

Determining the Malignancy of Breast Cancer Using a Combined Neural Network-Support Vector Machine Scheme

Bahareh Hemmati^a, Hosein Jafarkarimi^{a,*}

^a Department of Computer, Damavand Branch, Islamic Azad University, Damavand, Iran

* Corresponding author email address: karimi@damavandiau.ac.ir

Abstract

Data mining has been useful in medical diagnosis area. The current research proposes a new combined classification technique based on support vector machine and neural network to determine malignancy of breast cancer tumours. The proposed method uses a meta-heuristic algorithm to find the best weights for combining the results of two aforementioned techniques. The numerical and statistical results confirm the superiority of our model in comparison to the previous methods. Following the results, the best accuracies are 98.23% and 96.85% for proposed and the previous methods, respectively. The method has potential to be implemented as a decision support system for breast cancer diagnosis.

Keywords: Classification, Breast cancer, Artificial Neural Network, Support Vector Machine, Wilcoxon hypothesis test

1. Introduction

Breast cancer diagnosis would be verified by biopsy of lump. After diagnosis of cancer, further test will be done to determine whether the cancer has spread to other parts of the body and what therapies might not react to disease (Polat & Günes, 2007).

According to World Cancer Research Fund (Stewart & Wild), breast cancer is the most common type of cancer among women, in 2012, 1.7 million cases of breast cancer were reported. On average, 37.3% of the patients regain their health completely; this rate of success is the result of early diagnosis in addition to correct evaluation of severity of malignancy (Salama, Abdelhalim, & Zeid, 2012). In this regard, detection of malignancy severity of a cancerous lump is the main subject of this research. In this content, the present study, by focusing on the breast cancer-related data, is intended to combine two training methods- Artificial Neural Network (ANN) and Support Vector Machine (SVM). Different kernel functions in SVM and layers and neurons in each layer of neural network lead to flexibility of these two methods and application of a combinational approach could improve the prediction precision in comparison with the state in which these two methods were applied alone.

The objectives of this research are:

- i. Developing a model for determining the malignancy of a cancerous lump in the breast.

- ii. Increasing the prediction precision by combining ANN and SVM.

The rest of the article is organized as follows. Support Vector Machine and Artificial Neural Network algorithms are introduced in Section 2. Section 3 presents a literature review on researches about breast cancer. Section 4 is devoted to explanation and implementation of the proposed method. Experimental results are addressed in Section 5. The recommendations for future studies are provided in Section 6.

2. Background

2.2 Artificial Neural Network

ANN is one of machine training methods inspired by biological neural network which is used for function estimation. This network is generally made in the form of interconnected neurons which transfer messages. Neural network connections have weights adjusted to obtain maximum efficiency (McCulloch & Pitts, 1943). In neural network, layer implies number of adjacent neurons at the same level; in addition, neurons are the nodes in the structure by which the data move from input to output along the layers. Any neuron has a weight, sum of these weights affect the performance of neural network. Neural network training means adjustment of these weights to achieve the best performance (Tu, 1996).

2.3 Support Vector Machine

SVM was presented by Vapnik, which is one of the most common algorithms in data classification field, in which data are precisely classified by making support vectors as criteria (Mishra, Ananth, Shelke, Sehgal, & Valadi, 2015). Support vectors are the set of points in n-dimensional space which determines the boundaries of classes. In two-dimensional space, support vectors are a line, in three-dimensional space they are a plane and in an n-dimensional space, they are a hyperplane. The hyperplane in the SVM plays the role of decision function and is calculated in such a way that it maximizes the margin of classification which is equal to maximizing the classification precision. Margin of classification means the distance of the nearest point to the classifier hyperplane. If this classifier hyperplane does not have proper accuracy, then it is necessary to map the data to the other space called feature space. Such mapping is achieved by kernel function (Vapnik, 1998).

