Neural Network Ensemble for the Prediction Of Pathological Complete Response After Neoadjuvant Chemotherapy for Breast Cancer

Raghvi Bhardwaj, Nishtha Hooda

Abstract*: Neoadjuvant Chemotherapy is given intravenously during the treatment of breast cancer. Before the surgery, doctors suggest chemotherapy to deflate the large size of encroaching tumor. This research work propounds Neural Network Ensemble Machine Learning framework, which accomplishes ensemble of Machine Learning algorithms for constructing an systematized solution for anticipating the pathological complete response of the patients after Neoadjuvant Chemotherapy. Performance score is calculated by considering ten evaluation metrics namely, Accuracy, Mean Absolute Error, Root Mean Square Error, TP Rate, FP Rate, Precision, Recall, F-Measure, MCC, and, ROC. The outcomes are verified using K Fold cross validation technique and achieving an accuracy of 97.20%. When the execution of the proposed framework is accumulated with the accomplishment of state-of-the-art classifiers such as Bayes Net, Naïve Bayes, Logistic, Multilayer Perceptron, SMO, Voted Perceptron, etc, the results are quite promising. Machine learning can play a key role in saving lives, in the field of cancer detection.

***Index Terms*: machine learning, neural network, prediction, neoadjuvant chemotherapy.**

I. INTRODUCTION

Breast cancer is the utmost usual type of cancer detected in women all over the world. In India, 23% of breast cancer reports of all the cancers in females. Breast cancer is caused due to lumps in the breasts [1]. Lumps are very common, they occur in young women as they go through puberty and they occur in women even as they enter into their 80s. This becomes more concerned about lumps in the breast in women after menopause because they do not have hormonal changes that would be typically causing benign lumps but in younger women most of those lumps are benign in there is very common lumps called fibroadenomas that occur in young women in their late teens to early 20s [2]. Most of the time for early-stage breast cancer, patient gets chemotherapy if she needs it after surgery. The stage of the breast cancer is the main factor. There is not so much threat that early stage breast cancer will be repeated so it has a more favorable prognosis. Breast cancer detected at a later stage has a huge uncertainity of recurrence, so it has a less suitable prognosis [3].In this research work, various machine learning algorithms are trained after pre-processing the data to check the performance of classification. The execution of the most prominent classifier is also compared with classifiers like SVM, Naïve Bayes, Bayes Net, etc. Instead of just focusing on the accuracy of the proposed framework, ten evaluation metrics are calculated and comprehensive performance score is evaluated using K fold cross validation technique. The ten evaluation metrics namely, Accuracy, Mean Absolute Error, Root Mean Square Error, TP Rate, FP Rate, Precision, Recall, F-Measure, MCC, and, ROC are used to calculate the performance score.

This paper ahead is methodized as follows: Section 2 comprises a brief description of the classification methods that are used in the proposed framework. Section 3 contains the discussion on the data, experimental setup and its features. Section 4 is the summary of the observational results and performance comparison. Ultimately, in Section 5, we discuss the conclusion.

II. RELATED WORK

Researcher has compared proficiency and effectiveness of various stratagems concerning accuracy, precision, sensitivity and specificity to discover the foremost categorization accuracy [4]. In this paper, machine learning for mpMRI of the breast empowers the initial conjecture of the PCR to NAC and consequently may give profitable predictive knowledge to escort medication decisions [5]. An outstanding tendency of big data analytics and machine learning will optimistically construct elevated property confirmation in radiation oncology. Reduction in computer power value, rationalization of EHR, and approaches in artificial intelligence algorithms will instinct alterations in this domain [6]. This study is also foremost because four various ML methods are equated. As a result, the acquired accuracy rate cannot be considered as very huge. This study explored the utility of suchlike data with ML techniques in breast cancer identification [7]. Researcher has explored five DNA viruses in this paper i.e HSV-1, EBV, CMV, HPV, and HHV-8 – influencing the breast cancer discovered by utilizing support vector machines. The results released that the SVM based model has better presentation in detecting

