Quest Journals
Journal of Software Engineering and Simulation
Volume 9 ~ Issue 2 (2023) pp: 09-22
ISSN(Online) :2321-3795 ISSN (Print):2321-3809
www.questjournals.org

Research Paper

Decision Support System For Breast Cancer Detection Based on AI Using Mammography

Biswaranjan Senapati [1],Awad Bin Naeem[2]
[1]*Department of Computer Science, National College of Business Administration &Economics, Multan, Pakistan.*
[2]*Department of Computer Science (CTU), Univeristy of Arkansas, ARKANSAS, USA*

Abstract:*In modern era, machine learning has become pivotal for the diagnoses of different cancers. It has proven its effectiveness with the highest classification accuracy for breast cancer detection and diagnosis. In this work, we used five different classifiers for breast cancer which includes: Artificial Neural Network (ANN), Random Forest (RF), K-Nearest Neighbor (KNN), Support Vector Machine (SVM) and Logistic Regression (LR) that are implemented on Wisconsin Diagnostic Breast Cancer (WDBC) dataset. The goal of this comparative study was to identify the most reliable machine learning algorithms with optimized dataset that could serve as a best diagnostic tool for breast cancer. Feature selection plays a significant role for diagnostic tool. For the data analysis heat map is used for the selection of features that have less similarity and more uniqueness. The proposed hybrid approach use best of the classifiers for the reduction of features in Support vector machine (SVM) because SVM can be used for superior accuracy with minimum features and less computation time. Comparative analysis of these algorithms is done on the basis of Recall, Precision, F1_score, Error rate, and Specificity. In broader perspective, this can be useful enough for medical practitioners to make right decisions.*

Keywords:*Decision Support System,CNN, Mammography Images, Feature Extraction, Breast Cancer*

Received 05 Feb., 2023; Revised 13 Feb., 2023; Accepted 15 Feb., 2023 © The author(s) 2023. Published with open access at www.questjournals.org

I. Introduction

The breast cancer death rate is relatively high, Over 1.5 million women worldwide are affected by breast cancer each year, according to the World Health Organization. 570,000 women died in 2015 from breast cancer, which is around 15% of all cancer-related deaths in women. In America, there are estimated to be 40,610 breast cancer deaths in 2017 and 252,710 newly diagnosed cases. In Asia, Indian and Pakistan has one of the highest rates of breast cancer, with 90 000 cases recorded annually and a death rate of over 40,000. By using proper screaming and diagnosis techniques early on, before serious physical signs started to appear on the body, the death rate from breast cancer can be decreased. Breast cancer detection methods have included using ANN, Support vector machines (SVM), and other methods(Abbas et al., 2019). The most popular and highly successful method for the early identification of breast cancer is mammography. Even extremely slight changes in the body are picked up by it. Mammograms are examined by medical professionals, who may advise a biopsy if anomalies are discovered. Breast cancer is typically found through a biopsy, which is an expensive, time-consuming, and unpleasant process. At this point, the radiologist's advice is crucial; otherwise, the patient would have to undergo an unneeded biopsy. By automating this analysis, the radiologist can increase the accuracy of his diagnoses; such a system can act as a second reader. The classification of mammograms into their appropriate classes, such as benign (not damaging to the body and does not spread to other parts of the body) or malignant, is made possible by a breast cancer detection system based on AI employing mammography pictures (cell spreads to other part of the body and cause to death). For this reason, machine learning approaches may replace some difficult manual tasks performed by doctors, such as text and speech processing used to identify emotions in healthcare workers and their reactions. According to research, a patient's emotions can best be used to gauge their health and provide better results in the future(Azavedo et al., 2012).

Seems to be the communal cancer in women is breast cancer (BC). Breast cancer has strong prevalence and death rate. According to current cancer figure new cases register are 25% and death rate is 15%(Bargalló et al., 2014). As per international health organization this is 8.8 million of death worldwide per year triggered by this disease. Cancer is responsible of many deaths Americans society has reported that cancer rate is increase about

39% from 1989 to 2015. Breast cancer is higher number of cancer deaths in Malaysia also. Approximately women in Malaysia are at the risk of 5% and records show that united states are at the risk of 12.5 percent.Breast cancer is top cancerous growth in Indian and Pakistan regions. Based on current report the main three malignancies in both age rangesof all sexes mixed were breast cancer, leukemia of lip and oral cavity cancer. In 2017 40,000 people died of breast cancer and 90.000 additional cases were reported in same year. Genetic inclusion, life style and environmental influence is most common reason of breast cancer which may affect about 8 to 9 women in India and Pakistan(Broeders et al., 2012).

Breast cancer is still taking the highest death toll even in advance countries. An efficient algorithm or combinations of algorithms have to be found to detect breast cancer, with higher accuracy and precision. Another problem is to keep the noise low while finding sufficient features and a sufficient sized dataset. Making dataset extensive may cause learning efficient and adding more to experience, same happens with adding more features but this may bring lot of computational overheads and issues related to accuracy. Another required characteristic in the propose technique is a certain level of accuracy, precision, recall and reliability, so the proposed algorithm/technique can be deployed successfully at lab scale. Women have a very high probability for being infected by breast cancer but mostly undetected. The origin of this condition relies on several variables and cannot be easily determined. In comparison the diagnoses process that decide whether the cancer is benign or malignant involve an exceptional amount of work on the behalf of doctors and physicians and number of evaluations are included such as size of cell and its uniformity, clump thickness may often create an issue for the doctors. In addition, this has led to an increase in the usage of machine learning classifiers and comparing it as study tool. Advanced recovery rooms often rely on eligible devices on the other side using other aspect of artificial intelligence to call breast cancer is not unacceptable it has been confirmed that breast cancer is prevalent cancer that is affecting the girls and women . The main goals of doing this research are:

- Finding an adequate set of features for the selected classifiers to improve the detection of breast cancer.
- Finding a suitable dataset which can be used for accurate cancer detection with less computational overheads.
- Computing recall, Fl score precision and accuracy for reliable detection process using reduced features while having sufficient level values required for the breast cancer detection.
- Analysis of classifiers (as in objective 3) with reduced features for quality objectives.

This work is organized as follows: Section 2 contains the literature review. Section 3 presents the methodology of the study. Section 4 discusses the results. In last, the conclusion and future work is described in section 5.

II. Literature Review

Screening used for breast cancer symptoms detection in women's, different techniques and tools used for breast cancer detection such as Mammography and Medical Breast.Test, this analysis is done by health professionals(Bunkhumpornpat et al., 2009). In Data Science, different methods used for the prediction of various cancer and is the best approach for detection (Casal-Guisande et al., 2022). Machine learning is consistent with tools utilized for the design and valuation of algorithms that simplify prediction, pattern recognition, and classification. Machine learning contains four stages(Cerqueiro-Pequeño et al., 2021). Collections of records choose the model, training the model, and testing the model. Machine learning methods used broadly for classification of breast cancer and predictions and Machine learning approaches are used for detection and diagnosing. In this survey, use three ML techniques Decision Tree, Random Forest, and SVM. For breast cancer ordering(Chawla et al., 2002). Use Wisconsin. breast cancer datasets, and use this approach to detect cancer and get the best accuracy(Smith et al., 2010). These classifiers also used to detect and examine breast cancer, so that the syndrome can be identified at a premature stage(Weber et al., 2016). The initial detection of symptoms of breast cancer is noble for health specialists and physicians to detect cancer at the initial stage and is helpful for patient's recovery. Uses, Wisconsin datasets, which is, diagnose data and to present evolving actual machine learning techniques for malignance classification for malignance classification using three classifiers in a Wisconsin diagnose data set(Comesaña-Campos et al., 2020). In actually all of this effort done to automate the predictions of tumors with acceptable accuracy. In recent decade's breast cancer is the most frequently selected topic(van Bommel et al., 2017). In data mining using machine learning is also detected but in Image processing using cellular automata predicts cancer with the help of CA methods and some image processing techniques such as noise detection and removal and spot detection on images of cancer. Same as in data mining breast cancer is projected and categorized with the aid of machine learning algorithms (Cunningham & Delany, 2020).

A brand-new strategy for detecting breast cancer that made use of the multi-layer perceptron, probabilistic neural networks, and radial basis function as classification algorithms (Yue et al., 2018). This system calculates accuracy and productivity. For training and testing, MLP achieved accuracy of 97.80% and 97.66%, respectively(T. G. Debelee et al., 2020). The k-nearest neighbors (Taye Girma Debelee et al., 2020)algorithm is one of the most used machine learning algorithms. With the use of CAD, mammography techniques can be used to detect breast cancer and do so at an early stage. (Shler Farhad & Adnan Mohsin, 2021). More and more doctors are using machine learning to aid in cancer diagnosis as a result of advances in computer power and the growth of associated big data in medicine in recent years.(Ogunleye et al., 2022). A mixture of statistics, probability, and optimization techniques are used by machine learning algorithms to enable computers to "learn" from the input and recognise patterns that are challenging for humans to perceive from enormous volumes of unstructured data. Machine learning can be used to assess cancer prognosis and recurrence in addition to cancer diagnosis. (Fenton et al., 2007).

Predictive analytical techniques and artificial intelligence advancements in recent years have made it possible to apply quantitative testing, quantitative tests, and machine-learning procedures that use automatic processes for unbiased evaluations. In this paper, we propose a classifier for cancer illness prognosis and diagnosis that uses Support Vector Machines (SVMs) rather than Bayesian classifiers and Artificial Neural Networks. (Rimmer, 2017). The details of the breast cancer death application are laid forth in the study, along with the corresponding results for all the classifiers considered. With regard to both the prognosis and the diagnostic issue, a number of comparative application along with studies were carried out, proving the superiority of the proposed SVM algorithm in terms of sensitivity, accuracy, and precision(Fenton et al., 2013). The second leading cause of death for women worldwide is cancer. Early detection and screening are key to breast cancer prevention since they will increase clinical efficacy, preserve lives, and reduce costs because the disease's cause is unknown(Samulski et al., 2010). Ultrasound imaging 15 one of the medical devices most used to identify and diagnose breast anomalies. Computer-aided diagnosis (CAD) program is a powerful and useful tool for identifying and classifying breast cancer to reduce user dependency and increase diagnostic accuracy(Trister et al., 2017). This study sought to determine the accuracy of computer-aided detection (CAD) software in detecting cancer in breast cancer patients undergoing preoperative automated ultrasound (ABUS) and to assess the factors associated with false-negative results.(Fernandes et al., 2010).

Isotonic separation is a linear programing method which is based on fundamental premises that tells weather the interpretation is correct and accurate or not. Many patients shows different symptoms which are more complicated using same dataset(Fusco et al., 2016). The most popular method which is used for diagnoses of cancer is aforementioned which examine the data and indicate the problem .researcher combine this method with machine learning technique for better prediction(Gilbert et al., 2008). Breast cancer using artificial neural network based on BRI-Rads Standardize Lexicon in 1995(Gromet, 2008). They achieve 89.00% with the help of Area under the curve technique on Wisconsin Dataset. They use 133 samples on both male and female, but the majority sample was female. They did this on the basis of two parameters which is benign and malignant(Jalalian et al., 2017). Survey paper with four Algorithms SVM, ANN, RVM, ELM. In this they said that SVM and RVM are the best predictors all the time for breast cancer problem(Hupse et al., 2013). Research on breast cancer with the help of four algorithms of wachine learning names are. SVM, K-NN, Decision Tree, Naiive Bayes. The highest accuracy achieved by SVM (97.13%). They use WEKA tool for data mining(Kooi et al., 2017). The primary goal is to evaluate the accuracy of data classification in terms of each category's efficiency and effectiveness in terms of accuracy, precision, sensitivity, and specificity(Kunar, 2022). Diagnoses and prognosis the breast cancer SVM, k-NN. Decision Tree and ANN (Lehman et al., 2015). The Iwo step SVM produced the highest result (99.10) with 10 cross validations. Judge the two cases—benign and malignant—based on the results. Their main sources of information are the two benign and malignant breast instances from Wisconsin. Their main source of information is the Wisconsin Breast Cancer Database (WBCD), which serves as the standard-setting database for comparing the outcomes of various algorithms. Here, a dataset was analysed using statistical methods like cross selection and linear regression to learn more about the distribution and variation of the data(Litjens et al., 2017).

Different machine learning algorithms for breast cancer on the basis of blood analysis data were used. Algorithms are SVM, ELM, ANN, KNN and hyper parameter Optimization. They use MATLAB environment for ELM classification and graphical representation. Here ELM is the highest achiever with 80% accuracy in 0.0075 seconds time. Team members have discussed Breast cancer (Rimmer, 2017) with the help of picture data with CNN algorithm and SVM .SVM accuracy was 87.2%, but with use of AUC the accuracy was 94%. They mine their data with the help of pictures(Magny et al., 2022). Hybrid approach for the prediction of BC is used in Oct 2018). They choose ANN, SVM, RF, and navie-bayes with PCA rule. They Analyze result on the basis of accuracy, sensitivity, specificity and prevalence .ANN is the highest achiever with 97% accuracy and 95% sensitivity and 98% specificity(Ogunleye et al., 2022). Cancer Survival Analysis Using Machine Learning in(Saleh et al., 2011) with SVM, ANN, NB and DT .Feature selection is applied on the Wisconsin Breast cancer dataset. They proved that if we detect Cancer on stage 1 there is 99% chances of treatment and alive rate and if detect on early stage 4 there is 30% chance of living. ANN achieved 96% accuracy (Fernandes et al., 2010). The implication for the future of healthcare systems embracing ML and big data to detect cancer is that the combination of various techniques with multidimensional heterogeneous data, while focusing on feature classification and selection, is likely to produce more accurate results(Rimmer, 2017).

III. Materials and Methods

This section contains the methodology of the study.

3.1. Proposed Methodology

In proposed work from (Figure 1) workflow Diagram describes the straightforward analysis. In this work we are using ANN+KNN+SVM+RF+LR for cancer detection with reduced of features. A supervised learning strategy is used while keeping accuracy, precision, Material and Methods score, recall well above the dataset with numerical are recall well above the required level of detection of breast cancer. We have attained a ut with numerical values of different examples and split the data into train and ratio of 20. In the order to test and train we use five classifier of machine learning: Artificial neural work (ANN). Random forest (RF), K-nearest neighbor (KNN), logistic regression (LR), Support vector machine (SVM). Feature selection is performed with the assistance of Python library Scikit -Learn with One Encoder. Then we trained model and took results and analyze with expected results and optimize them.

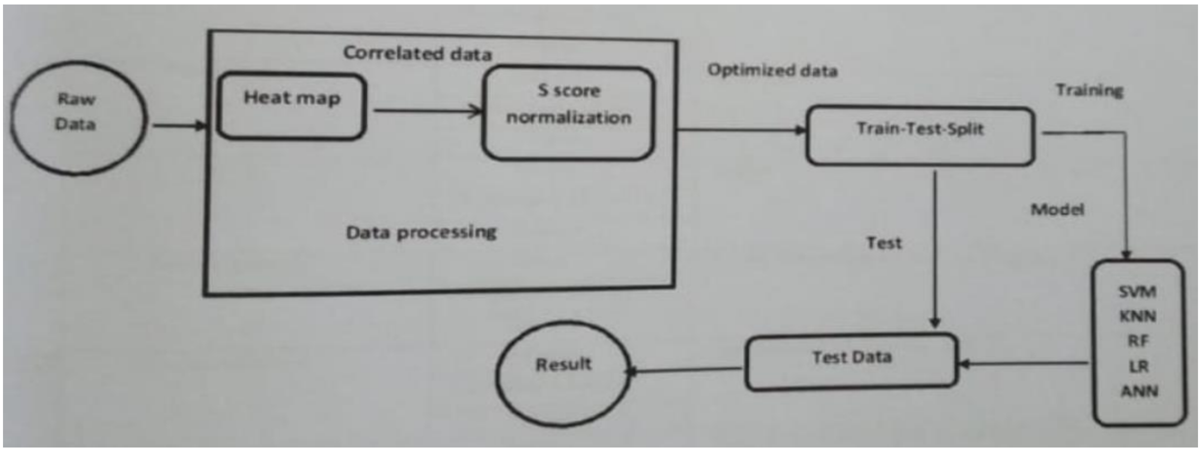

Figure 1. Propose Methodology Diagram

3.1. Data Collection

The dataset cast-off is available and remained bent by Dr. William H. Wolberg, doctor at the University Of Wisconsin Hospital at Madison, Wisconsin, United States of America. To generate the dataset, Dr. Wolberg formulate fluids tasters, occupied of women with dense breast tumors and extremely simple touse digital computer application called Xevt. That effectively accomplishes cytological feature investigation focusing on a digital scan. The sequencer use a curve-fitting technique to determine the structures from each cell in the model, after which it analyses the extreme value, mean value, and standard error of each feature for the im-age, repeatedly iterating a 30 real-valued vector. The dataset includes 57 cases of benign tumours and 212 cases of malignant tumours. It has 32 columns, the first of which is a patient ID column. The remaining columns contain the tumour detection findings, which include whether a tumour is benign or malignant.

There were no misplaced values in the dataset: information of dataset includes.

1. ID of patient
2. Diagnosis (M = malignant, B = benign)

3.2 Feature Selection:

During the analysis, the supervised method ANN+KNN+SVM+RM+LR is employed for feature selection. Automatic selection of data aspects that are most pertinent to the challenge of predictive modelling. Dimensionality reduction and feature selection are entirely independent processes.From (figure 2) Each strategy asks for a decrease in the number of features in the dataset; however, dimensionality reduction methodology is utilised, creating new combinations of features; in contrast, feature selection technique accepts and omits attributes present in the data without ever modifying the attributes. Feature selection techniques produce a precise predictive model. The best or better accuracy will be achieved with less data by assisting in feature selection. The feature selection method can be applied to identify and remove. Algorithms for feature selection that are often used include wrapper methods, filter techniques, and embedding methods. Filtering technique: A statistical method for assigning ratings to each feature used in filter feature selection techniques. The characteristics are either chosen to be unbroken or excluded from the dataset and are organised hierarchically by the score. The approaches are often accurate and consider the characteristic multiple times, or with reference to the number of attributes. Wrapper approach Consider group selection and options as a search downside in this method when completely diverse features are ready to be measured and analysed with various combinations. A prediction model employs conventional evaluations to rank a variety of combinations and provide a score supporting pattern precision. Unlike a best first search, the hunt procedure is planned. It can use random hill climbing algorithms or heuristics like forward and backward to present and remove options. Another approach, known as the embedded method, is used to find the solution that improves accuracy and performance when the model is twisted. The most common approach is regularisation method, which is also known as the common form of embedded function. The regularisation technique, sometimes referred to as the penalization strategy, such as the bias model with reduced complexity, imposes additional restrictions on the use of predictive algorithms.

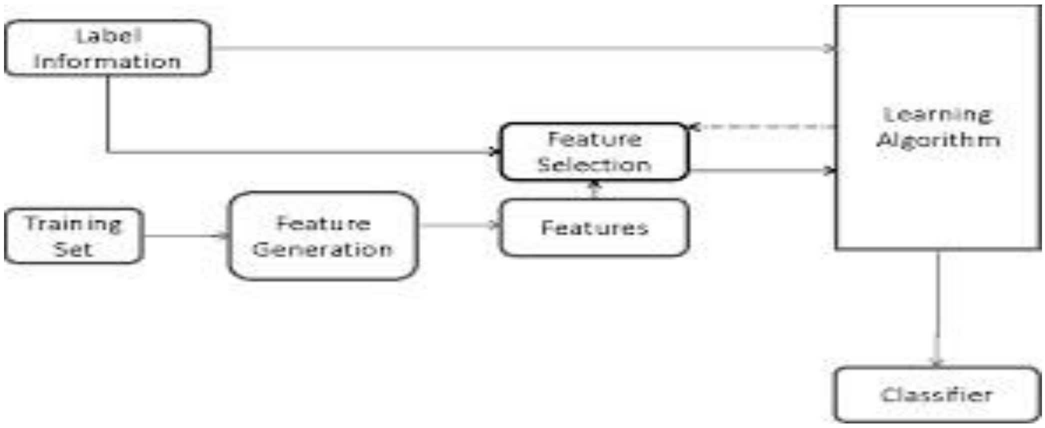

Figure 2. Feature Selection

3.3 Data Analysis:

We used various approaches of machine learning for data analysis such as principle component analysis PCA, which is a mathematical technique. use orthogonal transformation to turn series of information of theoretically associated variables I nto a collection of linearly uncorrelated variables called PCA in which dataset is composed of a variety of variables linked to each other it can be large or tiny through keeping the variability present inside the dataset to the fullest degree possible identification is done by renovating variables to be the set of variables to be built original variables are present in initial variables when we continue to go down in sequence thus throughout this process the first key item is kept the main aspect are Manfred Eigen of covariance matrix, and thus orthogonal. The dataset of PCA methodology to be used will be distributed the findings are rather responsive. This technique is used for summarization of information imagine a wine bottles on a board each wine describes by its characteristics such as color intensity etc.Actually it is dimensional reduction technique which consists of related features. In the order to reduce large dataset and to be presented within few variable. Data analysis techniques applied to correlate attributes gives better results. In this study, we used heat map which is old but powerful technique to find correlation between features. The key role of heat map is to detecting the similarity in the sample and identifying correlations and variation between individual attributes on the basis of thresholds value. After heat map was applied, highest correlated features were selected and other features which were less correlated we dropped them from dataset. We find correlation in two different threshold values.

