# A Review of the Automatic Methods of Cancer Detection in Terms of Accuracy, Speed, Error, and the Number of Properties (Case Study: Breast Cancer)

Jalilvand Farnaz[1]

[1] Tehran University

## Abstract

The purpose of this article is a review of the automatic methods of cancer detection in terms of accuracy, speed, error, and the number of properties and we have selected the breast cancer as the subject of the case study.The data used in this academic study area courtesy of the UCI in California. This database is called The Wisconsin Breast Cancer Datasets and includes 699 data units divided into benign and malignant classes. Ten properties were assigned to each datum. Four types of algorithms are used in this article, namely, classification algorithms, vector machine algorithms, neural networks algorithms, and data mining algorithms. Each category was evaluated separately and the best method in each category was identified in terms of accuracy, speed, error, and the number of properties.

*Index terms*— automatic methods of cancer detection, breast cancer, classification algorithms, vector machine algorithms, neural network algorithms, dat ining algor

# 1  I. Introduction

reast cancer is a common type of cancer. According to the calculations of the U.S. National Cancer Institute one of every eight women may be afflicted with cancer throughout of her life **??**1].

Generally, this disease is a result of the convergence of some risk factors that cause the disease. The breast cancer is a cancer with high mortality rate among women such that it is the most common cause of death in the female community [2]. Timely detection of breast cancer (maximum 5 years after the first cancer cell division) increases the chance of the patient survival from 56% to over 86% [3]. So, it seems essential to have a precise and reliable system for timely diagnosis of benign or malignant breast tumors [4]. In or country, the breast cancer ranks second after lung cancer and ranks first among female cancers and is responsible for a fifth of the women's deaths caused by cancer [5]. Traditionally, this type of cancer is diagnosed by biopsy via surgery and this Author: Faculty of Fine Arts Department of Industrial Design, Tehran university, Iran -Tehran. e-mail: Farnazjalilvand01@gmail.com method is the most accurate one among the existing methods but it is an expensive and invasive procedure and raises emotional and psychological concerns and anxiety for the patient [6]. However, because of biological similarity between benign and malignant patterns, the diagnostic specificity is not reasonable yet and differentiating between benign and malignant through mammography is highly dependent on the radiologist's skill. But it is not always possible to find a skilled radiologist for the interpretation of mammography images; therefore there is a strong urge to have an experienced radiologist at hand because it is the only way that one may raise the bar for both diagnostic sensitivity and specificity at the same time. But since access to a qualified radiologist for interpretation of the images is not possible in all parts of Iran, and even in the case of availability there is the possibility of human errors, therefore designing a computer model as a rapid and cost effective and easily available diagnostic method could be in order. Clearly if the mass is correctly diagnosed as benign, there is no need for surgery and that will dismiss many concerns imposed on the patient as a result of

such operations [7][8][9][10]. In this study, the researcher intends to choose the best model by comparing various techniques.

# 2   II. Materials and Methods

# 3   a) Introduction to algorithms i.   Identification via support vector machine (SVM)

In 1965, the Russian researcher Vladimir Pnik came up with the idea of minimizing the risk and it was a very important step in the design of classifiers. The SVM is a binary classifier that separates two classes using a linear border or super plane so that the maximum margin of the super plane may result. Maximizing the margin of the super plane will result in maximum separation between the classes. The training points that are closest to the maximum super plane margin are called support vectors. Only these vectors (points) are used to determine the boundary between the classes. In this method, the boundary line between the two classes can be calculated in a way that:

? All instances of class +1 are on one side of the border and all instances of class -1 are on the other side of the border. ? The decision boundary must be such that the distance between the closest training instances of both classes in the direction perpendicular to the decision making boundary is maximized as much as possible. ? The general form of decision making border line can be written as equation (1):W.x+b=0 (1)

where, X is a point on the decision making border line and W is an n-dimensional vector perpendicular to the decision making border line. The distance from the origin to the decision making border line and w.x indicates the internal multiplication of the two vectors w, x [11].

# 4   b) K-means algorithm

In the clustering technique is a famous data mining method in which an automatic process is used to classify samples in a given data space into distinct categories based on their characteristics and each category is called a cluster. Therefore, the cluster is a collection of objects that contains objects with the highest similarity together and they have the lowest level of similarity to the objects in other clusters. A number of criteria may be used to define the similarity including the parameter of distance, i.e. nodes that have the minimum distance with one another are grouped in one cluster and such clustering method is called distance-based clustering ??12].

