2020 23^{rd} International Conference on Computer and Information Technology (ICCIT), 19-21 December, 2020

Breast Cancer Detection using Deep Neural Network

Md. Omaer Faruq Goni
Department of Electrical and Computer Engineering
Rajshahi University of Engineering and Technology
Rajshahi-6204, Bangladesh
omaerfaruq0@gmail.com

Fahim Md. Sifnatul Hasnain
Department of Electrical and Computer Engineering
Rajshahi University of Engineering and Technology
Rajshahi-6204, Bangladesh
sifnatul2475@gmail.com

Md. Abu Ismail Siddique
Department of Electrical and Electronics Engineering
Rajshahi University of Engineering and Technology
Rajshahi-6204, Bangladesh
saif101303@gmail.com

Oishi Jyoti
Department of Electrical and Computer Engineering
Rajshahi University of Engineering and Technology
Rajshahi-6204, Bangladesh
oj.ruet@gmail.com

Md. Habibur Rahaman
Department of Electrical and Computer Engineering
The University of Alberta
Edmonton,Canada
mdhabibu@ualberta.ca

2020 23rd International Conference on Computer and Information Technology (ICCIT) | 978-1-6654-2244-4/20/$31.00 ©2020 IEEE | DOI: 10.1109/ICCIT51783.2020.9392705

***Abstract*—In recent years, the number of breast cancer patients has increased rapidly. To increase the survival rate, the early diagnosis of breast cancer is very important. Hence, a trustworthy diagnosis and detection process is required. Automatic detection processes will be very helpful for medical practitioners. There are numerous proposed methods for timely sensing of breast cancer This study has suggested a deep neural network with feature selection techniques to predict breast cancer. The appraisal of the suggested strategy is performed on different evaluation benchmarks like train accuracy, test accuracy, precision, recall, specificity, sensitivity, f measure and MCC. Simulation result of the proposed method is very promising (accuracy 99.42%). Based on the findings of experimental simulations and analysis of the statistical data, the suggested method is efficient, beneficial and more accurate to predict breast cancer. As stated by the presented literature review, It is like setting an example among the prevailing machine learning strategies.**

***Index Terms*—Deep Neural Network, Wisconsin Breast Cancer (Diagnostic), Machine Learning**

I. Introduction

The amount of death due to breast(carcinoma) cancer around the world increases rapidly. As per the statistical review of the world health organization (WHO), 1.5 million women are affected around the world by carcinoma cancer [1]. In Bangladesh, due to less awareness, the number of deaths due to this cancer is high [2]. Data technologies are developed day by day, and the amount of data increases day by day. Among those data, find the most relevant data and classify those become more difficult for medical practitioners [3]. Thus, a well-developed, trustworthy and efficient automatic classification process can reduce the difficulties of medical practitioners, reduce the hassle of the patient, and decrease the death rate.

Numerous researchers proposed different machine learning approaches to make predictions accurately. In the present study, a feature selection technique is applied to finding the most relevant attributes to predict more accurately with less amount of time. A DNN with selected attributes is proposed. A comparative study with existing ML approaches on different performance metrics is performed. It shows that DNN gives the highest performance (100% and 99.42% train-test accuracy). There is less bias and variance, so the DNN model is stable.

II. Literature Review

There are frequent researches on breast cancer diagnosis based on modern technologies such as machine learning, fuzzy system, genetic algorithm.

In [4], authors have compared the performance between two MLP algorithms (Levenberg-Marquardt (LM), along with Scaled Conjugate Gradient (SCG)) to classify the WDBC dataset. This analysis showed that LM, and SCG have 95%, and 93% maximum accuracy.

In research [5], authors have mentioned a comparative study of SVM, and Multilayer Perceptron (MLP) with principal component analysis (PCA) technique to reduce the dimension of the dataset to diagnosis carcinoma cancer on Wisconsin Original Breast Cancer (WOBC) dataset.

