Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online): 0493-2137
E-Publication: Online Open Access
Vol:55 Issue:07:2022
DOI10.17605/OSF.IO/9FYTS

EARLY DETECTION AND CLASSIFICATION OF BREAST CANCER USING MACHINE LEARNING AND DEEP LEARNING TECHNIQUES

SYED PERVEZ HUSSNAIN SHAH,
Ph.D. Computer Science from Superior University Lahore, Pakistan, and working as Lecturer IT at Lahore Leads University, Pakistan.

Dr. ARFAN JAFFAR,
Dean Computer Science and Information Technology Superior University Lahore, Pakistan.

SARFRAZ NAWAZ,
Ph.D. (Scholar) Computer Science, Superior University, Lahore, Pakistan, and working as Lecturer CS at Govt. Graduate College, Kamoke, Gujranwala.

WISHAL ARSHAD,
Ph.D. (Scholar) Computer Science, Superior University Lahore, Pakistan.

MUHAMMAD IZHAR,
Ph.D. Scholar computer science, Superior University Lahore, Pakistan and working as a Subject Specialist (CS) at Govt. Higher Secondary School. Rajunpur, Pakistan.

AMNA IQBAL,
Ph.D. (Scholar) Computer Science, Superior University Lahore, Pakistan

Abstract

In this paper, the researchers have empirically used the various Machine Learning (ML) and Deep Learning (DL) Algorithms and analyze the findings of different Machine Learning and Deep Learning algorithms on the very well-known Wisconsin Diagnostic Breast Cancer Data-Set (WDBC). This study assessed the degree of their capacity to accurately order the sample images as "malignant" or "benign". The separate utilizing of these algorithms was decided on the grounds of different assessment measurements as accuracy is the main factors of datasets. From the trial results, we rational that the deep learning approaches have given preferable outcomes on assessment grounds over the machine learning algorithms. In quantitative terms, CNN performed most reliably among every one of the considered methodologies for the given breast cancer dataset with an accuracy of CNN Deep Learning Model is 99.48 % and MLP 99.45% individually. The ML algorithm SVM has the betters testing accuracy 97.13% and 98.36% training accuracy. In the consequences of finding breast cancer can be predicted on early basis using the Machine Learning and or Deep Learning Models effectively and efficiently. Early detection of breast cancer (BC) will treat well and save many breast cancer patients. As a result, the BC patient's rate and death rate can be reduced.

Keywords: Breast Cancer detection, machine learning, deep learning algorithms, classifiers, cancer prediction, convolutional neural network, AlexNet, features extraction, Wisconsin dataset. Benign and malignant images.

Nomenclature

WDBC	Wisconsin Diagnostic Breast Cancer
BC	Breast Cancer
ML	Machine Learning
CNN	Convolutional Neural Network
MR	Mean Radius
DL	Deep Learning
FNA	Finite Needle Aspirate
SVM	Support Vector Machine
NB	Naïve Bayes
M	Malignant
B	Benign
TP	True Positive
TN	True Negative
FP	False Positive
FN	False Negative
DT	Decision Tree
LR	Logistic Regression
KNN	K-Nearest Neighbor
MLP	Multi-Layer Perceptron
RF	Random Forest
WHO	World Health Organization
MRI	Magnetic Resonance Imaging

INTRODUCTION

Breast Cancer

Abnormality in breast tissue is referred as Breast Cancer (BC). BC is the cause of mortality, according to the WHO [1]. The BC is the highest common cancer that affects women. The chance of rescuing a BC patient is highly dependent on the position of the first stage and the start of the timely treatment. The highest death rate from the BC is due to a lack of understanding about the early identification of indicators of the different forms of breast cancers and a lack of preparedness in detecting BC symptoms. Fig. 1 depicts the some types of BC histopathological images [2]

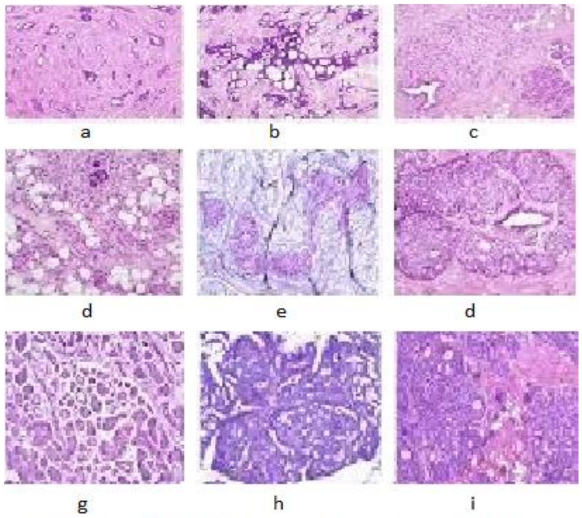

Fig. 1. Histopathological BC images; a) Tubular carcinoma, b) cribriform carcinoma, c) classic invasive lobular carcinoma, d) pleomorphic invasive lobular carcinoma, e) mucinous carcinoma, f) neuroendocrine carcinoma, g) micro-papillary

1.1 Breast Cancer Data Set

The Wisconsin Breast Cancer Diagnostic Data Set (WDBC) is a feature set derived from digitized images of fine needle aspirates (FNAs) of breast tumors. These features explain the main features of nucleus cells in breast images in the mammary gland regions of breast cancer images.

Digitized FNA images of breast masses are used for characterization. They determine the nuclear features of the imaged cells. The Midwest Society of Artificial Intelligence and Cognitive Sciences' 4th Annual Conference, An extensive search in the space of 1-4 features and 1-3 separation planes was used to find relevant features.

K. P. Bennett and O. L. Mangasarian: "Robust Linear Programming Discrimination of Two Linearly Inseparable Sets," Optimization Methods and program (Software- 1), 1992, 23-34 describe the actual linear program used to produce the separation plane in the 3-D. space. The dataset is available freely and can be down load from the given link. *"https://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+%28Diagnostic%29"*[3] *,and also available at "https://archive.ics.uci.edu/ml/datasets/breast+cancer+wisconsin+(diagnostic)"* [4]

The Table 1 shows the details of the Wisconsin Breast Cancer Diagnostic Data Set (WDBC)

TABLE 1

Description of WDBC Dataset

Instances	569
Attributes	32
Information of Attributes	ID number Diagnosis (M= Malignant, B = Benign) 32 Attributes Columns
Actual Valued Features are calculated for every cell-nucleus	1) Radius (mean of distances from center to points on the perimeter) 2) Texture (The standard deviation SD of gray-scale values) 3) Area 4) Compactness (perimeter^2 / area - 1.0) 5) Perimeter 6) Concave points (number of concave portions of contour) 7) Smoothness (local variation of radius lengths) 8) Concavity (extremity of concave portions of the contour) 9) Symmetry 10) Fractal dimension ("coastline approximation" - 1)
Attribute (Missing)	None
Class Distribution	Benign-357, Malignant - 212
Findings of Dataset	1) No. of predicting field = 2 2) Diagnosis: B = benign and M = malignant 3) All 30 input features may be used to segregate sets linearly. 4) The best prediction accuracy was obtained in 3-D space of the Worst Area, Smoothness and Mean Texture by employing one separation plane. Using repeated 10-fold cross validations, the accuracy was estimated to be 97.5 percent. 5) As of November 1995, Classifier had properly diagnosed 176 new cases.

The computation of these properties is described in depth in several of the works linked above. The standard error, mean and "worst" and the worst (mean of the three greatest values) of these attributes were determined the every image, and resultant in 30 features. The field -3 denotes Mean Radius (MR), field 13 shows Radius SE, and the field-23 denotes Worst Radius, for example. Four significant digits have been recoded for all feature values. This paper is consisted at various sections and sub sections that cover various topics such as a literature review, the research methods used, experiments, results and findings, and lastly, the conclusion.

LITERATURE REVIEW

Breast cancer is the second leading cause of mortality, according to the WHO [1] . Despite the fictitious universe's long-term technical developments, there are a few distinct challenges. Breast cancer is now perhaps the most severe threat to the women all over the world. Breast cancer is defined as the uncontrolled cell growth in the breast tissue. 2.3 million women will be diagnosed with BC worldwide, with 685000 deaths in the year of 2020 . By the end of 2020, breast cancer will be the most common illness, in the previous five years; the 7.8 million women were diagnosed BC disease[5].

Every year, a diverse range of new breast cancers is discovered. Breast cancer-related deaths increased at a 13.82 percent annual rate from 2008 to 2012. Because it is limiting network expansion and confidence, the sharp increase in cancer risk and dying must be treated in a range of ways. If early identification and adequate treatment are delivered at the correct time, breast cancer death rates can be lowered. Recognizing the signs and symptoms of breast cancer early on makes it easier to disclose the disease. Although the exact cause of breast cancer is unknown, research have shown that changes in quality, radiation exposure, modernization of nutrition, heritage, lack of activity, cigarette and potentially alcohol intake are all possible causes [6]

The breast is an expert organ that comprises milk pipes, lobules, greasy tis-sues, areola, and areola. It is tied on the chest dividers by pectorals muscles. The circular shapes in the breast that see milk transportation are known as milk channels. Lobules are a network of canals that link the milk-producing glandular structures. The oily tissue's primary function is to maintain the breast's energy balance. The areola is a non-greasy tissue that forms the breast's final milk move point. Areola is the pigmented skin that surrounds the areola and is implanted with Montgomery organs. Breast tumors of various types have an initial impact on milk conduits, lobules, and areolas. The breast cancer in this case exemplifies the underlying detrimental consequences [7]

Breast cancers are commonly divided into two types: benign and malignant. The most prevalent breast cancer types are depicted in Fig. 2.