(1) We set threshold value 1.0

(2) In second case, we set value 0.7.

We find correlation among all features and most correlated features was dropped out from dataset. In above paragraph Heat map is used to find correlation and then on the bases of correlation features selected and rejected define 16 selected features ama dataset with thresholds value 1.0.

3.4 Train-Test Split:

Machine learning data in most cases are broken down into either testing of data or training data or often into three practices, validate, and check. Makes our model fit to train the data. Test data is part of real dataset; algorithms are required to be trained. The model identified and learns from results, for testing o to the testing dataset. Testing of data is optimal standard used for validation the pattern s used to deliver an objective review of final model match to the testing dataset. Testing of data is optimal standard used for validation the pattern and used after the completion of model training.

Dataset is divided into training, the validate check may be divided into two groups. Initially it depends on how large the overall amount of samples in the data. Secondly it checks particular depends on how large model that the user is training. Most of the models require powerful or significant data to train.in his scenario, one would be able to prepare with wide training sets. Some models can be easily validate and adjust and can reduce the size validation size moreover considering that model has several hyper parameter, and user demands for the validation which include large datasets. In this thesis the percentage of dividing dataset is 80 percent of trained data and 20 percent of testing dataset 400 instances in dataset is of training dataset and 169 instances of testing. Training of our algorithms in machine learning is crucial. From 80 percent of research we hold 73 percent for training and 7 percent for cross validation. The cross testing phase consist of splitting the data segment into complementary subset, conducing an experiment on one subset (training groups) evaluating an experiment on the other subset called testing of data. To minimize uncertainty, several rounds of cross validation are conducted using separate division of several methods and validation test are summed and tis round provide estimation of the model's predictive efficiency.

IV. Results and Discussions

Whole set of five machine learning classification algorithms are rummage-saleand K Nearest Neighbors, artificial neural network (ANN), Support Vector Machines (SVM), Random Forest, Logistic Regression (LR),Classifier have been practical on the dataset. For each experiment, the performance of the algorithms is measured using, Precision, F1 Score, Accuracy, Recall Specificity and error rate**.**

4.1 Confusion Matrix of SVM

We trained model with SVM algorithms and obtained result analysis done with confusion matrix.
(Table 1 and table 2) explain confusion matrix and classification reportof SVM.

Table 1. Confusion matrix of SVM

	Actual Negative	Actual Positive
Predicted Negative	66	1
Predicted Positive	1	46

Table 2. Classification Report of SVM

Accuracy	Precision	Recall	F1_Score	Specificity	Error Rate
98.24%	0.98	0.98	0.98	0.9850	0.0176

4.1.1 Confusion Matrix of K-NN

Accuracy of the k-NN is regulating with the help of Confusion matrix under the recall, precision, f1_score and specificity parameter. (Table 3 and table 4) explain K-nearest neighbor of Confusion matrix and and classification reportbelow here.

Table 3. Confusion matrix of KNN

	Actual Negative	Actual Positive
Predicted Negative	67	0
Predicted Positive	5	42

Table 4. Classification Report of KNN

Accuracy	Precision	Recall	F1_Score	Specificity	Error Rate
95.61%	0.97	0.95	0.95	1	0.0439

4.1.2 Confusion Matrix of Logistic regression

Logistic regression resembles a linear regression but the major difference is logistic regression when we want classification on the basis of Independent and dependent variable. (Table 5 and table 6) we evaluate the logistic regression generates the confusion matrixand classification reportwhich is shown below.

Table 5. Confusion Matrix of LR

	Actual Negative	**Actual Positive**
Predicted Negative	65	2
Predicted Positive	2	45

Table 6. Classification report of LR

Accuracy	**Precision**	**Recall**	**F1_Score**	**Specificity**	**Error Rate**
96.49%	0.96	0.96	0.96	0.9701	0.0351

4.1.3Confusion Matrix of Random Forest

Another important algorithm in machine learning is random forest or we can say that it is backbone. Because it is performs in both classification and regression problems.(Table 7 and table 8) we useconfusion matrix and classification reportof the random forest.

Table 7. Confusion Matrix of RF

	Actual Negative	**Actual Positive**
Predicted Negative	66	1
Predicted Positive	2	45

Table 8. Classification report of RF

Accuracy	**Precision**	**Recall**	**F1_Score**	**Specificity**	**Error Rate**
97.36%	0.97	0.97	0.97	0.9850	0.264

4.1.4 Confusion Matrix of Artificial Neural Network

Ann is stimulated by the humanoid mind. As same in the human mind, Ann has the same function. Here we have not Confusion matrix because in the ANN we have continuous value like regression so confusion matrix has not the ability to take both parameter binary and continuous that's why we use accuracy graph here.Other parameters remain as in the previous algorithms. (Table 9 and table 10) shows the clear picture of ANN accuracy and tells us that how much accurate and explain Confusion matrix of ANN.

Table 9. Confusion Matrix of ANN

	Actual Negative	**Actual Positive**
Predicted Negative	67	0
Predicted Positive	2	45

Table 10. Classification report of ANN

Accuracy	**Precision**	**Recall**	**F1_Score**	**Specificity**	**Error Rate**
98.245%	0.98	0.98	0.98	0.9701	0.0175

4.2 Comparative Analysis of ML Algorithms

The objective of our thesis is done in this section to analyze the results of all implemented models. Here we will do a comparison on the basis of accuracy, precision, recall, f1-Score, specificity, error rate. To check the performance of a model the only accuracy is not a single parameter on which we focused. We want a perfect model which is a good predictor to get this kind of model we focus on different parameters which we discussed above. (Table 11) and (figure 3) show the clear results of all the a fall the Algorithms which we are implemented in this thesis.

