Breast Cancer Classification Algorithms: A Review

N. A. Abdulkadir.
Department of Computer Science
Federal University of Technology, Minna
Minna, Nigeria
abdulkadirnafisahabidemi@gmail.com

K.M.Muhammad.
Department of Computer Science
Federal University of Technology, Minna
Minna, Nigeria
muhammad_kudu@futminna.edu.ng

A.M. Aibinu.
Department of Mechatronics Engineering
Federal University of Technology, Minna
Minna, Nigeria
abiodun.aibinu@futminna.edu.ng

Abstract— Review of application of some selected artificial intelligence approach for classification of breast cancer is presented in this paper. The work firstly introduces cancer, a leading cause of death worldwide responsible for over 7.4 million deaths in recent time. Then, shortly followed by brief justification for breast cancer, one of the deadliest and most frequent form of cancer. Results obtained by the application of Support Vector machine(SVM), Relevance Vector Machine(RVM) and Artificial Neural Network(ANN) in correctly classifying breast cancer upon application of the approach on one of the popularly known dataset thus ends the contribution in this paper.

Keywords—Artificial Neural Network, Breast Cancer, Classification, Machine Learning, Relevance Vector Machine, Support vector machine

I. INTRODUCTION

A leading cause of death around the world is Cancer with about 7.4 million cases recorded in 2014 [1]. Cancer refers to the uncontrolled multiplication of a group of cells in a certain location of the body. A group of rapidly dividing cells may form a lump, micro calcifications or architectural distortions which are usually referred to as tumors [2]. The main types of cancer include but not limited to: cancer of the lungs, cancer of the stomach, cancer of the colorectal, cancer of the liver and cancer of the breast which accounted for about 1.3million death, 803,000 deaths, 639,000 deaths, 610,000 deaths and 519,000 deaths per year respectively[1].

Health factors that increases the rate of cancer includes; inappropriate use of alcohol and tobacco, insufficient consumption of fruits and vegetables, excessive weight, environmental risks and even risks form occupation amongst others[1].The treatment of cancer involves series of interventions such as: chemotherapy, radiotherapy and even surgery. This treatment could be aimed at prolonging the lives of the victims or even curing this disease[1].

Cancer of the Breast in human occurs both in men and women. This do occur in form of: lumps in the breast, discharge from the nipples (blood), and sometimes a change in the shape and texture of the breast and nipple [3]. Cancer of the breast can be classified in various ways and the most important of these classifications are binary classification called benign and malignant breast cancer.

One of the areas of applications of machine learning is in supervised or/and unsupervised classification. It consist of splitting the data into non-overlapping segments. Thus, classification is way of looking for a different models that helps to define and separate class labels of the data object [4]. Classification can be done with the use of some machine learning algorithms which includes Decision Tree, Support Vector Machine, Boosting, Neural Network, Nearest Neighbor Algorithms, Naïve Bayes, and Random Forests amongst others [4].

An algorithm which can be used to properly classify breast cancer is the Relevance vector machine (RVM). RVM makes use of Bayesian inference to acquire minimal solutions and it is mainly used for regression and classification based on probability. This innovation is often used for detection and classification. It is based on a sparse kernel model whose functional form is quite identical to Support Vector Machine but the difference is aimed at the probabilistic classification approach [5].The decision function is computed with the use of only few of the training examples (relevance vectors) and this makes it analogous to SVM [6]. It is reported that in several benchmark studies RVM can yield nearly identical performance to, if not better than, that of Support Vector Machine (SVM) while using far fewer relevance vectors than the number of support vectors for SVM. In comparison to SVM, the regularization parameter does not require tuning during the training phase of RVM [7].

This paper is aimed at reviewing the application of some related artificial intelligence approaches for classifying breast cancer and also to show some performance metrics adopted.

The continuing sections in this paper is structured as follows: motivation for breast cancer classification is provided in the second section and detailed review of various applications of RVM to breast cancer classification is provided in the third section, the fourth section shows different performance metrics adopted and results obtained in literature and the paper ends with conclusion in the last section.