In study [6], authors have represented a hybrid method based on K-means along with SVM algorithms to classify benign together with malignant tumours on WDBC dataset. The proposed method has obtained 97.38% accuracy.

In paper [7], authors have proposed a dimension reduction technique based on independent component analysis (ICA) that applied on WDBC data set by ANN, KNN, radial basis function neural network (RBFNN), and SVM. Using this technique, ANN, KNN, SVM, and RBFNN have obtained 97.53%, 91.03%, 95.25%, and 90.49% accuracy respectively. Research [8] have illustrated a comparative survey of six algorithms on WDBC dataset that includes Gated Recurrent Unit (GRU) and SVM ensemble model, Linear Regression, MLP, Nearest Neighbor (NN) search, Softmax Regression, and SVM and MLP has shown the highest accuracy of 99.04%. Paper [9] has represented a performance analysis of SVM, KNN, and Decision tree (DT) to categorize the WDBC data set. SVM, KNN along with DT have obtained 97.9%, 96.7%, and 93.7% accuracy respectively.

III. Methods & Materials

Classification is a very important task. Experimentation has been performed according to the workflow in Fig. 1

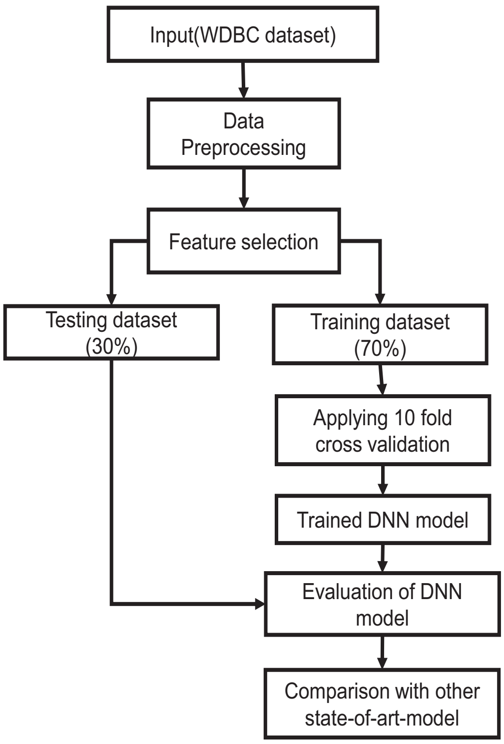

Fig. 1. Proposed methodology

A. Dataset

This study is conducted on Wisconsin Breast Cancer (Diagnostic) dataset. [10] it contains a total of 569 instances with 32 columns of real-valued feature, among them benign carcinoma cancer has 357 (62.7%) instances, and malignant carcinoma cancer has 212(37.3%) instances ie. benign count is 1.684 times more than malignant. It is an unbalanced dataset. There are no features in the dataset that have missing value, and hence it does not require any imputation techniques.

B. Feature Selection

Feature selection has an immense effect on the performance of a classifier. It is an automatic or manual process to choose the best feature that gives more and non-redundant information [11]. By withdrawing the irrelevant feature, it improves the efficiency of the learning task, prediction accuracy and reduces the problem of over-fitting [12], [13].