Revised Manuscript Received on June 5, 2019
Raghvi Bhardwaj, M.E. student in Computer Science Engineering Department, Chandigarh University, Mohali, India.
Nishtha Hooda, Assistant Professor at the Department of Computer Science Engineering, Chandigarh University, Mohali, India.

breast cancer in accordance to their dataset [8]. Researcher has mainly focused on the ensemble algorithms in this research. In classification and prediction, a carefully arranged ensemble algorithm generally offer more accuracy and solidity than a single algorithm can attain [9]. Researcher examined the ideas of machine learning while defined their prosecution in cancer prediction/ prognosis. Most of the studies which have been profound the previous years and concentrated on the evolution of predictive models using supervised machine learning methods and classification al- gorithms pointing to forecast authentic disease results [10]. Researcher has studied that after Neoadjuvant Chemotherapy, patients who attain pathological complete response have fundamental improvements in specific survival and comprehensive survival [11]. Researcher has developed the various clinical decision support systems for analyzing gene expression data that can improve the prediction of survival of the patient and cancer patient's prognosis [12]. Researcher has proposed an SVM build weighted AUC ensemble machine learning exemplary for the analysis of the breast cancer. Five fusion programmes are determined to combine the opinions from various base models to differentiate with schemed WAUCE model which reveal that the proposed WAUCE model can be remarkably expand cancer diagnosis performance [13].

III. METHODS

This section summarizes the machine learning classifiers and presents the detailed view of the proposed Neural Network Ensemble Machine Learning framework.

A. Machine Learning Classifiers

i. **Bayes net:** This classifier is Directed Acyclic Graphs (DAG) whose nodes represent variables in the Bayes Net perception. Each node corresponds to a possibility function which takes, as input, a certain integrity for node's established variables, and relinquish the possibility grouping of the variable presented by node [14].

ii. **Naïve bayes:** These classifiers are extremely scalable, requiring a number of parameters straight in the amount of variables in a learning problem [15].

iii. **Logistic:** This is generally the detailed arrangement and execution of a complex action. This is the mainframe of the flow of things amidst the point of provenance and the point of utilization [16].

iv. **Multilayer Perceptron Neural Network:** This is a class of feed forward artificial neural network. A Multilayer Perceptron incorporated, at least, three layers of nodes: an input layer, a hidden layer and an output layer [17].

v. **Support vector machine or SMO:** It is a portrayal of the prototypes as points in space, plotted to create distinct categories, divided by a clear gap. New samples are then plotted in that same area and then forecasted to pertain to a class build on the side of the space they fall [18].

vi. **Voted Perceptron Neural Network:** It is a network or orbit of neurons, or in other words, an artificial neural network, collected of artificial nerve impulses or intersections [19].

B. Proposed Framework

In this research, Neural Network Ensemble Machine Learning framework is designed, implementing optimized version of Machine Learning Algorithm for the prediction of Pathological Response after Neoadjuvant Chemotherapy for Breast Cancer as described in the abstract view in Figure 1.

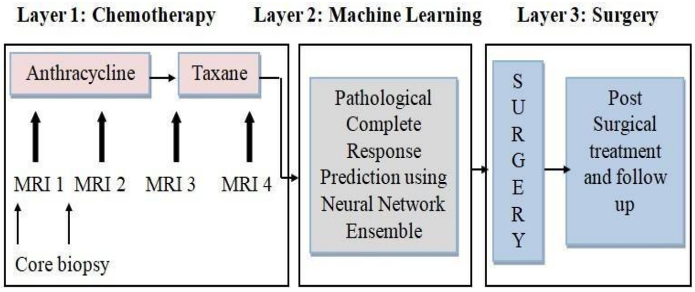
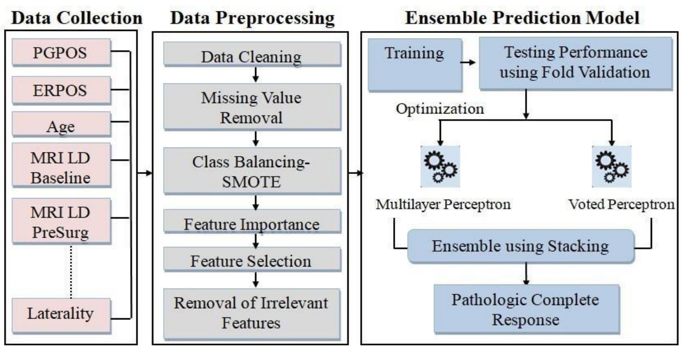


Figure 1. Abstract view of Neural Network Ensemble Machine Learning Framework

The abstract view of Neural Network Ensemble Machine Learning Framework is presented in Figure 1. This complete framework is divided into three layers as explained below:

i. Layer 1: Chemotherapy is the layer 1, in this layer drugs like Anthracycline and Taxane are used in cancer chemotherapy which blocks the cancerous cell growth.

ii. Layer 2: Machine Learning is the second layer in which Neural Network Ensemble Machine Learning framework is implemented for the prediction of Pathological complete response (PCR)



iii. Layer 3: In the layer three, Post Surgical treatment and follow up is done.

Figure 2. Detailed view of Neural Network Ensemble Machine Learning Framework

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

The detailed view of the Neural Network Ensemble framework is presented in Figure 2. The complete framework is split into three segments as explained below:

i. **Data Collection:** The ISPY-1-Trial dataset consist of 222 cases medicated for breast cancer, for which data was derived from the cancer imaging attestation and the breast imaging research project at UCSF. In this data is congregated and computed information on different variables namely, PGPOS, ERPOS, Age, MRI LD Baseline, MRI LD PreSurg and, Laterality, in an esteemed system. These variables then allow to analyze the end results.

ii. **Data Preprocessing:** It is a significant step in Machine Learning projects. In this section, data assembled from ISPY-1-Trial is cleaned. By data cleaning, it means the process of recognizing and improving the inappropriate records from a database or a dataset. After the process of data cleansing, missing values from the dataset is removed. The conceptualization of the missing values is major to perceive in order to successfully oversee the data. Removal of missing values leads to the unbalancing of the dataset, therefore, after the step of missing value removal the balancing of the unbalanced dataset is done using Synthetic Minority Oversampling Technique (SMOTE) which endeavors to set equilibrium in the dataset by constructing synthetic instances. Following this process, miscellaneous important features are selected and the rest of the irrelevant features are removed.

iii. **Prediction Model:** We used nine types of different classification models: Bayes Net, Naïve Bayes, Logistic, Multilayer Perceptron Neural Network, SMO, Voted Perceptron Neural network, etc. By virtue of these models, Training and Testing are performed. Various machine learning classification models are trained to monitor the execution of classification. Testing is performed by using K Fold cross validation technique and the outcomes are demonstrated. By optimizing, two ensemble classification models namely, Multilayer Perceptron, Voted Perceptron, etc, are used to calculate the various performance metrics like Accuracy, TP Rate, FP Rate, ROC. These classifiers successfully achieved satisfactory results. By product of which, we ensembled the two classifiers using the stacking and achieved an accuracy of 97.20%.

IV. EXPERIMENTAL INVESTIGATION

This section describes the dataset and experimental settings of the framework.

A. Dataset

The data of 222 patients which are treated for breast cancer is collected from breast imaging research project at UCSF with the alliance with ACRIN, CALGB, the I-SPY TRIAL and TCIA. The description of every variable and the target class is presented in Table 1.

Table 1. Variable Description

PGPOS	Progestrone Receptor status, pre-treatment
ERPOS	Estrogen receptor status, pre-treatment
AGE	Patient Age
MRI LD Baseline	Massive tumor extent at baseline evaluated by MRI.
MRI LD PreSurg	Massive tumor extent before surgery estimated by MRI.
Laterality	Breast with major or single tumor

B. Experimental Setting

The different model building techniques are implemented using WEKA tool. The main objective isto compute the accuracy of prediction of this classifier. For evaluation of the performance of the proposed framework, various discrete parameters, which are Accuracy, Mean Absolute Error, Root Mean Square Error, TP Rate, FP Rate, Precision, Recall, F- measure, ROC, MCC, and, PRC area, etc are used.

V. RESULTS AND DISCUSSIONS

This section summarizes the execution analysis of the proposed framework and the experimental results validated using K Fold cross validation technique.