3. Previous studies

Review of the previous works showed that application of data mining-based methods and machine training theory for prediction of breast cancer and also determination of the disease severity are very common. Quinlan (Quinlan, 1996), after presentation of C4.5 algorithm for construction of decision tree, used his proposed method for prediction of cancer lump malignancy and reached to precision of 94.74%. Wang et al. (P. Wang, Hu, Li, Liu, & Zhu, 2016), processed the cell nucleus images by SVM and proposed a method by which the alternation of cell nucleus shape and possibility of cancer could be predicted. They also determined the effective variables in the process in which the cell nucleus become cancerous. In another study, Desampaio et al. (de Sampaio, Silva, de Paiva, & Gattass, 2015), by use of SVM, employed data obtained from mammography images for determination of malignant lump, they reached to 92.99% accuracy. Sweilam et al. (Sweilam, Tharwat, & Moniem, 2010) proposed an active set method to increase the speed of support vector machine and then implemented it on the data related to breast cancer. The results showed the superiority of the proposed method. Liu et al. (Liu et al., 2003) revealed that SVM surpass neural network in term of breast cancer prediction. Stoean et al. (Stoean & Stoean, 2013) applied support vector machines for feature selection, malignancy and reoccurrence determination of breast cancer. Korkmaz et al. (Korkmaz & Poyraz, 2015) employed Minimum Redundancy Maximum Relevance- Least Square Support Vector Machine (MRMR-LSSVM) to analyze the microscope medical images. In this method, the most effective factor would be selected by evaluation of the variables' relation with cancerous cells and then determined the probability of the cell becoming cancerous. By application of genetic algorithm and combining that with neural network, Rouhi and Jafari (Rouhi & Jafari, 2016) determined the most important features in malignancy of a tumor, and then performed classification.

Their results indicated the superiority of the combinational method over the traditional ones. Rouhi et al. (Rouhi, Jafari, Kasaei, & Keshavarzian, 2015) applied neural network for diagnosis of malignancy of a lump and determination of the growth site to find the best place for surgery and removal of cancerous cells. In another study, Marcano et al. (Marcano-Cedeño, Quintanilla-Domínguez, & Andina, 2011) presented a new algorithm based on Beez theory for training the neural network and employed that for malignancy determination with 94.51% accuracy. Majid et al. (Majid, Ali, Iqbal, & Kausar, 2014) combined SVM with k-nearest neighbor and presented a method for breast cancer prediction. They also used this method for colon cancer and showed that combination of these two methods would increase the prediction accuracy. Jiao et al. (Jiao, Gao, Wang, & Li, 2016) used neural network for classification and detection of effective factors in order to choose the feature. In this study 9 factors among the 17 available factors were selected as effective ones and classification via them resulted in accuracy up to 96.7%. Naushada et al. (Naushada et al., 2016) reached to accuracy of 94.2% for analysis of genetic factors of breast cancer by application of forward neural network which showed 11.3% improvement in comparison with the conventional methods. Wang et al. (D. Wang, Shi, & Heng, 2009) used SVM to detect existence of lump in the breast from the medical images and reached to 90.7% accuracy. In this study, second-order polynomial kernel functions and radius bases were employed. Lisboa et al. (Lisboa et al., 2010) used neural network and logistic regression for predicting the consequences of breast cancer. Their results showed that neural networks were more accurate. McLaren et al. (McLaren, Chen, Nie, & Su, 2009) compared neural network and logistic regression in prediction of malignant breast lesions through MRI. Their results indicated the superiority of neural network with accuracy of 0.76 over logistic regression method. Chen et al. (Chen et al., 2009) used artificial neural network, support vector machine and logistic regression for determination of whether the breast tumors are benign or malignant. Analyzing the Receiver Operating Characteristic (ROC) showed that there was no significant difference among these three methods in terms of tumor diagnosis. Sattlecker et al. (Sattlecker, Baker, Stone, & Bessant, 2011) employed several support vector machines and bootstrap technique for detecting malignancy in data of breast cancer and reached to accuracy of 88.9%. Song et al. (Song, Venkatesh, Conant, Arger, & Sehgal, 2005) in a comparative study through application of logistic regression and ANN for detection of breast lumps, showed that ANN had better performance in determination of the type of lump. Utomo et al. (Utomo, Kardiana, & Yuliwulandari, 2014) investigated the performance of various types of neural networks on breast cancer-related datasets and, in best condition, reached to accuracy of 93.4%. According to the mentioned research, effective features in determination of lumps' malignancy are lump thickness, uniformity of cell shape, marginal shape of the lump, cells mitosis and normal cell nucleus. Arevalo et al. (Arevalo, González, Ramos-Pollán, Oliveira, & Lopez,