Thereafter, some useful information may be extracted by checking and comparing the data in each cluster. K-means is one of the well-known clustering algorithms. This algorithm is one of the cluster-center based algorithms with the following order of execution [13]:

? Obtaining points as the clusters' centers at random ? Attributing each data sample to a cluster such that the data sample has the lowest distance to the center of the cluster.

In the standard presentation of the algorithm, first some points are randomly selected corresponding to the number of the required clusters (K). Then the data are assigned to one of the clusters based on the level of proximity (similarity) and thus, new clusters are generated.

The same procedure can be repeated and each time by averaging the data, new centers can be calculated for the data and the data can be assigned to new clusters again. This process continues until the clusters do not change anymore. Usually, the condition for algorithm termination and its convergence criterion is the lack of change in the existing clusters, achieving a predetermined number of iterations, or the expiration of the deadline.

The precise choice of the initial cluster centers highly influences the convergence of this algorithm and the optimal final clusters. These centers must be selected carefully and should be located at appropriate distances relative to each other. c) Imperialist competitive algorithm (ICA)

The evolutionary algorithms are a subset of the evolutionary calculations that belong to the branch of artificial intelligence and include algorithms wherein the search takes place from several points in the problem space. These algorithms are based on random search and exemplify natural biological evolution and work on potential answers that have superior characteristics and are capable of providing a better approximation of the optimal answer.

The ICA algorithm, like other evolutionary optimization methods, starts out with an initial population. In this algorithm, each data point is called a country. Countries are divided into two categories: colony and imperialist. Each imperialist country, depending on the extent of its power, dominates several colony countries and controls them. This algorithm is based on the colonial competition and absorption policy. In the provision of this algorithm, such policy is implemented via movement of an empire's colonies in accordance with a specific relationship [14-15].

# 5   d) Neural network algorithm

The neural network design has two main aspects of architecture and learning algorithm. The target neural network has a multilayer perceptron (MLP) structure with a better performance in comparison with other methods. MLP structure is a standard combination of inputs, linear and nonlinear neural units, and outputs ??15] ??16].

The output of all processing units of each layer is fed to all of the processing units of the next layer as input. The input layer processing units are all linear but in the hidden layers, especially in the output layer, nonlinear

neurons with hyperbolic or sigmoid tangent function or any other possible continuous and derivable nonlinear function can be used.

Neural networks have this ability to learn from the past, experience and environment; and improve their behavior during learning.MLP neural network's learning method uses a supervised learning method for training. In the supervised learning method, a set of paired data called training samples are defined as A = (Xi, ti) wherein Xi is the input and ti is the desired network output corresponding to the input.

After feeding the Xi input to the neural network, the actual output of Yi network is compared to ti and the learning error is calculated and used later to set the parameters of the network in such a way that if the next time the same Xi input is fed to the network, the network There are two computational paths in the above algorithm. The departure path in which the stimulating functions act on each neuron and the return path in which the sensitivity vectors (error vectors) are returned back to the first layer. Finally, the information obtained from the above two paths is used to set up the network's weight matrices and Bias vectors of MLP network. In order to stop the repetition of the post-errorpropagation algorithm, we can use the mean squared error (MSE)in the form of the equation( **??**):?????? = ? ????? ? ????(????)?2/?? 0 ??=1 (2)

The important point in the neural network is the sound choice of the weights and if necessary, the network's bias quantities. The method of weights selection is called learning algorithm training and a significant part of the network differences with one another lies in the method by which the parameters are set up[17-20].

# 6 e) Classification methods

Four classification algorithms were used for this study, namely, Perceptron artificial neural network method, Kohonen, fuzzy Artmap, and tree classification.

# 7 f) Artificial neural network classification

The artificial neural network concept was adopted from the human nervous system. These networks classify input patterns based on two unsupervised and supervised methods. In the unsupervised method, the output patterns are not introduced and the system itself classifies them on the basis of the similarity of the input patterns. But in the supervised method, the input is fed to the system and the difference between the desired output and the output of the network is used to change and adapt the weights. The training phase is the precursor of the classification phase in which the network carries out the training process by changing and adjusting the linking weights between the layers. In this study, we have used the Perceptron, Kohonen, and fuzzy Artmap artificial neural networks.

# 8 g) Perceptron artificial neural network

The Perceptron neural network is the first applied network in the artificial neural network history and a symbol of supervised pre-fed neural network and its design consists of one input layer, at least one hidden layer, and an output layer. Each layer is made up of nonlinear processing units called neurons (the nerves) and the connections between neurons in the successive layers carry the associated weights.