II. THE NEED FOR BREAST CANCER CLASSIFICATION

In a survey carried out by World Health Organization (WHO), the most common cancer in women all around the world is Breast Cancer and in 2012; about 1.7 million new cases were diagnosed with Belgium having the highest rate, and Denmark as the second and then the Netherlands as the third [8]. Events of life which includes early menarche (before the age of 12), late natural menopause (after the age of 55)and not bearing children (pregnancy)after the age of 30 are vital risk influences for cancer of the breast. They intensify lifetime exposure to estrogens and progesterone and the threat of breast cancer [9].In developed and developing countries, a more prevailing cancer in women is Breast Cancer [10]. This cancer continues to prevail in developing countries due to the fact that there is an increase in life expectancy, increased development and western life adoption [11]. Although, these threats could be reduced with proper prevention, this strategy cannot be achieved in countries with low and middle income where the diagnosis of this cancer occurs in very late stages. The cornerstone of breast cancer control lies in early detection so as to improve the outcome and survival rate [12]. It is also of utmost importance to comprehend that breast lumps are either benign (non-cancerous) or malignant (cancerous/harmful).Proper classification of this cancer can help to achieve an appropriate treatment. One of the methods of early detection is the use of SVM classification technique which suffers from; Lack of an output that is probabilistic, the need to approximate a tradeoff parameter, the necessity to apply a kernel function, hence the need for another approach [5]. Other machine learning approaches have been used in literature; however the use of RVM has not been widely reported.

III. REVIEW OF APPLICATION OF RVM, SVM & ANN IN BREAST CANCER

In this section, review of some related paper will be presented. Thus showing application of RVM, SVM and ANN in breast cancer detection.

Reference [13] presented the detection of breast cancer with the use of SVM and Artificial Neural Network (ANN) classification approach. A feed forward ANN with 9, 121 and 2 neurons for the input, hidden and outer layers respectively was implemented. Several kernels in SVM which includes Basis function, Polynomial function, linear kernel, multilayer Perceptron and quadratic kernel were tested for accuracy. It was observed that SVM provided a high advantage over other classifiers and they are free of optimization issues of neural network as it presents a convex programming problem. Making use of only the relevant vectors and thus much faster to evaluate than density estimators. It was observed that ANN makes use of all the data points irrespective of its importance to the decision boundary.

The detection of breast cancer with the use of RVM and other machine learning approaches was presented by [14]. RVM was applied exclusively on the Wisconsin breast cancer dataset to detect breast cancer. The features were reduced with the use of LDA which is a feature selection algorithm. A GUI was designed to take in user's inputs and provide a result. RVM reduces the computational cost which makes its accuracy far better than other algorithms.

Reference [15] carried out a comparative analysis of SVM and Neural Network (NN) for classifying breast cancer. A multiclass SVM, which makes use of Crammer and Singer method and also a Neural Network which consist of: Trainbfg, Traincgb, Traingdx and Trainlm were used for the analysis on a Wisconsin dataset. MATLAB Neural Network Toolbox command newff was used to generate the MLPN neural network and sim command was used to check how well the MLP netl estimates the data. SVM is based on mathematics and it evaluates into a simple way and a very powerful algorithm and it solves a simple convex optimization problem.

Reference[16] proposed the use of a mix of K-means and support vector machine in the diagnosis of breast cancer which is based on feature extraction. A dataset from Wisconsin diagnostic breast cancer dataset was used; the dataset contains 32 features in 10 categories for the cell nucleus and three indicators were measured which includes mean value, standard error and maximum value and they are treated as different features in the dataset. Abstract malignant and benign tumor pattern were extracted separately before training the original data to obtain the classifier and the patterns are recognized using feature extraction. The extracted patterns were reconstructed as the new abstract features for the training phase and SVM was used to obtain a precise classification. This technique saves a lot of time during the training phase and feature selection algorithm helps reduces feature dimension and it eliminates noisy information for prediction. It was observed that feature selection has some issues for small scale and sparse data in a high dimensional space.