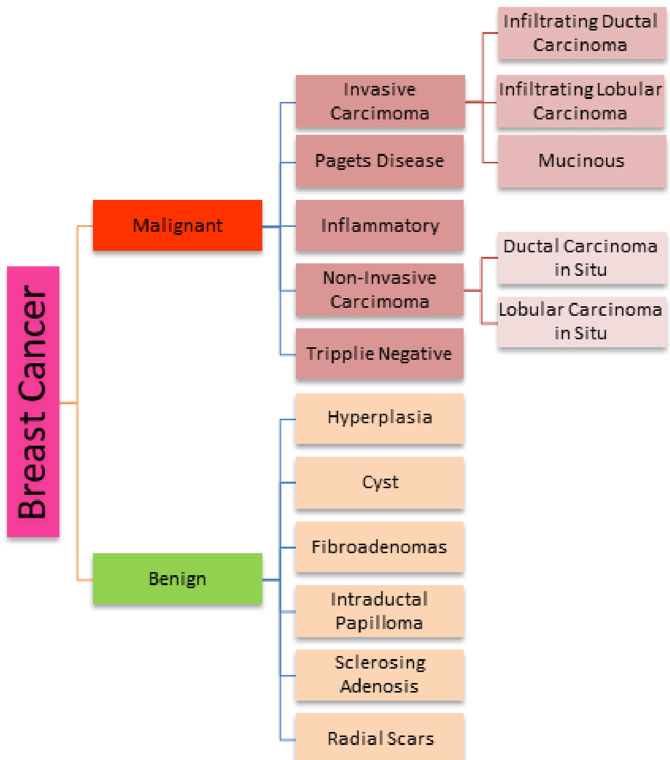

Fig. 2 Shows the Diversities of Breast Cancers as major is Benign and Malignant

Inflammatory breast cancer, which inhibits the lymph nodes, is an exceptional and severe type of breast cancer disease. This kind of BC is characterized by edema, redness, and pinkish skin produced by fluid build-ups. The nipple ducts typically impact it, and it covers the areola. The capacity to distinguish between heterogeneous breast tumors will be required for early identification and administration of suitable therapy. This will necessitate pinpointing the tumor's specific location, size, and meta-sizes. The most frequent way for achieving this goal is BC staging [8]. Cancer has spread to the lungs, bone, kidney, liver, brain, and other important internal organs. Despite the fact that it is incurable, timely treatment can save patients' lives. After two weeks, a self-breast examination (BSE) should be advised.

To detect breast cancer in its early stages and preserve a patient's life, a good pre-diagnosis method is required. Breast cancer is diagnosed using the gold standard procedure of biopsy. It's a needle operation that removes a tiny sample of tumors for further analysis and the imaging methods aid in the avoidance of unnecessary biopsies. Imaging methods assist the radiologist in determining the stage of breast cancer.

According to the findings of the experiments, SVM has the best accuracy (97.13 percent) and the lowest error rate. All of the tests are run in a simulated environment using the WEKA data mining tool[9]. The most recent implementation of the model on the training dataset yielded a 99.67% accuracy. With the Wisconsin Breast Dataset, researchers demonstrated that deep learning technologies may help diagnose breast cancer[10].

For hard voting, the authors utilized the normal of probabilities, result of probabilities, the limit of probabilities, and least of probabilities voting mechanisms, and for delicate voting, they utilized the larger part-based voting system. Hard voting (larger part-based voting) outflanks the cutting-edge WBCD algorithm, with a 99.42 percent achievement rate. [11]. This study proposed a hybrid approach for breast cancer diagnosis that includes using linear discriminant analysis (LDA) to reduce the high dimensionality of characteristics and then using SVM to apply the new reduced feature dataset. According to the receiver operating characteristic (ROC) curve, the recommended approach has 98.82 percent accuracy, 99.07 percent specificity, 98.41 percent sensitivity, and a 0.99 percent area [12].

Researchers adopt a data-level strategy, which entails resampling the data to reduce the impact of class imbalance. A tenfold cross-validation procedure is used to assess the results.[13]. The use of SVM and the Bayesian classifier in giving diagnostic and prognostic information for breast cancer has been explored in this research. The SVM is superior to the Naive Bayes in terms of prediction accuracy [14]. Authors will be able to categorize and forecast whether a cancer is benign or malignant using machine learning techniques [15].

The acclaimed hybrid deep -features (AlexResNet+) based model has the ultimate classifying accuracy of 95.87 percent, specificity 0.9621, precision 0.9760, sensitivity 1.0, AUC of 0.960 % , and F-Measure 0.9878, when tested on DDMS mammography BC tissue images [16]. BC recognition suggests that CNN can achieve better detection rate than hand-crafted feature descriptors, but at the cost of increased system complexity, which requires more training time and particular knowledge to fine-tune the CNN's architecture [17].

The results demonstrate that all of the given ML algorithms scored well on classification job as above 90% test accuracy). With a 99.04 percent test accuracy, and MLP algorithm distinct among the applied algorithms [18]. Breast tumor identification in ultrasonic imaging, ovarian cancer diagnosis, and heart sound diagnostics are all common uses for DT classifiers[19]. SVM is applied in two experiments, and the AlexNet model which is pre trained is fine-tuned on BC images in the third trial. The results of the experiments demonstrate that the fine-tuned Alex-Net improved, with the first experiment's results being superior to the second[20].

ML has emerged as a novel tool for predicting breast cancer survival, and there is still significant possibility for advancement and model development. Current prediction models have limitations due to inadequate of data preparation stages, excessive variations in samples features selection, and validation and advancement concerns [21].

METHODOLOGY

The methodology adopted in this paper to evaluate the performance results of various types of Machine Learning (ML) and Deep Learning (DL) Algorithms on the Wisconsin

Breast Cancer Diagnostics Dataset (WBCD) is first collect the data set and relevant information, study and download concerned research papers, articles, reports and books from the different web sources. Secondly data set is prepared for the processing. Thirdly Machine--learning and Deep—Learning--Algorithms are applied in practical environment in the Jupyter Notebook (Python 3) and Matlab Tool on the Cancer datasets (csv file and image dataset). Features are also extracted using AlexNet for deep learning algorithms. Then the results are compiled and later compared all results with ML and DL classifiers and before and after features extraction. Finally select and predict the best classifier in term of accuracy.

