Engineering and Technology Journal e-ISSN: 2456-3358
Volume 08 Issue 12 December -2023, Page No. – 3156-3163
DOI: 10.47191/etj/v8i12.06, I.F. – 7.136

Deep and Machine Learning for Improving Breast Cancer Detection

Qabeela Q. Thabit
Basrah Education Directorate, Ministry of education, Basrah, Iraq

ABSTRACT: Breast cancer is the most common type of cancer in the world, as the number of people infected with it reached 2.2 million women in 2020, and World Health Organization reports indicated that the incidence of it is 1 to 12 women, that is, one woman out of every woman. every 12 women. As a result, it is crucial to have high cancer-predictive accuracy to update patient survival criteria and treatment options. Research on machine learning and deep learning, whether using traditional neural networks or using convolutional neural networks, has spread widely and has proven to be a useful technology. It can be very helpful in early detection and prognosis of breast cancer. According to the six machine learning algorithms used in this study and based on the Wisconsin breast cancer diagnostic dataset, they are as follows: Naive Bays (NB), K-Nearest Neighbors (KNN), Random Forest (RF), Decision Tree (DT), Logistic Regression (LR), and Support Vector Machine (SVM), we reached an accuracy of 99.1% with SVM that surpassed all competitors and achieved the highest accuracy. As for deep learning, we have reached an accuracy of up to 99.9%, and this is a reliable result for analysis purposes. In the presented work, the Anaconda environment (Jupyter platform) was used, which uses the Python programming language in all work.

KEYWORDS: Data preparing, feature and diagnosis, algorithm of machine learning, deep learning, neural networks, accuracy, validation data.

1. INTRODUCTION With over 2.2 million cases expected in 2020, breast cancer is the most prevalent type of cancer. In her lifetime, breast cancer will affect one in every twelve women. With approximately 685,000 women expected to die from it in 2020, breast cancer is the leading cause of cancer death in women [1]. Breast cancer is regarded as the second leading cause of death for women, following lung cancer [2]. From this critical importance to the lives of millions of women, it required looking for practical solutions that would advance the development of disease detection. Researches appeared in the field of breast cancer detection, the most important of which focused on two filed of machine learning represented by algorithms that depend on an official database documented by health institutions, the first field: machine learning in classification algorithms that are in the form of Tables including features(inputs) and outcome or label(output) [3,4], Many researchers have applied machine learning algorithms for classification to find optimal solutions in training to reach high accuracy, many of these algorithms are classifiers: Logistic Regression, Decision Tree, Random Forest, K-Nearest Neighbors (KNN), Naïve Bays (NB), Gradient Boosting Machine (GB), Xgboost, and Support Vector Machine(SVM)[5,6], and the second field: deep learning neural networks which uses a large set of Mammogram images as a database, is also prepared by official health institutions [7,8]. The other field deals with deep learning using neural network algorithms, which are of different types as follows: a deep belief network DBN [9], long short-term memory (LSTM) and the gated recurrent unit (GRU) networks[10], recurrent neural networks RNN[11], and the most one used in deep learning is convolutional neural network CNN[12] which consists of three layers(convolution, pooling and fully connected) and its take large implementation in different network models such as: AlexNet [13], VGG-16[14],Inception-v1 and Inception-v3[15], ResNet-50[16], and Inception- ResNet-V2[17].

This paper compares the performance of six classifiers and deep neural networks, which the research community considers among the most influential: logistic regression, decision tree, random forest, k-nearest neighbors (KNN), naïve bays (NB), and support vector machine (SVM). The top 10 data mining algorithms include data collection algorithms [3,4]. Use machine learning algorithms. The performance of each classifier in terms of the confusion matrix will be used to determine the most effective algorithm in predicting and diagnosing breast cancer, and the degree of accuracy, precision, sensitivity or recall. After that, deep learning was used using neural networks, and the rest of the parts are arranged as follows: The second section relates to the topic Research from previous works, Section Three: Proposed work methodology, Section Four: Presentation of results and discussion of results, and finally, Section Five: Conclusion of the paper: Conclusions and future recommendations.

