

Breast Cancer Prediction using Machine Learning

Prof. Dr. Prakash. H. Patil¹, Prachi Andure², Anushka Patil³, Anshul Kumar⁴, Snehal Sawant⁵

Department of Information Technology^{1,2,3,4,5}

Dr. D Y Patil College of Engineering, Ambi, Pune, India

Abstract: Women are seriously threatened by breast cancer with high morbidity and mortality. The lack of robust prognosis models results in difficulty for doctors to prepare a treatment plan that may prolong patient survival time. Hence, the requirement of time is to develop the technique which gives minimum error to increase accuracy. Four algorithm SVM, Logistic Regression, Random Forest and KNN which predict the breast cancer outcome have been compared in the paper using different datasets. All experiments are executed within a simulation environment and conducted in JUPYTER platform. Aim of research categories in three domains. First domain is prediction of cancer before diagnosis, second domain is prediction of diagnosis and treatment and third domain focuses on outcome during treatment. The proposed work can be used to predict the outcome of different technique and suitable technique can be used depending upon requirement. This research is carried out to predict the accuracy. The future research can be carried out to predict the other different parameters and breast cancer research can be categories on basis of other parameters. **Keywords** — Breast Cancer, machine learning, feature selection, classification, prediction, KNN, Random Forest, ROC.

Keywords: Cancer.

REFERENCES

- [1] Wang, D. Zhang and Y. H. Huang “Breast Cancer Prediction Using Machine Learning” (2018), Vol. 66, NO. 7.
- [2] B. Akbugday, "Classification of Breast Cancer Data Using Machine Learning Algorithms," 2019 Medical Technologies Congress (TIPTEKNO), Izmir, Turkey, 2019, pp. 1-4.
- [3] Keles, M. Kaya, "Breast Cancer Prediction and Detection Using Data Mining Classification Algorithms: A Comparative Study." Tehnicki Vjesnik - Technical Gazette, vol. 26, no. 1, 2019, p. 149+.
- [4] V. Chaurasia and S. Pal, “Data Mining Techniques: To Predict and Resolve Breast Cancer Survivability”, IJCSMC, Vol. 3, Issue. 1, January 2014, pg.10 – 22.
- [5]] Delen, D.; Walker, G.; Kadam, A. Predicting breast cancer survivability: A comparison of three data mining methods. Artif. Intell. Med. 2005, 34, 113–127.
- [6] R. K. Kavithal, D. D. Rangasamy, “Breast Cancer Survivability Using Adaptive Voting Ensemble Machine Learning Algorithm Adaboost and CART Algorithm” Volume 3, Special Issue 1, February 2014
- [7] P. Sinthia, R. Devi, S. Gayathri and R. Sivasankari, “Breast Cancer detection using PCPCET and ADEWNN”, CIEEE’ 17, p.63-65
- [8] Vikas Chaurasia and S.Pal, “Using Machine Learning Algorithms for Breast Cancer Risk Prediction and Diagnosis” (FAMS 2016) 83 (2016) 1064 – 1069
- [9] N. Khuriwal, N. Mishra. “A Review on Breast Cancer Diagnosis in Mammography Images Using Deep Learning Techniques”, (2018), Vol. 1, No. 1.
- [10] Y. Khourdifi and M. Bahaj, "Feature Selection with Fast Correlation-Based Filter for Breast Cancer Prediction and Classification Using Machine Learning Algorithms," 2018 International Symposium on Advanced Electrical and Communication Technologies (ISAECT), Rabat, Morocco, 2018, pp. 1-6.
- [11] R. M. Mohana, R. Delshi Howsalya Devi, Anita Bai, “Lung Cancer Detection using Nearest Neighbour Classifier”, International Journal of Recent Technology and Engineering (IJRTE), Volume-8, Issue-2S11, September 2019

- [12] Ch. Shravya, K. Pravalika, Shaik Subhani, "Prediction of Breast Cancer Using Supervised Machine Learning Techniques", International Journal of Innovative Technology and Exploring Engineering (IJITEE), Volume-8 Issue-6, April 2019.
- [13] Haifeng Wang and Sang Won Yoon, "Breast Cancer Prediction Using Data Mining Method", Proceedings of the 2015 Industrial and Systems Engineering Research Conference,
- [14] Abdelghani Bellaachia, Erhan Guven, "Predicting Breast Cancer Survivability Using Data Mining Techniques
- [15] Juhyeon Kim, Hyunjung Shin, Breast cancer survivability prediction using labeled, unlabeled, and pseudo-labeled patient data, Journal of the American Medical Informatics Association, Volume 20, Issue 4, July 2013, Pages 613–618.
- [16] N. Khuriwal and N. Mishra, "Breast cancer diagnosis using adaptive voting ensemble machine learning algorithm," 2018 IEEMA Engineer Infinite Conference (eTechNxT), New Delhi, 2018, pp. 1-5.
- [17] M. Amrane, S. Oukid, I. Gagaoua and T. EnsarĪ, "Breast cancer classification using machine learning," 2018 Electric Electronics, Computer Science, Biomedical Engineering's Meeting (EBBT), Istanbul, 2018, pp. 1-4.
- [18] M. R. Al-Hadidi, A. Alarabeyyat and M. Alhanahnah, "Breast Cancer Detection Using K-Nearest Neighbor Machine Learning Algorithm," 2016 9th International Conference on Developments in eSystems Engineering (DeSE), Liverpool, 2016, pp. 35-39.
- [19] Kibeom Jang, Minsoon Kim, Candace A Gilbert, Fiona Simpkins, Tan A Ince, Joyce M Slingerland "WEGFA activates an epigenetic pathway regulating ovarian cancer initiating cells" Embo Molecular Medicines Volume 9 Issue 3 (2017)
- [20] Joseph A. Cruz and David S. Wishart "Applications of Machine Learning in cancer prediction and prognosis Cancer informatics" 2(3):59-77 · February 2007
- [21] SA Medjahed, TA Saadi, A Benyettou "Breast cancer diagnosis by using k-nearest neighbor with different distances and classification rules" International Journal of Computer Applications 62 (1), 2013