PREDICTION OF BREAST CANCER USING NOVEL MULTI LAYER PERCEPTRON IN COMPARISON WITH SUPPORT VECTOR MACHINE TO IMPROVE ACCURACY

S. Iyswarya Lakshmi[1], N. Bharatha Devi[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 the novel multi layer perceptron comparative with Support Vector Machine.
Material and Methods: The accuracy and loss are performed with dataset from the github library. The total sample size is 20. The two groups Novel Multi Layer Perceptron (N=10) and Support Vector Machine (N=10) were proposed for predicting the accuracy of Breast Cancer prediction.
Result: The results proved that the Novel Multi Layer Perceptron with better accuracy of 90.45% is obtained than the Support Vector Machine of 86.45%. The Support Vector Machine appears significantly better than the Novel Multi Layer Perceptron. The Statistical significance difference between the Support Vector Machine algorithm and Novel Multi Layer Perceptron was found to be 0.194 ($p<0.05$).
Conclusion: The results proved that Support Vector Machine helps predicting Breast Cancer with high accuracy.

Keywords: Breast Cancer, Support Vector Machine, Novel Novel Multi Layer Perceptron, Machine Learning, Accuracy.

[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 is a type of Tumor that occurs in the tissue of the breast(Ang et al. 2022).It is the most common type of cancer formed in Women around the world and it is among the leading cause of death in Women Breast Cancer is the leading mahogany Morality with over 40,000 Women who succumbed to this disease. Breast Cancer became one of the deadliest cancers in Women. It occurs when the growth of cells in Breast tissue becomes out of Control. Cells are building Blocks for the organs and tissues in the body (Du et al. 2022; Bychkov et al. 2022). When the growth of New cells are controlled then they build-up a mass of tissue called a tumor. The Tumors are Categorized into benign and Malignant tumors (Balasubramanian 2021)) 1. Breast cancer health line 2. The cancer therapy advisor 3. Belong Beating Advisor. SVM has been successfully used in many applications such as image recognition, medical diagnosis, and text analytics (*Breast Cancer Prediction System* 2006).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)

Breast Cancer Detection is carried out by researchers 15 related research articles in IEEE digital Xplore and 35 (Chen et al. 2022) articles are published in the research gate Many researchers have put their efforts on breast cancer diagnoses and prognoses, every technique has a different accuracy rate and it varies for different situations (Salem Abdull 2011), tools and datasets being used (Yu 2017). Our main focus is to comparatively analyze different existing MachineLearning and Data Mining techniques in order to find out the most appropriate method that will support the large dataset with good accuracy of prediction (Lee et al. 2022).

The main purpose of this review is to highlight all the previous studies of machine learning algorithms that are being used for breast cancer prediction and this article provides all the necessary information to the beginners who want to analyze the machine learning algorithms to gain the base of deep learning (Lalami et al. 2022). Risk-based breast cancer screening is a cost-effective intervention for controlling breast cancer in China, but the successful implementation of such intervention requires an accurate breast cancer prediction model for Chinese women (Bychkov et al. 2022)**.**

2. Materials and Methods

The research work is carried out in the Machine Learning laboratory lab at Saveetha School of Engineering.The number of groups 2, Group1 Algorithm is Novel MultiLayer Perceptron Group 2 AlgorithmSupport Vector Machine (SVM) and Sample Size 10. Dataset we used the Breast Cancer Data set in 340 instances and 33 attributes. Novel Multi Layer Perceptron (MLP) comparison with Support Vector Machine (SVM)is a class of Support Vector Machine (SVM). It consists of 3 nodes: Input layer , hidden layer and output layer. Except input that each node is a hidden layer and output layer. Except input, each nord is a neuron that uses a nonlinear activation function.(Salem Abdull 2011) MLP is applied using different numbers of layers and neurons to Examine the effect of the number of layers of number classification Accuracy.

Novel Multi Layer Perceptron

Breast cancer (BC) is a standout disease of the most well-known cancers among women around the world. The analysis and prediction of BC leads to early management of the disease and protects the patients from further medical complications. In the light of its noticeable focal points in basic highlights identification from complex BC datasets, Machine Learning (ML) is generally perceived as the technique of decision in BC design order and gauge displaying. Because of the high performance of the Multi-layer Perceptron (MLP) algorithm as one of the ML techniques, we conducted experiments in order to enhance the accuracy rate of MLP by tuning its hyper-parameters along with studying the effect of feature selection methods and feature reduction of MLP. As feature selection results indicated that an increase in the number of input parameters tends to reduce the error associated with the estimator model. The tuned MLP proposed in this paper, based MLP best fit hyper-parameters along with feature selection, is applied for breast cancer classification using Wisconsin Diagnostic Breast Cancer (WDBC) dataset.