2. RELATED WORKS

In the recent past, specifically the past decade until this time, intensive research has been and is still in this field, there ara of machine learning algorithms have been used to predict breast cancer by classification method, meaning the presence or absence of the disease (benign or malignant) of these algorithms: Decision Tree, K-Nearest Neighbors (KNN), Logistic Regression, Random Forest, Naïve Bays (NB), and Support Vector Machine. These completed researches were practically applied using a several datasets such as mammogram images, Wisconsin dataset, SEER dataset, in addition to a set of data from different hospitals. As examples for related works, in 2016 Hiba Asria et al. [3] introduced the accuracies for machine learning algorithms DT is 95.13 %, SVM is 97.13%, NB is 95.99%, K-NN is 95.27%, while Vincent Peter C. Magboo and Ma. Sheila A. Magboo in 2021[6] presented algorithms of machine learning accuracies: SVM is 74%, K-NN is 71%, NB is 65% and LR is 75%. In 2022 Sadia Safdar et al. [18] proposed detection breast cancer for algorithms LG, SVM and K-NN in accuracies 94.0%,97.7% and 97.0%, respectively.

3. DATASET DESCRIPTION

We conducted the research using a database for predicting breast cancer, which is a data set available on the Kaggle website that contains 30 columns and 570 rows [19]. The last column represents the output column and contains a zero, and a one in the case of zero there is no spray, while one is an indication of the presence of the disease. The database includes 357 cases of a benign tumor, meaning there is no breast cancer, which is represented in the table by zero. The remaining 212 cases represent a case of a malignant tumor, meaning the presence of the disease, which is represented by 1.

4. METHODOLOGY

In the Wisconsin breast cancer diagnostic dataset, we experimented first applying machine learning classifiers: logistic regression (LR), decision tree (DT), random forest (RF), K-nearest neighbors (KNN), and Naive Bays (NB), And Support Vector Machine. (SVM) and evaluated the results in two stages [20], before and after adjusting the data to determine the model that provides the highest accuracy. The second experiment is also using deep learning in stages, represented by a change in the number of layers and a change in the parameters used to obtain the best results, because the primary goal of our experiment is to find a reliable and accurate algorithm for detecting breast cancer. Figure 1 provides details about the proposed architecture [6,21].

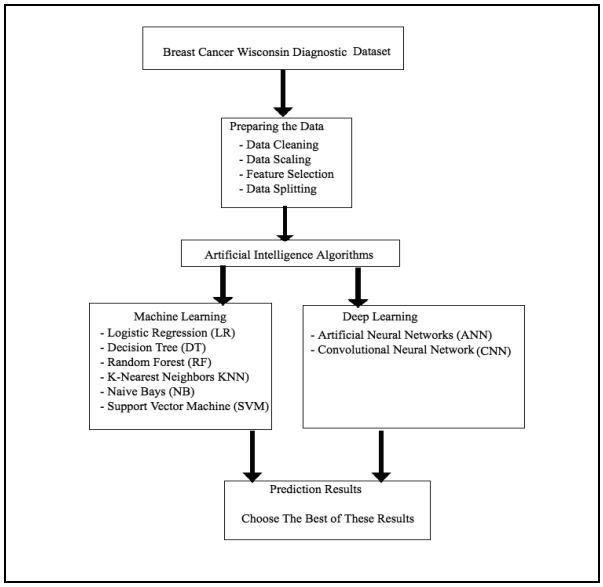

Figure 1. Procedure steps.

Data gathering is the first step in our methodology. The process of preparing the data then begins, consisting of four steps: feature extraction, attribute identification, target role assignment, and data cleaning. Initialized data is used to create machine learning algorithms that can predict breast cancer for a new set of measurements. We display brand-new model data for which we have labels in order to assess the performance of the algorithms.

The classified data we gathered is split into two parts using the Train test split method in the majority of works in this area. Our machine learning model is built using (75–80%) of the data, also known as training data or training set. The test data, also known as the test set, will be used to determine how well the model performs using 20–25 %. As for our presented work, we will do this and add to it another step in dividing the data, where we adopt the division that depends on three parts. The first part is the training data 60%, the second part is the control data 20%, and the third is the test data 20%. We have achieved with this division higher percentages than the previous division. I will mention it in the results section. In order to select the algorithm that offers the highest accuracy and determine the most predictive algorithm for the detection of breast cancer, we compare the results after testing the models.

