immune Recognition System (AIRS) and presented results for Cancer diagnostic with some critics and remarks of these methods.

Ludwig and Roos [10] investigated the prognosis of breast cancer using a machine learning approach, in particular genetic programming, where it is a method takes a digitized image of a patient and automatically generates the prediction of the time to recur as well as the disease-free survival time.

3. Wisconsin Breast Cancer Dataset

The Wisconsin Breast Cancer datasets from the UCI Machine Learning Repository is used, to distinguish malignant (cancerous) from benign (non-cancerous) samples. They have been collected by Dr. William H. Wolberg (1989–1991) at the University of Wisconsin–Madison Hospitals. A brief description of these datasets is presented in table 1. Each dataset consists of some classification patterns or instances with a set of numerical features or attributes.

3.1 Wisconsin Diagnosis Breast Cancer (WDBC)

The details of the attributes found in WDBC dataset: ID number, Diagnosis (M = malignant, B = benign) and ten real-valued features are computed for each cell nucleus: Radius, Texture, Perimeter, Area, Smoothness, Compactness, Concavity, Concave points, Symmetry and Fractal dimension. These features are computed from a digitized image of a fine needle aspirate (FNA) of a breast mass. They describe characteristics of the cell nuclei present in the image [13]. The mean, standard error, and "worst" or largest (mean of the three largest values) of these features were computed for each image, resulting in 30 features. For instance, field 3 is Mean Radius, field 13 is Radius SE, field 23 is Worst Radius.

3.2 Wisconsin Prognosis Breast Cancer (WPBC)

The details of the attributes found in WPBC dataset: ID number, Outcome (R = recur, N = non-recur), Time (R => recurrence time, N => disease-free time), from 3 to 33 ten real-valued features are computed for each cell nucleus: Radius, Texture, Perimeter, Area, Smoothness, Compactness, Concavity, Concave points, Symmetry and Fractal dimension. The thirty four is Tumor size and the thirty five is the Lymph node status. It's known from the previous lines that the diagnosis and prognosis has the same features yet the prognosis has two additional features [13]. The mean, standard error, and "worst" or largest (mean of the three largest values) of these features were computed for each image, resulting in 30 features. For instance, field 4 is Mean Radius, field 14 is Radius SE, field 24 is Worst Radius. Feature 34 Tumor size - diameter of the excised tumor in centimeters, and feature 35 Lymph node status - number of positive axillary lymph nodes observed at time of surgery.

Table 1: Description Of The Breast Cancer Datasets.

Dataset	No. of Attributes	No. of Instances	Class1	Class2
Wisconsin Diagnosis Breast Cancer(WDBC)	32	569	357 Benign	212 Malignant
Wisconsin Prognosis Breast Cancer(WPBC)	34	198	151 NonRecur	47 Recur

4. Breast Cancer Diagnosis Model

Feature reduction process can be viewed as a preprocessing step which removes distracting variance from a dataset, so that classifiers can perform better. In our proposed algorithm PCA transform used for dimensionality reduction which is commonly used step, especially when dealing with high dimensional space of features. PCA-based approaches improve system performances and a trained artificial immune system to diagnosis or prognosis the cancer. Following are the steps used in our algorithm:

4.1 Data Preprocessing

Normalization is used for data preprocessing, where the attribute data are scaled so as to fall within a small specified range such as -1.0 to 1.0 or 0.0 to 1.0. Normalizing the input values for each attribute measured in the training samples will help speed up the learning phase.

4.2 Feature Reduction by PCA

Feature reduction applies a mapping of the multidimensional space into a space of lower dimensions. Feature extraction includes features construction, space dimensionality reduction, sparse representations, and feature selection all these techniques are commonly used as preprocessing to machine learning and statistics tasks of prediction, including pattern recognition. Although such problems have been tackled by researchers for many years, there has been recently a renev interest in feature extraction. The feature space having redu features truly Contributes to classification that c... preprocessing costs and minimizes the effects of the 'peaking phenomenon' in classification. Thereby improving the overall performance of classifier based intrusion detection systems. The commonly used dimensionality reduction methods include supervised approaches such as linear discriminant analysis (LDA), unsupervised ones such as principal component analysis (PCA), and additional spectral and manifold learning methods [9].