Support Vector Machine (SVM)

A Support Vector Machine (SVM) is a binary linear classification whose decision boundary is explicitly constructed to minimize generalization error as shown in figure 4. It is a very powerful and versatile Machine Learning model, capable of performing linear or nonlinear classification, regression and even outlier detection. SVM is well suited for classification of complex but small or medium sized datasets.

SVMs don't output probabilities natively, but probability calibration methods can be used to

convert the output to class probabilities. In the binary case, the probabilities are calibrated using Platt scaling: logistic regression on the SVM's scores, fitted by an additional cross-validation on the training data.

The support vector machine (SVM) is a predictive analysis data-classification algorithm that assigns new data elements to one of labeled categories. SVM is, in most cases, a binary classifier; it assumes that the data in question contains two possible target values.
Another version of the SVM algorithm, multiclass SVM, augments SVM to be used as a classifier on a dataset that contains more than one class (grouping or category). SVM has been successfully used in many applications such as image recognition, medical diagnosis, and text analytics.

Statistical Analysis
For statistical implementation, the software to be used here is IBM SPSS V26. Statistical package for social sciences is used for calculating the statistical calculations such as mean, standard deviation, and also to plot the graphs etc. The independent variables are Label and the dependent variable is 'accuracy'. In SPSS, the dataset is prepared using 10 sample sizes for each group and accuracy is given as the testing variable.

3. Results

Table 1 Shows the simulation results of proposed algorithm Novel Multi Layer Perceptron and the existing system Support Vector Machine (SVM) Encryption were run at different times in the Colab with a sample size of 10.
Table 2 represents the T-test comparison of both Novel MultiLayer Perceptron andSupport Vector Machine (SVM). 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 than theSupport Vector Machine (SVM) with a value of and effect size. The mean accuracy of Novel Multi Layer Perceptron is 90.50% andSupport Vector Machine (SVM) is 86.45%.
Table 3 represents the independent sample T-test with confidence interval at 95% and level of significance as 0.05. It shows the statistical significance P<0.05 -tailed.
Figure 1 gives the comparison chart of Novel MultiLayer Perceptron andSupport Vector Machine (SVM) in terms of mean accuracy. The accuracy in Google Assistant of the Novel MultiLayer Perceptron is better than theSupport Vector Machine (SVM).

4. Discussion

Based on the above study, the Novel Multi Layer Perceptron has better accuracy 86.45% than the Support Vector Machine (SVM) with the accuracy of (Basmadjian et al. 2022; Yao, Tong, and Cheng 2022) 90.50 %. The statistical insignificant difference between the two algorithms is found to be p=0.194 (p<0.05).
In the existing system the accuracy for the Novel Multi Layer Perceptron and Support Vector Machine is 90.40% and 86.45% respectively. This analysis is the paper to make use of machine learning to predict breast cancer (Tschodu et al. 2022; Ang et al. 2022) . Support Vector Machine was implemented with an accuracy percentage of 86.45% . Novel Multi Layer Perceptron combined with Support Vector Machine (Lee et al. 2022) and predicts a time consumption of 78w% Breast Cancer is one of the most commonly diagnosed cancer types in women and automatically classifies breast cancer histopathological image .
The suggested work limits that Support Vector Machines are not a common method for classifying very large data sets. This is because training and testing such data is computationally intensive. Many researchers are trying to reduce Support Vector Machine training time by using sample reduction techniques. Early diagnosis greatly increases the chances of survival. The main challenge in future application detection is to classify tumors as malignant or benign. A tumor is considered malignant if the cells can grow into surrounding tissues or spread to remote areas of the body.

5. Conclusion

This work involves the Novel Multi Layer Perceptron algorithm to predict Breast Cancer with better accuracy of 90.40% when compared to Support Vector Machine accuracy of 86.45%. 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. From this we can conclude that the shopping cart algorithm has the highest accuracy for a particular 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 manuscript.

Acknowledgements
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

Ang, Boon Hong, Weang Kee Ho, Eldarina Wijaya, Pui Yoke Kwan, Pei Sze Ng, Sook Yee Yoon, Siti Norhidayu Hasan, et al. 2022. "Predicting the Likelihood of Carrying a or Mutation in Asian Patients With Breast Cancer." *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology*, February, JCO2101647.

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.

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.

Breast Cancer Prediction System. 2006.

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.

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.

Du, Yu, