Breast Cancer Detection using Machine Learning Techniques

Rahul Maheshwari
IIIT Delhi
MT19027
rahul19027@iiitd.ac.in

Surbhi
IIIT Delhi
MT19055
surbhi19055@iiitd.ac.in

Nikunj Agarwal
IIIT Delhi
MT19093
nikunj19093@iiitd.ac.in

Sumedha Bhatia
IIIT Delhi
MT19129
sumedha19129@iiitd.ac

***Abstract*—Breast Cancer is one of the most commonly found cancer among women and has a high mortality rate. One out of eight women suffers from it during their lifetime. The tests and diagnosis available, at present, take a long period of time to detect the disease and follow the treatment thereafter. In medical treatments, the earlier detection of diseases is directly proportional to the high probability of successful treatment. This has led to the approach of using Machine learning algorithms to classify and detect the cancerous tumour by using data mining techniques. Data mining techniques can help to discover hidden patterns and relationships between the features that help to detect the tumour. The usefulness of such procedures is their highly accurate prediction accuracy and the short amount of time it takes to perform the classification, as compared to conventional methods of diagnosis. It can also help medical practitioners to have a second opinion to estimate the severity of disease and diagnose the disease better. Significant works have been done for such a task to achieve high accuracy. This paper aims to achieve higher accuracy compared to existing systems and the use of such procedures in real world applications. All experiments and models have been performed by using the Wisconsin Breast Cancer (Diagnostic) data set. Six models including Support Vector Machine (SVM), Logistic Regression (LR), Random Forest (RF), K-Nearest Neighbor (KNN), Multilayer Perceptron (MLP) and FastAI-based neural network Machine learning models were applied with 4 variations along with sophisticated feature selection PPS (Predictive Power Score) and dimensionality reduction method of PCA (Principal Component Analysis. We applied several classifier variations for all classifiers, out of which FastAI (98.60%) performed the best. Classifier results were obtained by performing 5-fold cross-validation in order to obtain unbiased predictions. From the experiments and results, it is evident that FastAI (novel Artificial Neural Network provided by FastAI) is well suited for breast cancer detection.**

***Index Terms*—Classification, Machine learning, FastAI, RF, 5-fold cross-validation.**

I. Introduction

Breast cancer affects nearly 10% of the women across the world and is listed among the top 10 cause of deaths in women by the World Health Organization (WHO). It is claimed as the second highest cause of death in women after lung cancer in the United States by the American Society of Clinical Oncology (ASCO). Major studies have proved that the rate of survival in breast cancer is nearly 91% five years after it was first diagnosed and nearly 84% ten years after its diagnosis [16]. The high survival rates hence provide the importance for its early detection.

Diagnosis of breast cancer is done by classifying the tumour as benign or malignant. Three types of diagnostic techniques are available for the detection of breast cancer. First is the diagnostic **Mammography** where additional pictures of the breast are taken when a woman is experiencing severe symptoms like new lumps or nipple discharge. Second is the **Ultrasound** that uses sound waves to recreate an image of the breast tissue. Third is an **MRI**, which is also the most common diagnosis techniques for breast cancer. It utilizes the magnetic field to detect cancerous cells in lymph nodes. Such methods are costly, time-consuming and require frequent visits to the oncologist. To make things worsen, not all oncologists are experts in distinguishing between benign and malignant tumor. All these disadvantages have led the researchers to look more closely at the classification techniques available in data science to distinguish between the tumour categories based on patients' medical records. This can enhance the prediction and survival rate significantly so that patients can be informed to take clinical treatment at the earliest and avoid needless treatments that are otherwise harmful.

A literature review on available work done using the cancer data set available in UCI (University of California, Irvine) repository showed that techniques like SVM, kNN, etc. have been applied but not much emphasis has been made on feature selection. In [1], the authors have analyzed the performance of supervised learning classifiers only such as SVM-RBF kernel and Naive Bayes, which showed that SVM-RBF was more accurate. While in [2], the authors have compared the performance of Radial basis neural networks, Decision tress, Naive bayes, SVM and CART using WEKA. They also reported highest accuracy of 96.99% using SVM. While highest accuracy of 97.13% has been reported by [3] in their research paper.

This paper not only discusses the performance of supervised learning techniques like kNN and SVM, but also the performance of Random Forest Classifier and Artificial Neural Network using FastAI which is built over PyTorch. These models have been trained on the standard breast cancer (diagnostic) data set available in the UCI repository. The subsequent sections in this paper are divided as follows. Section 2 defines the materials used and methodologies applied. Section 3 provides all the collated results and

compares each of them on the basis of sensitivity, specificity and accuracy. Section 4 gives a discussion by providing an overall conclusion and limitations involved while Section 5 presents the contribution of each author involved.