A. Performance Evaluation

The proposed framework is tested with definite variables and experimental results are shown in Table 2. The ten evaluation metrics namely, Accuracy, Mean Absolute Error, Root Mean Square Error, TP Rate, FP Rate, Precision, Recall, F-Measure, MCC, and, ROC are used to calculate the performance score. In the prophecy of survival after Neoadjuvant Chemotherapy for breast cancer using ensemble machine learning.

Confusion matrix

A confusion matrix is a table that is usually used to illustrate the execution of a classification model

i. True Positives: True Positive rate refers to the positive tuples that are precisely categorized by the classifier.

ii. True Negatives: True Negative rate are the negative tuples that are precisely categorized by the classifier.

iii. False Positives: False positives are the negative tuples which are inaccurately categorized as positive.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

iv. False Negatives: False negatives are the positive tuples which are mislabeled as negative.

Performance Metrics

i. Accuracy: It is a degree in which the result of calculation observe to the standard value.

ii. Mean Absolute Error: It is a action of difference between two continual variables.

iii. Root Mean Squared Error: Root mean squared error is a measure of differences between predicted values and the actual values.

iv. TP Rate: It refers to the positive tuples which are precisely categorized by the classifier.

v. FP Rate: It refers to the negative tuples which are inaccurately categorized as positive.

vi. Precision: It is an illustration of unusual errors.

vii. Recall: It refers to the percentage of total relevant results accurately categorized by algorithm.

vii. F-measure: It is a measure of a test's accuracy and is defined as the weighted harmonic mean of the precision and recall of the test.

viii. MCC: It has been conceptualized to the multiclass case (for K different classes)

ix. ROC: It is a symbolic plot that embellishes the diagnostic capability of a binary classifier system as its differentiation threshold is diverged.

classifiers	Acc%	MAE	RMSE	TPR	FPR	Precision	Recall	F-M	MCC	ROC
Bayes Net	94.21	0.06	0.15	0.96	0.05	0.96	0.96	0.96	0.92	0.95
Naïve Bayes	95.24	0.03	0.14	0.96	0.04	0.97	0.96	0.97	0.94	0.96
Logistic	90.57	0.08	0.26	0.92	0.11	0.93	0.93	0.92	0.81	0.87
Multilayer Perceptron	93.27	0.04	0.18	0.95	0.05	0.94	0.96	0.95	0.90	0.95
SMO	94.31	0.03	0.17	0.96	0.03	0.96	0.96	0.96	0.92	0.92
Voted Perceptron	71.02	0.27	0.52	0.72	0.73	0.53	0.73	0.51	0.50	0.51
NN Ensemble	97.20	0.05	0.17	0.97	0.05	0.97	0.96	0.97	0.93	0.92

Table 2. Experimental results

In Figure 3, the performance comparison of Neural Network Ensemble Machine Learning framework using Accuracy, True Positive rate, False Positive rate, and, ROC curve with the standard classifiers like Bayes Net, Naïve Bayes, Logistic, Multilayer Perceptron, SMO, Voted Perceptron, etc. is presented graphically. It can be observed that the accuracy of the Neural Network Ensemble Machine Learning framework is the highest. Voted Perceptron has the lower accuracy which is below 70% and multilayer Perceptron has the highest accuracy of 94%. It can be true that the presentation of ensemble classifiers is superior than all the traditional machine learning algorithms, in a similar way ensemble classifiers are performing better in terms of True Positive Rate and False Positive Rate.

In Figure 4, the experimental results of K Fold cross validation testing technique are shown using Accuracy, TP Rate, FP Rate, and, ROC. This figure shows the stability of the various metrics with the Neural Network Ensemble Machine Learning framework. It can be clearly noticed that the proposed framework is quite stable in all ten iterations(folds).