The connections are directional and are only in the forward direction. The method of learning in the supervised algorithm is the post-error-propagation method. In this method, the network weight is adjusted according to the gradient technique. So that after the desired output value is compared with the actual network output, the network searches for the maximum descending gradient and in the next iterations, the network parameters can be set up based on the guideline of the descending error gradient. The parameters' set up is repeated throughout this process until the amount of network error reaches an acceptable value [21]. h) Kohonen neural network Kohonen network is dual-layer network with unsupervised training. The Kohonen network, is a selforganizing network that learns a mapping of samples introduced for learning [22]. The structure of a Kohonen network (similar to a single-layer Perceptron network) has one input layer and some output neurons.

Kohonen Network training with n inputs and m outputs is described as follows: 1. First, the initial values of the network weights are selected at random. 2. The training samples are introduced to the network. 3. The following values are calculated for each of the output layer neurons.?? ?????? = ????????? 1 = ? (?? ?? ? ?? ???? ) 2 ?? ??=1 , ?? = 1, ? , ???**(3)**

4. The winning output neuron is identified and the weights are correct by using a neighborhood function.

?? ???? (?? + 1) = ?? ???? (??) + ??(??)??(??) ??? ?? ? ?? ???? (??)? (4)

In the above equation, ?(t)is the training parameter and N(t)is the neighborhood function.

5. The value of t is added. 6. The algorithm is repeated from step 2. The number of iterations can be considered fixed or the iterations continue until the network is trained, i.e. the weight values change slightly [23].After the network was trained, it is necessary to introduce the samples to the network. The output of the network is based on the closest distance. Among the output neurons, the winner (of the network output) is the neuron that has the least Euclidean distance with the input sample [23].

The output of the Kohonen map is the topological mapping corresponding to the network input.

## 9   i) Fuzzy Artmap artificial neural network

This network was introduced in 1992 by Hansen et al. [24]. Fuzzy Artmp is a supervised network that combines two fuzzy Art networks: ARTa and ARTb.

In the following, the parameters of these two networks are defined. These two networks connect to each other by a series of connections between the F2 layers of the two networks that are called Map Fields abbreviated as Fab. Each of these connections has a weight Wij with a value between 0 and 1. The Map Fieldh as two parameters ? ab and ? ab and the output vector X ab . The input vector into ARTa is converted into vector A under the supplemental coding.

In the training phase of the fuzzy Artmap network, the input pattern vector into the ART a network and the desired output (B) associated with the input pattern A are presented to the ARTb network. ARTb, the care parameter (? b ), is set to 1 to distinguish the desired output vectors. After the presentation of the vectors B and A, the ARTa and ARTb networks enter the resonance phase.

At this stage, another care criteria defined according to the equation ( 3) is calculated to assess whether the winning neuron in ARTa is associated with the desired output vector in ARTb.??? ?? ??? ?? ???? ? ??? ?? ? ? ?? ????**(5)**

In equation ( 3), y b is the output vector in ARTb (pattern in F 2 b ) ,J is the subscript of the winner neuron in F 2 a , W J ab is the weights of Map Field connections with the Jth neuron in [?[0,1? ab F 2 a and the care parameter in the Map Field.If the above criterion is not met, the care parameter in ARTa is increased a certain amount until the fuzzy Artmap network selects another winning neuron.

The vector A will be entered into the network again and the process will be repeated until the care criterion is satisfied. At this time the weights of the Map Field connections are updated according to the following equation:?? ?? ???? = ?? ???? ?? ???? + (1 ? ?? ???? )?? ?? ????**(6)**

The initial value of the ? a is set by the basic care paramete (? a ???). After updating the weights, the basic care parameter in ARTa is set to this initial value again. After completion of the training phase, parameters ? a and ? a are initialized to a zero value. The Map Field output vector is defined as follows:?? ???? = ?? ?? ???? (7)

So that J is the subscript of the winning neuron in F 2 a . This equationshows that the Map Field assigns a classifying number to each neuron in the F 2 a layer.

## 10   j) Tree classification

In the IDRISI software, the tree classification is based on the algorithm [25]. In practice, this algorithm selects an attribute (such as the reflective band) by iteration, divides the samples that can be split into two groups, and minimizes the difference within each subgroup while maximizes the difference between the groups. The tree classification progresses by consecutive breakdown of the data within new middle nodes containing more homogeneous subsets of the training pixels. A newly created middle node, in a situation that the training pixel has only one class or one class predominates the pixels, may generate a leaf.

When there is not any middle nodes left for splitting, the final tree classification rules take shape. The IDRISI software makes use of 3 splitting algorithms: entropy, gain ratio, and Gini.