II. Materials and Methodology

Before explaining the various machine learning models applied for classification of tumour, the dataset used has been explained. We have obtained this dataset from [4] and have used Jupyter notebook to implement the code and Google Colab platform to combine the contribution of each author.

A. Dataset Description

The dataset mentioned consists of 569 samples with 32 attributes each. Out of these 32 attributes, one of them is the classifier output, Benign(B) or Malignant(M) and the other is the Case ID. Rest 30 attributes are numerical and correspond to the patients' medical records. Out of these 569 data points, 212 are malignant tumour cases while the rest are benign tumour cases.

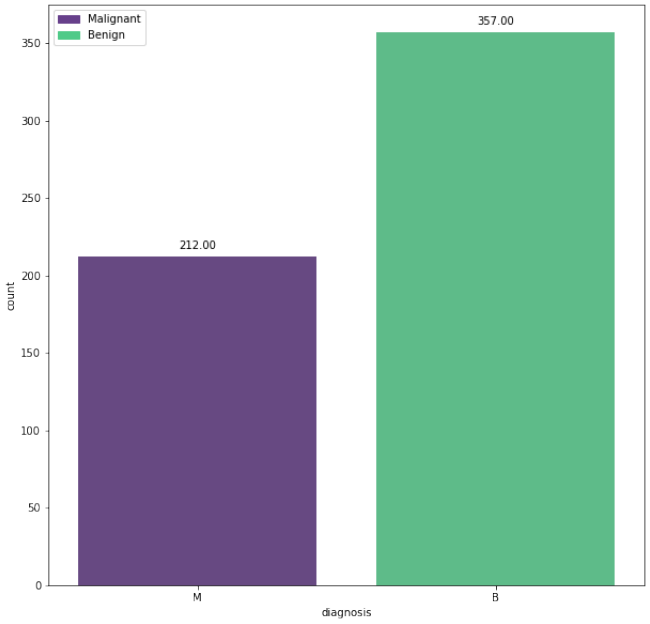

Fig. 1. Malignant and Benign Tumour sample counts.

Ten real valued attributes used for preparing the model are described as follows. They have been computed for each cell nucleus.

The mean, worst and the standard error values were calculated for each of the 569 samples which resulted in 30 attributes.

B. Data Analysis & Preprocessing

The given data has been analyzed so that relationship between various attributes can be determined. Thereafter, data preprocessing has been done to normalize the data.

TABLE I. Attribute information

S.No.	Attribute Name	Attribute definition
1	radius	mean of distances from center to points on the perimeter
2	texture	standard deviation of gray-scale values
3	Perimeter	perimeter of the cell nucleus
4	Area	Area of the cell nucleus
5	smoothness	local variation in radius lengths
6	compactness	$perimeter^2/area - 1.0$
7	concavity	severity of concave portions of the contour
8	concave points	number of concave portions of the contour
9	symmetry	symmetry in cell nucleus
10	fractal dimension	"coastline approximation" - 1

Finally, attribute selection and dimensionality reduction has been performed using two distinct methods. For model train and testing, the entire dataset has been split into 75:25 train-test ratio.

- **Correlation between features**: Correlation describes how close are two attributes to each other. It attempts at defining a linear relationship between them. Two attributes having linear relationship will be more correlated than attributes having non linear relationship between them. This score ranges from +1 to -1, where 0 indicates no relationship. If the correlation is higher than 0.9, one of them is dropped. Figure 2 is used to illustrate correlation between mean features(10 in number), but the correlation was plotted between all 30 features.

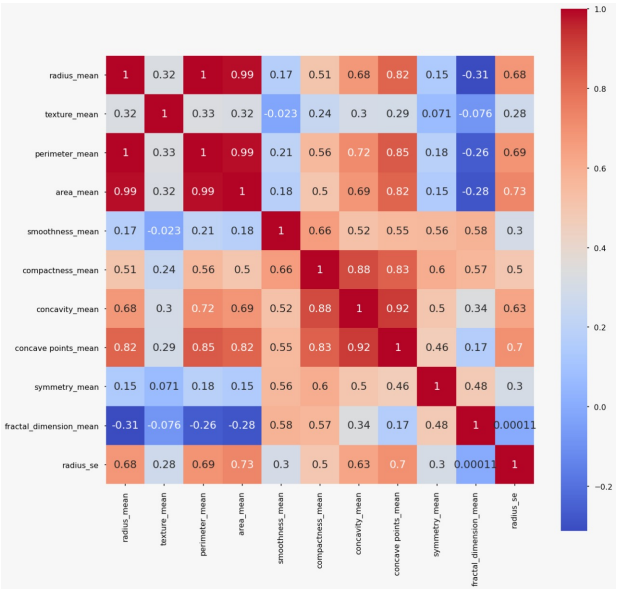

Fig. 2. Correlation between mean attributes

- **Normalization of features**: The selected features have been normalized using the MinMaxScaler of scikit-learn. This has been done to reduce all the features between the range 0-1. The MinMaxScaler subtracts the minimum value and then divides by the original range. This is done for each attribute, i.e., column wise in the given data. [5]