Table 11. Analysis of All Algorithms

Algorithms	**Accuracy**	**Precision**	**Recall**	**F1-Score**	**Specificity**	**Error rate**
SVM	98.24%	0.98	0.98	0.98	0.9701	0.0175
k-NN	95.61%	0.97	0.95	0.95	1	0.0439
LR	96.49%	0.96	0.96	0.96	0.9701	0.0351
RF	97.36%	0.97	0.97	0.97	0.9850	0.0264

ANN	98.245%	0.98	0.98	0.98	0.9701	0.0175

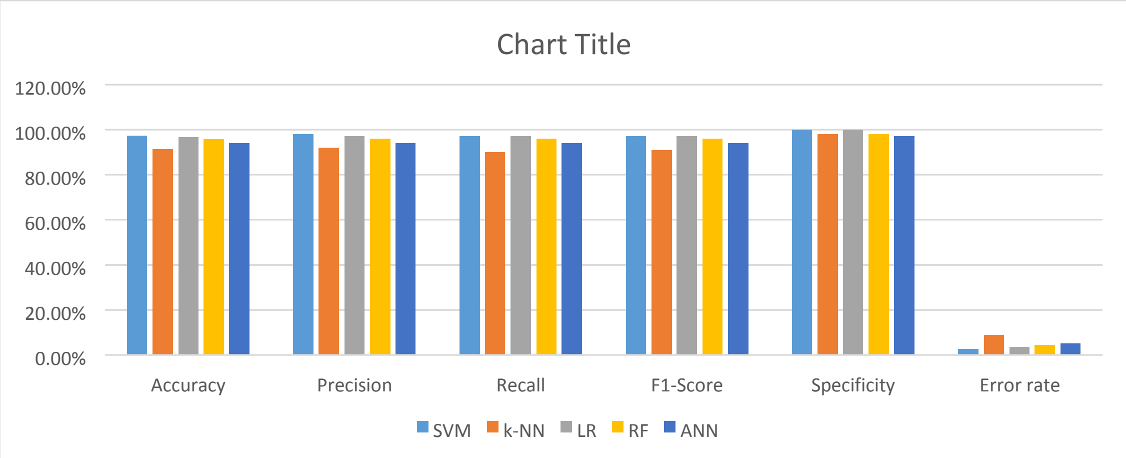

Figure 3. Accuracy and Precision comparison

4.3 Model Performances with less Features

Above models us the models use the whole dataset and use all the features which are 30. But in this section we measure the correlation between features and use those features that are unique and have less similarity. After finding the correlation the total features which we will be used in algorithms are16 instead of 30. We find a correlation by using heatmap. The first picture defines the correlation of 30 features. Color bar defines the correlation between features or input data.

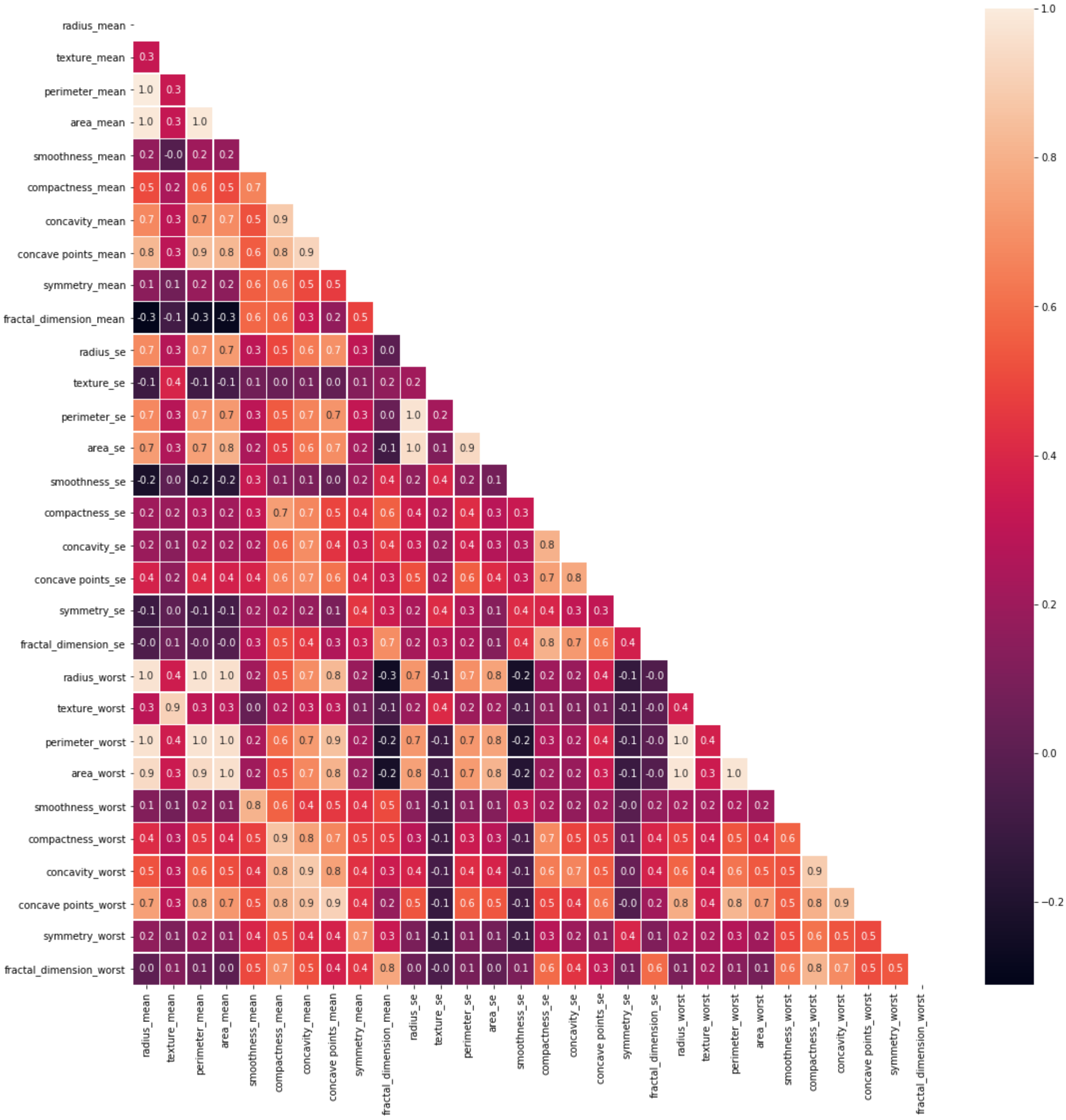

Figure 4. Heat map of 30 features

This (figure 4) explains the possible featured at which algorithms will work in this figure 30 features are extract to work and dark area sows the feature that are used moreover light black area indicates the features that are substitute.

4.4.1 Confusion Matrix of SVM

Confusion matrix is used for the measurement of our models how accurate are our models. Support vector machine performs both regression and classification. (Table 12) confusion matrix of SVM with less features shown below.

Table 12. Confusion Matrix of SVM with Selected Features

	Actual Negative	**Actual Positive**
Predicted Negative	66	1
Predicted Positive	2	45

(Table 13) Classification report of SVM with 16 features are shown below here.

Table 13. Classification Report of SVM with Selected Features

Accuracy	**Precision**	**Recall**	**F1_Score**	**Specificity**	**Error Rate**
97.36%	0.97	0.97	0.97	0.98	0.0264

4.4.2 Confusion matrix of KNN

In the turf of machine learning and exactly the problematic of statistical classification, a confusion matrix, also identified as an error matrix. A confusion matrix is a bench that is frequently used to designate the performance of a classification model (or ―classifier") on a usual of test data for which the true values are known. (Table 14) Confusion matrix of KNN is shown below.

