VFAST Transactions on Computer Sciences
http://vfast.org/journals/index.php/VTCS@ 2018 ISSN(e): 2308-8168; ISSN(p): 2411-6335
Volume 6, Number 1, January-December, 2018

pp. 78-86

A PARELLEL TWO STAGE CLASSIFIER FOR BREAST CANCER PREDICTION AND COMPARISON WITH VARIOUS ENSEMBLE TECHNIQUES

ALI TARIQ NAGI[1], AHMED WALI[1], ADNAN SHAHZADA[1], MOHAMMAD MASROOR AHMED[2]

[1]Department of Computer Science, University of Management and Technology, Lahore, Pakistan
[2]Department of Computer Sciences, Capital University of Science and Technology, Islamabad, Pakistan
Email: ali.tariq.nagi@gmail.com

ABSTRACT. *Life is a blessing but some diseases snatch human life away before even they are being diagnosed. One such horrifying disease is cancer. Among cancer, the most leading and common type is breast cancer. The actual problem lies in the fact that it is very hard and time consuming for even the most experienced medical specialist to detect the disease with high accuracy but the machines and modern computer science techniques have increased the accuracy and reduced the amount of time taken to diagnose cancer. In the subject paper, a new parallel machine learning technique called the two-stage classifier for identifying breast cancer is presented and compared with various existing techniques in terms of accuracy and percentage error reduction. The proposed technique, turns out to be better not only in terms of parallelism but also in terms of the evaluated metrics and reduced the error percentage to almost 50% in one of the cases.*

Keywords*:* Machine Learning, Breast Cancer, Neural Networks, SVM, Logistic Regression, KNN, Naïve Bayes, Decision Tree, MLP, Gradient Boosting Classifier, Random Forest, Wisconsin

1. **Introduction.** Breast Cancer is the second leading cause of deaths due to cancer in women and the chance of developing this disease for a woman is 1 in every 8 [1]. It accounts for 44% of the female illnesses with critical severity [2]. An early detection of the subject disease can lead to a full recovery treatment [3] but most of the time, cancer is diagnosed at the last stage [4]. Therefore, we see an increased importance of addressing this issue to correctly identify it at an early stage and save lives. With the increased computation power of modern computers and advancement in machine learning, modern computers can train themselves to identify illnesses when provided with historical data. Computers can identify minute and microscopic details of medical images which are not perceived by a human eye and can use these details to train themselves for identifying the new cases. Therefore, technology can be utilized to use machine learning techniques to correctly resolve this issue. Various machine learning algorithms have been developed and used to identify this disease.

Increasing the percentage improvement in accuracy by a fraction for machine learning algorithm for breast cancer detection can result in saving many lives. Various ensemble and multi-stage techniques are available which improve the accuracy for the unseen data by avoiding the over-fitting problem. The techniques [19,20] etc. for the subject disease and further, KNN, SVM and Naïve Bayes which have been used by many researchers [21-28] to diagnose breast cancer gave accuracies which showed us that a room for improvement exists. We developed a new parallel multi-stage classifier which combines the results from different techniques and tunes itself to perform the final, optimized prediction. We call it as a Two-Stage classifier, which is a blend of already developed algorithms and operates in two staged fashion for performing classification. The final accuracy turned out to be better than the intermediate ones in the first stage.

Our main goal is to help contribute for resolving the breast cancer identification issue by devising, identifying and classifying the machine learning algorithm better suited for breast cancer detection. Our research work includes the development of a new technique as well as comparison of the accuracies of various existing machine learning algorithms (ensemble techniques) for breast cancer classification. We have

made evaluations on two different data sets provided by Wisconsin and available on machine learning data repository [15]. The two datasets differ in number of instances and features and therefore, we check the results over different features for the same problem.

The paper is organized in the following manner: The related research for diagnosis of the subject is described in section II. Section III explains the machine learning techniques we have used and Section IV explains the proposed algorithm i.e., Two Stage Classifier. Methodology and experimental results are explained in Section V and conclusion is provided in Section VI.