PCA is a linear transformation with linear orthonormal basis vectors, it can be expressed by a translation and rotation. It convert a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables called principal components. The number of principal components is less than or equal to the number of original variables [12][13]. If we consider the two dimensional case then Figure 1 illustrates the basic principle of this transformation [12].

Figure (a) presents each X_{ith} sample, denoted the initial data set which is transformed into another representation (Figure (b)), denoted Y_{ith}. The main portion of the variance is stored in the first variable Y_1 . This means that if we ignore the second variable Y_2, as in figure (c), the main variance of the data is kept. Therefore, representing an initial data set with a new more compact space keeping much of the variance of the data in the new compact representation offers many facilities to interpret the data in a new reduced space. This example illustrates a reduction from two dimensions into one dimension. However, in reality the reduction might be performed over hundreds or thousands of variables into only 2 or 3 variables.

Principal components are particular linear combinations of the p random variables X1, X2, …, Xp, with three important properties: (1) the principal components are uncorrelated, (2) the first principal component has the highest variance, the second principal component has the second highest variance, and so on, and (3) the total variation in all the principal components combined equal to the total variation in the original variables X1, X2, …, Xp. The new variables with such properties are easily obtained from eigenanalysis of the covariance matrix or the correlation matrix of X1, X2, …, Xp [14].

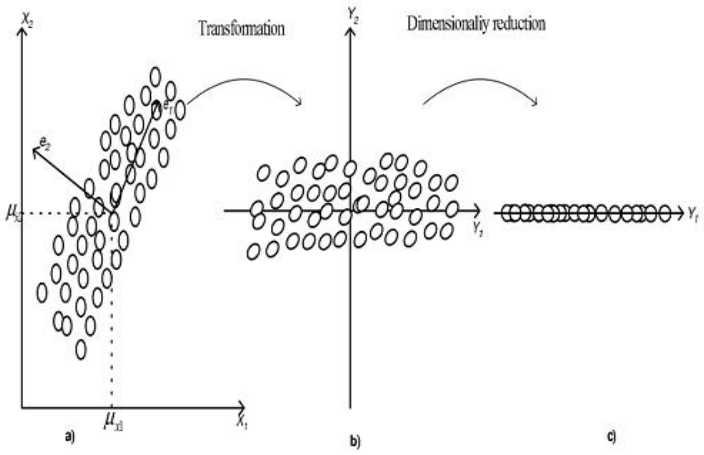

Figure 1: PCA basic principle transformation [12]

4.3 Classification Using Clonal Selection Algorithm

The clonal selection principle, or theory of (CLONA), is the algorithm used by the immune system to describe the basic features of an immune response to an antigenic stimulus. Clonal selection establishes the idea that only cells that recognize the antigens will proliferate where the rest will not. The most triggered cells selected as memory cells for future pathogens attacks and the rest mature into antibody secreting cells called plasma cells [11].

Clonal selection in AIS is the selection of a set of artificial lymphocytes ALCs with the highest calculated affinity with a non-self pattern. The selected ALCs are then cloned andmutated in an attempt to have a higher binding affinity with the presented nonself pattern. The mutated clones compete with the existing set of ALCs, based on the calculated affinity between the mutated clones and the non-self pattern, for survival to be exposed to the next non-self pattern.

The selection of a lymphocyte by a detected antigen for clonal proliferation, inspired the modeling of CLONALG. De Castro and Von Zuben presented CLONALG as an algorithm that performs machine-learning and pattern recognition tasks [4]. The affinity between an ALC and a non-self pattern is measured as the Hamming distance between the ALC and the non-self pattern. The Hamming distance gives an indication of the similarity between two patterns, i.e. a lower Hamming distance between an ALC and a non-self pattern implies a stronger affinity [4].