A. Algorithm of Machine Learning

In this work, we use six artificial intelligence algorithms to implement machine learning, which are succinctly described as follows [3,4,6,18]:

Logistic Regression(LR): Logistic regression, also known as logistic regression method, is a high effective modeling in classification tool. In order to determine whether a risk factor affects the likelihood of developing a disease detection or other health situation, logistic regression is used (and covariates). The relationship between the independent variable(s) (Xi), also known as exposure or predictor variables, and the dichotomous dependent variable (Y), also

known as the outcome or response variable, is evaluated using both simple and multiple logistic regression. Predicting dependent binary or multi-class variables is its main application.

Decision Tree(DT): is a predictive modeling tool which provides in different situations. It can be made using an algorithmic technique that separates the dataset in distinct manners based on different criteria.

Random Forest(RF): Random decision forests are one of the aggregative methods of classification and regression, as well as other tasks. It works on the principle of generating many trees at the time of training the data and outputting the category which represents the placement of the prediction categories or classification of regression from the trees.

K-Nearest Neighbors(KNN): An algorithm for supervised classification is used. It uses a set of labeled points to determine the names of the other points. The ranked closest points to the additional point, or its nearest one in neighbors, are considered when naming it, and those neighbors cast their votes.

Naïve Bays(NB): It is based on Bayes Theorem and is a supervised classifier. Its predictions are probabilistic based on repetition. It can be applied in both binary and multiple prediction. It is a relatively fast algorithm when compared to other learning algorithms.

Support Vector Machine (SVM): It is a classifier that makes predictions for both linear prediction and classification prediction, as it divides the data set into categories to get the upper bound of the hyper-marginal level across the nearest point of those points.

In this work, we use six artificial intelligence algorithms to implement machine learning, which are succinctly described as follows [3,4,6,18]:

Logistic Regression(LR): Logistic regression, also known as logistic regression method, is a high effective modeling in classification tool. In order to determine whether a risk factor affects the likelihood of developing a disease detection or other health situation, logistic regression is used (and covariates). The relationship between the independent variable(s) (Xi), also known as exposure or predictor variables, and the dichotomous dependent variable (Y), also known as the outcome or response variable, is evaluated using both simple and multiple logistic regression. Predicting dependent binary or multi-class variables is its main application.

Decision Tree(DT): is a predictive modeling tool which provides in different situations. It can be made using an algorithmic technique that separates the dataset in distinct manners based on different criteria.

Random Forest(RF): Random decision forests are one of the aggregative methods of classification and regression, as well as other tasks. It works on the principle of generating many trees at the time of training the data and outputting the category which represents the placement of the prediction categories or classification of regression from the trees.

K-Nearest Neighbors(KNN): An algorithm for supervised classification is used. It uses a set of labeled points to determine the names of the other points. The ranked closest points to the additional point, or its nearest one in neighbors, are considered when naming it, and those neighbors cast their votes.

Naïve Bays(NB): It is based on Bayes Theorem and is a supervised classifier. Its predictions are probabilistic based on repetition. It can be applied in both binary and multiple prediction. It is a relatively fast algorithm when compared to other learning algorithms.

Support Vector Machine (SVM): It is a classifier that makes predictions for both linear prediction and classification prediction, as it divides the data set into categories to get the upper bound of the hyper-marginal level across the nearest point of those points.

- Data Method Acquisition

The Wisconsin Breast Cancer Diagnostic Dataset from the University of Wisconsin Hospitals Breast Cancer Database is used in our study [4]. A digital image of a breast cancer sample obtained through a fine needle aspiration is used to calculate the data set's features (FNA). These features allow us to infer the characteristics of the cell nuclei visible in the image. Breast Cancer Wisconsin Diagnostic has 569 instances (357 benign and 212 malignant) as shown in Figure 2, 2 classes show (62.74 benign and 37.26 malignant), 11 integer-valued features (-Id -Diagnosis -Radius -Texture -Area -Perimeter -Smoothness), and more, concavity, compactness, and concave points Fractal dimension - Symmetry).