Reference [17] presented a genetically optimized Neural Network model for the diagnosis of Breast Cancer. A Wisconsin breast cancer dataset from UCI is used for classifying breast cancer. Out of this dataset (699 instances), 16 of them were found to have some missing attributes and were discarded thus only 683 instances were used for the classification. The ANN architecture was evolved using Genetic Programming for classifying the dataset and this was achieved by forming a GONN architecture in such a way to treat it like an ANN structure. The GONN structure was built following the steps of GP life cycle with a modification of the operators of mutation and crossover. The result of the GONN architecture represents an ANN with appropriate mapping of GP parameters to ANN learning parameters. The classification accuracy was done using a 10-fold cross validation technique. The introduction of a new mutation and crossover operator helps to counter he destructive nature of these operators.

Reference[18] presented an automated classification of breast cancer with the use of Deep Belief Networks. MATLAB 2014a and Palm DBN were used for the implementation. The Wisconsin dataset was trained with DBN and the weight matrix was transferred to native back propagation neural network with a similar architecture. The instances were reduced from 699 to 683 to enable result comparison with other techniques in literature. The pretrained back propagation neural network with unsupervised phase DBN achieves higher classification accuracy in comparison to a classifier with just one supervised phase. One of the weaknesses showed that Building a CAD scheme which is based on DBN using commercial hardware to assist medical professionals in the early detection of Breast abnormality wasn't an easy task.

An application of Deep Neural Networks for classifying Breast Cancer was proposed by [19]. The Wisconsin breast cancer dataset is first pre-processed to extract noise form the data and four best features selected with the application of logistic regression model. RFE is used for the iteration process and also to rank the features for the next level and they are chosen based on the outcome of the result from logistic regression and then classification is carried out with the use of Deep Neural Network. This method has a complex architecture and it has a lot of hidden layers with more neurons in each layer and its input can be propagated and pass through multiple layers. Therefore in every epoch the error rate of the system is gradually reduced by adjusting the weights of nodes and fine tuning the network values of nodes in every layer. This algorithm takes a longer time to train the data because its architecture has a complex structure.

Reference[20] presented an application of machine learning algorithms on the Diagnostic Dataset from Wisconsin to detect Breast Cancer. A comparison of six (6) machine learning algorithms which includes, GRU-SVM (a combination of Recurrent Neural Network, the Gated recurrent unit and SVM), Linear regression, Multilayer Perceptron, Nearest Neighbor search, Softmax Regression, and SVM were used to classify breast cancer and the machine learning algorithms implementation was done using Google Tensor Flow and other scientific computing libraries which includes MATPLOTLIB (a Mat lab function), numpy and scikit-learn. Wisconsin Diagnostic Breast Cancer dataset was used to train the machine learning algorithms to detect breast cancer. Scikit-learn was used to standardize the data. It was observed that most optimal hyper-parameter for the machine learning algorithms was not determined, the same hyper-parameter was used for all.

REVIEW OF OTHER MACHINE LEARNING APPROACHES FOR BREAST CANCER CLASSIFIER

The Use of Bayesian rough set classifier for breast cancer detection was proposed by [21]. Data preprocessing and machine learning tasks were applied on the dataset. At the pre-processing stage, data cleaning was first applied on the dataset to improve the accuracy of the analysis and the classification. Then, relevant features were selected with the use of CFS which was modified with the use of DI and BFS. Bayesian rough set was used in building the model for classification. One of the strengths recorded in this work is the use of Data cleaning and feature extraction as it helped to reduce the time complexity and increase its accuracy.

Reference [22] proposed the use of a fuzzy logic approach to develop knowledge based system for classifying breast cancer. After pre-processing the data, data clustering was carried out with the use of Expectation maximization and then, a reduction technique based on dimension called Principal Component Analysis (PCA) was used to address multi-collinearity issues and to extract the most important data from the datasets. Fuzzy logic was used for the classification and the fuzzy rules were generated automatically using Classification and Regression trees (CART) and then prediction models were generated by fuzzy rule-based method in each cluster. The proposed technique was evaluated on two datasets from UCI repository which includes the WBCD and Mammographic mass. The combination of all these techniques helped to create a hybrid technique for accurate classification of breast cancer. In addition to obtaining an unbiased result, there was an application of 10-fold cross validation in the process of clustering. The clustering technique that was used helped the classifier to learn the prediction models from the data better. The method used in this approach is a non-incremental data mining technique.