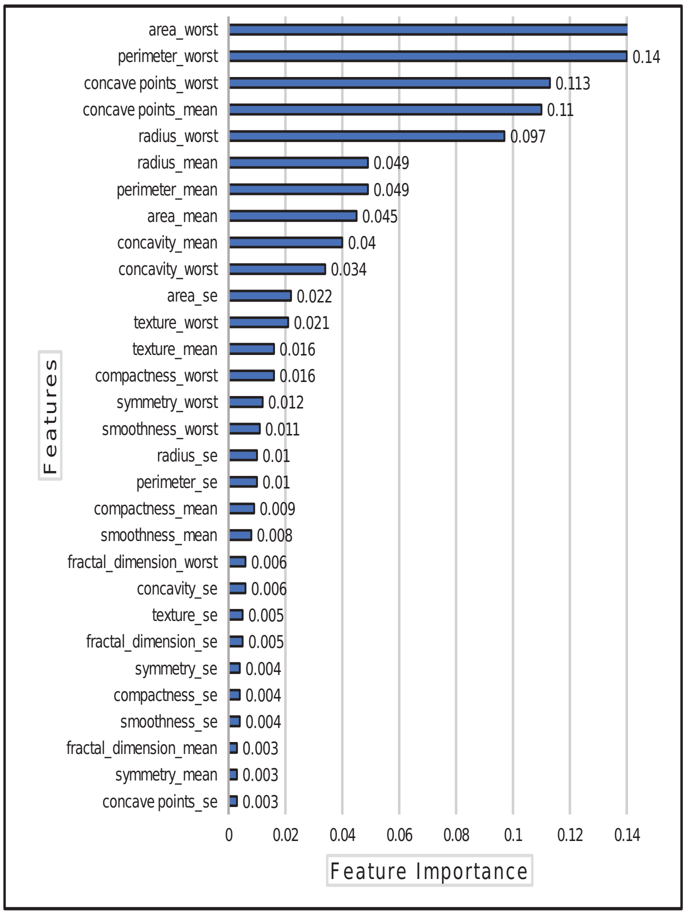

Fig. 2. Graphical representation of feature importance

One of the feature subset selection technique, choose the most important features accordingly and check whether it is best or not. In this study, feature importance is calculated using random forest (Fig. 2).To get the best feature subset, based on feature importance multiple feature subsets are created and tested. Among them, the best performing one is selected. In the experiments and results, the process have been discussed.

C. Deep Neural Network

A deep neural network (DNN) is an artificial neural network (ANN) with multiple hidden layers between the input and output layer. DNN is a computational network stimulated by biological neurons [12]. Fig. 3 represents a DNN. It can create a linear or complex non-linear mathematical model between input and output. Just like the human brain, DNN can be trained to do some tasks on pragmatic knowledge. It can be trained for classification, recognition and regression problems.

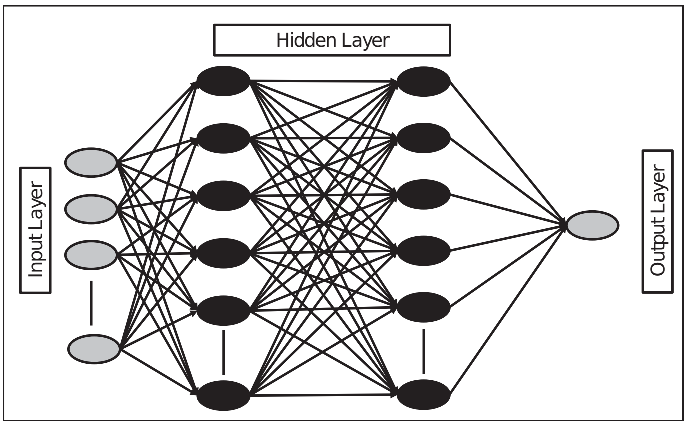

Fig. 3. An example of a DNN model

D. Performance Metrics

The mathematical expression of each performance metrics is given below.

$$Accuracy = \frac{(TP + TN)}{(TP + TN + FP + FN)} \times 100\% \quad (1)$$

$$Sensitivity = \frac{TP}{(TP + FN)} \times 100\% \quad (2)$$

$$Specificity = \frac{TN}{(FP + TN)} \times 100\% \quad (3)$$

$$Precision = \frac{TP}{(TP + FP)} \quad (4)$$

$$Recall = \frac{TP}{(TP + FN)} \quad (5)$$

$$F - measure = \frac{2 \times (Precision \times Recall)}{(Precision + Recall)} \quad (6)$$

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}} \quad (7)$$

To assess a binary classifier, confusion matrix (Table I) is an important benchmark. Its outcome can be classified into four categories. If a classifier predicts an instance as positive and actual prediction is also positive then it is called true positive (TP), in contrast, the actual prediction is negative then it is called false positive (FP). On the other hand, a prediction is called true negative (TN) if the prediction and actual result both are negative otherwise it is called false negative (FN).

