PREDICTION OF BREAST CANCER USING NOVEL MULTI LAYER PERCEPTRON IN COMPARISON WITH CART TO IMPROVE ACCURACY

C. Gnanendhra Reddy[1], S. Magesh Kumar[2*]

Article History: Received: 12.12.2022 **Revised:** 29.01.2023 **Accepted:** 15.03.2023

Abstract

Aim: The Objective of the work is to predict the Accuracy of Breast Cancer using a Novel Multi Layer Perceptron comparative with Cart Algorithm.
Materials and Methods: The accuracy and loss are performed with the dataset from the Github library. The total sample size is 48. The two groups of Novel Multi Layer Perceptron (N=10) and Cart Algorithm (N=10) were proposed by predicting the accuracy of 93.50% Breast Cancer prediction compared with Cart Algorithm 89.60%. **Results:** The results proved that Novel Multi Layer Perceptron achieves better accuracy than the Cart Algorithm. The Cart Algorithm appears significantly better than Feature Selection. The Statistical insignificance difference between the Novel Multi Layer Perceptron and Cart algorithm is found to be $p=0.523$ ($p<0.05$). **Conclusion:** The Results proved that Novel Multi Layer Perceptron helps to predict Breast Cancer Prediction and gives more accuracy.

Keywords: Screening, Breast Cancer, Machine Learning, Mammogram, Novel Multi Layer Perceptron, Cart Algorithm, Prediction.

[1]Research Scholar, Department of Computer Science and Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and technical Science, Saveetha University, Chennai, Tamil Nadu, India. Pincode: 602105.
[2*]Project Guide, Department of Computer Science and Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai, Tamil Nadu, India. Pincode: 602105.

1. introduction

Breast Cancer prediction is a type of tumor that occurs in the tissue of the breast. A mammogram is a picture of the breasts taken using x-ray imaging (Baum 1993). It is one of the most commonly used screening tools to detect breast cancer (Loke et al. 2022). In some cases, mammograms allow doctors to detect breast cancer up to three years before a tumor is palpable. A woman with an average risk of breast cancer can benefit from getting mammograms starting at age 40 (Bychkov et al. 2022). However, a woman with a higher risk, i.e., has a family history of breast cancer, may want to start screening earlier. It is very important to know the family history so that the best course of action can be taken. For instance, if a mother had breast cancer at age 45, then the doctor may suggest her children start the mammogram screenings ten years prior to her age of diagnosis, which is at 35 ("Risk Prediction in Breast Cancer" 2007).

The rapid development of machine learning, a family of machine learning techniques, has spurred much interest in its application to medical imaging problems. Here, we develop a machine learning algorithm that can accurately detect breast cancer on screening mammograms using an "end to end"(Balasubramanian 2021; Malla, Cs/it, and Dehradun 2017) training approach that efficiently leverages training datasets with either complete clinical annotation or only the cancer status of the image (Bartlett 2010). Machine learning techniques have shown significant performance in various healthcare applications over the traditional computer-aided system for disease diagnosis and patient monitoring (Basmadjian et al. 2022).Our team has extensive knowledge and research experience that has translated into high quality publications(Pandiyan et al. 2022; Yaashikaa, Devi, and Kumar 2022; Venu et al. 2022; Kumar et al. 2022; Nagaraju et al. 2022; Karpagam et al. 2022; Baraneedharan et al. 2022; Whangchai et al. 2022; Nagarajan et al. 2022; Deena et al. 2022)
In a previous study, the efficiency improvement of the Novel Multi Layer Perceptron in Data retrieval was not properly considered to increase accuracy (Darabi et al. 2012). To overcome this issue, a Novel Multilayer Algorithm is implemented to improve secured data retrieval with reduced time in machine learning (Balasubramanian 2021; Salem Abdull 2011).