2016) used convolution neural network to determine the lump from mammography results. Zheng et al. (Zheng, Yoon, & Lam, 2014), through a combinational approach, used SVM and K-means algorithms for clustering-classification. In this research, data were first clustered and then a classifier was used for each cluster. Accuracy and convergence speed enhancement were the results of this research. Uzer et al. (Uzer, Inan, & Yılmaz, 2013) employed forward and backward method for feature selection and classification with ANN to maintain the accuracy and increase classification speed, in which forward method led to more effective results. Addeh and Ebrahimzadeh (Addeh & Ebrahimzadeh, 2012) used ant colony algorithm and support vector machine for feature selection and reducing the data dimension which resulted in convergence speed improvement. With the aim of improving breast cancer prediction accuracy, Sivakami (Sivakami, 2015) used SVM and decision tree simultaneously and compared the combinational method with Instance-Based Learning (IBL), Sequential Minimal Optimization (SMO) and Nayo Beezine methods, the results show higher precision of the combinational method. Hassanien et al. (Hassanien, Moftah, Azar, & Shoman, 2014) used clustering method based on ant colony and neural network to detect malignancy or benign state of the breast lumps in MRI images; the results revealed increased accuracy of the combinational approach. In another study, Kalteh et al. (Kalteh, Zarbakhsh, Jirabadi, & Addeh, 2013) applied clustering-classification combinational approach through K-Modified Imperialist Competitive Algorithm (K-MICA) and ANN to improve the classification accuracy. The rest of available methods addressed accuracy improvement by application of their ideas. But presentation of a more accurate method has been yet addressed in the studies which resulted in increase of the number of publications in this regard.

Table 1

Variables in the dataset.

Row	Variable name	Description
1	Patient code	-
2	Lump thickness	Indicating the volume and thickness of the lump including the cancerous cells
3	Cell size uniformity	Indicating how the cell volume change through becoming cancerous
4	Cell shape uniformity	A scale for determining the cell shape alternation through the process of becoming cancerous
5	Marginal adhesion	Adhesion level of lymph node to each other
6	Size of Epithelial tissue cell	Covering tissues without blood vessels in the body
7	Bare nuclei	Size of cell without considering membranes, vacuoles, cytoplasm, etc.
8	Chromatin	Chromosomes of non-dividing cell
9	Normal nuclei	Nuclei size with vacuoles, membrane cytoplasm and ...
10	Mitosis	A cell division method in which sister chromosomes are separated from each other

4.2 Data preprocessing

In the first step of data preprocessing, missing values were completed by linear regression method. Linear regression investigates the relation of a dependent variable with a number of independent variables. Based on that, the derived equation can be applied for completing the missing values (Han, Pei, & Kamber, 2011).

4. Method

In this section, the method of combinational application of SVM and ANN which is the main idea of this research will be introduced and the calculation results will be presented. Fig.1 depicts a summary of methodology and research implementation.

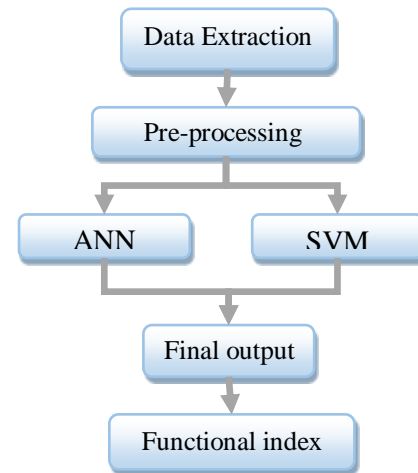


Fig. 1. Methodology and research implementation.