## 11   k) Data mining

The processing models described in this work are as follows:

Neural networks, Bayesian network, K-nearest neighbor, C5.0 tree, and CART tree. This section presents a brief introduction of these models [26].

Suppose you have a series of data in the set S. The C4.5 algorithm works as such: if all the items in S are for a collection or S is sufficiently small, a leaf with maximum collection in S is added to the tree.

Otherwise, a test is selected based on a single distribution with one or two outputs. This experiment as the root of a tree is considered as a test by placing each of the outputs.

The S set is divided into S1, S2, etc. subsets based on the corresponding output. This process is applied to all subsets.

In 1997, The C 4.5 algorithm was replaced with a business system entitled See5/c 5.0. The applied changes resulted in remarkable capabilities that include:

? Variation of the boosting process which is a collection of classifiers that will be considered the final classification at a later stage. The boosting process usually achieves significant prediction accuracy. ? New data types, irrelevant values, cost of incorrectly classified variables, and the methods used to prefilter the domain. ? Set of the chaotic rules when a class is classified, all the relevant rules are pinpointed and voted thereto. This enhances the interpretation of the set of rules. Moreover, it increases the precision of their prediction rate. The decision tree and the set of rules have progressed in an ascending fashion. The rate of climbing increases through segmentation. C 5.0 can run on multi-processor computers as well [7].

The CART decision tree is a binary recursive segmentation method. The Data are randomly selected. Throwing away is not necessary and not recommended. The trees grow to their maximum size without the use of the stop law. Then the regressive pruning phase back to the root commences (part by part isolation).

The next section which to be pruned is the part that has the lowest performance relative to the whole in the training data tree (it should be pointed out that more than one section can be removed at any stage). This In fact, the CART algorithm mechanism is intended to generate not one, but a consecutive collection of pruned trees which are all optimal candidate trees. The right size tree is identified and selected by evaluating the performance of each tree in the pruning stage [27].

Rote is one of the easiest and almost insignificant classifications which saves all the conducted training data. It runs the classification operation only when the test exactly matches a training example. A clear objection to this method is that a lot of test data will not be classified because they do not exactly match with any of the recorded training data.

A more advanced method in this regard is the k-nearest neighbor method. According to this algorithm, a group is pinpointed with k samples in the training dataset that are closest to the test data and the desired tag is selected based on the superiority of a specific class in its neighborhood [28].Based on the given dataset, each belongs to a class. Each of these classes has a vector. Our goal is to create a rule to assign each of the samples to a class in the future. This is carried out only through the vectors intended for the variables. Many ways have been introduced to create such rules. One of the most important methods is Bayes method. This algorithm is important for many reasons. This algorithm can be made very easily because it does not require complex iterative parameters.

This means the algorithm can be used in some cases that the there is a high volume of data. Neural networks are new systems and computational methods for machine learning and display of knowledge.

Also, its ultimate goal is the application of gained knowledge to predict output response of the complex systems [29,30].

The main idea of this kind of network is (partly) inspired by biological nervous system functions. The key element of this idea is the formation of new structures for the information processing system. This system consists of many interconnected processing elements called neurons.

They collaborate together to solve a problem and pass on the information through synapses. In these networks, if a cell is damaged, the rest of the cells may compensate for the lack of it and also contribute to its reconstruction.

# 12   l) Introduction of data

The data used in this academic study are a courtesy of the UCI in California. The title of this database is The Wisconsin Breast Cancer datasets 2 and includes 699 data items. The data are divided into two benign and malignant classes. Ten properties were attributed to each datum. These properties are presented in the table 1. The data properties were chosen with an eye to medical considerations. The thickness of malignant cells' masses is usually classified in single-layer groups while the cancer cells are usually placed in multi-layered groups.

There are various sizes and apparent abnormal shapes of the cancer cells. This is why these parameters are useful in diagnosis and detection of cancerous and non-cancerous cells. With regard to the rate of cell adhesion, normal and healthy cells usually get close to each other while the cancer cells do not show such solidarity among them.

As a result, lowered adherence also can be used to detect cancer cells. With regard to the single biological cell size, its size may vary as we explained earlier.

Biological cells that grow big can be considered as cancer cells. A bare nucleus is in fact a nucleus without any cytoplasm surrounding it.

Usually, these nuclei are observed in malignant tumors. In fact, chromatin is the expression of nonuniform nuclear material that is observed in malignant cancer cells. In the cancer cells, Chromatin is observed as follows.

Typical nuclei with tiny structures have been observed in healthy tissues. In cancer cells, the nuclei gradually swell and their number increase over time. Mitosis is a part of the nucleus that generates two identical daughter cells during prophase. Cell division takes place during this process. The pathologists can estimate the amount of cancer by counting the mitoses.