2. Literature review. Chetan Nashte et al [5] in 2017 explored the possibility of using a machine learning technique for breast cancer prediction in a Clinical Decision Support System (CDSS). The author performed an analysis on the feature set, used statistical tools and found out the most significant features for classification. Moreover, a comparison was made among Naïve Bayes, k-Nearest Neighbour, Decision Tree, SVM, Random Forest and Multi-Layer Perceptron. The accuracies were 92% for Naïve Bayes, 93% for K-NN, 95% for Decision Tree, 96% for SVM, 95% for Random Forest and 72% for MLP. SVM with kernel-based method showed the best performance. It was suggested that it is an ideal machine learning technique for breast cancer detection.

Desta Mulatu et al [6] in 2017 performed a research to identify the best algorithm for the prediction of recurrence of breast cancer. The author reviewed several data mining techniques. The reported highest accuracies were 96% for Naïve Bayes, 95.7% for SVM, 91.2% for ANN, 93.6% for Decision Trees and 91.1% for Logistic Regression. The author concluded that naïve Bayes, support vector machine and decision tree give more accurate results.

Walaa Gad et al [7] in 2016 merged SVM and K-means machine learning techniques and proposed a method, called SVM-Kmeans. The authors evaluated the technique after reducing the number of features by choosing only the significant ones. They used the dataset from UCI repository [15]. The accuracy turned out to be 99.8% which was greater than the accuracies of other techniques from the literature reported by the authors.

R Delshi Howsalya Devi et al [8] in 2016 proposed a custom algorithm based on FF (Farthest First) clustering, ODA (Outlier Detection Algorithm) and J48 decision tree. It also deployed a custom algorithm for feature extraction. Moreover, datasets from UCI repository [15] were operated upon. The custom algorithm showed the best accuracy, 99.9% for WBCD dataset and 99.6% for WDBC dataset as compared to other techniques accuracies in earlier literature.

Samuel Giftson Durai et al [9] in 2016 presented a review on research papers between year 2012 to the year 2016 for algorithms for breast cancer prediction. The survey demonstrated that decision trees was best in accuracy. The authors further verified it using MATLAB and found the accuracy to be 98%.

Mohammed Abdullah Hassan Al-Hagery [10] in 2016 performed a research where objective was to find the most accurate classifier for breast cancer detection. He utilized Breast Cancer dataset from UCI repository [15] and WEKA software. He compared J48, LMT, Bayes Net, Naïve Bayes, MLP and RBF. Bayes Net was found to beat others in terms of accuracy by achieving 95.63% accuracy.

Animesh Hazra et al [11] in 2016 did research to reduce the number of features that could produce highly accurate breast cancer detection. Some machine learning methods (SVM, Naïve Bayes and ensemble classifiers) were compared. Naïve Bayes was found to be the best classifier with the lowest time complexity.

Jaimini Majali [12] in 2014 proposed a method based on FP (Frequent Pattern mining) growth algorithm. The author also concluded that Decision Tree classifier performs best on breast cancer dataset from UCI repository [15].

Jahanvi Joshi et al [13] in 2014 compared four clustering data mining techniques, namely, Farthest First (FF), Expectation Maximization (EM), Hierarchical Clusterer Method (HCM) and K-means. The dataset from UCI web data repository [15] and Weka tool were used for comparison. It was found that k-means and FF algorithm were suitable for the subject problem. HCM had a high error rate while EM was not able to diagnose 36% of the patients.

Vikas Chaurasia et al [14] in 2014 utilized Dataset from UCI repository [15] and Weka software to evaluate different machine learning methods. The results showed that Sequential Minimal Optimization (SMO) performs better in terms of accuracy (96.2%) than the techniques, IBK and BF Tree.

3. Techniques for classification. The subject section introduces the existing machine learning algorithms used inside the Two-Stage Classifier and other ensemble techniques which are compared with our proposed

algorithm.

Support Vector Machine (SVM) comes under supervised model and operates in a way that the points mapped in the space with different classes have a clear and wide gap between them. The separation between them is chosen such that the distance from the nearest data point to the line/plane is maximized. SVM also separates non-linear separation problem by mapping the points to a high dimensional plane where the points with different classes are linearly separable which is called the kernel trick. We used RBF (Radial Basis Function) kernel for SVM. The first reason was that the number of features of the data set were pretty large and secondly, it was giving better results than the ones obtained when polynomial function was deployed as kernel. The RBF is commonly used in SVM classification and is the recommended kernel.