Accuracy is very sensitive to the class imbalance and f-measure omits the result of true negative [14]. Due to this, we use a promising parameter Matthews correlation coefficient (MCC), that gives importance to all the members of the confusion matrix. It is a vital performance metric for a binary classification [15].

TABLE I
CONFUSION MATRIX

	Classified as Positive	**Classified as Negative**
Actually Positive	**TP**	**FN**
Actually Negative	**FP**	**TN**

IV. EXPERIMENT & RESULTS

A. Experimentation

To carry out the demonstration, we did two things, selection of best feature subset and determination of the required number of hidden layers and nodes. For the best feature subset selection, we selected features according to the feature importance given in Fig. 2 and calculated the accuracy of few ML methods. The accuracies are given in table II when significant changes occur.

TABLE II
COMPARISON OF ACCURACY OF THE CLASSIFIERS WITH SOME FEATURE SUBSET

No. of Features	**KNN**	**RF**	**LR**	**SVM**	**NB**	**DNN**
2	0.94	0.95	0.93	0.93	0.94	0.94
5	0.95	0.95	0.94	0.95	0.94	0.95
17	0.96	0.95	0.98	0.98	0.92	0.98
19	**0.95**	**0.98**	**0.98**	**0.98**	**0.92**	**0.99**
23	0.96	0.95	0.98	0.98	0.91	0.99

We tested four combinations to decide the best architecture of DNN. Table III represents the combinations and corresponding accuracy. All the models were trained and tested with 10-fold cross-validation on UCI WDBC dataset. The experiment was performed using python, popular python modules for data mining, TensorFlow 2.0 and Google Colaboratory.

TABLE III
COMBINATIONS FOR DNN WITH CORRESPONDING ACCURACY

No. of hidden layers	**No. of hidden nodes**	**Accuracy**
2	50	97.66
2	60	98.25
3	50	98.83
3	**60**	**99.42**

B. Result Analysis

In the present study, we have implemented a DNN model with feature selection on WDBC dataset from UCI to classify malignant and benign. The best outcome of DNN has been achieved using 19 selected feature subset with 3 hidden layers and 60 hidden units in each layer. The proposed method has shown 100% train and 99.42% test accuracy. It is the highest accuracy among the existing methods. In terms of bias-variance trade-off, it is also a stable model. The suggested model has accomplished sensitivity and specificity scores of 97.05% and 100% accordingly.

TABLE IV
COMPARISON BETWEEN THE DNN MODEL AND OTHER CLASSIFIERS IN TERMS OF PERFORMANCE METRICS.

Classifier	Train Accuracy	Test Accuracy	Sensitivity	Specificity	f- measure	MCC	Precision	Recall
Soft Voting	98.99	98.83	100	96.83	0.991	0.975	0.982	1
KNN	97.74	95.32	96.3	93.65	0.963	0.899	0.963	0.963
RF	100	97.66	97.22	98.41	0.981	0.95	0.991	0.972
LR	98.24	97.66	99.07	95.24	0.98	0.95	0.97	0.99
SVM	98.92	98.25	99.07	96.83	0.986	0.962	0.982	0.991
NB	93.96	91.81	92.59	90.48	0.935	0.825	0.943	0.926
ET	97.98	95.32	98.15	90.48	0.964	0.899	0.946	0.981
AdaBoost with SVM	98.92	98.25	99.07	96.83	0.986	0.962	0.982	0.991
AdaBoost with DT	99.75	98.25	100	95.24	0.986	0.963	0.973	1
AdaBoost with NB	98.99	98.83	98.15	100	0.991	0.975	1	0.981
AdaBoost with LR	97.74	97.08	98.15	95.24	0.977	0.937	0.972	0.981
Bagging with DT	99.25	98.25	98.15	98.41	0.986	0.963	0.991	0.981
Bagging with KNN	97.99	96.49	98.15	93.65	0.972	0.924	0.964	0.981
Bagging with LR	96.23	95.32	99.07	88.89	0.964	0.9	0.939	0.991
Bagging with SVM	98.24	97.66	98.15	96.83	0.981	0.95	0.981	0.981
Bagging with NB	93.97	91.81	92.59	90.48	0.935	0.825	0.943	0.926
DNN	**100**	**99.42**	**99.07**	**100**	**0.995**	**0.987**	**1**	**0.991**