All patterns in the training set are seen as non-self patterns. The algorithm below summarizes the general CLONALG for pattern recognition tasks. When applied to pattern matching, a set of antigens, G, to be matched. The task of CLONALG is to then produce a set of memory ALC M , that match the members in G.

input: G = set of antigens to be recognized, *n* the number of worst elements to select for removal
output: M = set of memory ALCs capable of classifying unseen patterns
begin

Create an initial random set of ALCs, *A*
forall *antigens in G* do
Determine the affinity with each ALC in *A*.
Generate clones of a subset of the ALC in *A* with the highest affinity.
The number of clones for an ALC is proportional to its affinity.
Mutate attributes of these clones to the set *A*, and place a copy of the highest affinity ALC in *A* into the memory set, *M*.
Replace the *n* lowest affinity ALCs in *A* with new randomly generated ALCs.
end
end

Figure 2 depicts the functional block diagram of the Breast Cancer Diagnosis and Prognosis Model. It consists of two phases namely: training and testing phases. The training phase includes four steps: acquisition, preprocess, feature selection, Abs generating by CLONALG, and store the best Abs in the immune memory, whereas the testing phase includes the same first three steps in the training phase in addition to the classification step.

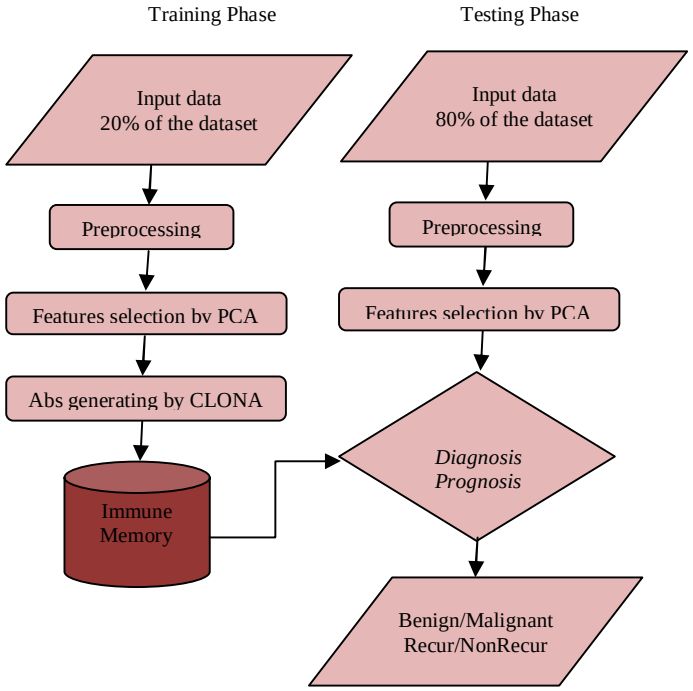

Figure 2. Immune Breast Cancer Diagnosis and Prognosis Model.

5. Experimental setup and Results

We used C# programming for implementation. To evaluate the proposed model several measurements were used like (Detection Rate D.R., False Alarm Rate F.A.R., Accuracy) all of them depend on the calculation of confusion matrix (TP, TN,FP,FN). The first two features in the dataset WDBC are the ID number and Diagnosis (M or B), and for WPBC are the ID number and Outcome (R , N), so these two features not encounter in the experiments. These experiments were performed as follow:

A. Experiments on Diagnosis Training data which is 20% from WDBC (89 Benign, 53 Malignant), show figure 3. These Experiments depend on the number of features selected by PCA, as follow:
- **A.1.** Table 2, show that PCA select 10 from 30 features, the average of D.R. and F.A.R. for five training are 0.99 and 0.1 continuity.
- **A.2.** Table 3, show that PCA select 15 from 30 features, the average of D.R. and F.A.R. for five training are 0.96 and 0.6 continuity.
- **A.3.** Table 4, show that PCA select 21 from 30 features, the average of D.R. and F.A.R. for five training are 0.94 and 0.55 continuity.
- **A.4.** Table 5, show that all 30 features were used, the average of D.R. and F.A.R. for five training are 0.78 and 0.44 continuity.