- **Predictive Power Score (PPS)**: While correlation matrix assumes symmetrical relationships between attributes, it is not generally true for all kinds of data. PPS score varies between 0 and 1 and is helpful for finding specific patterns in the data, feature selection, detects information leakage between variables as well as data normalization. From the analysis, area_mean, perimeter_mean, perimeter_worst and area_worst were dropped. [6]
- **Principle Component Analysis (PCA)**: This is perhaps one of the most popular algorithm for dimensionality reduction based on the co-variance matrix of the given features. It is used to overcome the curse of dimensionality when a lot of linearly related attributes are present in the data set. In our algorithm, PCA has been fit on the train dataset to find out the best features, which is then used to transform the test data [7].

C. Machine Learning Techniques

Scikit-learn library provides a number of supervised and unsupervised models to be applied for the classification task. We have used Random Forest, K-Nearest neighbor, Logistic regression, Support Vector Machine, Multi-layer Perceptron and Neural Network for tabular data using FastAI for the given classification task at hand.

- **Logistic Regression Classifier**: This is a naive classification method where the classifier utilizes the sigmoid function. It is given by the equation $1/(1+e^{-value})$. It is used to plot the odds of being a case based on the attributes in the dataset. Hence, it gives probability as

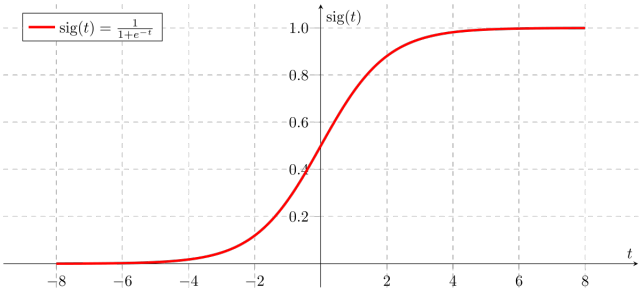

Fig. 3. Sigmoid Function

output and to predict which class a prediction belongs to, a threshold(usually 0.5) can be set. We have implemented logistic regression using L2 regularization here.
In our implementation, accuracy was found out without and with feature selection. Following table II is the classification report for the same. The accuracy for this technique was reported as 97.90% with & without feature selection.

- **Support Vector Machine**: SVM is a very strong predictive analysis and maximum-margin classifier. It tries to figure out a hyper-plane that can distinctly classify the data points. The dimensions of the hyper-plane depend on the total number of features in the

Table II. Classification report for Logistic Regression with & without feature selection

	Precision	Recall	f1-score
Class 0	0.99	0.98	0.98
Class 1	0.96	0.98	0.97
Accuracy			0.98
Micro avg	0.98	0.98	0.98
Macro avg	0.97	0.98	0.98

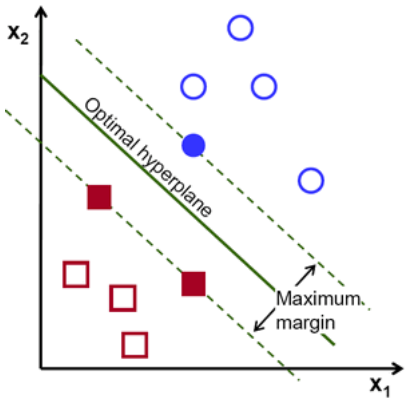

Fig. 4. Hyper-plane detection using SVM

dataset. Hence, each data point in the SVM model is represented in an n-dimensional space. Classification is done by maximizing the margin separating classes with the objective of minimizing the classification errors. Support vectors are points on the plane that can help decide the orientation of the hyper-plane. The hyper-plane equation used here is $W^TX=0$, as opposed to y=ax+b which is used in logistic regression classifier to separate the data points.
In our implementation, the best parameters for SVM were selected as C=10, gamma=1 and kernel = linear using the GridSearchCV method available in scikit-learn which helps to fine tune SVM hyper-parameters. For feature selection, however, the kernel selected was radial basis as opposed to linear while the rest of the parameters were same.
Following is the classification report for SVM with and without feature selection.