# 13   m) Preprocessing

Data preprocessing is the first step in performing the task. According to the information that the database put in our disposal, there are 16 forgotten data in its 7th column. To fix this problem, a data The outlying data are the next problem that will be dealt with during the statistical tasks. Of course, the data used in this study had been previously checked by the related center and the outlying data have been removed from them. In other words, the database has presented the users with neat data. However, some processes were carried out to ensure there are no outlying data around. n) Selection of the properties One of the most important steps that can greatly help the processing to get better results is selecting the right features. If these properties are not selected correctly, there is a high potential for synergy.

Synergy means trouble at the prediction stage. The large number of features can drastically increase the number of calculations in the event that it does not appreciably change the authenticity. So, a good choice of the right properties can increase the processing efficiency.

To do this, the property correlation matrix was calculated. Since the correlation threshold of 0.95 was selected, no property correlation was observed. Therefore, all of the properties can be used in the prediction process.

We used the random selection procedure to do justice to the training data sample selection and testing and to eliminate the dependence of the training on a particular order of data. The highlight of performance assessment of the models presented in this research is this same point that the evaluation results are not dependent on the particular order of data and each model that runs using a random selection from the dataset, 70 percent are allocated to training and 30 percent to testing. So, the results are not dependent on a particular order of the data.

# 14  III. Results

To carry out the modeling based on the artificial neural network, first we fed the data (dataset) to the Excel software. The dataset is randomly (MATLAB software's Rand function) divided into two parts: training (learning) dataset and testing dataset.

In order to do this, about 30 percent of the data was used for validation and testing According to the final results, their recommended k-SVM procedure is more accurate than SVM for the diagnosis of benign and malignant tumors and the new input structures and patterns with the membership of new patterns based on the main data.

The results of the classification using methods such as artificial neural network (Perceptron neural networks, Kohonen, and fuzzy Artmap) and the tree classification are presented in the figure. According to the figure, the tree classification method with three branching techniques(gain ratio, entropy, and Gini) has obtained the average total accuracy of 90 and Kappa coefficient 0.88, respectively; while the neural network methods had an accuracy of 92 and the Kappa coefficient of 0.90, respectively (except the Kohonen method).Thus, the neural networks classification methods (with the average total accuracy of 2 and the average Kappa coefficient of 2%) were more accurate than the tree classification method (with three branching methods) for the series of data used in the study.

In addition, when the different methods of the neural network were analyzed, it was clear that the fuzzy Artmap neural network method was more accurate than Perceptron and Kohonen methods (with the total accuracy enhancement of 2% and 22% and the Kappa coefficient enhancement of 3% and 24%).Finally, it can be said that there was no significant difference among the three branching methods employed in this study (fig **??**). , clearly the Bayesian network has predicted the classes more accurately and can be used with more confidence for predictions. Also, its distribution authenticity is much less than other models. But these cases are quite the reverse in the neural network, i.e. its distribution is much more and its mean authenticity is lower than all other models.

# 15  IV. Conclusion

In the case of diagnosis by SVM from the K-SVM algorithm based on the properties it can be competitively compared with any of the old data mining methods of cancer detection. In the property extraction phase, the older methods are not used for extracting useful information. Properties with below 50% accuracy may not be appropriate for classification purposes. No single property is highly effective for mass classification applications. The properties must be combined to achieve a better performance. After evaluating each property individually, it is advisable to check a combination of these properties. Both geometry and texture properties are useful for detection of mass. The selection of a classification property that may be able to improve the accuracy will prove useful. In the case of diagnosis by neural network, the algorithm based on gradient that considers the Levenberg-Marquardt algorithm and post-error-propagation method. One of its disadvantages is late convergence and stopping at optimal local points, as well as the optimal choice of number of layers, the number of neurons in each layer, and the type of stimulation function of each neuron in the data with high complexity which is not a simple task. Therefore, the smart optimization algorithms such as mass particles algorithm, genetic algorithm, imperialist competitive algorithm, etc. that we used three types of them for network modeling in this study. The results showed that the imperialist competitive method is significantly capable of correct determination of neural network weights for its training and error elimination.

With regard to classification methods, all pixels could correctly identify the agricultural lands. be a result of its distinctive spectroscopic characteristics in comparison with other types of coating. When different methods of artificial neural networks were analyzed, it was clear that the fuzzy Artmap artificial neural network provides more accurate results relative to Perceptron and Kohonen artificial neural networks (with an overall accuracy enhancement of 2% and 22% and the Kappa coefficient enhancement of 3% and 24%).