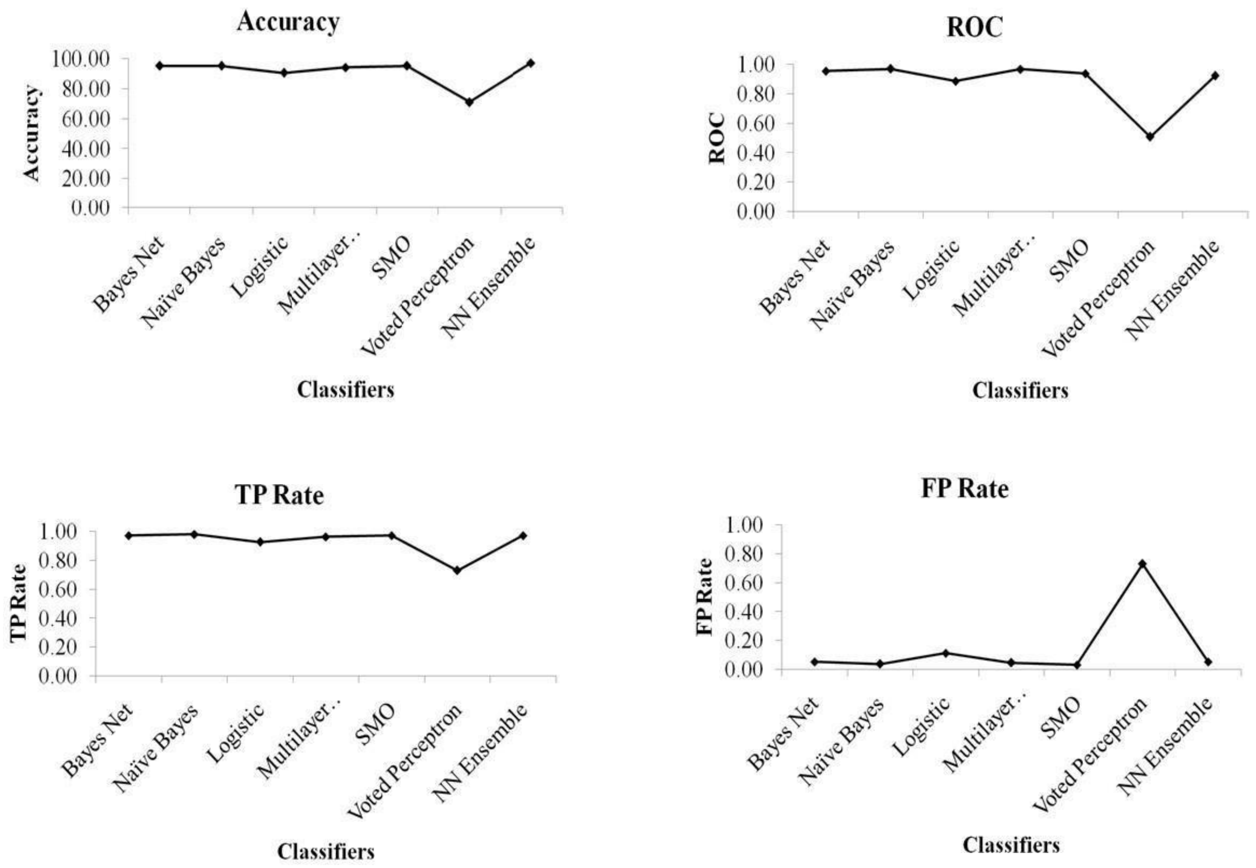


Figure 3. Performance comparison of Neural Network Ensemble ML framework with standard classifiers using Accuracy, ROC, TP Rate, FP Rate