2. Materials and Methods

The research work was carried out in the Machine Learning laboratory lab at Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai. The number of groups is 2 in which Group 1 Algorithm is Novel Multi Layer Perceptron Group 2 Algorithm is Cart with Sample Size 20. The dataset used was the Breast Cancer prediction Data set in 260 instances and 33 attributes. Novel Multi Layer Perceptron is compared with Cart Algorithm (Walker and Eeles 2007). It consists of 3 nodes: Input layer, hidden layer, and output layer. Except for input, each node is a hidden layer and output layer. Except for input, each node is a neuron that uses a nonlinear activation function. Multi Layer Perceptron is applied using different numbers of layers and neurons to examine the effect of the number of layers of classification

Novel Multi Layer Perceptron
Breast cancer forms in breast cells and is considered a very common type of cancer in women. Breast cancer is also a very life-threatening disease of women after lung cancer. A Novel Multi Layer Perceptron method is proposed in this study to boost the automatic identification of breast cancer by analyzing hostile ductal carcinoma tissue zones in whole-slide images (WSIs). The paper investigates the proposed system that uses various Novel Multi Layer Architectures to automatically detect breast cancer prediction, comparing the results with those from Machine Learning. Multi Layer Perceptron models have been effectively used in wide-ranging computer vision fields for years.

Cart Algorithm
Cart Algorithm a type of classification algorithm is used to predict and the probability of target variable. Feature selection is a supervised learning classification algorithm used to predict the probability of a target variable. The nature of the target or dependent variable is dichotomous, which means there would be only two possible classes mathematically; a Cart Algorithm model predicts P (y=1) as the function of X.
The software tool to evaluate Multi Layer Perceptron and Cart algorithms is colab in Python programming language, The hardware configuration includes an Intel i5 processor with a RAM size of 8GB. The system used was a 64-bit Windows 10 operating system.

Statistical Analysis
SPSS software is used for statistical analysis of Multi Layer Perceptron andCart algorithm methods. The independent variable is Multi Layer Perceptron accuracy and the dependent variable is log loss. The independent T test analyses were carried out to calculate the accuracy of the Multi Layer Perceptron for both methods.

3. Results

Table 1 shows the simulation results of the proposed algorithm Novel Multi Layer Perceptron and the Cart Algorithm. Encryption was run at different times in the Colab with a sample size of 20.

Table 2 represents the T-test comparison of both Novel Multi Layer Perceptron and MultiLayer Perceptron. The mean Standard Deviation and Standard Error mean were calculated by taking an independent variable T-test among study groups. The Novel Multi Layer Perceptron produces a significant difference from the Cart Algorithm with a value of 0.712 and effect size of -1.414. The mean accuracy of the Novel Multi Layer Perceptron is 93.50% and the Cart Algorithm is 89.60%.

Table 3 represents the independent sample T-test with a confidence interval at 95% and a level of significance of 0.05. It shows the statistical significance P<0.05 -tailed. The Statistical insignificance difference between the Novel Multi Layer Perceptron and Cart algorithm is found to be p=0.523.

Figure 1 gives the comparison chart of the Novel Multi Layer Perceptron and Cart Algorithm in terms of mean accuracy. The accuracy of the Novel Multi Layer Perceptron is better than the Cart Algorithm.

4. Discussion

Based on the above study, the Novel Multi Layer Perceptron has better accuracy 93.50% than the Cart Algorithm with an accuracy of 89.60%. There is a statistically 2-tailed insignificant difference in accuracy for two algorithms that is p=0.523 (p<0.05) by performing an independent sample T-test.

In this paper, the accuracy for the Novel Multi Layer Perceptron and Cart algorithm is found to be 93.50% and 89.60% respectively. The analysis of this paper is to make use of machine learning to predict breast cancer. Multi Layer Perceptron was implemented with an accuracy percentage of 89.60% (Philippe 2011; Chen et al. 2022). Novel Multi Layer Perceptron combined with Cart Algorithm predicts a time consumption of 88 w%. Breast Cancer prediction is one of the most commonly diagnosed cancer types in women and automatically classifies breast cancer histopathological image (Malla, Cs/it, and Dehradun 2017). Research efforts have reported with increasing confirmation that the multi layer perceptron has greater accurate diagnosis ability (Miller et al. 2020). This work uses the prediction of breast cancer analysis to compare the algorithms of Novel Multi Layer Perceptron and Cart algorithm of accuracy (93.50% and 89.60%). The attributes are mainly concentrated on decreasing the time consumption of women. Novel multi Layer Perceptrons are compared with multi layer perceptrons that are discussed in previous research articles (Taleghamar et al. 2022).

The suggested work limitation is that small changes to the dataset can cause the algorithm to destabilize the tree structure and cause discrepancies. If some classes are imbalanced, decision tree learners create poorly serviced trees. Therefore, it is advisable to balance the dataset before adapting it to the decision tree. The future scope of the project is that one in 28 women in India is susceptible to breast cancer because early detection techniques for the presence of breast cancer still lack accurate predictions of breast cancer. In addition, lack of consciousness, precautionary measures, and treatment facilities increase the risk of survival. Early detection of the syndrome can be aimed at overcoming breast cancer through appropriate treatment.