Naïve Bayes comes under supervised learning algorithm and uses probabilistic theory for classification. It assumes that every feature is independent of the other and is based on Bayes Theorem. Let us suppose we have a vector, x of independent variables for labelling and a dependant variable y then Bayes theorem is:

$$P(y \mid x\_1,.....,x\_n \quad)= \quad (P(y) \quad P(x\_1,....., \llbracket \ x \rrbracket \_n \quad \dashv \mid \quad y \ ))/(P(x\_1,....., \llbracket \ x \rrbracket \_n \quad))$$

When the input variables are continuous, Gaussian distribution is assumed.

Logistic Regression models a relationship between the independent vector x and the dependant variable y using a polynomial equation. It generates the coefficients of the equation depending upon the dataset and provides a prediction of the outcome.

K-NN (K-Nearest Neighbour) is a non-parametric model and works by using a user-defined integer K which represents the number of neighbours to be considered for predicting the output class of the new data. Given an unseen observation, it runs through the whole dataset and computes K points that are closest to the new data and based on the class with majority points in the vicinity, it assigns a label to the new data.

Decision Tree works by inferring decision rules through the attributes of the domain set and creates a tree like structure to predict the output. The final leaves of the tree structure present the output labels. It comes under the category of supervised learning.

Multi-Layer Perceptron (MLP) is an Artificial Neural Network (ANN) which is based on the structures consisting of three different types of layers. The inputs have to pass through hidden layers before getting through output layer. Each layer contains processing elements called neurons and is responsible for transforming data according to some parameters and pass it to the next layer. This structure is used for Regression and Classification problems in machine learning.

Gradient Boosting Classifier is an ensemble technique and use various stages to classify the data. It takes various algorithms in stages. Every next algorithm is developed to be tuned for the data-set which was not predicted correctly in the previous stage. It then combines all the algorithms to give the final output.

Random Forests is also an ensemble machine learning technique that use two stages for classification. In the first stage, N Decision Trees are used to predict the outputs. Then the mode of the outputs from the previous stage is forwarded to the final output. This allows generalization improving the accuracy for the unseen values.

4. Proposed Algorithm. We propose a novel, multi-class generic classification technique to improve the results for breast cancer prediction. Our model works in two stages. The first stage comprises of different algorithms working independently to produce intermediate classification results (Pre-prediction phase). All the techniques in stage-1 can run in parallel and work independently on the input features resulting in a highly parallel nature of the technique. The results produced from stage-1 are then passed on to stage-2 where a neural network is deployed. The neural network transforms the intermediate results into a final output. The number of features of the domain set in the final stage is equal to the number of algorithms deployed in stage-1. The whole picture is presented in the below figure.

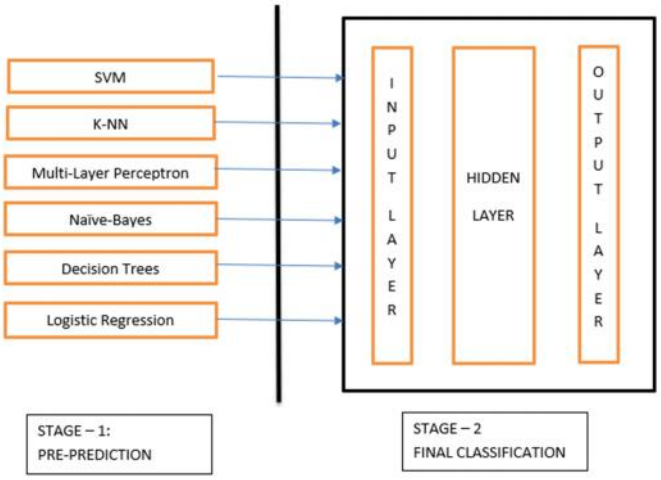

Fig. 1 Two-Stage Classifier

Stage-1 – Pre-Prediction

We call this stage as Pre-prediction phase. In this phase, we used six algorithms namely, SVM (RBF-Kernel), Naïve Bayes, Logistic Regression, KNN, Decision Trees and Multi-Layer Perceptron (MLP) to predict the type of breast cancer.

Stage-2 – Final Classification

The second stage is a Multi-Layer Perceptron (MLP) which takes as input the prediction matrix from the algorithms in the first stage and identifies the final classification of the tumour. As clear in figure 1, stage-1 consists of algorithms which can run in parallel and independently of one another. Hence, the stage-1 can be easily parallelized. Stage-2 gets the attributes' count equal to the algorithms' count used in stage-1 which can be significantly less than the number of attributes of the input data. This greatly reduces the computation in stage-2.