Tools and Equipment

The empirical study is carried out in Jupyter Notebook (Python-3) and Matlab Tool (R2020a) on Dell Laptop Intel I5 with processor (Intel (R) Core (TM) i5-2430M CPU @ 2.4GHz 2.4 GHz), RAM 6GB and 64--bit operating system , x64—based--processor and Windows 10 Professional. All the programs codes are consisted both in Python language and the Matlab codes to produce the desired outcomes.

Flow Diagram of Methodology

The Fig. 2 shows the flow diagram of methodology that is used in this paper. Each step is clearly mentioned in sequential order for better understandings. As starting from identifying the problem then dataset collection and relevant information gathering, and after that review the literature. Then data preprocessing and normalization carried out, and apply ML and DL algorithms. After features extraction of fc6, fc7, and fc8 using AlexNet then applying once again ML Algorithms and producing results. Finally making the recorded results and identifying the best model on the basis of accuracy.

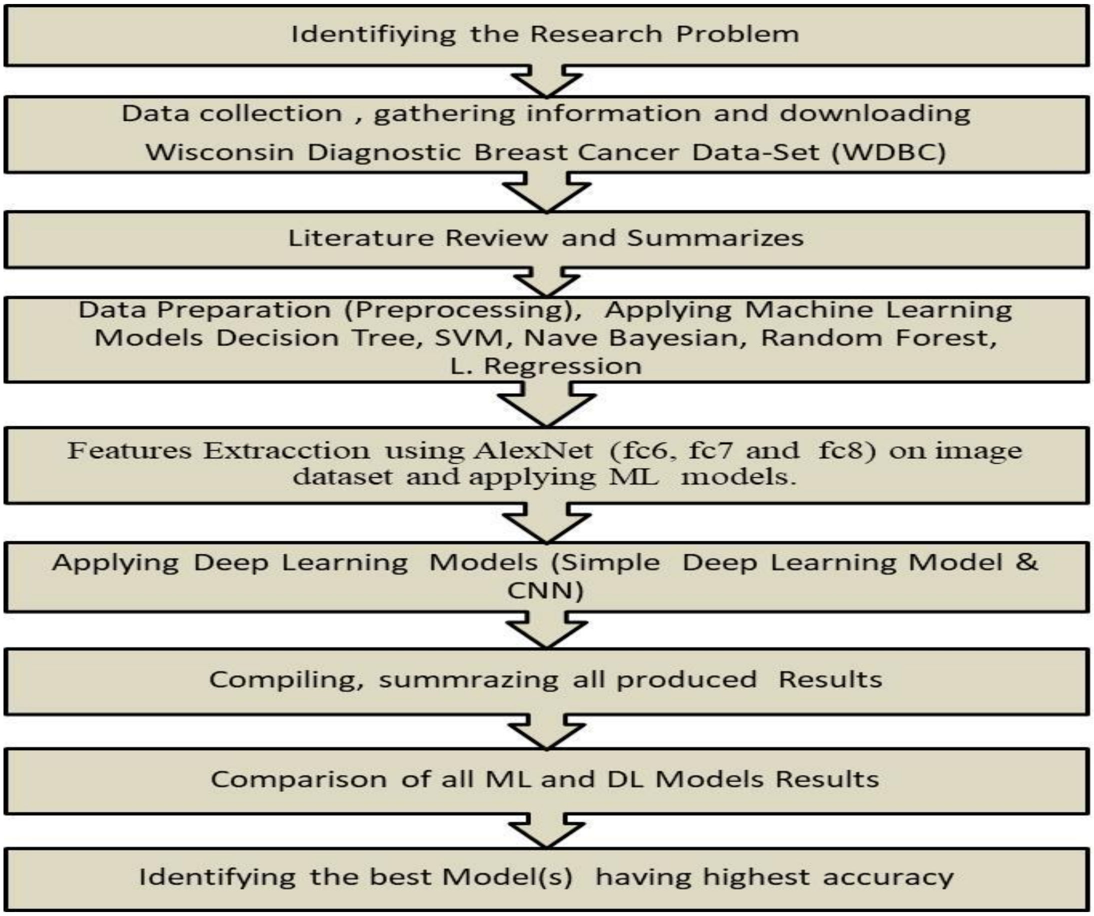

Fig. 2. Flow Diagram of Methodology of this research work.

1.1 The Machine Learning (ML) Classifiers

The following Machine Learning classifier are used to implement on cancer dataset Decision Tree, Logistic Regression (LR), Support Vector machine (SVM), Nave Bayesian (NB), Random Forest (RF) and Multilayer--Perceptron (MLP) and K--Nearest--Neighbors Algorithm (KNN).

1.2 Deep Learning Classifiers

The Simple-Deep-Learning-Models (SDLM) as Multilayer-Perceptron (MLP) and Convolutional-Neural-Network (CNN) Model are implemented on the cancer dataset to measure the accuracy of models.