5. Conclusion

The work involves the Novel Multi Layer Perceptron algorithm to predict breast cancer with reduced time to be proved with better accuracy of 93.50% when compared to cart algorithm accuracy of 89.60%. The purpose of this comparative analysis was to find the most accurate machine learning algorithm that could serve as a tool for diagnosing breast tissue according to the predicted results. It can be concluded that the Novel Multi Layer Perceptron has the highest accuracy for the given dataset.

Declarations

Conflicts of interest

No conflicts of interest in this manuscript.

Authors Contribution

Author SIL was involved in data collection, data analysis, manuscript writing. Author NBD was involved in conceptualization, data validation and critical review of the manuscript.

Acknowledgments

The authors would like to express their gratitude towards Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences (Formerly Known as Saveetha University) for providing the necessary infrastructure to carry out this work successfully.

Funding

We thank the following organization for providing financial support that enabled us to complete the study.

1. Cyclotron Technologies.
2. Saveetha University.
3. Saveetha Institute of Medical and Technical Science.
4. Saveetha School of Engineering.

6. References

Balasubramanian, Sanjana. 2021. *Breast Cancer Prediction Using Machine Learning*.

Baraneedharan, P., Sethumathavan Vadivel, C. A. Anil, S. Beer Mohamed, and Saravanan Rajendran. 2022. "Advances in Preparation, Mechanism and Applications of Various Carbon Materials in Environmental Applications: A Review." *Chemosphere*. https://doi.org/10.1016/j.chemosphere.2022.134596.

Bartlett, J. 2010. "Biomarkers for the Diagnosis and Prediction of Therapeutic Response in Clinical Breast Cancer." *Breast Cancer Research*. https://doi.org/10.1186/bcr2490.

Basmadjian, Robert B., Shiying Kong, Devon J. Boyne, Tamer N. Jarada, Yuan Xu, Winson Y. Cheung, Sasha Lupichuk, May Lynn Quan, and Darren R. Brenner. 2022. "Developing a Prediction Model for Pathologic Complete Response Following Neoadjuvant Chemotherapy in Breast Cancer: A Comparison of Model Building Approaches." *JCO Clinical Cancer Informatics* 6 (February): e2100055.

Baum, M. 1993. "Prognosis and Prediction for Early Breast Cancer." *Adjuvant Therapy of Breast Cancer IV*. https://doi.org/10.1007/978-3-642-84745-5_12.

Bychkov, Dmitrii, Heikki Joensuu, Stig Nordling, Aleksei Tiulpin, Hakan Kücükel, Mikael Lundin, Harri Sihto, et al. 2022. "Outcome and Biomarker Supervised Deep Learning for Survival Prediction in Two Multicenter Breast Cancer Series." *Journal of Pathology Informatics* 13 (January): 9.

Chen, Jieying, Xiaolian Su, Tingting Xu, Qifeng Luo, Lin Zhang, and Guangyu Tang. 2022. "Stratification of Axillary Lymph Node Metastasis Risk with Breast Magnetic Resonance Imaging in Breast Cancer." *Future Oncology* , February. https://doi.org/10.2217/fon-2021-1559.

Darabi, Hatef, Kamila Czene, Wanting Zhao, Jianjun Liu, Per Hall, and Keith Humphreys. 2012. "Breast Cancer Risk Prediction and Individualised Screening Based on Common Genetic Variation and Breast Density Measurement." *Breast Cancer Research*. https://doi.org/10.1186/bcr3110.

Deena, Santhana Raj, A. S. Vickram, S. Manikandan, R. Subbaiya, N. Karmegam, Balasubramani Ravindran, Soon Woong Chang, and Mukesh Kumar Awasthi. 2022. "Enhanced Biogas Production from Food Waste and Activated Sludge Using Advanced Techniques – A Review." *Bioresource Technology*. https://doi.org/10.1016/j.biortech.2022.127234.

Karpagam, M., R. Beaulah Jeyavathana, Sathiya Kumar Chinnappan, K. V. Kanimozhi, and M. Sambath. 2022. "A Novel Face Recognition Model for Fighting against Human Trafficking in Surveillance Videos and Rescuing Victims." *Soft Computing*. https://doi.org/10.1007/s00500-022-06931-1.