A collective approach of RS and Extreme Learning Machine (ELM) to develop a new intelligent classifier for diagnosing breast cancer was presented by [23]. Relevant features were extracted with the use of RS approach. ELM, a feed forward neural network with a single hidden layer classification model classified the reduced features obtained by the RS. The hybrid model testing was done with the use of the Wisconsin dataset obtained from UCI repository. The optimal attributes which was selected with the use of RS before classifying the cancer increased the classification success rate. The process of learning takes more time and the error can be focused on a local point.

Reference[24] proposed the use of a probabilistic modeling approach to diagnose breast cancer by measurement.BN model was constructed with the use of a learning algorithm called K2 and computation methods based on statistics. The system testing was done with the use of datasets obtained from a clinical ultrasound dataset gotten from a Chinese local hospital and a fine needle aspiration cytology (FNAC) dataset from UCI repository. The structure of a BN is based on data which is discretized and so, the features of the data from the Chinese hospital were discretized based on their conforming meanings and the discretization of FNAC resulted into benign or malignant. Then, an algorithm called Information Gain was applied to rank each feature that proved to be relevant to the result being diagnosed. The learning algorithm K2 was used to learn the structure of the BN for the Dataset that was given with the IGs-ordered features and 10-fold cross validation was used to validate. The implementation was done on a PC with the use of Weka and GeNle software.

An innovative approach which makes use of data mining techniques for detection of breast cancer was proposed by[25]. A toolkit called WEKA was used on the Wisconsin dataset with three different data mining classification techniques. Missing values were removed from the dataset and so only 638 instances were used for classifying. The experiments were carried out using libraries from WEKA machine learning environment and 10-cross validation was done for all the classifiers. SMO shows a concrete result with breast cancer disease of patient's records and it helps to get a better result with accuracy, low error rate and performance.

A computerized approach of classifying breast cancer with the use of Gaussian Naïve Bayes was proposed by [26]. The Wisconsin Original dataset was loaded, and the missing values were replaced using the median model. The dataset was partitioned into two equal parts which was used for training and testing. Relevant features were selected with the use of LDA. The selected features were calculated using Naïve Bayes classifier and the result was stored in a variable. The data is then classified with the use of the stored result and pre-calculation rule and if these two conditions are satisfied, the data is classified as either benign or malignant.

In this section, other techniques have been discussed which includes Fuzzy Logic, Naïve Bayes, Data mining approaches, Probabilistic modelling, Extreme learning approach and they also made significant contributions.

IV. PERFORMNCE EVALUATION

Performance metrics used in evaluating various reported approaches is presented in this section.

This gives a brief explanation of the performance metrics and states its related equations. Also, papers that made use of these metrics will also be presented and the details are as contained in Table I.

ACCURACY: Accuracy is the degree of right prediction of a model, it is the simplest measure to assess a classifier [27].In problems of classification, Accuracy is a measure of the number of accurate forecasts that the tested model makes over other classes of predictions made. Accuracy is best used when there is a close balance between the target variable classes in the data. When the target variable class in the data are a majority of a single class, Accuracy should not be used as a measure [28].

$$\text{Accuracy}=\frac{TP+TN}{TP+FN+FP+TN} \quad (1)$$

SENSITIVITY OR RECALL: This is a measure that shows the proportion of accurate forecast of the algorithm of a certain disorder in relation to the actual result [29]. TP and TN shows the people with the disorder while TP shows the people the system diagnoses as having that disorder [28]. It measures the fraction of positive patterns that are classified correctly [30].

$$\text{Sensitivity}=\frac{TP}{TP+FN} \quad (2)$$

SPECIFICITY: This is the reverse of sensitivity, it tells us in what proportion the model predicts that a disorder does not exist in relation to the actual result. FP and TN shows the people without this disorder while TN shows the result of the model as not having the disorder [28]. It measures the fraction of negative patterns that are classified correctly [30].

$$\text{Specificity}=\frac{TN}{TN+FP} \quad (3)$$

True Positives (TP): True positives are instances where the predicted and the actual data point's classes are True (1).