Platform & Implementation Details

We used Python and scikit-learn [17] libraries to deploy the techniques and evaluate their performances.

Data Splitting

Data was divided in test and train set in the percentages, 66.66% and 33.33%.

Optimizations

Each of the used algorithm was optimized in terms of parameters and by scaling the domain values. Accuracy for training set and test set was checked and overfitting problem, wherever observed, was removed. The optimized parameters for MLP in the second stage of the Two-Stage Classifier are hidden layers = 400, learning rate = 0.001, learning rate = adaptive, iterations = 1000. We used SVM (RBF kernel function) with scaled inputs. Optimized results for K-NN were obtained at 5 neighbours. Decision Trees used 4 levels depth which avoided over-fitting problem. MLP in stage-1 used scaled inputs which gave highly optimized results. The complexity parameter, C for Logistic Regression was set to 100.

5. Evaluation.

Methodology

We compared the results of three techniques, Gradient Boosting Classifier, Random Forests and Two-Stage Classifier for the classification of breast cancer. The details of the used datasets, metrics for evaluation and implementation are explained in the below sections.

Data Sets

Experiments are performed on two different datasets of Diagnostic Wisconsin Breast Cancer Database. The datasets are present in the UCI repository [15]. These datasets are described below.

Dataset 1 - WDBC (Wisconsin Diagnosis Breast Cancer)

This dataset contains 569 instances and 10 main attributes for each instance. Mean, standard error and worst value were calculated for each of the features and finally, inputs with 30 attributes were fed into the algorithms. Digitized image of a Fine Needle Aspirate (FNA) was obtained and features were calculated from

that the image. The attributes present the characteristics of the cell nuclei in the digitized image [16]. Each of the instance is classified as either 'malignant' or 'benign'. The list of the attributes [16] is given below in table I.

TABLE 1: FEATURES OF DATASET 1

Features
Radius (mean of distances from center to points on the perimeter)
Texture (standard deviation of gray-scale values)
Perimeter
Area
Smoothness (local variation in radius lengths)
Compactness (perimeter^2 / area - 1.0)
Concavity (severity of concave portions of the contour)
Concave points (number of concave portions of the contour)
Symmetry
Fractal dimension ("coastline approximation" - 1)

Dataset 2 - Wisconsin Breast Cancer

This data set consists of 699 instances where each instance has 9 features. 16 instances were deleted from the data since they had incomplete values for some features. The below table shows the descriptive statistics of this dataset [8].

TABLE 2: DESCRIPTIVE STATISTICS OF DATASET 2

Attribute	Min	Max	Mean	StdDev
Clump Thickness	1	10	4.418	2.816
Cell Size Uniformity	1	10	3.134	3.051
Shape Uniformity	1	10	3.207	2.972
Marginal Adhesion	1	10	2.807	2.855
Single Epi Cell Size	1	10	3.216	2.214
Bare_Nuclei	1	10	3.545	3.644
Bland_Chromatin	1	10	3.438	2.438
Normal_Nucleoli	1	10	2.867	3.054
Mitoses	1	10	1.589	1.715
Class	1	10	Benign (458)	Malignant (241)

Evaluation Metrics

For evaluation, we took the metrics accuracy and percentage error reduction. Accuracy is defined below:

$$\text{Accuracy} = \frac{\text{True\_Predictions}}{\text{Total\_Predictions}} \quad (2)$$

It shows the percentage of correct outputs. Percentage Error Reduction is defined as below:

$$\text{Error Reduction} = \frac{\text{Accuracy}_{\text{New}} - \text{Accuracy}_{\text{Old}}}{\text{Percentage\_Error}_{\text{Old}}} \quad (3)$$

Which shows the percentage of reduction in error w.r.t to the previous occurred error by using our proposed classifier [29, 30].

Experiments And Results

The methods, Gradient Boosting Classifier, Random Forests and the two-stage classifier were tested over the presented datasets in section V. The input data distribution for training and testing was kept the same i.e.,

66.66% for training and 33.33% for testing for each of the techniques. The evaluation metrics were calculated and are presented in the below sections.