Table III. Classification report for SVM Classifier with & without feature selection

	Precision	Recall	f1-score
Class 0	1.0	0.99	0.99
Class 1	0.98	1.0	0.99
Accuracy			0.99
Micro avg	0.99	0.99	0.99
Macro avg	0.99	0.99	0.99

The overall test accuracy of SVM was calculated as 99.30% with & without feature selection.

- **Random Forest Classifier**: Random Forest is an

ensemble learning method for classification and consists of of a large number of decision trees. In data science, this phenomenon is known as "wisdom of the crowd", where a crowd is believed to outperform the individual components. Below figure 5 can be used to interpret this definition in a better way. As the figure illustrates, there are 9 total decision trees out of which 6 predict the data point as belonging to class 1 and rest 3 predict as class 0. Hence, the overall decision made by the random forest classifier for that particular data point is that it belongs to class 1. In the random forest classifier implementation,

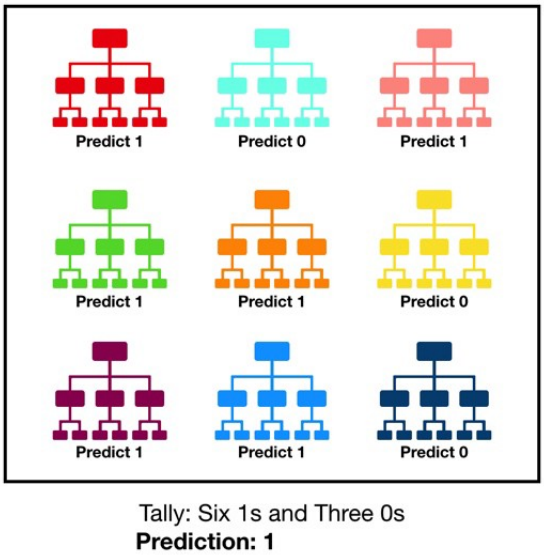

Fig. 5. An ensemble of many decision trees

the number of estimators(n_estimators) were taken as 10 and the criterion for classification was taken as 'entropy'. Classification report for without feature selection is shown as below. The overall accuracy for

Table IV. Classification report for Random Forest Classifier without feature selection

	Precision	Recall	f1-score
Class 0	0.99	0.99	0.99
Class 1	0.98	0.98	0.98
Accuracy			0.99
Micro avg	0.98	0.98	0.98
Macro avg	0.99	0.99	0.99

this technique was reported as 98.60% without feature selection. However, with feature selection, the accuracy was reported as 93.01% and hence its classification report is not being shown here.

- **K-Nearest Neighbor Classifier**: K-Nearest Neighbor is a non-parametric classification algorithm where the result of the test data point depends upon the class of the 'k' nearest train data points. If k=1, then the test data is simply allotted the same class as its nearest train data point. The nearness or proximity in this algorithm is captured using Euclidean distance, which gives the distance between 2 points (x1,y1) and (x2,y2) as

$$\sqrt{(x2-x1)^2+(y2-y1)^2} \quad (1)$$

Other distance functions like Manhattan and Minkowski are also very popular. The following classification report is for KNN classifier without feature selection. The overall accuracy reported for this process was 98.60%. The parameters set for this implementation

Table V. Classification report for KNN without feature selection

	Precision	Recall	f1-score
Class 0	0.99	0.99	0.99
Class 1	0.98	0.98	0.98
Accuracy			0.99
Micro avg	0.98	0.98	0.98
Macro avg	0.99	0.99	0.99

were n_neighbors = 7, weights = 'distance' and n_jobs = 4.

- **Multi Layer Perceptron**: MLP or popularly known as Feed Forward Neural Network, consists of more than one linear layers of neurons. Each layer of neurons is associated with an activation function, which helps in describing the relationships between input-output in a non-linear way. The model training consists of 3 major steps: forward pass, error calculation and backward pass. In forward pass, the input is simply multiplied by weights and bias is added at each layer. When we get some prediction at the output layer after forward pass, error is calculated between ground truth and the predicted value. This happens in the second step of model training. Finally, the loss is used to calculate the gradient, which is back-propagated to update the weights at each layer.