1.3 Features Extraction

In this research work authors have prepared the image data set of the breast cancer images. After the preprocessing of the image data set, features are extracted using AlexNet architecture model. The features are extracted from the FC6, FC7 and FC8 layers of CNN. Then applied the ML Models on the newly extracted features and analyzed the findings.

1.4 Performance Factor

In this analysis work the effectiveness and efficiency of the applied models are measured for the prediction of cancer classification is Accuracy in percentages.

$$Testing\ Accuracy\ (TA) = (TP + TN) / (TP + TN + FN + FP) \qquad (1)$$

EXPERIMENTS AND RESULTS

In order to apply the above mention ML and DL classifiers, first classify the data into two parts, as first part is for training data-set and the second is for testing dataset. In this practical work the 70 % of dataset is split for training purposes and 30 % is for testing purpose.

ML Classifiers

The Table 2 describes the ML Testing and Training Accuracy in percentage of the following Algorithms; DT, LR, RF, SVM, NB, and KNN. Under the study, the SVM has the highest accuracy as shown in Table 2.

TABLE 2

Machine learning Algorithm Accuracy

Name of Algorithm	Testing Accuracy %(value)	Training Accuracy %(value)
Decision Tree	95.13	100.0
Logistic Regression	95.10	99.06
Random Forest	96.50	99.53
SVM	97.13	98.36
Naive Bayes	95.99	95.07
KNN	95.00	95.00

```
# Getting all of the models
model = models(X_train, Y_train)
```

```
[0]Logistic Regression Training Accuracy: 0.9906103286384976
[1]Decision Tree Training Accuracy: 1.0
[2]Random Forest Training Accuracy: 0.9953051643192489
[3]Support Vactor Machine Training Accuracy: 0.9835680751173709
[4]Naive Bayes Training Accuracy: 0.9507042253521126
[5]Multi-layer Perceptron Training Accuracy: 0.9906103286384976
```

Fig. 3. Shows the actual experimental results of the Training Accuracy of Machine Learning Models (ML)

```
Model 0
[[86  4]
 [ 3 50]]
Testing Accuracy = 0.951048951048951

Model 1
[[83  7]
 [ 2 51]]
Testing Accuracy = 0.9370629370629371

Model 2
[[87  3]
 [ 2 51]]
Testing Accuracy = 0.965034965034965

Model 3
[[88  2]
 [ 3 50]]
Testing Accuracy = 0.965034965034965

Model 4
[[87  3]
 [ 5 48]]
Testing Accuracy = 0.9440559440559441

Model 5
[[87  3]
 [ 4 49]]
Testing Accuracy = 0.951048951048951
```

Fig. 4. Shows the actual experimental results of the Testing Accuracy of Machine Learning Models (ML)

The Fig.5 graph shows the Testing and Training accuracy of Machine Learning algorithms.

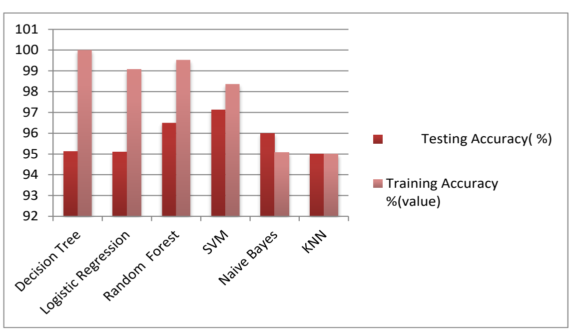

Fig. 5 Graph of Testing and Training accuracyof ML Algorithms

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online): 0493-2137
E-Publication: Online Open Access
Vol:55 Issue:07:2022

AlexNet Features fc6, fc7, fc8

The AlexNet fc6, fc7, fc8 Matlab results of ML Models accuracy are shown in Fig. 6, the recorded results may vary while execution. However, some values are listing here. 100% is the ideal accuracy of any ML model.

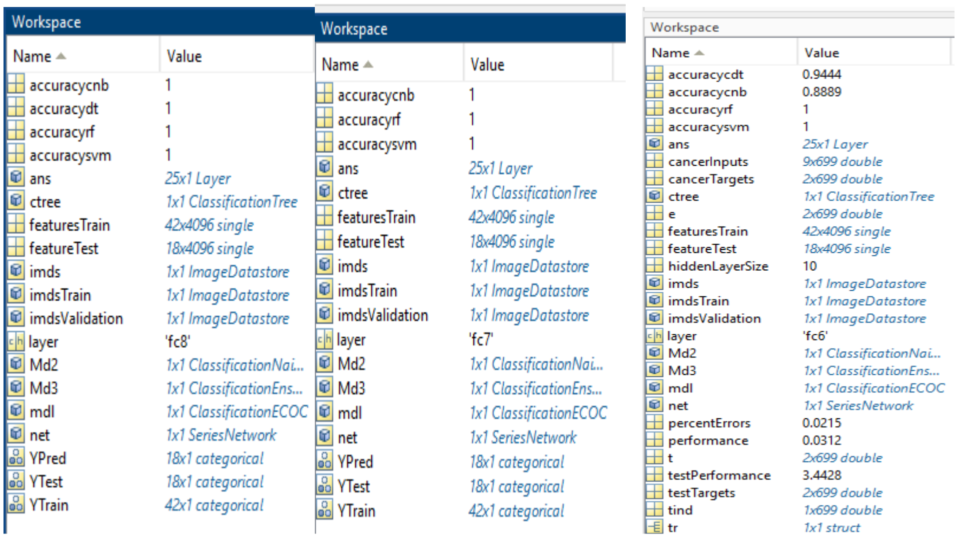

Fig. 6. FC6, FC7, FC8 accuracy results ML Models of NB, DT, RF, and SVM

TABLE 3
Accuracy of Machine Learning Algorithms of Fc6 using Alex-Net

Name of Algorithm	Accuracy %(value)
Decision Tree	94.44
Random Forest	100
SVM	100
Naive Bayes	94.44

The Fig. 7, graph shows the accuracy of machine learning algorithms of FC6 using AlexNet.