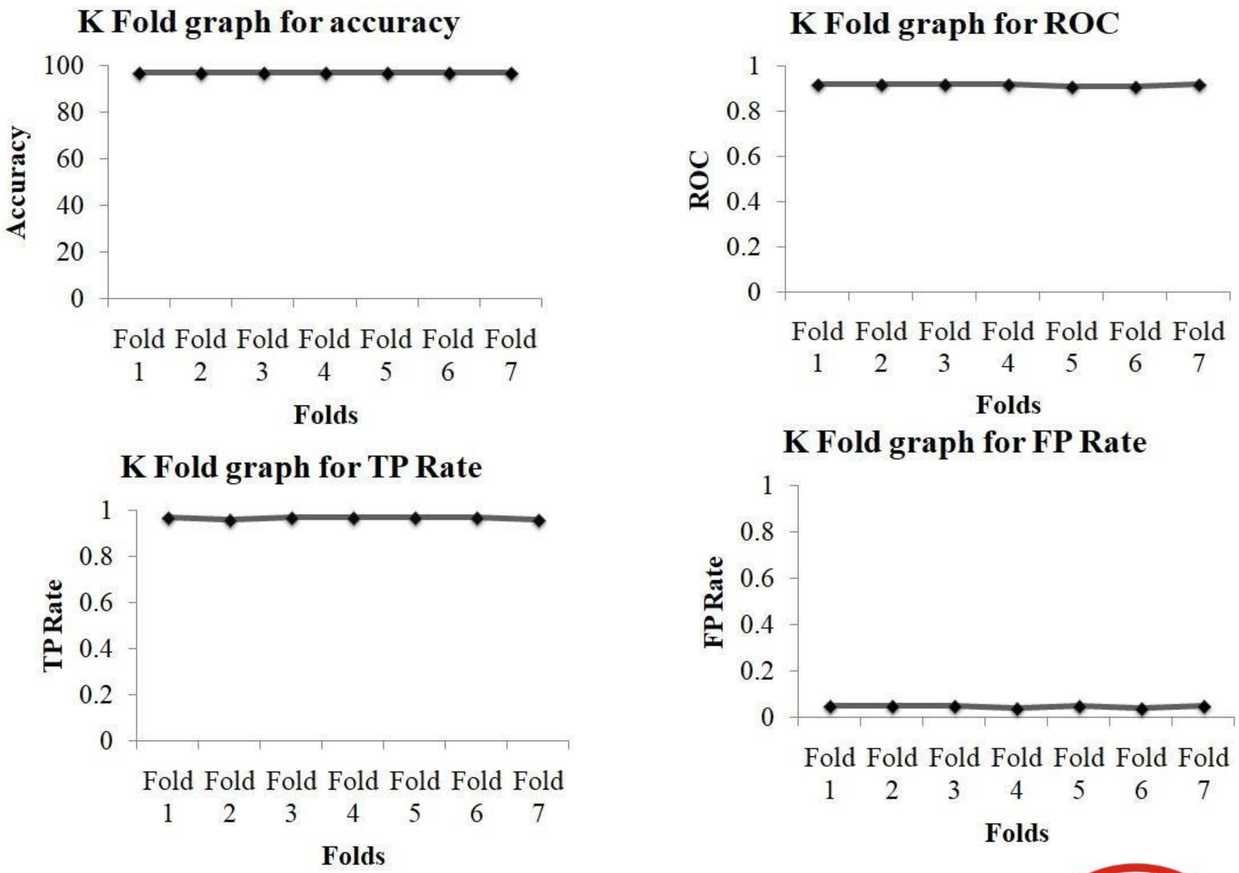


Figure 4. Experimental results of K Fold cross validation testing using Accuracy, ROC, TP Rate, FP Rate

VI. CONCLUSION

In this research work, a coherent ensemble machine learning derived with the name Neural Network Ensemble is propound to speculate the Pathological Complete Response after Neoadjuvant Chemotherapy. For this ensemble establishment, trial and error is executed by combinations of neural network machine learning classifiers, and after miscellaneous iterations, a final ensemble model is constructed for the divination. The logical framework overtops other standard state-of-the-art classifiers by showing an Accuracy, TP Rate, FP Rate, ROC of 97.20%, 0.97, 0.05, 0.92 respectively. The proposed system helps researchers before the surgery, to predict the Pathological Complete Response of the patient at early stage of the cancer.

ACKNOWLEDGMENT

This research work is supported by center of excellence at Chandigarh University, India.