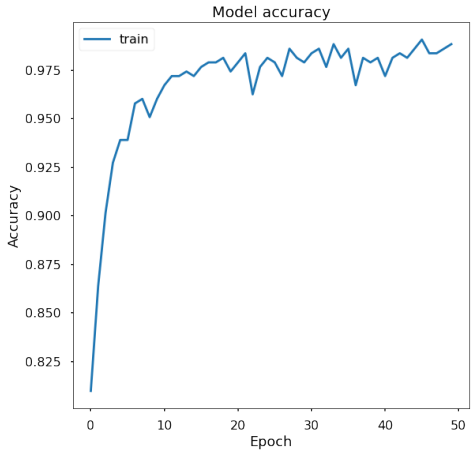

Fig. 6. Accuracy v/s epoch for MLP

In our implementation of MLP, a three-layered Sequential classifier was initialized with all the layers as Dense. First and the second layer had activation functions as ReLu (Rectified Linear Unit) with 64 neurons in input layer, followed by a single hidden layer with ReLu activation function consisting of 64 neurons, while the output layer consists a single neuron having Sigmoid activation function. The model was trained for 50 epochs and batch size was taken as 10. Same methodology was applied for feature selection, which generated a higher accuracy in this case. Plot in figure 6 illustrates the accuracy obtained in this case with the corresponding epochs. The classification report for this model is given in the following table. The accuracy without & with feature selection was found out to be 97.90%.

Table VI. Classification report for MLP with feature selection

	Precision	Recall	f1-score
Class 0	1.00	0.97	0.98
Class 1	0.94	1.00	0.97
Accuracy			0.98
Micro avg	0.97	0.98	0.98
Macro avg	0.98	0.98	0.98

- **FastAI classifier for Tabular Data**: The FastAI is built on PyTorch and provides simplified training of fast and accurate neural networks for vision, text, collaborative filtering and tabular data models. The TabularTranform class available in FastAI handles categorical data itself by assigning a unique ID to them before it is passed through an embedding layer. The continuous variables are automatically normalized before being passed to the model. Apart from this, NaN values are also handled by it. All these functionalities have proven it to be an excellent model and explains why it is being widely used at Google and Pinterest.

  In our implementation, we have used 2 layers with no.

epoch	train_loss	valid_loss	accuracy	time
0	0.352797	0.359494	0.964789	00:00
1	0.208346	0.124768	0.985915	00:00
2	0.146803	0.064235	0.978873	00:00
3	0.122709	0.043100	0.992958	00:00
4	0.109433	0.065139	0.985915	00:00
5	0.091162	0.084181	0.971831	00:00
6	0.081900	0.092581	0.971831	00:00
7	0.072660	0.075714	0.992958	00:00
8	0.063405	0.065535	0.992958	00:00
9	0.056832	0.065717	0.992958	00:00

Fig. 7. Accuracy & losses with epochs in FastAI

of neurons 2000 and 500, with dropout as 0.1. The model has been trained for 10 epochs and learning rate has been chosen as e^{-3}. The overall accuracy without feature selection (99.30%) has been found out to be greater as compared to the model with feature selection (98.59%). The following figure 7 displays the increase in accuracy with increase in epochs.

The learner.record.plot_losses() function was also used to plot the losses with the batches processed in FastAI. Figure 8 shows the same.

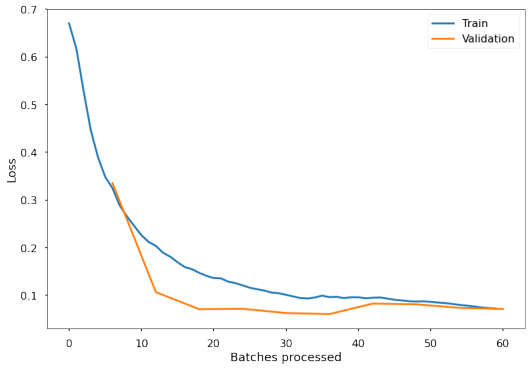

Fig. 8. Loss v/s batches processed in FastAI

III. Results

We have applied six different models to solve the presented classification problem and have compared results for each model. Within each model four variations have been presented.

- Model without feature selection without cross validation
- Model without feature selection with cross validation
- Model with feature selection without cross validation
- Model with feature selection with cross validation

We have presented FastAI as the novel method for classification of breast cancer, along with feature selection using predictive power score.

Hence, we have applied a total of 6*4 = 24 models for the classification problem at hand. Plot showing the accuracy of first variation of the models is as follows. For this variation of

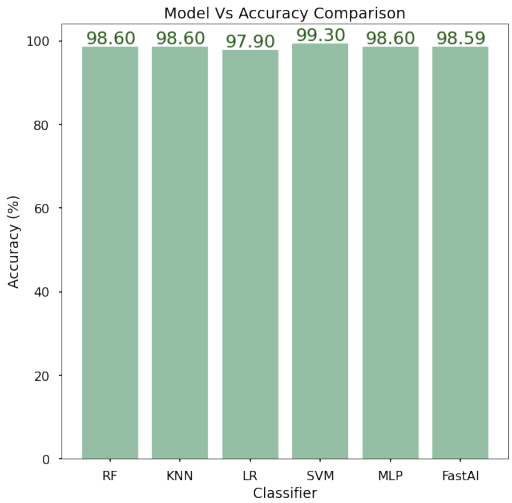

Fig. 9. Models applied without feature selection and without cross-validation

the models, we can identify that highest accuracy of 98.59% has been obtained for SVM. This verifies the claim given in previous works, where SVM has been identified as the most powerful models for classification of breast cancer.

If we compare the second variation of models, which has been

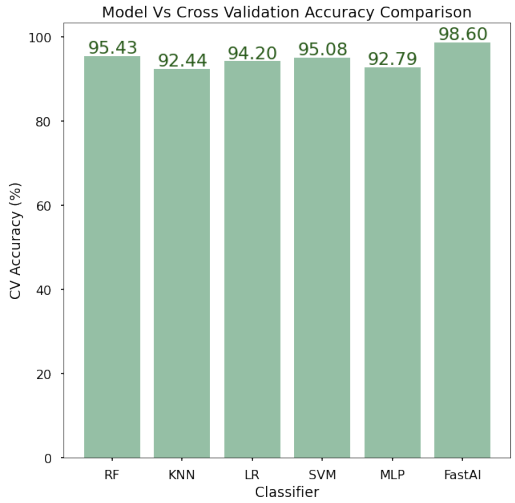

Fig. 10. Models applied without feature selection and with cross-validation

applied without feature selection but using cross validation, we can identify that FastAI performs the best in this case(with an accuracy of 98.60%). Cross validation has been applied using 5 folds of 80:20 ratio of train and test set. Fig 10 illustrates the same. When feature selection was performed, 4 attributes

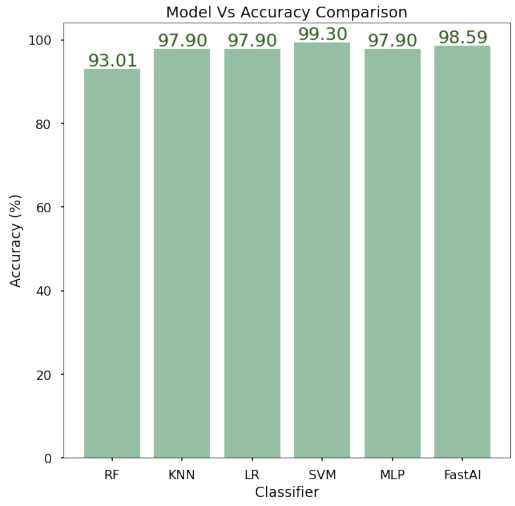

Fig. 11. Models applied with feature selection and without cross-validation

were removed using the predictive power score(PPS) and PCA was applied to transform the test and train data accordingly thereafter. While comparing the models applied with feature selection and without cross validation, it can be observed that SVM has performed the highest(having accuracy of 99.30%). Apart from this, all the other models have also performed better compared to their counterparts without feature selection.

If we compare results for the fourth variation of our models, we can find that FastAI performs better than all of the other models(having and accuracy of 97.72%). In this variation we can identify that accuracy of SVM drops significantly. This can be attributed to the fact that manual split might have caused SVM to give a better accuracy. Since less number of data points are present, cross validation forces to learn on smaller and different train sets. With small data size, this can easily lead to over fitting, and hence test accuracy gets reduced. k-fold validation can decrease over-fitting, but it cannot eliminate it completely. Since, amongst all the model variations FastAI

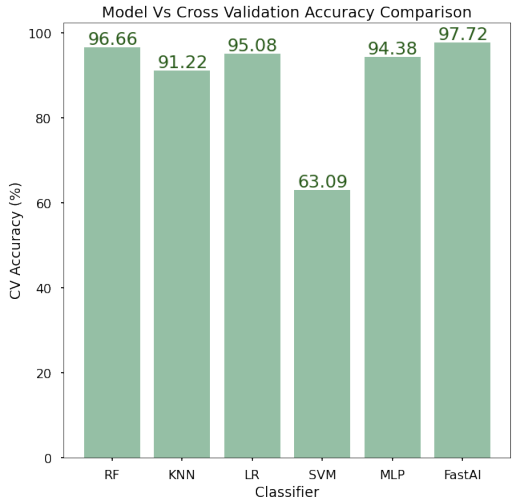

Fig. 12. Models applied with feature selection and with cross-validation

and SVM have performed consistently, we can look at the results of FastAI with a much closer outlook. The following graph shows the accuracy v/s batches processed for a FastAI model with variation 1, i.e., without feature selection and without cross validation. The graph clearly explains how the