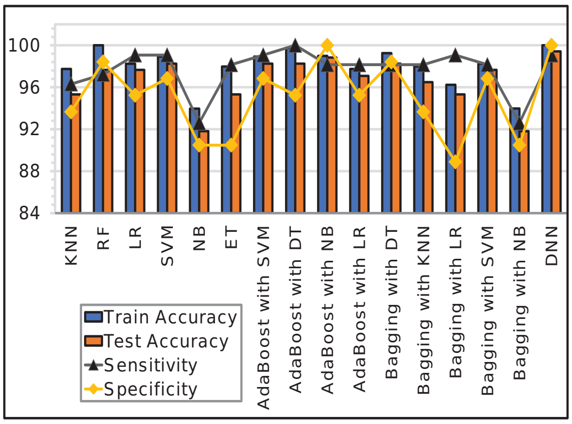

Fig. 4. Classifiers comparison in respect of Train-Test accuracy, Sensitivity, and Specificity

Fig. 4 represents the classifiers comparison in respect of Train-Test accuracy, Sensitivity, and Specificity graphically.

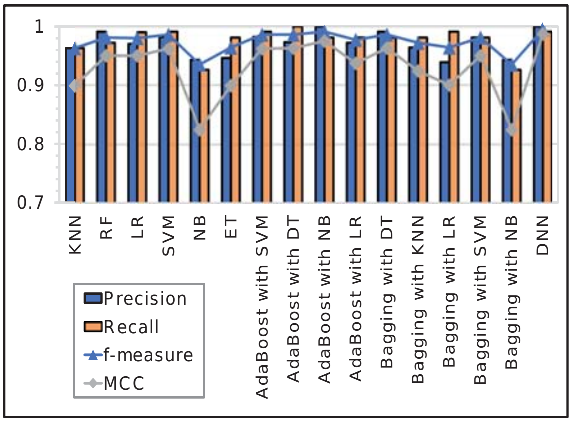

Fig. 5. Classifiers comparison in respect of F-measure, MCC, Precision, and recall.

The proposed model has achieved the highest f-measure and MCC score of 0.995 and 0.987 accordingly. Precision & recall achieved by this model are 1 and 0.991 respectively. Fig. 5 represents the classifiers comparison in respect of F-measure, MCC, Precision, and recall graphically. Table V represents the comparison of the proposed DNN model and the previous works.

TABLE V
COMPARISON OF PROPOSED DNN MODEL AND EXISTING METHODS.

Reference No.	Methods	Accuracy
Ref. 4	LM	95%
	SCG	93%
Ref. 6	Hybrid model of k-means & SVM	97.38%
Ref. 7	ANN	97.53%
	KNN	91.03%
	SVM	95.25%
	RBFNN	90.49%
Ref. 8	MLP	99.04%.
Ref. 9	SVM	97.90%
	KNN	96.70%
	DT	93.70%
-	**Proposed DNN model**	**99.42%**

In conformity with the statistical analysis presented above on this section, it can be said that the proposed DNN model has shown an optimistic result. It can classify breast cancer patients more accurately compared to other existing methods.