B. Experiments on Diagnosis Testing data which is 80% from WDBC (268 Benign, 159 Malignant), show figure 4. These Experiments depend on the number of features selected by PCA and the number of generated Benign ALCs and Malignant ALCs from experiments A, as follow:
- **B.1.** Table 6, show depend on the number of BenALCs and MalALCs from exp. A.1., the average of D.R. and F.A.R. for five testing are 0.96 and 0.08 continuity.
- **B.2.** Table 7, show depend on the number of BenALCs and MalALCs from exp. A.2., the average of D.R. and F.A.R. for five testing are 0.93 and 0.08 continuity.
- **B.3.** Table 8, show depend on the number of BenALCs and MalALCs from exp. A.3., the average of D.R. and F.A.R. for five testing are 0.93 and 0.11 continuity.
- **B.4.** Table 9, show depend on the number of BenALCs and MalALCs from exp. A.4., the average of D.R. and F.A.R. for five testing are 0.68 and 0.56 continuity.

C. Experiments on Prognosis Training data which is 20% from WPBC (45 NonRecur, 14 Recur), show figure 5. These Experiments depend on the number of features selected by PCA, as follow:
- **C.1.** Table 10, show that PCA select 10 from 32 features, the average of D.R. and F.A.R. for five training are 0.93 and 0.2 continuity.
- **C.2.** Table 11, show that PCA select 15 from 32 features, the average of D.R. and F.A.R. for five training are 0.87 and 0.33 continuity.
- **C.3.** Table 12, show that PCA select 21 from 32 features, the average of D.R. and F.A.R. for five training are 0.83 and 0.57 continuity.
- **C.4.** Table 13, show that all 32 features were used, the average of D.R. and F.A.R. for five training are 0.52 and 0.94 continuity.

D. Experiments on Prognosis Testing data which is 80% from WPBC (106 NonRecur, 33 Recur), show figure 6. These Experiments depend on the number of features selected by PCA and the number of generated Benign ALCs and Malignant ALCs from experiments C, as follow:

D.1. Table 14, show depend on the number of BenALCs and MalALCs from exp. C.1., the average of D.R. and F.A.R. for five testing are 0.94 and 0.26 continuity.

D.2. Table 15, show depend on the number of BenALCs and MalALCs from exp. C.2., the average of D.R. and F.A.R. for five testing are 0.83 and 0.55 continuity.

D.3. Table 16, show depend on the number of BenALCs and MalALCs from exp. C.3., the average of D.R. and F.A.R. for five testing are 0.59 and 0.81 continuity.

D.4. Table 17, show depend on the number of BenALCs and MalALCs from exp. C.4., the average of D.R. and F.A.R. for five testing are 0.16 and 0.99 continuity.

Table 2: Results of Experiments on Diagnosis Training data (89 Benign, 53 Malignant), No of selected features 10.

	Ben ALCs	Mal ALCs	TP	TN	FP	FN	D.R.	F.A.R.	ACY
1	5597	80	88	50	1	3	0.97	0.02	0.97
2	11013	159	88	52	1	1	0.99	0.02	0.99
3	6253	114	88	51	1	2	0.98	0.02	0.98
4	8821	353	89	53	0	0	1	0	1
5	11269	304	89	52	0	1	0.99	0	0.99

Table 3: Results of Experiments on Diagnosis Training data (89 Benign, 53 Malignant), No of selected features 15.

	Ben ALCs	Mal ALCs	TP	TN	FP	FN	D.R.	F.A.R.	ACY
1	4950	26	86	48	3	5	0.95	0.06	0.94
2	4828	25	87	49	2	4	0.96	0.04	0.96
3	5340	12	81	51	8	2	0.98	0.14	0.93
4	4943	13	85	49	04	4	0.96	0.08	0.94
5	5692	15	88	49	1	4	0.96	0.02	0.96

Table 4: Results of Experiments on Diagnosis Training data (89 Benign, 53 malignant), No of selected features 21.

	Ben ALCs	Mal ALCs	TP	TN	FP	FN	D.R.	F.A.R.	ACY
1	107	10	86	49	3	4	0.96	0.06	0.95
2	128	11	77	44	12	9	0.9	0.21	0.85
3	57	10	86	50	3	3	0.97	0.06	0.96
4	534	10	84	48	5	5	0.94	0.09	0.93
5	241	11	82	48	7	5	0.94	0.13	0.92

Table 5: Results of Experiments on Diagnosis Training data (89 Benign, 53 Malignant), No of selected features 30.

	Ben ALCs	Mal ALCs	TP	TN	FP	FN	D.R.	F.A.R.	ACY
1.	20	10	60	36	29	17	0.78	0.45	0.68
2.	20	10	65	34	24	19	0.77	0.41	0.7
3.	41	10	57	34	32	19	0.75	0.48	0.64
4.	40	10	60	38	29	15	0.8	0.43	0.69
5.	26	10	63	35	26	18	0.78	0.43	0.69

Table 6: Results of Experiments on Diagnosis Testing data (268 Benign, 159 Malignant), No of selected features 10.

	Ben ALCs	Mal ALCs	TP	TN	FP	FN	D.R.	F.A.R.	ACY
1.	5597	80	267	147	1	12	0.96	0.01	0.97
2.	11013	159	266	133	2	26	0.91	0.01	0.93
3.	6253	114	266	152	2	7	0.97	0.01	0.98
4.	8821	353	267	153	1	6	0.98	0.01	0.98
5.	11269	304	268	151	0	8	0.97	0	0.98

Table 7: Results of Experiments on Diagnosis Testing data (268 Benign, 159 Malignant), No of selected features 15.

	Ben ALCs	Mal ALCs	TP	TN	FP	FN	D.R.	F.A.R.	ACY
1.	4950	26	261	123	7	36	0.88	0.05	0.9
2.	4828	25	256	137	12	22	0.92	0.08	0.92
3.	5340	12	245	151	23	8	0.97	0.13	0.93
4.	4943	13	255	145	13	14	0.94	0.08	0.94
5.	5692	15	262	137	6	22	0.92	0.04	0.93

Table 8: Results of Experiments on Diagnosis Testing data (268 Benign, 159 Malignant), No of selected features 21.

	Ben ALCs	Mal ALCs	TP	TN	FP	FN	D.R.	F.A.R.	ACY
1.	107	10	254	135	14	24	0.91	0.09	0.91
2.	128	11	243	124	25	35	0.87	0.17	0.86
3.	57	10	252	148	16	11	0.96	0.1	0.94
4.	534	10	254	148	14	11	0.96	0.09	0.94
5.	241	11	249	140	19	19	0.93	0.12	0.91

Table 9: Results of Experiments on Diagnosis Testing data (268 Benign, 159 Malignant), No of selected features 30.

	Ben ALCs	Mal ALCs	TP	TN	FP	FN	D.R.	F.A.R.	ACY
1.	20	10	161	88	107	71	0.69	0.55	0.58
2.	20	10	184	50	84	109	0.63	0.63	0.55
3.	41	10	167	87	101	72	0.7	0.54	0.59
4.	40	10	174	88	94	71	0.71	0.52	0.61
5.	26	10	184	71	84	88	0.68	0.54	0.6

Table 10: Results of Experiments on Prognosis Training data (45 NonRecur, 14 Recur), No of selected features 10.

	Ben ALCs	Mal ALCs	TP	TN	FP	FN	D.R.	F.A.R.	ACY
1.	51	3	44	10	1	4	0.92	0.09	0.92
2.	80	3	42	11	3	3	0.93	0.21	0.9
3.	200	4	41	9	4	5	0.89	0.31	0.85
4.	85	6	42	12	3	2	0.95	0.2	0.92
5.	348	3	42	12	3	2	0.95	0.2	0.92

Table 11: Results of Experiments on Prognosis Training data (45 NonRecur, 14 Recur), No of selected features 15.