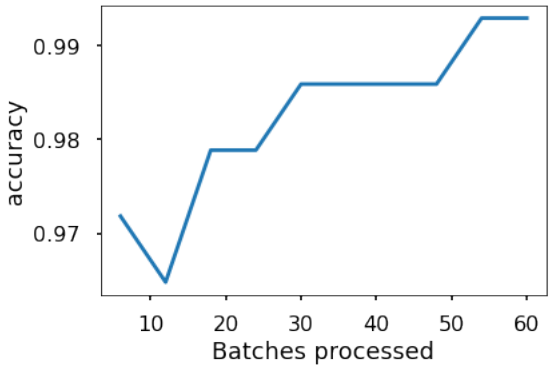

Fig. 13. Accuracy v/s batches processed in Neural Net using FastAI

accuracy of the model increases as more number of batches are processed by the model.

IV. Discussions

A tabular collection of the results produced in the study is also present which states which model performed the best

and for which variation. The highest accuracy of SVM is

Table VII. Comparison of best accuracies of each model

S. No.	Model Name	Accuracy	Variation No.
1	Logistic Regression	97.90%	1 & 3
2	Support Vector Machine	99.30%	1 & 3
3	Random Forest Classifier	98.60%	1
4	K-Nearest Neighbor	98.60%	1
5	Multi layer Perceptron	98.60%	1
6	FastAI classifier	98.60%	2

a clear indication of the potential of SVM to be used for such medical testing purposes. However, these results have been produced without cross-validation. The accuracies of the model with cross-validation is lesser than what is produced with a single train-test split. The more credible results are however, when the models have been applied with cross validation since the model has been trained at all the data points. The cross-validation models used here belong to variation 2 & 4, and the only model performing good in these is the FastAI classifier. FastAI gives a constant and almost similar accuracy for single training and Cross validation training. Deep learning is conventionally only applied to unstructured data like images or audio, and is not said to be credible for tabular data according to the industry standards. However, FastAI beats this notion due to its introduction of embedding layers for the categorical data and the way it normalizes the continuous data in the dataset. However, the FastAI model could also have been improvised using the Optuna framework for hyper-parameter optimization.

Also, there are some limitations to the studies conducted. Few of the critical attributes like *S-phase fraction* and *DNA index* were not available in the dataset [21]. Their inclusion in the dataset may enhance the predictive capabilities of the machine learning models which can be considered as the future scope of this study.

Overall, the studies discussed in this paper have focused on development and advancement of predictive models using supervised machine and deep leaning techniques that can be used by medical experts for early detection of breast cancer in patients in a more economical manner than the conventional diagnosis methods. The analysis of these results signify that the integration of multidimensional data along with different classification, feature selection and dimensionality reduction techniques to build accurate and computationally efficient classifiers for medical applications.

V. Contribution

All the members mentioned in the project have coordinated to produce a viable solution to this classification problem. They have performed data pre-processing in the same manner for all the models applied. However, the contribution can be outlined as follows for each member in the project:

- Rahul Maheshwari – Performed data visualization and analysis of features, plotted correlation between mean features, fine tuned the SVM (Support Vector Machine) using GridSearchCV, performed documentation by providing appropriate comments in the code, performed cross-validation for all models and finally plotted the model vs accuracy plot.
- Surbhi – Applied numerous machine learning techniques like KNN (K-Nearest Neighbor), LR (Logistic Regression) and RF (Random Forest) along with hyper-parameter tuning and plotted results of loss and accuracy for each model.
- Nikunj Agarwal – Outlined the novel feature selection technique of PPS (Predictive Power Score) and PCA (Principle Component Analysis), plotted correlation for all numeric features, applied ML models after PPS and PCA.
- Sumedha Bhatia – Performed data pre-processing and preparation for FastAI, implemented FastAI with hyper-parameter tuning, implemented MLP with hyper-parameter tuning and plotted model vs cross validation plot.

References

[1] Chaurasia, Vikas Pal, Saurabh. (2014). Data mining techniques: To predict and resolve breast cancer survivability. 3. 10-22.

[2] Aruna, S Rajagopalan, Dr Nandakishore, L. (2011). Knowledge based analysis of various statistical tools in detecting breast cancer. Computer Science Information Technology. 2. 10.5121/csit.2011.1205.