True Negatives (TN): True negatives are instances where the predicted and the actual data point's classes are False (0).

False Positives (FP): False positives are instances where the predicted data point class is True (0) while the actual data point class is False (1).

False Negatives (FN): False negatives are instances where the predicted data point class is False(1) while the actual data point class is True(0) [28].

V. CONCLUSION

The plethora of the works presented herewith shows different applications of classification algorithm for breast cancer diagnosis. In each case, the process of application, strength and weakness of the classification algorithm has been succinctly reviewed and presented. Also, from Table I, it was observed that the equations used for calculating the accuracy, sensitivity and specificity is same in all the reviewed papers and the only thing that seems to differentiate them is the way they have been expressed.

REFERENCE

[1] WHO. (2010). Cancers. NMH Fact Sheet, 1–2. https://doi.org/10.1056/NEJMra0804588

[2] K.Kipli, A. Z.Kouzani &I. R. A. Hamid (2013). Investigating Machine Learning Techniques for Detection of Depression Using Structural MRI Volumetric Features. International Journal of Bioscience, Biochemistry and Bioinformatics, 3(5), 444–448. https://doi.org/10.7763/IJBBB.2013.V3.252

[3] S.Şahan,K. Polat, H. Kodaz, &S. Güneş (2007). A new hybrid method based on fuzzy-artificial immune system and k-nn algorithm for breast cancer diagnosis. Computers in Biology and Medicine, 37(3), 415–423. https://doi.org/10.1016/j.compbiomed.2006.05.003

[4] A.Chaudhary, S. Kolhe, &R. Kamal (2013). Machine Learning Classification Techniques : A Comparative Study. International Journal of Advance Computer Theory Engineeringnal of Advance Computer Theory Engineering, 2(4), 2319–2526.

[5] M. E. Tipping (2000). The Relevance Vector Machine. Advances in Neural Information Processing Systems 12, (x), 652–658. https://doi.org/10.1.1.34.4986

[6] R. Zhang, G. B. Huang, N. Sundararajan, &P. Saratchandran, (2007). Multi-category classification using an Extreme Learning Machine for microarray gene expression cancer diagnosis. IEEE/ACM Trans Comput Biol Bioinfrm, 4(3), 485–495. https://doi.org/10.1109/tcbb.2007.1012

[7] L. Wei, Y. Yang, R. M. Nishikawa, &Y. Jiang (2005). A study on several machine-learning methods for

classification of malignant and benign clustered microcalcifications. IEEE Transactions on Medical Imaging, 24(2), 371–380. https://doi.org/10.1109/TMI.2004.842457

[8] A.Balekouzou, P. Yin, C. M. Pamatika, C. E. Bekolo, S. W. Nambei,M. Djeintote, B. Koffi, (2017). Reproductive risk factors associated with breast cancer in women in Bangui: A case-control study. BMC Women's Health, 17(1), 1–9. https://doi.org/10.1186/s12905-017-0368-0

[9] G. V. Dall&K. L. Britt (2017). Estrogen Effects on the Mammary Gland in Early and Late Life and Breast Cancer Risk. Frontiers in Oncology, 7(May), 1–10. https://doi.org/10.3389/fonc.2017.00110

[10] M. M. Rivera-Franco &E. Leon-Rodriguez (2018). Delays in breast cancer detection and treatment in developing countries. Breast Cancer: Basic and Clinical Research, 12. https://doi.org/10.1177/1178223417752677

[11] A. Suleiman (2014). Awareness and attitudes regarding breast cancer and breast self-examination among female Jordanian students. Journal of Basic and Clinical Pharmacy, 5(3), 74. https://doi.org/10.4103/0976-0105.139730

[12] C. Cruwys &J. Pushkin (2017). Breast density and impacts on health, 9–12.