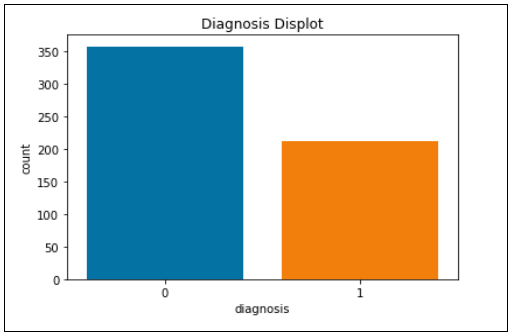

Figure 2. Diagnosis column (malignant & benign).

- Preparation Dataset

All the machine learning used algorithms in this research were implemented using the Scikit-Learn library, the Python programming language, and the Jupyter platform. It is known as an acronym sklearn and it is free and available as it is known because the programming language Python is open source. We also used the libraries in Python pandas. numpy, matplotlib. pyplot, seaborn and others. The process of initializing the data includes several steps before starting to work with the algorithms, which are as follows:

1) Data Cleaning

It means the process of searching for cells that have no data or missing data, i.e. empty, and treating this defect with the

filling of the cells, by using the Impute model, which fills the empty cells with the statistical measurement strategy such as mean, median, most, frequent or constant value [4,6].

It also includes the process of deleting unnecessary data in the machine learning process and does not affect the output and here the column of the identifier was deleted because it does not affect the infection of the disease and the deletion of unknown data (Nan) to the program. We checked the data used in the experiment here, and the result was no missing data, after that convert each symbol M(malignant) into 0 and B (benign) into 1 in the column titled diagnosis and delete columns "id" with "unknown", the data shown for only 5 rows in the following Figure 3. dataset includes (569 row,31 columns), from these 30 columns are features and one column is diagnosis.

	diagnosis	radius_mean	texture_mean	perimeter_mean	area_mean	smoothness_mean	compactness_mean	concavity_mean	concave points_mean	symmetry_mean
0	1	17.99	10.38	122.80	1001.0	0.11840	0.27760	0.3001	0.14710	0.2419
1	1	20.57	17.77	132.90	1326.0	0.08474	0.07864	0.0869	0.07017	0.1812
2	1	19.69	21.25	130.00	1203.0	0.10960	0.15990	0.1974	0.12790	0.2069
3	1	11.42	20.38	77.58	386.1	0.14250	0.28390	0.2414	0.10520	0.2597
4	1	20.29	14.34	135.10	1297.0	0.10030	0.13280	0.1980	0.10430	0.1809

...	radius_worst	texture_worst	perimeter_worst	area_worst	smoothness_worst	compactness_worst	concavity_worst	concave points_worst	symmetry_worst	fractal_di
...	25.38	17.33	184.60	2019.0	0.1622	0.6656	0.7119	0.2654	0.4601	
...	24.99	23.41	158.80	1956.0	0.1238	0.1866	0.2416	0.1860	0.2750	
...	23.57	25.53	152.50	1709.0	0.1444	0.4245	0.4504	0.2430	0.3613	
...	14.91	26.50	98.87	567.7	0.2098	0.8663	0.6869	0.2575	0.6638	
...	22.54	16.67	152.20	1575.0	0.1374	0.2050	0.4000	0.1625	0.2364	

Figure 3. Data wisconsin breast cancer after data cleaning.

2) Data Scaling

It is specific to the process of scaling all kinds of data and it comes from the preprocessing module, it is important in data settlement and it is also called normalization or standardization [5].

When the features are various numbers in the training data, that means each feature has a specific range of numbers that differs from the rest, and this difference causes a slowdown in the processing process, and from this aspect we turned to the normalization of all data in a specific range, for example, all from 0 to 1 and here the processing will be much faster.