[3] Hiba Asri, Hajar Mousannif, Hassan Al Moatassime, Thomas Noel, Using Machine Learning Algorithms for Breast Cancer Risk Prediction and Diagnosis, Procedia Computer Science, Volume 83, 2016, Pages 1064-1069, ISSN 1877-0509, https://doi.org/10.1016/j.procs.2016.04.224.

[4] https://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+(Diagnostic)

[5] https://towardsdatascience.com/scale-standardize-or-normalize-with-scikit-learn-6ccc7d176a02

[6] https://towardsdatascience.com/rip-correlation-introducing-the-predictive-power-score-3d90808b9598

[7] Joliffe, I., & Morgan, B. (1992). Principal component analysis and exploratory factor analysis. Statistical Methods in Medical Research, 1(1), 69–95.

[8] Svante Wold, Kim Esbensen, Paul Geladi, Principal component analysis, Chemometrics and Intelligent Laboratory Systems, Volume 2, Issues 1–3, 1987, Pages 37-52, ISSN 0169-7439, https://doi.org/10.1016/0169-7439(87)80084-9.

[9] M. Gupta and B. Gupta, "A Comparative Study of Breast Cancer Diagnosis Using Supervised Machine Learning Techniques," 2018 Second International Conference on Computing Methodologies and Communication (ICCMC), Erode, 2018, pp. 997-1002, doi: 10.1109/ICCMC.2018.8487537.

[10] Vanneschi, L., Farinaccio, A., Mauri, G. et al. A comparison of machine learning techniques for survival prediction in breast cancer. BioData Mining 4, 12 (2011). https://doi.org/10.1186/1756-0381-4-12

[11] S.Kharya et al "A Comparative Study of Breast Cancer Diagnosis Using Supervised Machine Learning Techniques" (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 4 (6) , 2013, 1023-1028

[12] Pedro Henriques Abreu, Miriam Seoane Santos, Miguel Henriques Abreu, Bruno Andrade, and Daniel Castro Silva. 2016. Predicting Breast Cancer Recurrence Using Machine Learning Techniques: A Systematic Review. ACM Comput. Surv. 49, 3, Article 52 (December 2016), 40 pages. DOI:https://doi.org/10.1145/2988544

[13] A. Osareh and B. Shadgar, "Machine learning techniques to diagnose breast cancer," 2010 5th International Symposium on Health Informatics and Bioinformatics, Antalya, 2010, pp. 114-120, doi: 10.1109/HIBIT.2010.5478895.

[14] Ch. Shravya, K. Pravalika, Shaik Subhani, "Prediction of Breast Cancer Using Supervised Machine Learning Techniques", International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-8 Issue-6, April 2019.

[15] Hiba Asri, Hajar Mousannif, Hassan Al Moatassime, Thomas Noel, Using Machine Learning Algorithms for Breast Cancer Risk Prediction and Diagnosis, Procedia Computer Science, Volume 83, 2016, Pages 1064-1069, ISSN 1877-0509, https://doi.org/10.1016/j.procs.2016.04.224.

[16] https://www.cancer.net/cancer-types/breast-cancer/diagnosis

[17] Sivakami, K.. "Mining Big Data: Breast Cancer Prediction using DT-SVM Hybrid Model." (2015).

[18] U. K. Kumar, M. B. S. Nikhil and K. Sumangali, "Prediction of breast cancer using voting classifier technique," 2017 IEEE International Conference on Smart Technologies and Management for Computing, Communication, Controls, Energy and Materials (ICSTM), Chennai, 2017, pp. 108-114, doi: 10.1109/ICSTM.2017.8089135.

[19] M. M. Islam, H. Iqbal, M. R. Haque and M. K. Hasan, "Prediction of breast cancer using support vector machine and K-Nearest neighbors," 2017 IEEE Region 10 Humanitarian Technology Conference (R10-HTC), Dhaka, 2017, pp. 226-229, doi: 10.1109/R10-HTC.2017.8288944.

[20] Lingxi Peng, Wenbin Chen, Wubai Zhou, Fufang Li, Jin Yang, Jiandong Zhang, "An immune-inspired semi-supervised algorithm for breast cancer diagnosis", Computer Methods and Programs in Biomedicine, Volume 134, 2016, Pages 259-265, ISSN 0169-2607, https://doi.org/10.1016/j.cmpb.2016.07.020.

[21] Eshlaghy, A.T. Pourebrahimi, Alireza Ebrahimi, Mansour Razavi, A.R. Ghasem Ahmad, Leila. (2013). Using three machine learning techniques for predicting breast cancer recurrence. Journal of Health Medical Informatics. 4. 124-130.