	Ben ALCs	Mal ALCs	TP	TN	FP	FN	D.R.	F.A.R.	ACY
1.	9	2	41	9	4	5	0.89	0.31	0.85
2.	9	2	40	8	5	6	0.87	0.38	0.81
3.	9	2	41	10	4	4	0.91	0.29	0.86
4.	9	2	43	9	2	5	0.9	0.18	0.88
5.	10	2	41	4	4	10	0.8	0.5	0.76

Table 12: Results of Experiments on Prognosis Training data (45 NonRecur, 14 Recur), No of selected features 21.

	Ben ALCs	Mal ALCs	TP	TN	FP	FN	D.R.	F.A.R.	ACY
1.	12	2	36	7	9	7	0.84	0.56	0.73
2.	9	2	36	6	9	8	0.82	0.6	0.71
3.	9	2	37	4	8	10	0.79	0.67	0.69
4.	10	2	37	8	8	6	0.86	0.5	0.76
5.	9	2	36	8	9	6	0.86	0.53	0.75

Table 13: Results of Experiments on Prognosis Training data (45 NonRecur, 14 Recur), No of selected features 32.

	Ben ALCs	Mal ALCs	TP	TN	FP	FN	D.R.	F.A.R.	ACY
1.	9	2	13	2	32	12	0.52	0.94	0.25
2.	9	2	12	2	33	12	0.5	0.94	0.24
3.	9	2	12	2	33	12	0.5	0.94	0.24
4.	9	2	14	2	31	12	0.54	0.94	0.27
5.	9	2	15	2	30	12	0.56	0.94	0.29

Table 14: Results of Experiments on Prognosis Testing data (106 NonRecur, 33 Recur), No of selected features 10.

	Ben ALCs	Mal ALCs	TP	TN	FP	FN	D.R.	F.A.R.	ACY
1.	51	3	100	24	6	9	0.92	0.2	0.89
2.	80	3	93	28	13	5	0.95	0.32	0.87
3.	200	4	97	28	9	5	0.95	0.24	0.9
4.	85	6	90	28	16	5	0.95	0.36	0.85
5.	348	3	100	26	6	7	0.93	0.19	0.91

Table 15: Results of Experiments on Prognosis Testing data (106 NonRecur, 33 Recur), No of selected features 15.

	Ben ALCs	Mal ALCs	TP	TN	FP	FN	D.R.	F.A.R.	ACY
1.	9	2	88	19	18	14	0.86	0.49	0.77
2.	9	2	84	11	22	22	0.79	0.67	0.68
3.	9	2	81	19	25	14	0.85	0.57	0.72
4.	9	2	88	18	18	15	0.85	0.5	0.76
5.	10	2	92	11	14	22	0.81	0.56	0.74

Table 16: Results of Experiments on Prognosis Testing data (106 NonRecur, 33 Recur), No of selected features 21.

	Ben ALCs	Mal ALCs	TP	TN	FP	FN	D.R.	F.A.R.	ACY
1.	12	2	64	13	42	20	0.49	0.76	0.55
2.	9	2	56	14	50	19	0.67	0.78	0.5
3.	9	2	79	9	27	24	0.57	0.75	0.63
4.	10	2	76	5	30	28	0.5	0.86	0.58
5.	9	2	68	5	38	28	0.56	0.88	0.53

Table 17: Results of Experiments on Prognosis Testing data (106 NonRecur, 33 Recur), No of selected features 32.