Table 14. Confusion Matrix of KNN with Selected Features

	Actual Negative	**Actual Positive**
Predicted Negative	65	2
Predicted Positive	2	45

(Table 15)classification report of KNN with selected features:

Table 15. Classification Report of KNN with Selected Features

Accuracy	**Precision**	**Recall**	**F1_Score**	**Specificity**	**Error Rate**
93.85%	0.94	0.93	0.94	0.97	0.0615

4.4.3 Confusion matrix of Logistic regression

Logistic regression is comparable to linear regression but the logistic is used for classification problem. (Table 16)confusion matrix of logistic regression is shown below.

Table 16. Confusion Matrix of LR with Selected Features

	Actual Negative	**Actual Positive**
Predicted Negative	65	2
Predicted Positive	2	42

(Table 17) explain Classification report of logistic regression.

Table 17. Classification Report of LR with Selected Features

Accuracy	**Precision**	**Recall**	**F1_Score**	**Specificity**	**Error Rate**
96.49%	0.96	0.96	0.96	0.97	0.0351

4.4.4 Confusion matrix of Random forest

Random forest works on the basis of ensam bling learning. In this we make decision on majority base. (Table 18) explain Confusion matrix of random forest.

Table 18. Confusion Matrix of Random Forest with Selected Features

	Actual Negative	**Actual Positive**
Predicted Negative	66	1
Predicted Positive	5	42

(Table 19) explain classification report of logistic regression.

Table 19. Classification Report of LR with selected Features

Accuracy	**Precision**	**Recall**	**F1_Score**	**Specificity**	**Error Rate**
94.73%	0.95	0.94	0.95	0.98	0.0527

4.4.5 Confusion Matrix of Artificial Neural Network

Ann inspire from biological neural network. It performs both regression and classification problem. Confusion matrix of neural network is hard to plot because we have continuous value during testing and testing so binary plot is difficult. So we have (Table 20) explain Confusion matrix of ANN.

Table 20. Explain confusion Matrix of ANN with Selected Features

	Actual Negative	**Actual Positive**
Predicted Negative	63	4
Predicted Positive	1	46

4.5 Comparative Analysis of ML Algorithms

The objective of our thesis is done in this section to analyze the results of all implemented models. Here we will do a comparison on the basis of accuracy, precision, recall, f1-Score, Specificity, error rate. To check the performance of a model the only accuracy is not a single parameter on which we focused. We want a perfect model which is a good predictor to get this kind of model we focus on different parameters which we discussed above. The (table 21) shows the clear results of all the Algorithms which we are implemented in this thesis.

Table 21. Analysis of All Algorithms with 16 selected Features

Algorithms	Accuracy	Precision	Recall	F1-Score	Specificity	Error rate
SVM	97.3%	0.97	0.97	0.97	0.98	0.0264
k-NN	93.85%	0.94	0.93	0.94	0.97	0.0615
LR	96.49%	0.96	0.96	0.96	0.97	0.0351
RF	94.73%	0.95	0.94	0.95	0.98	0.0527
ANN	96.61%	0.96	0.96	0.96	0.94	0.049

Graphical Representation of All Algorithms with 16 Selected features is show in (figure 4)

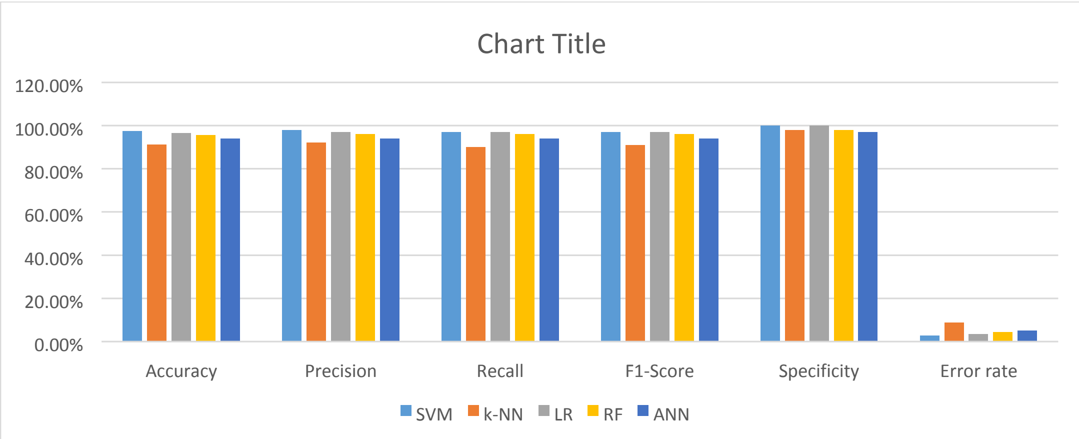

Figure 4. Comparison of algorithms on all quality measures

Table 22. Analysis of All Algorithms with 16 selected Features

Algorithms	Accuracy	Precision	Recall	F1-Score	Specificity	Error rate
SVM	97.36%	0.98	0.97	0.97	1	0.0264
k-NN	91.22%	0.92	0.90	0.91	0.98	0.0878
LR	96.49%	0.97	0.97	0.97	1	0.0351
RF	95.61%	0.96	0.96	0.96	0.98	0.0439
ANN	93.85%	0.94	0.94	0.94	0.97	0.049

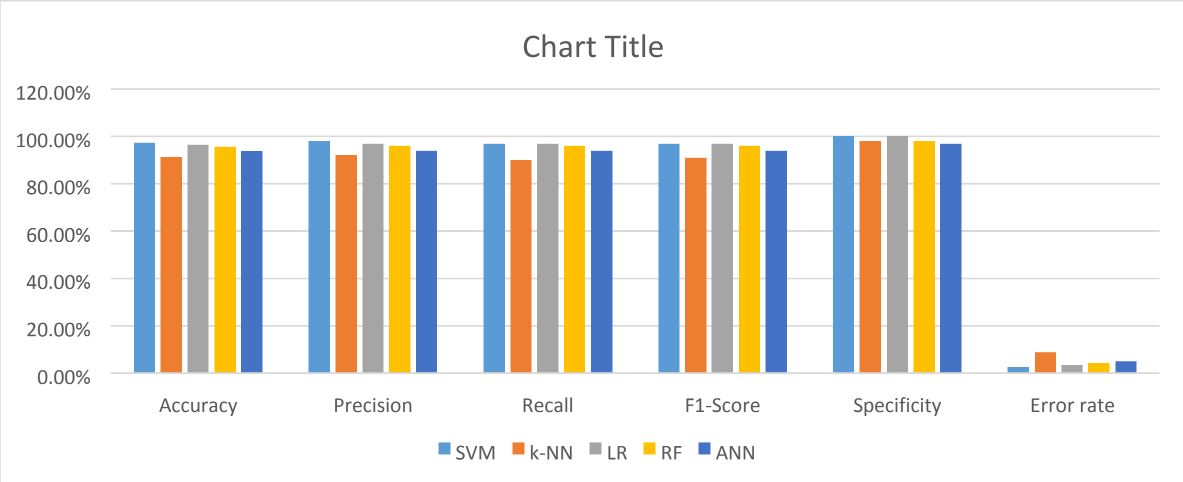

Figure 5. Comparison of quality scores

The proposed hybrid approach use best of the classifiers for reduced features as Support vector machine (SVM) for accuracy and precision and produced unchanged accuracy with reduction of features hence SVM can be used for superior accuracy with less features and less computation time. And KNN shows less accuracy after the reduction of features. For F1 score logistic regression (LR) is showing best results for reduced features. Although error rate is increased for most of the classifiers still it remains acceptable limit for SVM and LR.