3) Feature Selection

It is specific to selecting the required and influential features and excluding the rest of the features, and the features are selected based on the extent to which they are related to the output, which in the case of our paper, the disease is benign or malignant and is done through the module feature selection. Sometimes it is preferable to let the algorithm learn and train all the features, and sometimes some features and leave the rest, because there are features that are not related to training, for example, if I have the name of the patient or the patient's health number, it does not differentiate whether the disease is present or not, then they are unimportant traits, and therefore their presence works to mislead the work of the algorithm. Sometimes the presence of many features is also incorrect in training because the algorithm focus on them becomes less, so it can be a very important trait and another less important feature, here appears the role of the step of selecting traits in the training procedure. The Figure 4 below shows the relationship between the attributes in the database[3].

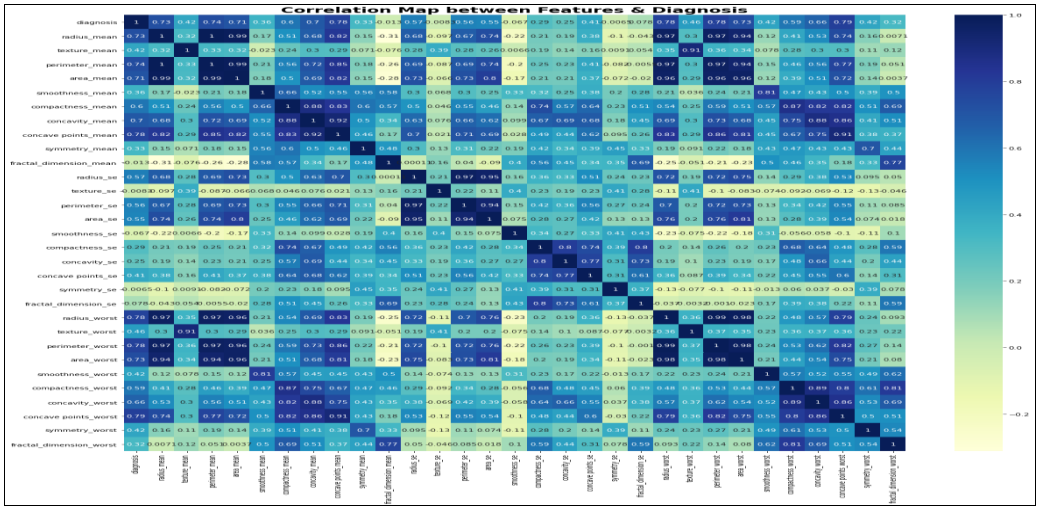

Figure 4. Correlation between features and diagnosis

4) Data Splitting

It is concerned with dividing the data into a special section for training and a section for examination, using the selection model. In some cases, the division is 80% training data part to 20% examination or testing data part [3,4].

B. Algorithm of Deep Learning

Deep learning is widely used in scientific research in various fields. In recent years, the focus on deep learning has taken a very large role in natural language processing, voice recognition, computer vision, disease prediction, games, etc. In general, deep learning includes three types of networks [22], which are neural networks, convolutional neural networks, and recurrent neural networks. The deep nerve network consists of several hidden layers in order to improve the performance of it, as back propagation is used to train the network and reduce the difference between the required and actual output. Figures 5 and 6. represents the implementation of the work using a two models to predict neural networks in different number layers, and the change was in the number of epochs until we reach the best accuracy in implementation. After that, we reach to best result between the two neural network was implemented by increasing the number of hidden layers as shown in Figure 6, so we obtained higher results in terms of accuracy and the rest of the metrics.

```
model = Sequential ()
model.add(Conv1D(filters=16,kernel_size=2,activation='
relu',input_shape= (30,1)))
model.add(BatchNormalization())
model.add(Dropout(0.2))

model.add(Conv1D(32,2,activation='relu'))
model.add(BatchNormalization())
model.add(Dropout(0.2))

model.add(Flatten())
model.add(Dense(32,activation='relu'))
model.add(Dropout(0.2))

model.add(Dense(1,activation='sigmoid'))
```

Figure 5. Model-1

```
model = Sequential ()
model.add(Conv1D(filters=16,kernel_size=2,activation='rel
u',input_shape= (30,1)))
model.add(BatchNormalization())
model.add(Dropout(0.2))

model.add(Conv1D(32,2,activation='relu'))
model.add(BatchNormalization())
model.add(Dropout(0.2))

model.add(Conv1D(32,2,activation='relu'))
model.add(BatchNormalization())
model.add(Dropout(0.2))

model.add(Flatten())
model.add(Dense(32,activation='relu'))
model.add(Dropout(0.2))

model.add(Dense(1,activation='sigmoid'))
```