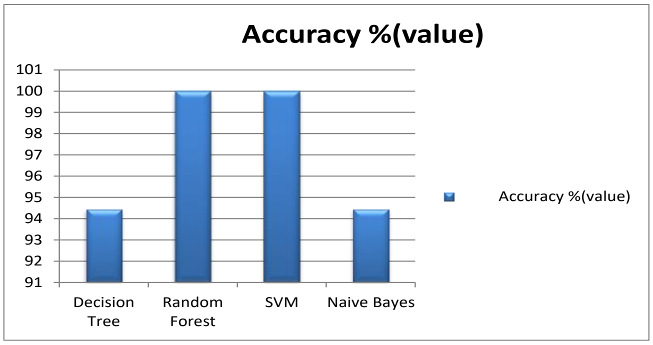

Fig. 7. Graph of FC6 using AlexNet

TABLE 4
Accuracy of Machine Learning Algorithms of Fc7 using AlexNet

Name of Algorithm	Accuracy %(value)
Decision Tree	100
Random Forest	94.44
SVM	100
Naive Bayes	94.44

Fig. 8 graph 3 shows the accuracy of machine learning algorithms of FC7 using AlexNet

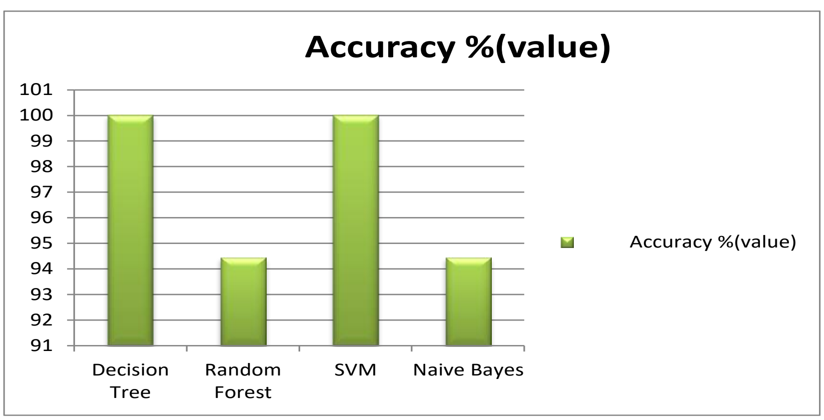

Fig. 8. Graph of Fc7 using AlexNet

FC8 results are showed in Table 4, the accuracy is recorded highest as 100 %, however it may varies 99% to 100 %.

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online): 0493-2137
E-Publication: Online Open Access
Vol:55 Issue:07:2022

TABLE 4
Accuracy of Machine Learning Algorithms of Fc8 using AlexNet

Name of Algorithm	Accuracy %(value)
Decision Tree	1
Random Forest	1
SVM	1
Naive Bayes	1

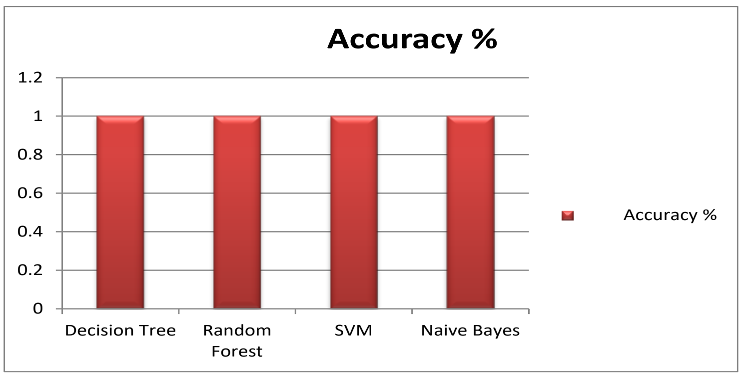

Fig. 9. Graph of FC8 Accuracy using AlexNet

Deep Learning Classifier

Table shows the deep learning classifier accuracy, simple deep learning model accuracy 94.45% testing and 98.45% is training accuracy. The CNN has the 99.48% testing accuracy and 96.89% is training accuracy. MLP is higher the as 99.48% as shown in Table 5 than CNN.

TABLE 5
Deep Learning Algorithm Accuracy

Name of Algorithm	Testing Accuracy %(value)	Training Accuracy %(value)
MLP	99.45	98.45
CNN	99.48	96.89

The Fig. 10 Graph shows the Testing and Training Accuracy of MLP and CNN. Reddish color bar represents MLP and pink shade colour bar represents CNN.