Kumar, P. Ganesh, P. Ganesh Kumar, Rajendran Prabakaran, D. Sakthivadivel, P. Somasundaram, V. S. Vigneswaran, and Sung Chul Kim. 2022. "Ultrasonication Time Optimization for Multi-Walled Carbon Nanotube Based Therminol-55 Nanofluid: An Experimental Investigation." *Journal of Thermal Analysis and Calorimetry*. https://doi.org/10.1007/s10973-022-11298-4.

Loke, Johnny, Ishraq Alim, Sarah Yam, Susan Klugman, Li C. Xia, Dorota Gruber, David Tegay, Andrea LaBella, Kenan Onel, and Harry Ostrer. 2022. "Prediction of Breast Cancer Risk Based on Flow Variant Analysis of Circulating Peripheral Blood Mononuclear Cells." *HGG Advances* 3 (2): 100085.

Malla, Younus Ahmad, Student Cs/it, and Bfit Dehradun. 2017. "A Machine Learning Approach for Early Prediction of Breast Cancer." *International Journal Of Engineering And Computer Science*. https://doi.org/10.18535/ijecs/v6i5.31.

Miller, Eric A., Paul F. Pinsky, Brandy M. Heckman-Stoddard, and Lori M. Minasian. 2020. "Breast Cancer Risk Prediction Models and Subsequent Tumor Characteristics." *Breast Cancer*. https://doi.org/10.1007/s12282-020-01060-9.

Nagarajan, Karthik, Arul Rajagopalan, S. Angalaeswari, L. Natrayan, and Wubishet Degife Mammo. 2022. "Combined Economic Emission Dispatch of Microgrid with the Incorporation of Renewable Energy Sources Using Improved Mayfly Optimization Algorithm." *Computational Intelligence and Neuroscience* 2022 (April): 6461690.

Nagaraju, V., B. R. Tapas Bapu, P. Bhuvaneswari, R. Anita, P. G. Kuppusamy, and S. Usha. 2022. "Role of Silicon Carbide Nanoparticle on Electromagnetic Interference Shielding

Behavior of Carbon Fibre Epoxy Nanocomposites in 3-18GHz Frequency Bands." *Silicon*. https://doi.org/10.1007/s12633-022-01825-1.

Pandiyan, P., R. Sitharthan, S. Saravanan, Natarajan Prabaharan, M. Ramji Tiwari, T. Chinnadurai, T. Yuvaraj, and K. R. Devabalaji. 2022. "A Comprehensive Review of the Prospects for Rural Electrification Using Stand-Alone and Hybrid Energy Technologies." *Sustainable Energy Technologies and Assessments*. https://doi.org/10.1016/j.seta.2022.102155.

Philippe, Becuwe. 2011. "Antioxidant Enzymes as New Biomarkers for Prediction of Tumor Progression in Breast Cancer." *Breast Cancer - Recent Advances in Biology, Imaging and Therapeutics*. https://doi.org/10.5772/23004.

"Risk Prediction in Breast Cancer." 2007. *Hereditary Breast Cancer*. https://doi.org/10.3109/9781420020366-6.

Salem Abdull, Mohamed A. 2011. *Data Mining Techniques and Breast Cancer Prediction: A Case Study of Libya*.

Taleghamar, Hamidreza, Seyed Ali Jalalifar, Gregory J. Czarnota, and Ali Sadeghi-Naini. 2022. "Deep Learning of Quantitative Ultrasound Multi-Parametric Images at Pre-Treatment to Predict Breast Cancer Response to Chemotherapy." *Scientific Reports* 12 (1): 2244.

Venu, Harish, Ibham Veza, Lokesh Selvam, Prabhu Appavu, V. Dhana Raju, Lingesan Subramani, and Jayashri N. Nair. 2022. "Analysis of Particle Size Diameter (PSD), Mass Fraction Burnt (MFB) and Particulate Number (PN) Emissions in a Diesel Engine Powered by Diesel/biodiesel/n-Amyl Alcohol Blends." *Energy*. https://doi.org/10.1016/j.energy.2022.123806.

Walker, Lisa, and Rosalind Eeles. 2007. "Risk Prediction in Breast Cancer." *Hereditary Breast Cancer*. https://doi.org/10.3109/9781420020366-3.