Figure 6. Model-2

5. DISCUSSION RESULT

The six algorithms (LR, DT, RF, KNN, NB, and SVC) were implemented on the Breast Cancer Wisconsin after performing the previous series of steps, starting from cleaning the data to splitting it. In the first part of the experiment, data split into two sets: 80% training data set and 20% testing data set, while the second part of experiment, data split into three 60% training data set, 20% validation data set, and 20% test data set , the confusion matrix was obtained with two cases of experiment and as shown in Table 1.

Table 1. Confusion matrix obtained by six algorithms

	Train and Test Data		Train, validation and Test Data		
	Malignant	Benign	Malignant	Benign	
LR	74	1	74	1	Malignant
	2	37	2	37	Benign
DT	72	3	72	3	Malignant
	3	36	1	38	Benign
RF	72	3	74	1	Malignant
	2	37	2	37	Benign
KNN	73	2	75	0	Malignant
	1	38	2	37	Benign
NB	71	4	73	2	Malignant
	4	35	2	37	Benign
SVC	74	1	75	0	Malignant
	2	37	1	38	Benign

In both experiment cases, the confusion matrices produced from the data indicated True Negative (TN), False Positive (FP), False Negative (FN), and True Positive (TP). In order to find the accuracy of the experiment in the first and second tiers, classification reports were calculated for all methods, which are as follows two parts (without validation and with validation data) results in the Figure 7 and Figure 8, respectively.

Report Classification (Train Data & Test Data):

LR:	precision	recall	f1-score	support
0	0.97	0.99	0.98	75
1	0.97	0.95	0.96	39
accuracy			0.97	114
macro avg	0.97	0.97	0.97	114
weighted avg	0.97	0.97	0.97	114

DT:	precision	recall	f1-score	support
0	0.96	0.96	0.96	75
1	0.92	0.92	0.92	39
accuracy			0.95	114
macro avg	0.94	0.94	0.94	114
weighted avg	0.95	0.95	0.95	114

RF:	precision	recall	f1-score	support
0	0.97	0.96	0.97	75
1	0.93	0.95	0.94	39
accuracy			0.96	114
macro avg	0.95	0.95	0.95	114
weighted avg	0.96	0.96	0.96	114

KNN:	precision	recall	f1-score	support
0	0.97	0.99	0.98	75
1	0.97	0.95	0.96	39
accuracy			0.97	114
macro avg	0.97	0.97	0.97	114
weighted avg	0.97	0.97	0.97	114

NB:	precision	recall	f1-score	support
0	0.99	0.96	0.97	75
1	0.93	0.97	0.95	39
accuracy			0.96	114
macro avg	0.96	0.97	0.96	114
weighted avg	0.97	0.96	0.97	114

	precision	recall	f1-score	support
0	0.99	0.95	0.97	75
1	0.90	0.97	0.94	39
accuracy			0.96	114
macro avg	0.95	0.96	0.95	114
weighted avg	0.96	0.96	0.96	114

Figure 7. Report classification without validation data

Report Classification (with validation data):

LR:	precision	recall	f1-score	support
0	0.97	0.99	0.98	75
1	0.97	0.95	0.96	39
accuracy			0.97	114
macro avg	0.97	0.97	0.97	114
weighted avg	0.97	0.97	0.97	114

DT:	precision	recall	f1-score	support
0	0.99	0.95	0.97	75
1	0.90	0.97	0.94	39
accuracy			0.96	114
macro avg	0.95	0.96	0.95	114
weighted avg	0.96	0.96	0.96	114

RF:	precision	recall	f1-score	support
0	0.97	0.96	0.97	75
1	0.93	0.95	0.94	39
accuracy			0.96	114
macro avg	0.95	0.95	0.95	114
weighted avg	0.96	0.96	0.96	114

KNN:	precision	recall	f1-score	support
0	0.97	1.00	0.99	75
1	1.00	0.95	0.97	39
accuracy			0.98	114
macro avg	0.99	0.97	0.98	114
weighted avg	0.98	0.98	0.98	114