4.1 Dataset

The data were derived from database of theory machine training data of university of California (Wisconsin Breast Cancer Dataset, 2014). This dataset includes 699 data some of which have missing values. In addition, there are 10 variables among them. These variables are more associated with the shape of lump and cancerous cell based on which we can decide about malignancy or benign state of lump. Table 1 listed the mentioned variable.

Among the data of the mentioned dataset, only 16 values are missing which are related to bare nuclei values. For completing them, linear regression, coefficient determination and the adjusted coefficient determination were used. Eq. (1) and Eq. (2) are related to these indices:

$$R^2 = 1 - \frac{SSE}{SST} \quad (1)$$

$$R_{adj}^2 = 1 - \frac{SSE/N - n - 1}{SST/N - 1} \quad (2)$$

where N is the total number of data and n is the number of independent variables. On the other hand, SSE and SST are sum of square errors and sum of square total which can be calculated by Eq. (3) and Eq. (4).

$$SSE = \sum_{i=1}^N (y_i - \hat{y}_i)^2 \quad (3)$$

$$SST = \sum_{i=1}^N (y_i - \bar{y}_i)^2 \quad (4)$$

where y_i is the real value, \bar{y}_i is the average value and \hat{y}_i is the predicted value for the dependent variable.

The results of calculations are listed in Table 2. As, overall there was only 16 missing values, the predicted value of each of these cases are listed in the table. It must be noted that for completing the missing values, the first variable, code of patient, was omitted from the data.

Table 2
Result of missing values.

Row	Number of data	Predicted value
1	24	7.20
2	41	3.41
3	140	1.20
4	146	1.58
5	159	1.27
6	165	1.44
7	236	1.96
8	250	1.41
9	276	1.62
10	293	6.34
11	295	1.22
12	298	1.00
13	316	2.01
14	322	1.41
15	412	1.20
16	618	1.05

Prediction-related indices were also calculated as $R^2 = 70.27\%$ and $R_{adj}^2 = 69.87\%$.

In the next step single-class support vector machine was used for discovery of the outer values which can be calculated by Eq. (5).

$$\min R^2 + \frac{1}{vm} \sum_{i=1}^m \xi_i \quad (5)$$

$$\|\phi(x_i) - c\|^2 \leq R^2 + \xi_i$$

$$\xi_i \geq 0$$

where R , ξ_i and c are decision variables of the model and indicate radius. Data deviation extends from coverage range and the center of the coverage range, respectively. Moreover, the level of trade-off between coverage radius and the number of data out of the coverage range can be determined by If $v \in \{0,1\}$ the value of this parameter is close to zero and hence, more importance will be given to the point out of the range. Therefore, R will be larger. Also, if v tends to 1, the importance of the out-of-range data would be lower and R will be smaller. The best value for v is the value which is able to balance the trade-off between these two variables. In this regard, a value of 0.1 was selected for v and 44 data were omitted.

4.3 Proposed model

The proposed method, along with combination of the outputs of SVM and ANN, determined a discretizing boundary by the proposed algorithm. It also tried to improve prediction accuracy in comparison with the traditional methods in which algorithms were implemented

separately. Fig. 2 shows the method of combining the SVM and ANN.

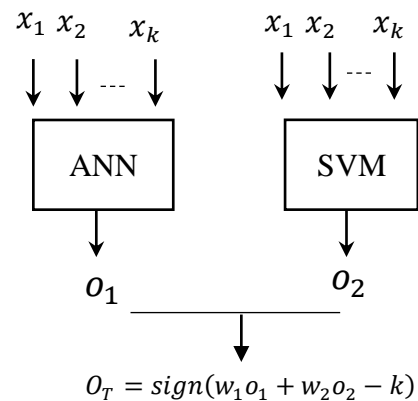


Fig. 2. Method of combining the SVM and ANN.

In this content, output of the each two methods was combined through w_1 and w_2 weights and then a separation boundary, namely k , was determined for them. Optimized values of these three variables were determined through an innovative iterative algorithm to attain the highest possible accuracy. The flowchart of the proposed method is depicted in Fig. 3.