	Ben ALCs	Mal ALCs	TP	TN	FP	FN	D.R.	F.A.R.	ACY
1.	9	2	8	1	98	32	0.2	0.99	0.06
2.	9	2	8	0	98	33	0.2	1	0.06
3.	9	2	5	0	101	33	0.13	1	0.04
4.	9	2	4	0	102	33	0.11	1	0.03
5.	9	2	7	0	99	33	0.18	1	0.05

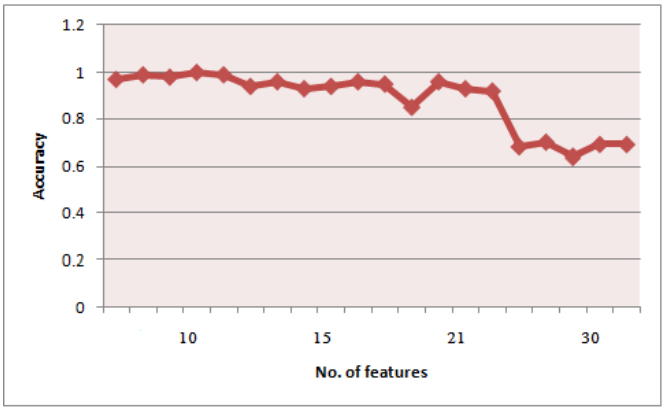

Figure 3: Experiments A.

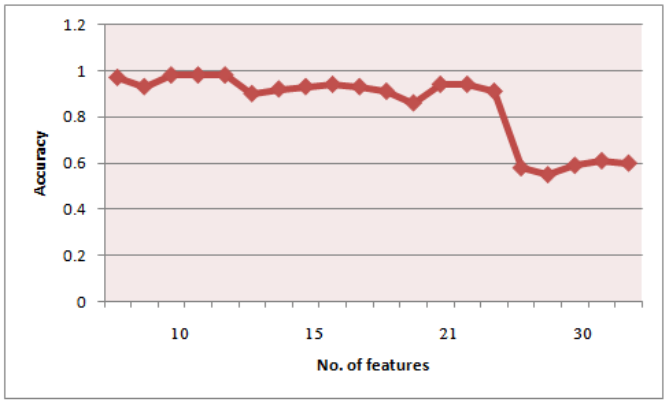

Figure 4: Experiments B.

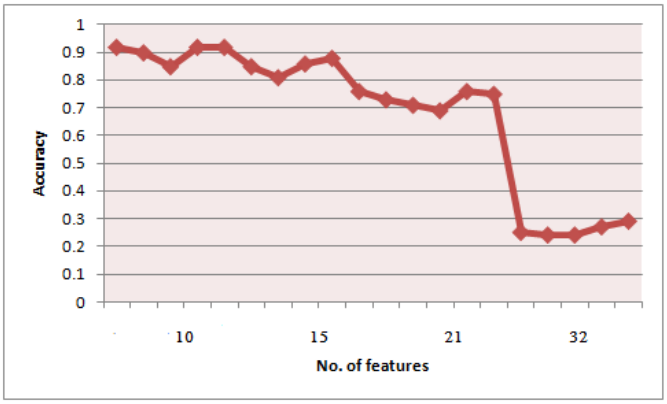

Figure 5: Experiments C.

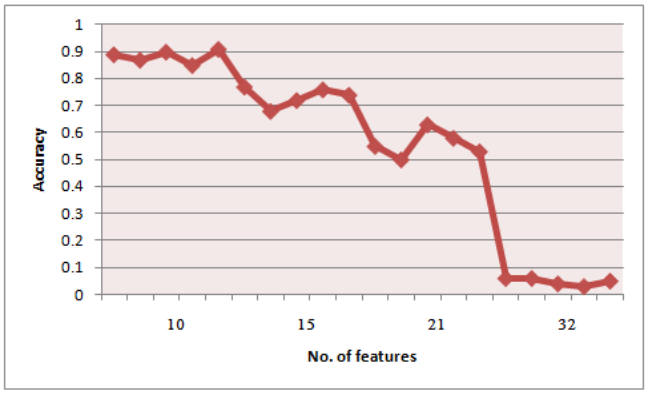

Figure 6: Experiments D.

6. Conclusions

This paper presented Breast Cancer diagnostic and prognostic results for an immerge between immuno-computing and features reduction. Where an immuno-computing is one of the newest directions in bio-inspired machine learning and has a very fruitful successes in different area. The clonal selection theory is one of the first applied theories in AIS, and in this paper supported with features reduction technique PCA as a first step before the start of immune defense.