NB:	precision	recall	f1-score	support
0	0.97	0.97	0.97	75
1	0.95	0.95	0.95	39
accuracy			0.96	114
macro avg	0.96	0.96	0.96	114
weighted avg	0.96	0.96	0.96	114

SVC:	precision	recall	f1-score	support
0	0.99	1.00	0.99	75
1	1.00	0.97	0.99	39
accuracy			0.99	114
macro avg	0.99	0.99	0.99	114
weighted avg	0.99	0.99	0.99	114

Figure 8. Report classification with a validation data

The accuracy was calculated for each of the above-mentioned methods in both cases of the experiment, and the results in the second case of data division showed higher accuracy, as shown in Table2.

Table 2. The Accuracy of the implemented algorithms

Algorithm	Accuracy (%) Without Validation Data	Accuracy (%) With Validation Data
LR	97.3	97.3
DT	94.7	96.4
RF	95.6	97.3
KNN	96.9	98.2
NB	93.1	96.4
SVC	97.3	99.1

Table 3. Deep Learning Results in different number Layers

Layers No.	Accuracy	Precision	Recall	F1-Score	AUC
Model1	0.9523	0.9201	0.9327	0.9333	0.9606
Model2	0.9995	0.9927	0.9855	0.9700	0.9961

6. CONCLUSIONS

We implemented six main algorithms: Logistic Regression, Decision Tree, Random Forest, K-NN, NB, and SVM, used the Wisconsin Breast Cancer Diagnostics Database (WBCD) to calculate, compare, and evaluate the effectiveness of each algorithm. based on confusion matrix, precision accuracy, recall, and f1-score to choose the best machine learning algorithm with the desired accuracy and reliability, and to find higher accuracy. All algorithms have been operating in Python by using Jupyter platform and the Anaconda environment. Following a precise comparison of our models, we discovered that Support Vector Machine outperformed all other algorithms and had a higher accuracy of 99.1 percent, precision of 99 percent, and recall of 100 percent.

This leads us to the conclusion that the Support Vector Machine has demonstrated its effectiveness in the prediction and precise diagnosis of breast cancer because it has produced the best results in terms of accuracy and precision. The process of conducting machine learning of the data has gone through several stages, namely, initializing the data after obtaining it in terms of cleaning, scale, choosing effective features, and also splitting the data was influential and a helping factor here, where we noticed the difference between the split in the traditional way and the method used here is the method of adjustment by adding a new split of the data. We have obtained results with high accuracy, but we aspire in the future to expand the database to a large base or use another database larger than the one used in this work and compare the results of both, in addition to taking into consideration the consideration of algorithms for machine learning and comparing them with the results of these algorithms. After that, the implementation was done using different neural networks in terms of the number of hidden layers and the number of attempts completed, so we reached a high accuracy result of almost 99.9%, so it is better to apply deep learning. It is worth noting that in the future, machine learning can be combined with deep learning to obtain an optimal accuracy result.

REFERENCES

1. https://www.who.int/news-room/fact-sheets/detail/breast-cancer.
2. U.S. Cancer Statistics Working Group. United States Cancer Statistics: 1999–2008 Incidence and Mortality Web-based Report. Atlanta (GA): Department of Health and Human Services, Centers for Disease Control and Prevention, and National Cancer Institute; 2012.
3. Hiba Asria, Hajar Mousannif, Hassan Al Moatassimec,and Thomas Noeld, "Using Machine Learning Algorithms for Breast Cancer Risk Prediction and Diagnosis," The 6th International Symposium on Frontiers in Ambient and Mobile Systems (FAMS 2016), vo.83,pp. 1064-1069,2016.
4. M. Amine Naji et al., "Breast Cancer Prediction and Diagnosis through a New Approach based on Majority Voting Ensemble Classifier," International Workshop on Edge IA-IoT for Smart Agriculture (SA2IOT), vol.191, pp. 481-486, 2021.
5. Nosayba Al-Azzam, Ibrahem Shatnawi, PE, PMP, and PTOE, "Comparing supervised and semi-supervised Machine Learning Models on Diagnosing Breast Cancer," Annals of Medicine and Surgery, vol.62,pp.53-64,2021.
6. Vincent Peter C. Magboo and Ma. Sheila A. Magboo, "Machine Learning Classifiers on Breast Cancer Recurrences," Procedia Computer Science, vol. 192, pp.2742–2752, 2021.
7. Anji Reddy Vaka, Badal Soni, Sudheer Reddy K., "Breast cancer detection by leveraging Machine Learning," ICT EXPRESS, vol.6, pp. 320-324,2020.

