negatives.It define how often a positive class is predicted when the actual outcome is negative.

$$FPR = \frac{False\ Positives}{(False\ Positives + True\ Negatives)} \quad (2)$$

ROC curve shows the true positive rate, on the y-axis and is plotted against the false positive rate represent on the x axis [10]. The values of both x and y-axis are spread from Zero to One. The graph is generated from measuring true positive rate and false positive rate for each feasible classifier threshold value[11]. Figure 1 shows the receiver operating curves of each machine-learning algorithm. AUROC, which is,stands for "Area under the ROC Curve." measures the entire two-dimensional area cover by ROC curve. The area under this curve is calculated in which larger area covered by respective classifier represent the better the performance. In our work as shown in above fig. 1 the Support vector machine (SVM) have greatest AUC (Area Under the ROC) Curve score of 99.8%, followed by Logistic regression, which had an AUC of 99.3%. Also it is shown in Fig. 1, that the Decision tree has the lowest AUC of 95.6 percent.

IV. Conclusion

In this work we observed WDBC dataset and use various classifiers classify malignant and benign tumors. The results are comparing, calculate and evaluate based on confusion matrix, precision, sensitivity and accuracy[12]. The experiment is set up in Python, using NumPy library, pandas, SciKit learn, Matplot Lib. After simulating the program and comparing the different models it is found that SVM has demonstrated its accuracy and achieve best performance in prediction of breast cancer. SVM achieved a highest accuracy of 98.24% with AUC 0.998, 99 precession in Malignant and 96 in Benign, which is better than all other algorithm. All the results are obtained by using WDBC dataset, same algorithm and model can be use for other dataset in future to get better result. In future, we will work on latest dataset with more disease classes to obtain higher accuracy with another machine learning.

References

[1] International Agency for Research on Cancer, Press release December 2020, https://www.who.int/news/item/03-02-2021-breast-cancer-now-most-common-form-of-cancer-who-taking-action.

[2] L.G. Ahmad, A.T. Eshlaghy, A. Poorebrahimi, M. Ebrahimi and A.R. Razavi, “Using three machine learning techniques for predicting breast cancer recurrence,” (2013), J Health Med Inform 4: 124. doi:10.4172/2157-7420.1000124

[3] S. Nayak, D. Gope “Comparison of supervised learning algorithms for RF-based breast cancer detection,”, 2017 Computing and Electromagnetics International Workshop (CEM), Barcelona (2017)

[4] M. H. Memon, J. P. Li, A. U. Haq, M. H. Memon, and W. Zhou, ‘‘Breast cancer detection in the IOT health environment using modified recursive feature selection,’’ Wireless Commun. Mobile Comput., vol. 2019, pp. 1–19, Nov. 2019

[5] S. Alghunaim and H. H. Al-Baity, ‘‘On the scalability of machinelearning algorithms for breast cancer prediction in big data context,’’ IEEE Access, vol. 7, pp. 91535–91546, 2019.

[6] H. Asri, H. Mousannif, H. A. Moatassime, and T. Noel, “Using machine learning algorithms for breast cancer risk prediction and diagnosis,” Procedia Computer Science, vol. 83, pp. 1064–1069, 2016.

[7] B. M. Gayathri and C. P. Sumathi, “Comparative study of relevance vector machine with various machine learning techniques used for detecting breast cancer,” 2016 IEEE Int. Conf. on Computational Intelligence and Computing Research (ICCIC), pp 1-5, IEEE, 2016

[8] UCI Machine Learning Repository. [Online]. Available: https://archive.ics.uci.edu/ml/datasets/breast+cancer+wisconsin+(diagnostic) Accessed [August] [2021].

[9] Poonam Pandey and Radhika Prabhakar, “An analysis of machine learning techniques (J48 & Ada Boost) - for classification,” 2016 1st India Int. Conf. on Information Processing (IICIP), PP 1- 6, IEEE, 2016, India.

[10] L. F. Carvalho, G. Fernandes, M. V. O. De Assis, J. J. P. C. Rodrigues, and M. Lemes Proenca, “Digital signature of network segment for healthcare environments support,” Irbm, vol. 35, no. 6, pp. 299-309, 2014.

[11] Dana Bazazeh and Raed Shubair. “Comparative study of machine learning algorithms for breast cancer detection and diagnosis,” 2016 5th Int. Conf. on Electronic Devices, Systems and Applications (ICEDSA), 6-8 December 2016, Ras Al Khaimah, UAE.

[12] Zahra Nematzadeh, Roliana Ibrahim and Ali Selamat, “Comparative studies on breast cancer classifications with k-fold cross validations using machine learning techniques,” Proc. in 2015 10th Asian Control Conf. (ASCC), pp 1-6, IEEE, 2015.