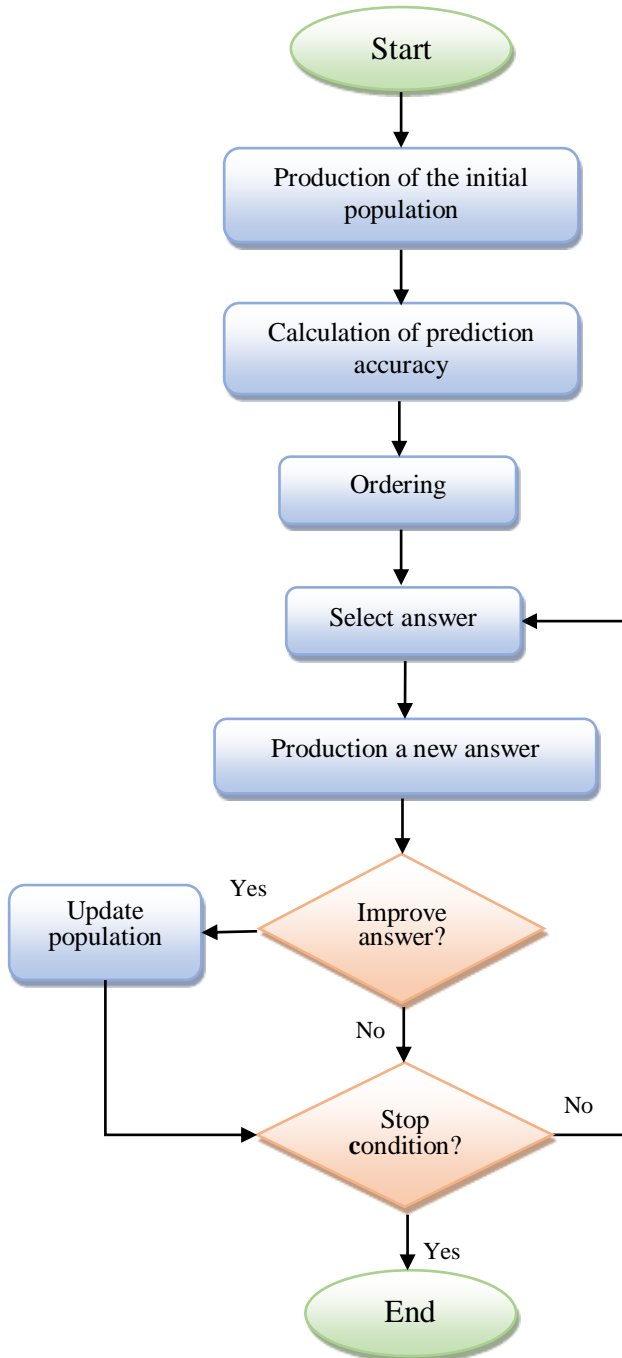


Fig. 3. Flowchart of the proposed method.

This algorithm was run through following steps:

1. An initial population of answers with the size of P was produced. Each answer included a triplet vector containing the weights and discretizing boundary.
2. Prediction accuracy would be calculated for each answer and labelled as f_i , where $i=1,2, \dots, P$.

3. The answers were regulated based on their accuracy in descending order.
4. An answer would be selected for production of a new answer. Each of the answers had a specific chance to be selected and this could be calculated via Eq. (6).

$$V_i = \frac{2(p-i+1)}{p(p+1)} \quad (6)$$

5. After selection of i^{th} answer, a new answer having normal distribution with average vector of $(w_1^{(i)}, w_2^{(i)}, k^{(i)})$ and variance of 1 would be produced.
6. The new answer would be substituted in the existing population.

The number of iterations could be adjusted based on observation of no improvement in sequential iterations or a predefined number. In this study, the number of algorithm iterations was adjusted on 1000 and this value was considered as a constant.

4.3.1 Solving the model and analysis of the proposed method results

In this section, the proposed combinational method was implemented on the existing dataset. In this regard, through application of a 10-fold cross-validation method, the available data were classified into two parts; training and test data. After 10 sessions of training and testing of the aforementioned model with different layers, the best values of indices were reported.