In this research, the fuzzy Artmap artificial neural network had the highest classification accuracy.

In the case of data mining techniques, it is clear in accordance with the table 3-6 that the Bayesian network predicted the classes more accurately in and can be used with more confidence in the predictions. Moreover, its distribution authenticity is much less than other models. But these cases are quite reversed in the neural network, i.e. its distribution authenticity is much more and its average accuracy is much lower than the other models. Among all models created by different training and testing data, only the neural network models were unable to provide a prediction in some cases. That is, they presented some data without any comments there to.
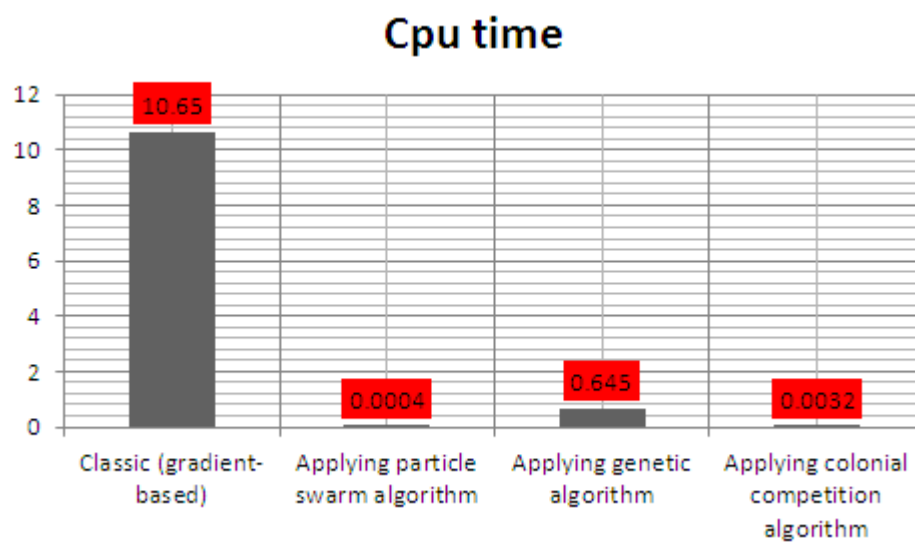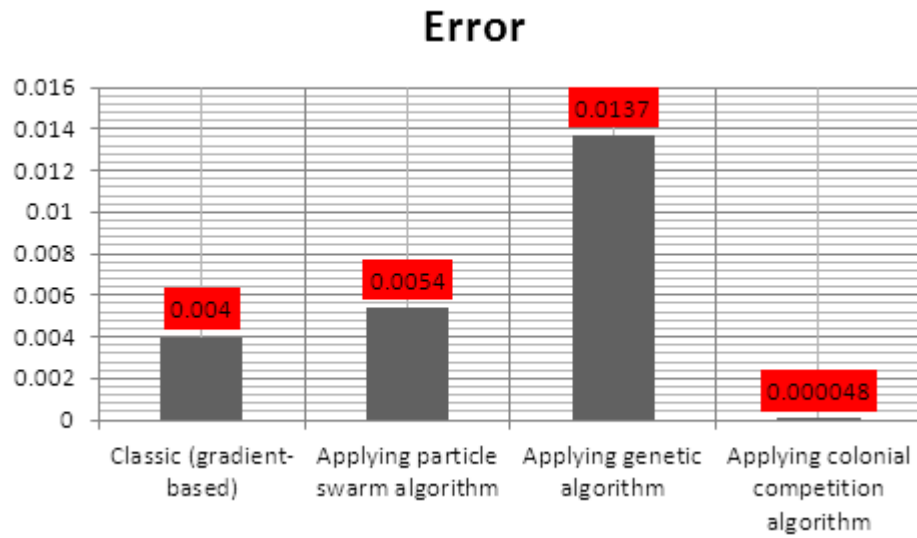
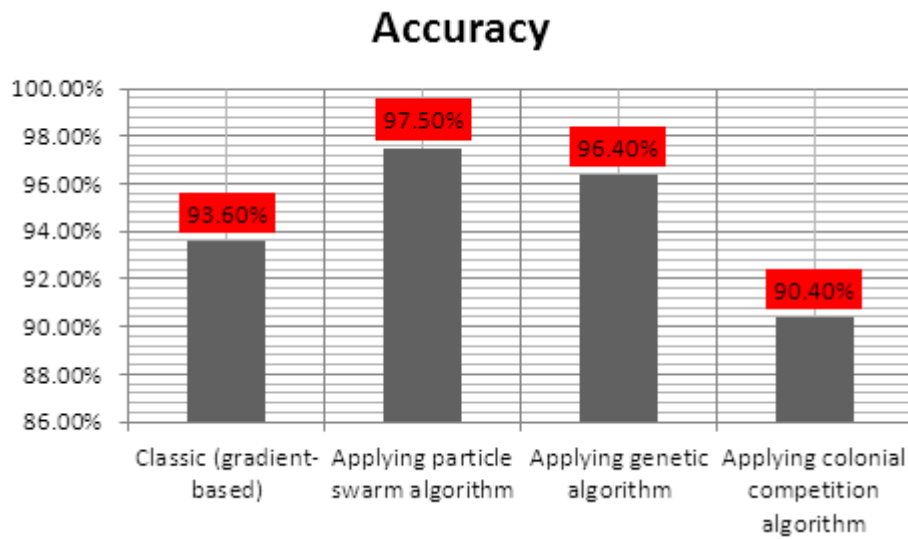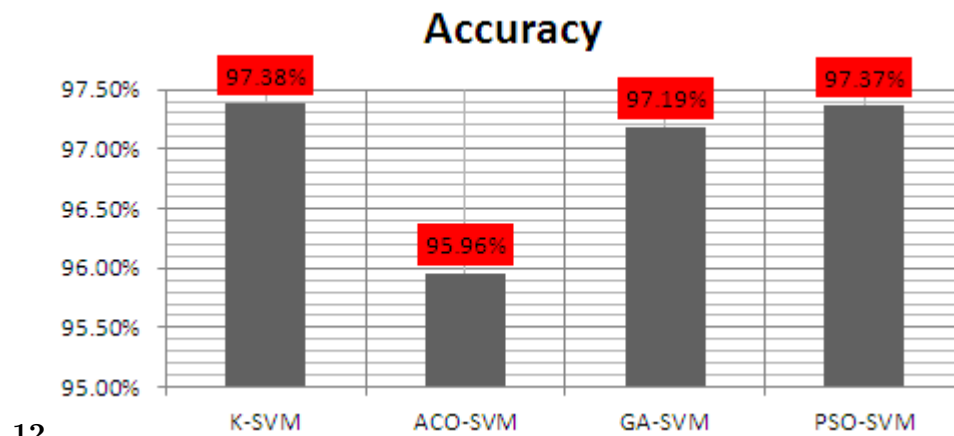Figure 1:



Figure 2:

Figure 3:



Figure 4:



Figure 5:  Figure 1 :Figure 2 :

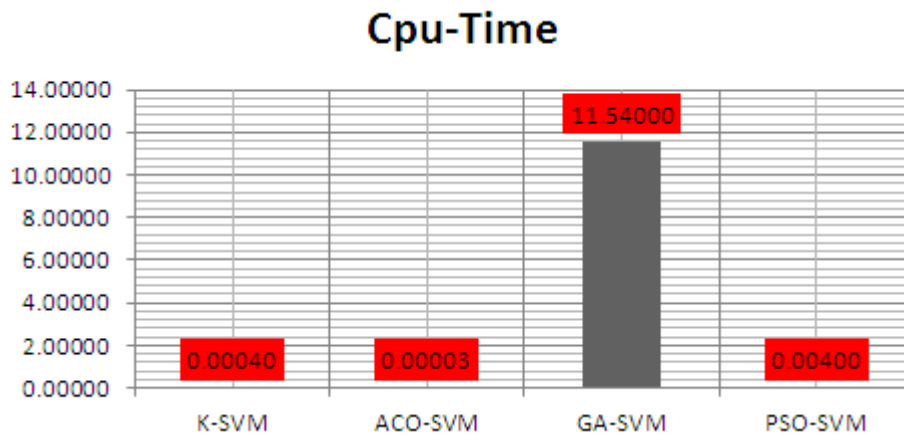## Cpu-Time



**34**

Figure 6: Figure 3 :Figure 4 :

## Feature Space Dimension



Figure 7:

**1**

| ID | Attribute | Domain | |
|----|-----------|--------|---|
| 1 | Sample code number | -10 | 1 |
| 2 | Clump Thickness | -1 | 10 |
| 3 | Uniformity of Cell Size | -1 | 10 |
| | Uniformity of Cell Shape | -1 | 10 |

Figure 8: Table 1 :

# 16   Volume XVI Issue I Version I

1 2 3 4

¹http://www1.ics.uci.edu/~mlearn/MLSummary.html [breast-cancerwisconsin]

²© 2016 Global Journals Inc. (US)

³A Review of the Automatic Methods of Cancer Detection in Terms of Accuracy, Speed, Error, and the Number of Properties (Case Study: Breast Cancer)

⁴http://www1.ics.uci.edu/~mlearn/MLSummary.html [breast-cancerwis consin]

335   [Mcclelland et al. ()] , Eds Mcclelland , Cambridge , Ma . 1986. MIT Press.