4.6 Computational Analysis:

(Table 23) is explain that time is improved by 50% which is a significant achievement especially when certain level of accuracy and precision is maintained to provide high reliability with fewer features. The proposed technique has given a good confidence with less features and improved time.

Table 23. Computational Time

Classifiers	9 Features	16 Features	Time saved
SVM	0.6	0.21	0.39
KNN	0.5	0.42	0.08
ANN	0.33	0.26	0.07
LR	0.71	0.31	0.4
RF	0.61	0.23	0.38

V. Conclusion and Future Work

In thus research paper ANN,KNN, SVM, RF,LR are used, compared and analysis with vision of computational gain and reduction of features to a sufficient quantity. A supervised learning framework is used while keeping accuracy, precision, F1 score, recall well above the required level of detection of breast cancer. The proposed hybrid approach use best of the classifiers for reduced features as Support vector machine (SVM) for accuracy and precision and produced unchanged accuracy with reduction of features hence SVM can be used for superior accuracy with less features and less computation time. And KNN shows less accuracy after the reduction of features. For F1 score logistic regression (LR) is showing best results for reduced features. Although error rate is increased for most of the classifiers still it remains acceptable limit for SVM and LR. The time is improved by 50% which is a significant achievement especially when certain level of accuracy, precision is maintained and provides high reliability with fewer features. The proposed technique has given a good confidence with less features and improved time.

In fact there is a shortage of dynamic structure included in the study while we have obtained stronger outcome with the model that use for more sophisticated models will capture dynamic interaction between function .while obtaining reliable results and consistency for five algorithm we used, we want to ensure that the findings we have produced are not skewed by the size of our dataset therefore as our dataset is quite outdated more statistical parameters and advanced technologies would have been accessible in order to acquire more reliable numerical results this will also check our research if we were able to define the right parameters from our present and potential datasets in the order to produce ROC curve. Actually comparison of the models that we have used, we will like to test other algorithms to evaluate and pursue our quest for strongest predictive model. Thus, the correct diagnosis of breast cancer and the classification of patients into malignant or benign groups is the subject of further research.

Funding: No funding was received.

Data Availability Statement: The authors declare that all data supporting the findings of this study are available within the article.

Conflicts of Interest: The authors declare no conflict of interest.

References

[1]. Abbas, M. R., Nadeem, M. S. A., Shaheen, A., Alshdadi, A. A., Alharbey, R., Shim, S.-O., & Aziz, W. (2019). Accuracy Rejection Normalized-Cost Curves (ARNCCs): A Novel 3-Dimensional Framework for Robust Classification. IEEE Access, 7, 160125-160143.

[2]. Azavedo, E., Zackrisson, S., Mejàre, I., & Heibert Arnlind, M. (2012). Is single reading with computer-aided detection (CAD) as good as double reading in mammography screening? A systematic review. BMC Med Imaging, 12, 22. https://doi.org/10.1186/1471-2342-12-22

[3]. Bargalló, X., Santamaría, G., Del Amo, M., Arguis, P., Ríos, J., Grau, J., Burrel, M., Cores, E., & Velasco, M. (2014). Single reading with computer-aided detection performed by selected radiologists in a breast cancer screening program. Eur J Radiol, 83(11), 2019-2023. https://doi.org/10.1016/j.ejrad.2014.08.010

[4]. Broeders, M., Moss, S., Nyström, L., Njor, S., Jonsson, H., Paap, E., Massat, N., Duffy, S., Lynge, E., & Paci, E. (2012). The impact of mammographic screening on breast cancer mortality in Europe: a review of observational studies. J Med Screen, 19 Suppl 1, 14-25. https://doi.org/10.1258/jms.2012.012078

[5]. Bunkhumpornpat, C., Sinapiromsaran, K., & Lursinsap, C. (2009). Safe-Level-SMOTE: Safe-Level-Synthetic Minority Over-Sampling TEchnique for Handling the Class Imbalanced Problem. PAKDD,

[6]. Casal-Guisande, M., Comesaña-Campos, A., Dutra, I., Cerqueiro-Pequeño, J., & Bouza-Rodríguez, J. B. (2022). Design and Development of an Intelligent Clinical Decision Support System Applied to the Evaluation of Breast Cancer Risk. J Pers Med, 12(2). https://doi.org/10.3390/jpm12020169

[7]. Cerqueiro-Pequeño, J., Comesaña-Campos, A., Casal-Guisande, M., & Bouza-Rodríguez, J.-B. (2021). Design and Development of a New Methodology Based on Expert Systems Applied to the Prevention of Indoor Radon Gas Exposition Risks. International Journal of Environmental Research and Public Health, 18(1), 269. https://www.mdpi.com/1660-4601/18/1/269

[8]. Chawla, N., Bowyer, K., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic Minority Over-sampling Technique. J. Artif. Intell. Res., 16, 321-357.

[9]. Comesaña-Campos, A., Casal-Guisande, M., Cerqueiro-Pequeño, J., & Bouza-Rodríguez, J.-B. (2020). A Methodology Based on Expert Systems for the Early Detection and Prevention of Hypoxemic Clinical Cases. International Journal of Environmental Research and Public Health, 17(22), 8644. https://www.mdpi.com/1660-4601/17/22/8644

[10]. Cunningham, P., & Delany, S. J. (2020). k-Nearest Neighbour Classifiers: 2nd Edition (with Python examples). ArXiv, abs/2004.04523.

[11]. Debelee, T. G., Kebede, S. R., Schwenker, F., & Shewarega, Z. M. (2020). Deep Learning in Selected Cancers' Image Analysis-A Survey. J Imaging, 6(11). https://doi.org/10.3390/jimaging6110121

[12]. Debelee, T. G., Schwenker, F., Ibenthal, A., & Yohannes, D. (2020). Survey of deep learning in breast cancer image analysis. Evolving Systems, 11, 143-163.