Accuracy: We calculated the accuracies of the techniques for the training set as well as for the test set which helped us to identify the overfitting phenomenon. The final accuracies, after the removal of overfitting phenomenon, for the testing and training data are presented in the table III and table IV. For Dataset I, the Two Stage Classifier performs better than Random Forests and Gradient Boosting Classifier by 0.5% and 1.2% respectively. For Dataset II, the Two Stage Classifier performs equally well as Random Forests and exceeds Gradient Boosting Classifier in accuracy by 0.5%. Similarly, Random Forests was found to perform better than Gradient Boosting Classifier in terms of accuracy for both the Datasets i.e., by 0.7% for Dataset I and by 0.5% for Dataset II.

TABLE II
PERCENTAGE ACCURACIES USING DATASET 1

Data Set 1		
Algorithm	**Accuracy (Test Set)**	**Accuracy (Training Set)**
Random Forests	97.9%	100%
Gradient Boosting Classifier	97.2%	99.5%
Two Stage Classifier	98.4%	98.7%

TABLE IV
PERCENTAGE ACCURACIES USING DATASET 2

Data Set 2		
Algorithm	**Accuracy (Test Set)**	**Accuracy (Training Set)**
Random Forests	97.8%	100%
Gradient Boosting Classifier	97.3%	96.7%
Two Stage Classifier	97.8%	96.7%

From the above results, it is clear that the proposed technique performed better than the other techniques by improving the accuracy by a maximum of 1.2%. Moreover, the performance of Random Forests ranked second among the three techniques.

Percentage Error Reduction

To get a clear picture of the comparison, we have plotted the percentage error reduction charts covering the results for the two datasets. The percentage error reduction presents the amount of error reduced when the Two-Stage Classifier was tested after the other technique. The figure 2 shows the comparison of the Two-Stage Classifier and Gradient Boosting Classifier while the figure 3 shows the comparison of the Two-Stage Classifier and Random Forests. From the figure 2, it can be seen that the Two-Stage Classifier reduced the amount of error introduced by Gradient Boosting Classifier up to at most 50% (almost) and as less as almost 20% which means that almost half of the amount of errors can be removed by our proposed technique when compared to the Gradient Boosting Classifier. Similarly, from the figure 3 it is clear that the Two-Stage Classifier reduced the amount of error introduced by Random Forests by up to 25%. Hence, it is clear that a significant amount of percentage error can be reduced by our proposed algorithm.

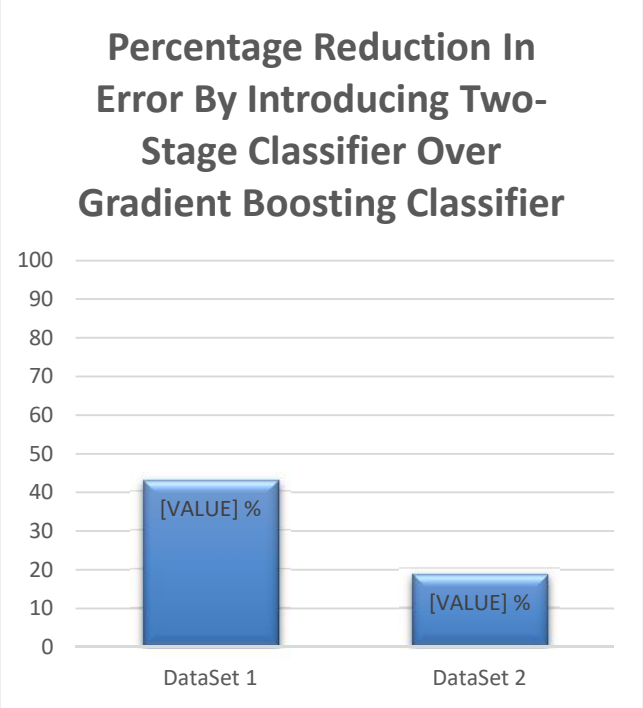

Fig. 2 Percentage Error Reduction: Two-Stage Classifier Vs. Gradient Boosting Classifier