[13] P. D. Shahare, &R. N. Giri (2015). Comparative analysis of artificial neural network and support vector machine classification for breast cancer detection. International Research Journal of Engineering and Technology, 2(9), 2114-2119.

[14] B. M. Gayathri&C. P. Sumathi (2016, December). Comparative study of relevance vector machine with various machine learning techniques used for detecting breast cancer. In Computational Intelligence and Computing Research (ICCIC), 2016 IEEE International Conference on (pp. 1-5). IEEE.

[15] E. Edriss, E. Ali, &W. Z. Feng (2016). Breast Cancer Classification using Support Vector Machine and Neural Network, 5(3), 1–6.

[16] B. Zheng, S. W. Yoon, &S. S. Lam (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.

[17] A. Bhardwaj&A. Tiwari (2015). Breast cancer diagnosis using genetically optimized neural network model. Expert Systems with Applications, 42(10), 4611-4620.

[18] A. M. Abdel-Zaher&A. M. Eldeib (2016). Breast cancer classification using deep belief networks. Expert Systems with Applications, 46, 139-144.

[19] S. Karthik,R. S. Perumal, &P. C. Mouli (2018). Breast Cancer Classification Using Deep Neural Networks. In Knowledge Computing and Its Applications (pp. 227-241). Springer, Singapore.

[20] A. F. Agarap (2017) On Breast Cancer Detection: An Application of Machine Learning Algorithms on the Wisconsin Diagnostic Dataset. arXiv preprint arXiv:1711.07831.

[21] A. R.Abbas,&M. A. Shihab (2017). Diagnosis the Breast Cancer using Bayesian Rough Set Classifier.

[22] M. Nilashi, O. Ibrahim, H. Ahmadi, &L. Shahmoradi (2017). A Knowledge- Knowledge - Based System for Breast Cancer Classification Using Fuzzy F uzzy Logic Method. Telematics and Informatics. https://doi.org/10.1016/j.tele.2017.01.007

[23] Y. Kaya (2013). A new intelligent classifier for breast cancer diagnosis based on a rough set and extreme learning machine: RS+ ELM. Turkish Journal of Electrical Engineering & Computer Sciences, 21(Sup. 1), 2079-2091.

[24] S. Liu, J. Zeng, H. Gong, H. Yang, J. Zhai, Y. Cao, &X. Ding (2018). Quantitative analysis of breast cancer diagnosis using a probabilistic modelling approach. Computers in biology and medicine, 92, 168-175.

[25] V. Chaurasia&S. Pal (2017). A novel approach for breast cancer detection using data mining techniques.

[26] B. M. Gayathri &C. P. Sumathi (2016). An Automated Technique using Gaussian Naïve Bayes Classifier to Classify Breast Cancer. International Journal of Computer Applications, 148(6).

[27] C. Ferri, J. Hernández-Orallo, &R. Modroiu (2008). An experimental comparison of performance measures for classification. Pattern Recognition Letters, 30, 27–38. https://doi.org/10.1016/j.patrec.2008.08.010

[28] M. Sunasra(2018). Performance Metrics for Classi cation problems inMachine Learning1. Confusion Matrix : Terms associated with Confusion matrix :,1–12.

[29] A. Godil, R. Bostelman, W. Shackleford, T. Hong&M. Shneier (2014). Performance Metrics for Evaluating Object and Human Detection and Tracking Systems. https://doi.org/10.6028/NIST.IR.7972

[30] M. Hossin, &M. N. Sulaiman (2015). A Review On Evaluation Metrics For Data Classification Evaluations. International Journal of Data Mining & Knowledge Management Process (IJDKP), 5(2). https://doi.org/10.5121/ijdkp.2015.5201

[31] M. Karabatak (2015). A new classifier for breast cancer detection based on Naïve Bayesian. MEASUREMENT, 72, 32–36. https://doi.org/10.1016/j.measurement.2015.04.028

[32] A. Taher& A. Shaimaa (2013). Performance analysis of support vector machines classifiers in breast cancer mammography recognition. https://doi.org/10.1007/s00521-012-1324-4

[33] W. Xie, Y. Li, &Y. Ma (2015). Author ' s Accepted Manuscript Breast mass classification in digital mammography based on Extreme learning machine Reference : To appear in : Neurocomputing. Neurocomputing. https://doi.org/10.1016/j.neucom.2015.08.048

[34] R. K. Tripathy (2013). An investigation of the breast cancer classification using various machine learning techniques (Doctoral dissertation).