[13]. Fenton, J. J., Taplin, S. H., Carney, P. A., Abraham, L., Sickles, E. A., D'Orsi, C., Berns, E. A., Cutter, G., Hendrick, R. E., Barlow, W. E., & Elmore, J. G. (2007). Influence of computer-aided detection on performance of screening mammography. N Engl J Med, 356(14), 1399-1409. https://doi.org/10.1056/NEJMoa066099

[14]. Fenton, J. J., Xing, G., Elmore, J. G., Bang, H., Chen, S. L., Lindfors, K. K., & Baldwin, L. M. (2013). Short-term outcomes of screening mammography using computer-aided detection: a population-based study of medicare enrollees. Ann Intern Med, 158(8), 580-587. https://doi.org/10.7326/0003-4819-158-8-201304160-00002

[15]. Fernandes, A. S., Alves, P., Jarman, I. H., Etchells, T. A., Fonseca, J. M., & Lisboa, P. J. G. (2010, 2010//). A Clinical Decision Support System for Breast Cancer Patients. Emerging Trends in Technological Innovation, Berlin, Heidelberg.

[16]. Fusco, R., Sansone, M., Filice, S., Carone, G., Amato, D. M., Sansone, C., & Petrillo, A. (2016). Pattern Recognition Approaches for Breast Cancer DCE-MRI Classification: A Systematic Review. J Med Biol Eng, 36(4), 449-459. https://doi.org/10.1007/s40846-016-0163-7

[17]. Gilbert, F. J., Astley, S. M., Gillan, M. G., Agbaje, O. F., Wallis, M. G., James, J., Boggis, C. R., & Duffy, S. W. (2008). Single reading with computer-aided detection for screening mammography. N Engl J Med, 359(16), 1675-1684. https://doi.org/10.1056/NEJMoa0803545

[18]. Gromet, M. (2008). Comparison of computer-aided detection to double reading of screening mammograms: review of 231,221 mammograms. AJR Am J Roentgenol, 190(4), 854-859. https://doi.org/10.2214/ajr.07.2812

[19]. Hupse, R., Samulski, M., Lobbes, M. B., Mann, R. M., Mus, R., den Heeten, G. J., Beijerinck, D., Pijnappel, R. M., Boetes, C., & Karssemeijer, N. (2013). Computer-aided detection of masses at mammography: interactive decision support versus prompts. Radiology, 266(1), 123-129. https://doi.org/10.1148/radiol.12120218

[20]. Jalalian, A., Mashohor, S., Mahmud, R., Karasfi, B., Saripan, M. I. B., & Ramli, A. R. B. (2017). Foundation and methodologies in computer-aided diagnosis systems for breast cancer detection. Excli j, 16, 113-137. https://doi.org/10.17179/excli2016-701

[21]. Kooi, T., Litjens, G., van Ginneken, B., Gubern-Mérida, A., Sánchez, C. I., Mann, R., den Heeten, A., & Karssemeijer, N. (2017). Large scale deep learning for computer aided detection of mammographic lesions. Med Image Anal, 35, 303-312. https://doi.org/10.1016/j.media.2016.07.007

[22]. Kunar, M. A. (2022). The optimal use of computer aided detection to find low prevalence cancers. Cognitive Research: Principles and Implications, 7(1), 13. https://doi.org/10.1186/s41235-022-00361-1

[23]. Lehman, C. D., Wellman, R. D., Buist, D. S., Kerlikowske, K., Tosteson, A. N., & Miglioretti, D. L. (2015). Diagnostic Accuracy of Digital Screening Mammography With and Without Computer-Aided Detection. JAMA Intern Med, 175(11), 1828-1837. https://doi.org/10.1001/jamainternmed.2015.5231

[24]. Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., van der Laak, J., van Ginneken, B., & Sánchez, C. I. (2017). A survey on deep learning in medical image analysis. Med Image Anal, 42, 60-88. https://doi.org/10.1016/j.media.2017.07.005

[25]. Magny, S. J., Shikhman, R., & Keppke, A. L. (2022). Breast Imaging Reporting and Data System. In StatPearls. StatPearls Publishing

[26]. Copyright © 2022, StatPearls Publishing LLC.

[27]. Ogunleye, A. Z., Piyawajanusorn, C., Gonçalves, A., Ghislat, G., & Ballester, P. J. (2022). Interpretable Machine Learning Models to Predict the Resistance of Breast Cancer Patients to Doxorubicin from Their microRNA Profiles. Adv Sci (Weinh), 9(24), e2201501. https://doi.org/10.1002/advs.202201501

[28]. Rimmer, A. (2017). Radiologist shortage leaves patient care at risk, warns royal college. Bmj, 359, j4683. https://doi.org/10.1136/bmj.j4683

[29]. Saleh, A., Barakat, S. I., & Awad, A. (2011). A Fuzzy Decision Support System for Management of Breast Cancer. International Journal of Advanced Computer Science and Applications, 2.

[30]. Samulski, M., Hupse, R., Boetes, C., Mus, R. D., den Heeten, G. J., & Karssemeijer, N. (2010). Using computer-aided detection in mammography as a decision support. Eur Radiol, 20(10), 2323-2330. https://doi.org/10.1007/s00330-010-1821-8

[31]. Shler Farhad, K., & Adnan Mohsin, A. (2021). BREAST CANCER DIAGNOSIS BASED ON K-NEAREST NEIGHBORS: A REVIEW. PalArch's Journal of Archaeology of Egypt / Egyptology, 18(4), 1927-1951. https://archives.palarch.nl/index.php/jae/article/view/6601

[32]. Smith, R. A., Cokkinides, V., Brooks, D., Saslow, D., & Brawley, O. W. (2010). Cancer screening in the United States, 2010: a review of current American Cancer Society guidelines and issues in cancer screening. CA Cancer J Clin, 60(2), 99-119. https://doi.org/10.3322/caac.20063

[33]. Trister, A. D., Buist, D. S. M., & Lee, C. I. (2017). Will Machine Learning Tip the Balance in Breast Cancer Screening? JAMA Oncol, 3(11), 1463-1464. https://doi.org/10.1001/jamaoncol.2017.0473

[34]. van Bommel, R. M. G., Weber, R., Voogd, A. C., Nederend, J., Louwman, M. W. J., Venderink, D., Strobbe, L. J. A., Rutten, M. J. C., Plaisier, M. L., Lohle, P. N., Hooijen, M. J. H., Tjan-Heijnen, V. C. G., & Duijm, L. E. M. (2017). Interval breast cancer characteristics before, during and after the transition from screen-film to full-field digital screening mammography. BMC Cancer, 17(1), 315. https://doi.org/10.1186/s12885-017-3294-5

[35]. Weber, R. J., van Bommel, R. M., Louwman, M. W., Nederend, J., Voogd, A. C., Jansen, F. H., Tjan-Heijnen, V. C., & Duijm, L. E. (2016). Characteristics and prognosis of interval cancers after biennial screen-film or full-field digital screening mammography. Breast Cancer Res Treat, 158(3), 471-483. https://doi.org/10.1007/s10549-016-3882-0

[36]. Yue, W., Wang, Z., Chen, H., Payne, A., & Liu, X. (2018). Machine Learning with Applications in Breast Cancer Diagnosis and Prognosis. Designs, 2(2), 13. https://www.mdpi.com/2411-9660/2/2/13