V. CONCLUSION

Automatic diagnosis and detection systems can do extraordinary help for medical practitioners. Advanced detection and accurate prediction can play a great role to reduce the death rate by dint of breast cancer. This paper represents a vigorous DNN model with feature selection techniques that shows a better result than other state-of-art models (like SVM, KNN, RF etc.) for WDBC dataset from UCI machine learning repository. It also outperforms the existing models accessible in the relevant research papers. For breast cancer diagnosis, the proposed model is very realistic and efficacious according to empirical outcomes & statistical assessment.

REFERENCES

[1] M. A. Mohammed, B. Al-Khateeb, A. N. Rashid, D. A. Ibrahim, M. K. Abd Ghani, and S. A. Mostafa, "Neural network and multi-fractal dimension features for breast cancer classification from ultrasound images," *Computers & Electrical Engineering*, vol. 70, pp. 871–882, 2018.

[2] S. F. Bellah, M. A. Salam, M. R. Karim, M. J. Hossain, and M. Ashrafudoulla, "Epidemiology of breast cancer among the female patients in bangladesh," *Oriental Pharmacy and Experimental Medicine*, vol. 16, no. 2, pp. 85–95, 2016.

[3] B. Gayathri, C. Sumathi, and T. Santhanam, "Breast cancer diagnosis using machine learning algorithms-a survey," *International Journal of Distributed and Parallel Systems*, vol. 4, no. 3, p. 105, 2013.

[4] N. Mohamad, F. Zaini, A. Johari, I. Yassin, and A. Zabidi, "Comparison between levenberg-marquardt and scaled conjugate gradient training algorithms for breast cancer diagnosis using mlp," in *2010 6th International Colloquium on Signal Processing & its Applications*, pp. 1–7, IEEE, 2010.

[5] S. Ghosh, S. Mondal, and B. Ghosh, "A comparative study of breast cancer detection based on svm and mlp bpn classifier," in *2014 First International Conference on Automation, Control, Energy and Systems (ACES)*, pp. 1–4, IEEE, 2014.

[6] B. Zheng, S. W. Yoon, and S. S. Lam, "Breast cancer diagnosis based on feature extraction using a hybrid of k-means and support vector machine algorithms," *Expert Systems with Applications*, vol. 41, no. 4, pp. 1476–1482, 2014.

[7] A. Mert, N. Kılıç, E. Bilgili, and A. Akan, "Breast cancer detection with reduced feature set," *Computational and mathematical methods in medicine*, vol. 2015, 2015.

[8] A. F. M. Agarap, "On breast cancer detection: an application of machine learning algorithms on the wisconsin diagnostic dataset," in *Proceedings of the 2nd International Conference on Machine Learning and Soft Computing*, pp. 5–9, 2018.

[9] O. I. Obaid, M. A. Mohammed, M. Ghani, A. Mostafa, and F. Taha, "Evaluating the performance of machine learning techniques in the classification of wisconsin breast cancer," *International Journal of Engineering & Technology*, vol. 7, no. 4.36, pp. 160–166, 2018.

[10] U. M. Learning, "Breast cancer wisconsin (diagnostic) data set," Sep 2016.

[11] R. Shaikh, "Feature selection techniques in machine learning with python," Oct 2018.

[12] M. Dash and H. Liu, "Feature selection for classification," *Intelligent data analysis*, vol. 1, no. 3, pp. 131–156, 1997.

[13] H. Bentz, M. Hagstroem, and G. Palm, "Selection of relevant features and examples in machine learning," *Neural Networks*, vol. 2, no. 4, pp. 289–293, 1997.

[14] L. Sisters, "Matthews correlation coefficient: when to use it and when to avoid it," May 2020.

[15] R. J. Manoj, M. A. Praveena, and K. Vijayakumar, "An aco–ann based feature selection algorithm for big data," *Cluster Computing*, vol. 22, no. 2, pp. 3953–3960, 2019.