[35] D. T. Sarabai (2016). Efficient Breast Cancer Classification Using Improved Fuzzy Cognitive Maps with Csonn, 11(4), 2478–2485.

TABLE I

1. SENSITIVITY

Paper title	**Reference**	**Equation**
A new classifier for breast cancer detection based on Naïve Bayesian	[31]	TP/(TP + FN)
Breast Cancer Diagnosis Using Genetically Optimized Neural Network Model	[17]	(TP/(TP + FN)) * 100
Performance analysis of support vector machines classifiers in breast cancer mammography recognition	[32]	(TP/ (TP + FN)) * 100%
Breast Cancer Classification Using Deep Belief Networks	[18]	100 * (TP/(TP + FN))
An Automated Technique using Gaussian Naïve Bayes Classifier to Classify Breast Cancer	[26]	TP/(TP + FN)
A Novel Approach for Breast Cancer Detection using Data Mining Techniques	[25]	TP/(TP + FN)
Breast mass classification in digital mammography based on Extreme learning machine	[33]	TP/(TP + FN)
An Investigation Of The Breast Cancer Classification Using Various Machine Learning Techniques	[34]	TP/(TP + FN)
Efficient Breast Cancer Classification Using Improved Fuzzy Cognitive Maps with Csonn	[35]	TP/(TP + FN)

2. SPECIFICITY

Paper title	**Reference**	**Equation**
A new classifier for breast cancer detection based on Naïve Bayesian	[31]	TN/(TN + FP)
Breast Cancer Diagnosis Using Genetically Optimized Neural Network Model	[17]	(TN/(TN + FP)) * 100
Performance analysis of support vector machines classifiers in breast cancer mammography recognition	[32]	(TN/(FP + TN)) * 100%
Breast Cancer Classification Using Deep Belief Networks	[18]	100 * (TN/(TN + FP))
A Novel Approach for Breast Cancer Detection using Data Mining Techniques	[25]	TN/(TN + FP)
An Investigation Of The Breast Cancer Classification Using Various Machine Learning	[34]	TN/(TN + FP)

Techniques		
Efficient Breast Cancer Classification Using Improved Fuzzy Cognitive Maps with Csonn	[35]	TN/(TN + FP)

3. ACCURACY

Paper title	**Reference**	**Equation**
A new classifier for breast cancer detection based on Naïve Bayesian	[31]	(TP + TN)/(TP + TN + FP +FN)
Breast Cancer Diagnosis Using Genetically Optimized Neural Network Model	[17]	((TP + TN)/(TP + TN + FP +FN)) * 100
Performance analysis of support vector machines classifiers in breast cancer mammography recognition	[32]	((TP + TN)/(TP + FP + FN +TN)) * 100%
An Automated Technique using Gaussian Naïve Bayes Classifier to Classify Breast Cancer	[26]	(TP + TN)/(TP + FP + FN +TN)
A Novel Approach for Breast Cancer Detection using Data Mining Techniques	[25]	(TP + TN)/(TP + FP + TN + FN)
Breast cancer diagnosis based on feature extraction using a hybrid of K-means and support vector machine algorithms	[16]	((TP + TN)/(TP + TN + FP +FN))
Breast mass classification in digital mammography based on Extreme learning machine	[33]	((TP + TN)/(TP + TN + FP +FN))
Breast Cancer Classification using Support Vector Machine and Neural Network	[15]	((TP + TN)/(TP + TN + FP +FN))
An Investigation Of The Breast Cancer Classification Using Various Machine Learning Techniques	[34]	((TP + TN)/(TP + TN + FP +FN))