The presented results are very good but the false alarm it must be improved using optimization algorithms. As future work it must be make more importance to the parameters values and propose a new method to search the best values of these ones in order to across the performance of these hybrid collection of AIS and features reduction techniques.

References

[1] Ali A., Shamsuddin S. M., Ralescu A. L., and Visa S., "*Fuzzy Classifier for Classification of Medical Data*", 2011 11th International Conference on Hybrid Intelligent Systems (HIS), 2011 IEEE, pp. 173-178.

[2] Anagnostopoulos I., Anagnostopoulos C., Rouskas A., Kormentzas G., and Vergados D., "*The Wisconsin Breast Cancer Problem: Diagnosis and DFS time prognosis using probabilistic and generalised regression neural classifiers*", DRAFT VERSION of paper to appear at the Oncology Reports, special issue Computational Analysis and Decision Support Systems in Oncology, last quarter 2005, IVSL.

[3] Balakrishnan S., Narayanaswamy R., and Paramasivam I., "*An Empirical Study on the Performance of Integrated Hybrid Prediction Model on the Medical Datasets*", International Journal of Computer Applications (0975 – 8887) Volume 29– No.5, September 2011, pp. 1-6, IVSL.

[4] Engelbrecht A. P., 2007, "Computational Intelligence An Introduction", Second Edition.

[5] Dasgupta D. and Niño L. F., 2009, "Immunological Computation Theory and Applications", Auerbach Publications.

[6] Gupta S., Kumar D., and Sharma A., "*Data Mining Classification Techniques Applied For Breast Cancer Diagnosis And Prognosis*", Indian Journal of Computer Science and Engineering (IJCSE), Vol. 2 No. 2 Apr-May 2011, pp.188-195, IVSL.

[7] Karabatak M., and Ince M. C., "*An expert system for detection of breast cancer based on association rules and neural network*", ScienceDirect, 2008 Elsevier, IVSL.

[8] Khelil H, and Benyettou A., "*Artificial Immune Systems For Illnesses Diagnostic*", Ubiquitous Computing and Communication Journal, Volume 3 Number 4, 2007, pp. 88-93.

[9] lakhina S., Joseph S., and Verma B., "*Feature Reduction using Principal Component Analysis for Effective Anomaly–Based Intrusion Detection on NSL-KDD*", International Journal of Engineering Science and Technology, Vol. 2(6), 2010, pp.1790-1799, IVSL.

[10] Ludwig S. A. and Roos S., *"Prognosis of Breast Cancer using Genetic Programming"*, Department of Computer Science, University of Saskatchewan, Canada, ludwig@cs.usask.ca .

[11] Nabil E., Badr A., Farag I., and Osama M., "*A Hybrid Artificial Immune Genetic Algorithm with Fuzzy Rules for Breast Cancer Diagnosis*", INFOS2008, March 27-29, 2008 Cairo-Egypt.

[12] RADI A., Kartit A., REGRAGUI 3B., El Marraki 4M., ABOUTAJDINE D., and RAMRAMI A.," *On the Three Levels Security Policy Comparison between PCA and SVM*", IJCSNS International Journal of Computer Science and Network Security, VOL.11 No.2, February 2011, pp.69-79.

[13] Salama G. I., Abdelhalim M.B., and Zeid M. A., "*Breast Cancer Diagnosis on Three Different Datasets Using Multi-Classifiers*", International Journal of Computer and Information Technology (2277 – 0764), Volume 01– Issue 01, September 2012, pp. 36-43.

[14] Shyu M., Chen S., Sarinnapakorn K., and Chang L.,"*A Novel Anomaly Detection Scheme Based on Principal Component Classifier*", in Proceedings of the IEEE Foundations and New Directions of Data Mining Workshop, in conjunction with the Third IEEE International Conference on Data Mining (ICDM'03), 2003, pp.172-179, IVSL.