A new answer was provided for each iteration of the algorithm. Any of these produced answers could appear in the form of a three-component vector, according to Eq. (7). In fact, the first component of this vector is related to output weight of the support vector machine; the second component is associated with the weight of neural network and finally; the third component is related to discretizing boundaries. For any of these answers, the final output would appear according to the relation illustrated in Fig. 1 whose comparison with the real outputs would determine the performance indices (accuracy, sensitivity and transparency).

$$[w_1, w_2, k] \quad (7)$$

The results of proposed model implementation on the available data are reported in Table 3. In this table, the various numbers of hidden layers in the neural network and different types of kernel function in the support vector machine were combined together and the accuracy index was reported for the proposed method.

Table 3

Accuracy of solving the model for population of 10 and 3 neuron per layer.

Kernel function of support vector machine					Numbers of layers in neural network
Linear	Grade 2	Polynomial	Radial Basis Function		
95.71	96.83	96.51	96.97	2	
97.21	96.61	95.97	97.04	3	
97.71	96.57	95.73	97.12	4	
97.12	96.11	95.39	97.01	5	
97.14	96.14	95.42	96.79	6	
97.02	96.17	95.40	96.56	7	

According to the results obtained for accuracy, sensitivity and transparency, the highest rates of accuracy were related to the combination of linear SVM and 4-layered ANN (in which there were three (3) neurons per layer).

Since any variation in neuron numbers can affect the model's performance, the model was trained to accommodate a different number of neurons (2-5) per layer. According to trial Taguchi L16 perpendicular design,

the number of neuron in the hidden layers of neural network was determined to be 4, 1, 4, and 3, respectively.

4.3.2 Comparison of the proposed method with traditional methods

The results from traditional support vector machine and artificial neural networks were also tested and shown in Table 4.

Table 4

Performance indicators of traditional methods.

Accuracy	Sensitivity	Transparency	Method
92.57	87.80	95.75	Neural Network
94.85	93.49	95.59	Linear kernel
93.14	87.82	95.74	Grade 2 kernel
94.57	87.93	97.86	Polynomial kernel
94.57	98.27	92.73	RBF kernel

For comparing the traditional methods with the proposed one, the linear support vector machine with the highest accuracy was selected and compared with the best state of the proposed method which is related to combination of

linear SVM and 4-layerd neural network with 3 neurons per layer. For this purpose, any of traditional or proposed methods were run for 10 iterations whose results are listed in Table 5.

Table 5

Accuracy of traditional and proposed methods.

Iterations	Traditional	Proposed Method
1	96	98.23
2	96.85	96.85
3	95.71	96
4	96.85	96.85
5	95.71	97.14
6	96.12	97.42
7	96.85	97.14
8	96.57	96.57
9	96	96.28
10	95.71	96.85

In order to compare the two methods, Wilcoxon Signed Rank test was employed for comparing two groups of variables. For example, in the case of the data in Table 6, the first group of data was related to traditional method and the second group was for the proposed one. Through such test, their differences can be determined.

P-value for the mentioned test is calculated as 0.0068 implying acceptable superiority of the proposed method at the confidence level of 0.9932.

5. Conclusion

The current research presents a new combination based on support vector machine and artificial neural networks for determining the malignancy of breast cancer. In this content, an innovative algorithm was developed to combine the two mentioned algorithms by which the discretizing boundaries of answers and

Their weight can be determined. The proposed algorithm improved the answers through application of an initial population in sequential iterations base on neighbour search, and reported the best answer after a definite number of iterations. According to the numerical experiments, the proposed method has the maximum accuracy of 98.23, while the accuracy of the traditional method was 96.85. Moreover, the calculated p-value of 0.0077 reflects the superiority of the proposed method over the traditional one by confidence level of 0.9923.

6. Suggestions for further studies

- i. Application of polar coordinate for improving the performance of the classifier models.
- ii. Application of multiple error rates and cost-sensitive methods which enhance the sensitivity of the model toward the errors with higher costs and prevent their occurrence.
- iii. Application of a combination of clustering-classification approaches by the following methods: first, the data are assigned to some clusters and then the proposed method of this study will be implemented in each cluster, separately.