Whangchai, Niwooti, Daovieng Yaibouathong, Pattranan Junluthin, Deepanraj Balakrishnan, Yuwalee Unpaprom, Rameshprabu Ramaraj, and Tipsukhon Pimpimol. 2022. "Effect of Biogas Sludge Meal Supplement in Feed on Growth Performance Molting Period and Production Cost of Giant Freshwater Prawn Culture." *Chemosphere* 301 (August): 134638.

Yaashikaa, P. R., M. Keerthana Devi, and P. Senthil Kumar. 2022. "Advances in the Application of Immobilized Enzyme for the Remediation of Hazardous Pollutant: A Review." *Chemosphere* 299 (July): 134390.

Tables and Figures

Table 1. Predicted Accuracy of Breast Cancer Prediction for 20 different sample sizes.

GROUP	Algorithms	Average accuracy
1	Novel Multi Layer Perceptron	93.50%
2	Cart Algorithm	89.60%

Table 2. Statistical analysis of Novel Multi layer perceptron and cart algorithm. Mean accuracy, standard deviation, and standard error values are obtained for 20 sample datasets.

	Algorithms	N	Mean	Std.Deviation	Std.Error Mean
Accuracy	Novel Multi Layer Perceptron	10	93.5000	0.23330	0.5109
	Cart Algorithm	10	89.6020	0.18030	0.03940

Table 3. Independent sample T-test with a confidence interval at 95% and level of significance as 0.05.

	Levene's Test for Equality of Variances	T-test for Equality of Means

	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std.Error Difference	95% Confidence Interval of the Differences Lower | Upper	
Equal Variances assumed	.425	.523	42.386	18	.000	3.95200	.09324	3.7561	4.14789
Equal Variances not assumed	-	-	42.386	16.924	.000	3.95200	.09324	3.7552	4.14878

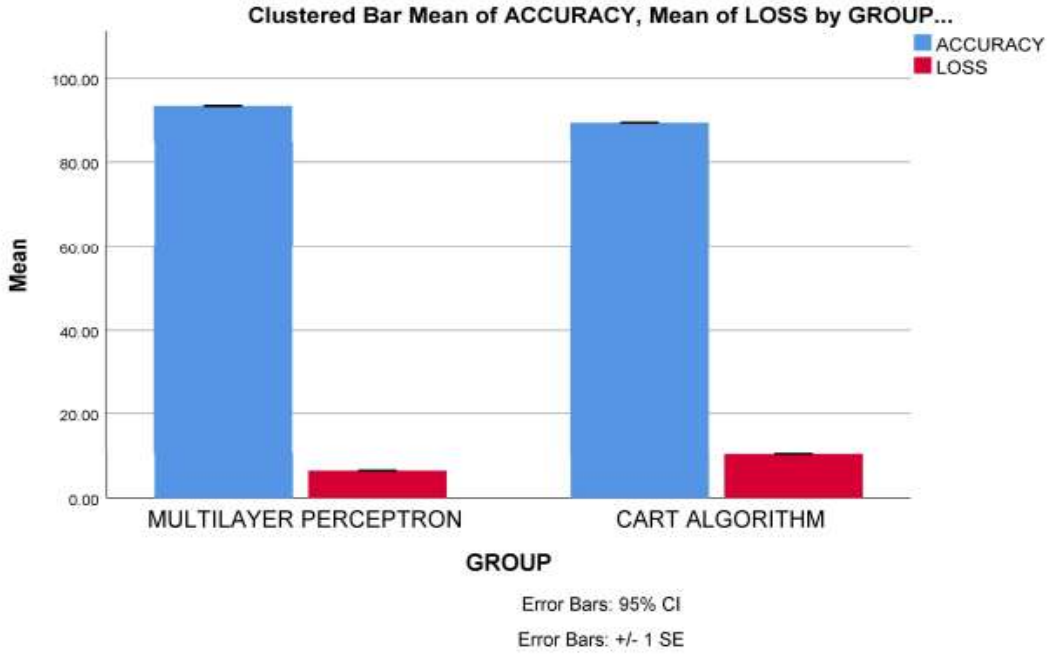

Fig. 1. Comparison of Novel-cart ks and cart algorithm in terms of means and accuracy. The mean accuracy of the Novel-cart is better than the cart algorithm. X-axis: Novel Multi Layer Perceptron vs Cart algorithm, Y-axis: Mean accuracy. Error Bar +/-1 SD with 95% Cl.