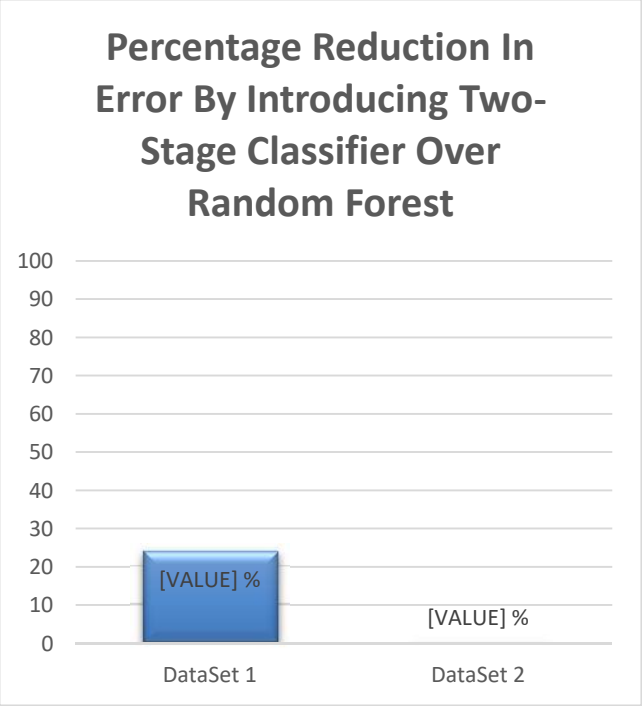

Fig. 3 Percentage Error Reduction: Two-Stage Classifier Vs. Random Forests

Conclusion. Our purpose was to propose a new classifier combining some of the existing machine learning techniques and to make a comparison between the performances of various ensemble techniques with our proposed approach. First of all, our proposed method provides a parallel implementation where multiple techniques can run in parallel in the first stage and the features' count in the second stage is reduced to the

number of algorithms used in the first stage which we do not find in the other two methods. Moreover, our proposed algorithm had a clear advantage in terms of accuracy and percentage error reduction as compared to the Gradient Boosting Classifier and Random Forests. The percentage error reduction by introducing our proposed technique over other methods went as high as 50% (approximately).

REFERENCES

[1] Burks, C. (2016). A COMPARISON OF BREAST CANCER CLINICAL TRIALS IN THE UNITED STATES VS INDIA: ANALYSIS OF REGULATORY REQUIREMENTS.

[2] Barszowski, S. (2016). The comparison of selected legal aspects of insurance products on the Polish and British markets.

[3] Delen, D., Walker, G., & Kadam, A. (2005). Predicting breast cancer survivability: a comparison of three data mining methods. Artificial intelligence in medicine, 34(2), 113-127.

[4] Anupama, Y. K. (2017). Survey on data mining techniques for diagnosis and prognosis of breast cancer. International Journal on Recent and Innovation Trends in Computing and Communication, 5(2), 33-37.

[5] Nashte, C., Nalavade, J., Darvemula, A., Singh, S., & Bhowmick, M. (2017). Differentiation of Malignant and Benign Breast Lesions Using Machine Learning Algorithms. Differentiation, 4(5).

[6] Delen, D., Walker, G., & Kadam, A. (2005). Predicting breast cancer survivability: a comparison of three data mining methods. Artificial intelligence in medicine, 34(2), 113-127.

[7] Gad, W. SVM-Kmeans: Support Vector Machine based on Kmeans Clustering for Breast Cancer Diagnosis.

[8] Gad, W. SVM-Kmeans: Support Vector Machine based on Kmeans Clustering for Breast Cancer Diagnosis.

[9] Delen, D., Walker, G., & Kadam, A. (2005). Predicting breast cancer survivability: a comparison of three data mining methods. Artificial intelligence in medicine, 34(2), 113-127.

[10] Al-Hagery, M. A. H. (2016). Classifiers' Accuracy Based on Breast Cancer Medical Data And Data Mining Techniques. International Journal of Advanced Biotechnology and Research (IJBR), 7.

[11] Hazra, A., Mandal, S. K., & Gupta, A. (2016). Study and Analysis of Breast Cancer Cell Detection using Naïve Bayes, SVM and Ensemble Algorithms. International Journal of Computer Applications, 145(2).

[12] Majali, J., Niranjan, R., Phatak, V., & Tadakhe, O. (2014). Data Mining Techniques For Diagnosis And Prognosis of Breast Cancer. International Journal of Computer Science and Information Technologies (IJCSIT), 5(5), 6487-6490.