Reference

- Addeh, J., & Ebrahimzadeh, A. (2012). Breast cancer recognition using a novel hybrid intelligent method. *medical signals and sensors*, 2 (2), 95-102.
- Arevalo, J., González, F., Ramos-Pollán, R., Oliveira, J., & Lopez, M. (2016). Representation learning for mammography mass lesion classification with convolutional neural networks. *Computer methods and programs in biomedicine*, 127, 248-257.
- Chen, S. T., Hsiao, Y. H., Huang, Y. L., Kuo, S. J., Tseng, H. S., Wu, H. K., & Chen, D. R. (2009). Comparative Analysis of Logistic Regression, Support Vector Machine and Artificial Neural Network for the Differential Diagnosis of Benign and Malignant Solid Breast Tumors by the Use of Three-Dimensional Power Doppler Imaging. *Korean Journal of Radiology*, 10, 464-471.
- de Sampaio, W., Silva, A., de Paiva, A., & Gattass, M. (2015). Detection of masses in mammograms with adaption to breast density using genetic algorithm, phylogenetic trees, LBP and SVM. *Expert Systems with Applications*, 42 (22), 8911-8928.
- Han, J., Pei, J., & Kamber, M. (2011). *Data mining: concepts and techniques*.
- Hassanien, A., Moftah, H., Azar, A., & Shoman, M. (2014). MRI breast cancer diagnosis hybrid approach using adaptive ant-based segmentation and multilayer perceptron neural networks classifier. *Applied Soft Computing*, 14, 62-71.
- Jiao, Z., Gao, X., Wang, Y., & Li, J. (2016). A deep feature based framework for breast masses classification. *Neurocomputing*, 197, 221-231.
- Kalteh, A., Zarbakhsh, P., Jirabadi, M., & Addeh, J. (2013). A research about breast cancer detection using different neural networks and K-MICA algorithm. *Journal of cancer research and therapeutics*, 9 (3), 456-466.
- Korkmaz, S., & Poyraz, M. (2015). Least Square Support Vector Machine and Minimum Redundancy Maximum Relevance for Diagnosis of Breast Cancer from Breast Microscopic Images. *Social and Behavioral Sciences*, 174, 4026-4031.
- Lisboa, P., Bourdes, V., Bonevay, S., Defrance, R., Perol, D., & Chabaud, S. (2010). Comparison of artificial neural network with logistic regression as classification models for variable selection for prediction of breast cancer patient outcomes. *Artificial Neural Systems*, 1, 12-39.
- Liu, H., Zhang, R., Luan, F., Yao, X., Liu, M., Hu, Z., & Fan, B. (2003). Diagnosing breast cancer based on support vector machines. *Chemical Information and Computer Sciences*, 43 (3), 900-907.
- Majid, A., Ali, S., Iqbal, M., & Kausar, N. (2014). Prediction of human breast and colon cancers from imbalanced data using nearest neighbor and support vector machines. *Computer methods and programs in biomedicine*, 113 (3), 792-808.
- Marcano-Cedeño, A., Quintanilla-Domínguez, J., & Andina, D. (2011). WBCD breast cancer database classification applying artificial metaplasticity neural network. *Expert Systems with Applications*, 38 (8), 9573-9579.
- McCulloch, W., & Pitts, W. (1943). A Logical Calculus of Ideas Immanent in Nervous Activity. *Bulletin of Mathematical Biophysics*, 5 (4), 115-133.
- McLaren, C., Chen, W., Nie, K., & Su, M. (2009). Prediction of malignant breast lesions from MRI features: a comparison of artificial neural network and logistic regression techniques. *Academic radiology*, 16 (7), 842-851.
- Mishra, G., Ananth, V., Shelke, K., Sehgal, D., & Valadi, J. (2015). Hybrid ACO Chaos-Assisted Support Vector Machines for Classification of Medical Datasets. Paper presented at the InProceedings of Fourth International Conference on Soft Computing for Problem Solving .
- Naushada, S.M., Ramaiaha, J., Pavithrakumaria, M., Jayapriya, J., Hussainb, T., Alrokayanc, S.A., Gottumukkalad, S.R., Digumartie, R., & Kutalaf, V.K. (2016). Artificial neural network-based exploration of gene-nutrient interactions in folate and xenobiotic metabolic pathways that modulate susceptibility to breast cancer. *Gene*, 580 (2): p. 159-168.
- Polat, K., & Günes, S. (2007). Breast cancer diagnosis using least square support vector machine. *Digital Signal Processing*, 17, 694-701.
- Quinlan, J. R. (1996). Improved use of continuous attributes in C4.5. *Artificial Intelligence Research*, 4, 77-90.
- Rouhi, R., & Jafari, M. (2016). Classification of benign and malignant breast tumors based on hybrid level set segmentation. *Expert Systems with Applications*, 46, 45-59.
- Rouhi, R., Jafari, M., Kasaei, S., & Keshavarzian, P. (2015). Benign and malignant breast tumors classification based on