8. Tanzila Saba, "Recent advancement in cancer detection using machine learning: Systematic survey of decades, comparisons and challenges," Journal of Infection and Public Health, vol.13, pp.1274-1289, 2020.
9. Yue Zhang and Fangai Liu, "An Improved Deep Belief Network Prediction Model Based on Knowledge Transfer," future internet, vol.12, pp. 1-18, 2020.
10. Balduíno César Mateus, Mateus Mendes, José Torres Farinha, Rui Assis andAntónio Marques Cardoso, "Comparing LSTM and GRU Models to Predict the Condition of a Pulp Paper Press," engirgies, vol. 14, pp. 1-21, 2021.
11. Soon Hoe Lim, N. Benjamin Erichson, Liam Hodgkinson and MichaelW. Mahoney, "Noisy Recurrent Neural Networks," 35th Conference on Neural Information Processing Systems (NeurIPS 2021).
12. Sakshi Indoliaa, Anil Kumar Goswami, S. P. Mishra, and Pooja Asopaa, "Conceptual Understanding of Convolutional Neural Network- A Deep Learning Approach," International Conference on Computational Intelligence and Data Science (ICCIDS 2018), vol. 132, pp. 679-688.
13. Hsing-Chung Chen et al., "AlexNet Convolutional Neural Network for Disease Detection and Classification of Tomato Leaf," electronics, vol.11, pp.1-17, 2022.
14. Prajakta Ganakwar, Ms. Saroj Date, "Convolutional Neural Network-VGG16 for Road Extraction from Remotely Sensed Images," International Journal for Research in Applied Science & Engineering Technology (IJRASET), vol. 8, issue VIII, pp.916-922, 2020
15. Jianfang Cao, Minmin Yan, Yiming Jia, Xiaodong Tian and Zibang Zhang, "Application of a modified Inception-v3 model in the dynasty-based classification of ancient murals," EURASIP Journal on Advances in Signal Processing, vol.49, pp.1-25,2021.
16. Fatima-Zohra Hamlili, Mohammed Beladgham, Mustapha Khelifi and Ahmed Bouida, "Transfer learning with Resnet-50 for detecting COVID-19 in chest X-ray images," Indonesian Journal of Electrical Engineering and Computer Science, vol. 25, no. 3, pp. 1458-1468, 2022.
17. Usman Nazir, Numan Khurshid, Muhammad Ahmed Bhimra and Murtaza Taj, "Tiny-Inception-ResNet-v2: Using Deep Learning for Eliminating Bonded Labors of Brick Kilns in South Asia," Computer Vision and Pattern Recognition, pp.1-6,2019.
18. Sadia Safdar et al., "Bio-Imaging-Based Machine Learning Algorithm for Breast Cancer Detection," diagnostics, vol.12, pp.1-18, 2022.
19. https://www.kaggle.com/datasets/uciml/breast-cancer-wisconsin-data
20. Q. Q. Thabit, Taqwa O. Fahad, Alyaa I. Dawood, Detecting Diabetes Using Machine Learning Algorithms, 2022 Iraqi International Conference on Communication & Information Technologies (IICCIT-2022), Basrah University ,Basrah , Iraq.
21. Alyaa Ibrahim Dawood , [22]. Qabeela Q. Thabit and Taqwa O. Fahad, Thyroid Disease Prediction with Machine LearningAlgorithms, Eurasian Research Bulletin,18, pp.229-237,2023.
22. Qabeela Q. Thabit, Alyaa Ibrahim Dawood and Bayadir A. Issa. Implementation three-step algorithm based on signed digit number system by using neural network, Indonesian Journal of Electrical Engineering and Computer Science, 24(3), pp. 1832-1839, 2021.