[13] Joshi, J., Doshi, R., & Patel, J. (2014). Diagnosis of breast cancer using clustering data mining approach. International Journal of Computer Applications, 101(10), 13-17.

[14] Chaurasia, V., & Pal, S. (2017). Performance analysis of data mining algorithms for diagnosis and prediction of heart and breast cancer disease.

[15] " Karaboga, D., & Ozturk, C. (2011). A novel clustering approach: Artificial Bee Colony (ABC) algorithm. Applied soft computing, 11(1), 652-657.

[16] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... & Vanderplas, J. (2011). Scikit-learn: Machine learning in Python. Journal of machine learning research, 12(Oct), 2825-2830.

[17] Xu, J., Chen, P., Yang, L., Meng, F., & Wang, P. (2017, November). LogDC: Problem Diagnosis for Declartively-Deployed Cloud Applications with Log. In e-Business Engineering (ICEBE), 2017 IEEE 14th International Conference on (pp. 282-287). IEEE.

[18] Martínez-Muñoz, G., & Suárez, A. (2005). Switching class labels to generate classification ensembles. Pattern Recognition, 38(10), 1483-1494.

[19] Lavanya, D., & Rani, K. U. (2012). Ensemble decision making system for breast cancer data. International Journal of Computer Applications, 51(17).

[20] Bäck, T., & Schwefel, H. P. (1993). An overview of evolutionary algorithms for parameter optimization. Evolutionary computation, 1(1), 1-23.

[21] Cong, J., Wei, B., He, Y., Yin, Y., & Zheng, Y. (2017). A Selective Ensemble Classification Method Combining Mammography Images with Ultrasound Images for Breast Cancer Diagnosis. Computational and mathematical methods in medicine, 2017.

[22] Albahri, O. S., Zaidan, A. A., Zaidan, B. B., Hashim, M., Albahri, A. S., & Alsalem, M. A. (2018). Real-Time Remote Health-Monitoring Systems in a Medical Centre: A Review of the Provision of Healthcare Services-Based Body Sensor Information, Open Challenges and Methodological Aspects. Journal of medical systems, 42(9), 164.

[23] Abdulnabi, M., Al-Haiqi, A., Kiah, M. L. M., Zaidan, A. A., Zaidan, B. B., & Hussain, M. (2017). A distributed framework for health information exchange using smartphone technologies. Journal of biomedical informatics, 69, 230-250.

[24] Kalid, N., Zaidan, A. A., Zaidan, B. B., Salman, O. H., Hashim, M., & Muzammil, H. (2018). Based Real Time Remote Health Monitoring Systems: A Review on Patients Prioritization and Related" Big Data" Using Body Sensors information and Communication Technology. Journal of medical systems, 42(2), 30.

[25] Zaidan, B. B., Haiqi, A., Zaidan, A. A., Abdulnabi, M., Kiah, M. M., & Muzamel, H. (2015). A security framework for nationwide health information exchange based on telehealth strategy. Journal of medical systems, 39(5), 51.

[26] Hussain, M., Zaidan, A. A., Zidan, B. B., Iqbal, S., Ahmed, M. M., Albahri, O. S., & Albahri, A. S. (2018). Conceptual framework for the security of mobile health applications on android platform. Telematics and Informatics, 35(5), 1335-1354.

[27] Kiah, M. M., Al-Bakri, S. H., Zaidan, A. A., Zaidan, B. B., & Hussain, M. (2014). Design and develop a video conferencing framework for real-time telemedicine applications using secure group-based communication architecture. Journal of medical systems, 38(10), 133.

[28] Hussain, M., Al-Haiqi, A., Zaidan, A. A., Zaidan, B. B., Kiah, M. L. M., Anuar, N. B., & Abdulnabi, M. (2015). The landscape of research on smartphone medical apps: Coherent taxonomy, motivations, open challenges and recommendations. Computer methods and programs in biomedicine, 122(3), 393-408.

[29] Khan, Y. D., Ahmad, F., & Anwar, M. W. (2012). A neuro-cognitive approach for iris recognition using back propagation. World Applied Sciences Journal, 16(5), 678-685.

[30] Khan, Y. D., Khan, S. A., Ahmad, F., & Islam, S. (2014). Iris recognition using image moments and k-means algorithm. The Scientific World Journal, 2014.