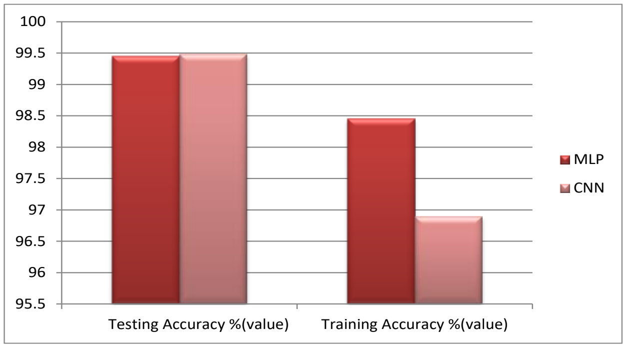

Fig. 10. Graph shows the Accuracy (testing & Training) of MLP and CNN

Comparison after Features Extraction

In the Table 6 all the testing accuracy results of ML Algorithms DT, RF, SVM, NB are compared with each other after features extraction fc6, fc7and fc8, the SVM shows distinction among others. As the accuracy 97% to 100 % is recorded during testing.

TABLE 6
Comparison after & before features extraction fc6, fc7 and fc8

Name of Algorithm	Testing Accuracy %(value)	Fc6-Testing Accuracy %	Fc7-Testing Accuracy %	Fc8-Testing Accuracy %
Decision Tree	95.13	94.44	100	100
Random Forest	96.50	100	94.44	100
SVM	97.13	100	100	100
Naive Bayes	95.99	94.44	94.44	100

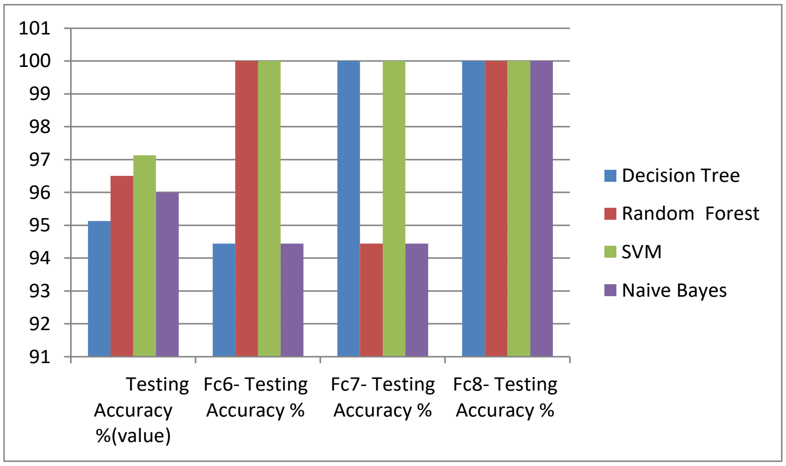

Fig. 11. Graph display the Accuracy of FC6, FC7 and FC8, the colours scheme is also mentioned in graph.

ML and DL Classifies Comparison

Table 7
Comparison of Machine Learning & Deep Learning

Machine Learning Algorithm	Testing Accuracy %(value)	Training Accuracy %(value)
Decision Tree	95.13	100.0
Logistic Regression	95.10	99.06
Random Forest	96.50	99.53
SVM	97.13	98.36
Naive Bayes	95.99	95.07
KNN	95.00	95.00
Deep Learning Algorithms		
MLP	99.45	98.45
CNN	99.48	96.89

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online): 0493-2137
E-Publication: Online Open Access
Vol:55 Issue:07:2022

The Fig.12 shows that Deep Learning Classifier accuracy results are better than the Machine Learning models.

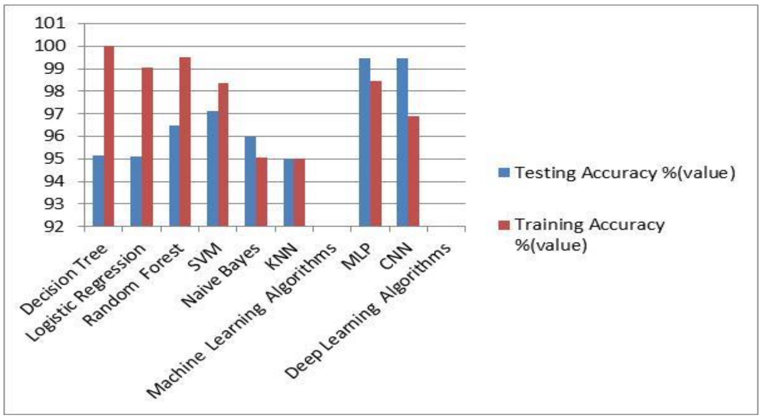

Fig. 12. The comparison graph between Machine and deep learning models

CONCLUSION

After the above analysis the results show that Deep Learning Models have produced better results than the ML algorithms applying on Wisconsin Breast-cancer (BC) dataset. The DL learning models produces the accuracy score of CNN and MLP in the form of 99.48 and 99.5 respectively. In the others hand ML models have produced accuracy score of Decision Tree, Naive Bayesian, Logistics Regression, SVM, KNN and Random Forest in the form of 95.13 %, 95.99 %, 95.10, 97.13, 95% and 96.50 respectively. In ML the SVM algorithm has the highest accuracy among the other ML algorithms.