REFERENCES

1. Asano Y et al. (2017). Prediction of survival after neoadjuvant chemotherapy for breast cancer by evaluation of tumor-infiltrating lymphocytes and residual cancer burden. BMC cancer. 17(1):888
2. BORCHERT G.H. et al. (1997). Elevated levels of prostate-specific antigen in serum of women with fibroadenomas and breast cysts. Journal of the National Cancer Institute, 89(8), pp.587-588.
3. Haffty B. et al (2011). Meta-analysis confirms achieving pathological complete response after neoadjuvant chemotherapy predicts favourable prognosis for breast cancer patients. European journal of cancer, 47(14), pp.2084-2090.
4. Asri H. et al. (2016). Using machine learning algorithms for breast cancer risk prediction and diagnosis. Procedia Computer Science, 83, pp.1064-1069.
5. Tahmassebi, A. et al. (2019). Impact of machine learning with multiparametric magnetic resonance imaging of the breast for early prediction of response to neoadjuvant chemotherapy and survival outcomes in breast cancer patients. Investigative Radiology, 54(2), pp.110-117.
6. Bibault J.E., Giraud P. and Burgun A., (2016). Big data and machine learning in radiation oncology: state of the art and future prospects. Cancer letters, 382(1), pp.110-117.
7. Aslan M.F. et al. (2018). Breast cancer diagnosis by different machine learning methods using blood analysis data. International Journal of Intelligent Systems and Applications in Engineering, 6(4), pp.289-293.
8. Huang C.L., Liao H.C. and Chen M.C., (2008). Prediction model building and feature selection with support vector machines in breast cancer diagnosis. Expert Systems with Applications, 34(1), pp.578-587.
9. Yang, P. et al. (2010). A review of ensemble methods in bioinformatics. Current Bioinformatics, 5(4), pp.296- 308.
10. Kourou K. et al. (2015). Machine learning applications in cancer prognosisand prediction. Computational and structural biotechnology journal, 13, pp.8-17.
11. Cain E.H. et al. (2019). Multivariate machine learning models for prediction of pathologic response to neoadjuvant therapy in breast cancer using MRI features: a study using an independent validation set. Breast cancer research and treatment, 173(2), pp.455-463.
12. Bashiri A. et al. (2017). Improving the prediction of survival in cancer patients by using machine learning techniques: Experience of gene expression data: A narrative review. Iranian journal of public health, 46(2), p.165.
13. Wang, H. et al. (2018). A support vector machine- based ensemble algorithm for breast cancer diagnosis. European Journal of Operational Research, 267(2), pp.687- 699.
14. Pearl Judea (2000). Causality: Models, Reasoning, and Inference. Cambridge University Press. ISBN 0-521- 77362-8. OCLC 42291253.
15. McCallum A. and Nigam K., (1998), July. A comparison of event models for naive bayes text classification. In AAAI-98 workshop on learning for text categorization (Vol. 752, No. 1, pp. 41-48).
16. Syarif A., Yun Y. and Gen M., (2002). Study on multi-stage logistic chain network: a spanning tree- based genetic algorithm approach. Computers & Industrial Engineering, 43(1-2), pp.299-314.
17. Orhan U., Hekim M. and Ozer M., (2011). EEG signals classification using the K-means clustering and a multilayer perceptron neural network model. Expert Systems with Applications, 38(10), pp.13475-13481.
18. Cortes C, Vapnik VN (1995). Support-vector networks. Machine Learning. 20 (3): 273–297. https://doiorg/10.1007/BF00994018.
19. Huang Y.M., Hung C.M. and Jiau H.C., (2006). Evaluation of neural networks and data mining methods on a credit assessment task for class imbalance problem. Nonlinear Analysis: Real World Applications, 7(4),pp.720- 74

AUTHORS PROFILE

Raghvi Bhardwaj is currently a M.E. student in Computer Science Engineering Department, Chandigarh University, Mohali, India. She did her B.Tech in Computer Science and Engineering from Himachal Pradesh Technical University. Her areas of interests are machine learning and data analytics. She can be contacted at

Dr. Nishtha Hooda is an Assistant Professor at the Department of Computer Science Engineering, Chandigarh University, Mohali, India. She has completed her Ph.D. degree from Thapar Institute of Engineering and Technology, Patiala, India. She received her Master of Engineering degree in Software Engineering from Thapar Institute of Engineering and Technology, Patiala, India. She did her B.Tech in Information Technology from Kurukshetra University, Haryana, India. Her areas of interests are Big Data, Data Analytics and Machine Learning. She can be contacted at