336   [Kumar et al. (2013)] 'An Efficient Prediction of Breast Cancer Data using Data Mining Techniques'. G Ravi
337      Kumar , Dr G A Ramachandra , K Nagamani . *International Journal of Innovations in Engineering and*
338      *Technology (IJIET)* August 2013. 2 (4) p. .

339   [Mangasarian et al. ()] 'Breast cancer diagnosis and prognosis via linear programming'. O Mangasarian , Nick
340      Street , W Wolberg , W . *J Operations Res* 1995. 43 p. .

341   [Salama et al. ()] 'Breast cancer diagnosis on three different datasets using multi-classifiers'. G I Salama , M
342      Abdelhalim , M A Zeid . *Breast Cancer (WDBC)* 2012. 32 (569) p. 2.

343   [Anderson et al. ()] 'Comparison of age distribution patterns for different histopathologic types of breast
344      carcinoma'. W F Anderson , R M Pfeiffer , G M Dores , M E Sherman . *Cancer Epidemiol Biomarkers*
345      *Prev* 2006. 15 (10) p. .

346   [Zhaohui et al. ()] 'Diagnosis of breast cancer tumor based on manifold learning and support vector machine'. L
347      Zhaohui , W Xiaoming , G Shengwen , Y Binggang . *Proc. IEEE Int. Conf. Information and Automation*,
348      (IEEE Int. Conf. Information and Automation) 2008. p. .

349   [Krishnaiah] *Diagnosis of Lung Cancer Prediction System Using*, V Krishnaiah .

350   [Ahmed et al. ()] 'Early Detection of Lung Cancer Risk Using Data Mining'. Kawsar Ahmed , Abdullah Al
351      Emran , Tasnuba Jesmin , Fatima Roushney , Md Mukti , Farzana Zamilur Rahman , Ahmed . *Asian Pacific*
352      *Journal of Cancer Prevention* 2013. 14 p. .

353   [López-Caneda et al. ()] 'Evolution of the binge drinking pattern in college students: Neurophysiological
354      correlates'. Eduardo López-Caneda , Socorro Rodríguez Holguín , Montserrat Corral , Sonia Doallo , Fernando
355      Cadaveira . *Alcohol* 2014. Elsevier. 48.

356   [Hansen et al. ()] 'Global land cover classification at 1 km spatial resolution using a classification tree approach'.
357      M C Hansen , R S De Fries , J R G Townshend , R Sohlberg . *International Journal of Remote Sensing* 2000.
358      21 p. .

359   [Lawrence ()] *Handbook of genetic algorithms*, D Lawrence . 1991. 1991. NY: Chapman and Hall. p. .

360   [Messadi et al. ()] 'Interpretable Aide Diagnosis System for Melanoma Recognition'. M Messadi , M Ammar , H
361      Cherifi , M A Chikh , A Bessaid . *Bioengineering & Biomedical Science* 2014.

362   [Kohonen ()] T Kohonen . *Self-organization and Associative Memory*, 1996. Springer-Velag. p. 312. (3rd Edition)

363   [Rumelhart et al.] 'Learning internal representations by error propagation'. D E Rumelhart , G E Hinton , R J
364      Williams . *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, D E Foundations,
365      J Rumelhart (ed.) 1.

366   [Media Centre for World Health Organization Fact sheet] 'Media Centre for World Health Organization'. *Fact*
367      *sheet* (297) p. 2007. (Cancer)

368   [Bishop ()] *Neural networks for pattern recognition*, C M Bishop . 1995. New York: Oxford University Press.

369   [Miklos ()] *Numerical control of kohonen neural network for scattered data approximation*, H Miklos . 2000.
370      Mathematics Subject Classification.

371   [Bellaachia and Guven ()] 'Predicting breast cancer survivability using data mining techniques'. A Bellaachia ,
372      E Guven . *Age* 2006. 58 (13) p. 110.

373   [Delen et al. ()] *Predicting breast cancer survivability: a comparison of three data mining methods. Artificial*
374      *intelligence in medicine*, D Delen , G Walker , A Kadam . 2005. 34 p. .

375   [Litigate ()] 'Predictive models for breast cancer susceptibility from multiple single nucleotidepolymorphisms'. J
376      Litigate . *J Clin Cancer Res* 2004. 10 p. .