- region growing and CNN segmentation. *Expert Systems with Applications*, 42 (3), 990-1002.
- Salama, G., Abdelhalim, M. B., & Zeid, M. (2012). Breast Cancer diagnosis on three different datasets using multi-classifiers. *International Journal of Computer and Information Technology*, 1, 36-43.
- Sattlecker, M., Baker, R., Stone, N., & Bessant, C. (2011). Support vector machine ensembles for breast cancer type prediction from mid-FTIR micro-calcification spectra. *Chemometrics and Intelligent Laboratory Systems*, 107 (2), 363-370.
- Sivakami, K. (2015). Mining Big Data: Breast Cancer Prediction using DT-SVM Hybrid Model. *International Journal of Scientific Engineering and Applied Science (IJSEAS)*, 1 (5).
- Song, J. H., Venkatesh, S., Conant, E. A., Arger, P. H., & Sehgal, C. M. (2005). Comparative Analysis of Logistic Regression and Artificial Neural Network for Computer-Aided Diagnosis of Breast Masses. *Academic Radiology*, 12, 487-495.
- Stewart, B., & Wild, C. (2014). World cancer report Retrieved from World Health Organization
- Stoian, R., & Stoian, C. (2013). Modeling medical decision making by support vector machines, explaining by rules of evolutionary algorithms with feature selection. *Expert Systems with Applications*, 40 (7), 2677-2686.
- Sweilam, N., Tharwat, A., & Moniem, N. (2010). Support vector machine for diagnosis cancer disease: A comparative study. *Egyptian Informatics Journal*, 11 (2), 81-92.
- Tu, J. (1996). Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes. *clinical epidemiology*, 49 (11), 1225-1231.
- Utomo, C., Kardiana, A., & Yuliwulandari, R. (2014). Breast cancer diagnosis using artificial neural networks with extreme learning techniques. *Advanced Research in Artificial Intelligence*, 3 (7).
- Uzer, M., Inan, O., & Yilmaz, N. (2013). A hybrid breast cancer detection system via neural network and feature selection based on SBS, SFS and PCA. *Neural Computing and Applications*, 23 (3-4), 719-728.
- Vapnik, V. (1998). *Statistical Learning Theory*. Wiley, New York.
- Wang, D., Shi, L., & Heng, P. (2009). Automatic detection of breast cancers in mammograms using structured support vector machines. *Neurocomputing*, 72 (13), 3296-3302.
- Wang, P., Hu, X., Li, Y., Liu, Q., & Zhu, X. (2016). Automatic cell nuclei segmentation and classification of breast cancer histopathology images. *Signal Processing*, 31 (122), 1-3.
- Wisconsin Breast Cancer Dataset (WBDC) (2014). (Original). <[https://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+\(Original\)](https://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+(Original))>.
- Zheng, B., Yoon, S., & Lam, S. (2014). Breast cancer diagnosis based on feature extraction using a hybrid of K-means and support vector machine algorithms. *Expert Systems with Applications*, 41 (4), 1476-1482.