This is also absorbed that after the new features extraction of fc6, fc7 and fc8 using AlexNet model and then applied the ML models on extracted features. Here, the accuracy of ML models has improved than before features extraction. The SVM algorithm produces the best result among the others simple machine learning models results.

Acknowledgement

We appreciate to Dr. Arfan Jaffar for such a great and useful supervision throughout this work.

Funding Statement

The authors have not got any particular financial aid for this research work.

Conflicts of Interest

The authors state that they have not conflicting interests to report in this work.

REFERENCES

1. Swathi, T.V., Krishna, S., Ramesh, M.V.: A Survey on Breast Cancer Diagnosis Methods and Modalities. In: 2019 International Conference on Wireless Communications Signal Processing and Networking (WiSPNET). pp. 287–292 (2019)
2. Weigelt, B., Geyer, F.C., Reis-Filho, J.S.: Histological Types of Breast Cancer: How Special Are They? Molecular Oncology. 4, 192–208 (2010). https://doi.org/10.1016/j.molonc.2010.04.004
3. Breast Cancer Wisconsin (Diagnostic) Data Set, https://www.kaggle.com/uciml/breast-cancer-wisconsin-data
4. UCI Machine Learning Repository: Breast Cancer Wisconsin (Diagnostic) Data Set, https://archive.ics.uci.edu/ml/datasets/breast+cancer+wisconsin+(diagnostic)
5. WHO official: Breast cancer, https://www.who.int/news-room/fact-sheets/detail/breast-cancer
6. Journal of Cancer Research and Therapeutics - Breast cancer statistics and markers, https://www.cancerjournal.net/downloadpdf.asp?issn=0973-1482;year=2014;volume=10;issue=3;spage=506;epage=511;aulast=Donepudi;type=2
7. Sharma, G.N., Dave, R., Sanadya, J., Sharma, P., Sharma, K.K.: VARIOUS TYPES AND MANAGEMENT OF BREAST CANCER: AN OVERVIEW. J Adv Pharm Technol Res. 1, 109–126 (2010)
8. Breast Cancer Stages, https://www.breastcancer.org/pathology-report/breast-cancer-stages
9. Asri, H., Mousannif, H., Moatassime, H.A., Noel, T.: Using Machine Learning Algorithms for Breast Cancer Risk Prediction and Diagnosis. Procedia Computer Science. 83, 1064–1069 (2016). https://doi.org/10.1016/j.procs.2016.04.224
10. Khuriwal, N., Mishra, N.: Breast Cancer Diagnosis Using Deep Learning Algorithm. In: 2018 International Conference on Advances in Computing, Communication Control and Networking (ICACCCN). pp. 98–103 (2018)
11. Assiri, A.S., Nazir, S., Velastin, S.A.: Breast Tumor Classification Using an Ensemble Machine Learning Method. J Imaging. 6, 39 (2020). https://doi.org/10.3390/jimaging6060039
12. Omondiagbe, D.A., Veeramani, S., Sidhu, A.S.: Machine Learning Classification Techniques for Breast Cancer Diagnosis. IOP Conf. Ser.: Mater. Sci. Eng. 495, 012033 (2019). https://doi.org/10.1088/1757-899X/495/1/012033
13. Analysis of Breast Cancer Detection Using Different Machine Learning Techniques | SpringerLink, https://link.springer.com/chapter/10.1007/978-981-15-7205-0_10
14. You, H., Rumbe, G.: Comparative Study of Classification Techniques on Breast Cancer FNA Biopsy Data. (2010). https://doi.org/10.9781/ijimai.2010.131
15. Jean Sunny, Nikita Rane, Rucha Kanade, Sulochana Devi, xavier institute of engineering: Breast Cancer Classification and Prediction using Machine Learning. IJERT. V9, IJERTV9IS020280 (2020). https://doi.org/10.17577/IJERTV9IS020280
16. S.h, S., Tiwari, D.H., C, D.D.V.: AlexResNet+: A Deep Hybrid Featured Machine Learning Model for Breast Cancer Tissue Classification. Turkish Journal of Computer and Mathematics Education (TURCOMAT). 12, 2420–2438 (2021). https://doi.org/10.17762/turcomat.v12i6.5686
17. Spanhol, F.A., Oliveira, L.S., Cavalin, P.R., Petitjean, C., Heutte, L.: Deep features for breast cancer histopathological image classification. In: 2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC). pp. 1868–1873 (2017)

18. Agarap, A.F.M.: 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. Association for Computing Machinery, New York, NY, USA (2018)
19. Doddipalli, L., Rani, K.: Performance Evaluation of Decision Tree Classifiers on Medical Datasets. International Journal of Computer Applications. 26, (2011). https://doi.org/10.5120/3095-4247
20. Deniz, E., Şengür, A., Kadiroğlu, Z., Guo, Y., Bajaj, V., Budak, Ü.: Transfer learning based histopathologic image classification for breast cancer detection. Health Inf Sci Syst. 6, 18 (2018). https://doi.org/10.1007/s13755-018-0057-x
21. Li, J., Zhou, Z., Dong, J., Fu, Y., Li, Y., Luan, Z., Peng, X.: Predicting breast cancer 5-year survival using machine learning: A systematic review. PLOS ONE. 16, e0250370 (2021). https://doi.org/10.1371/journal.pone.0250370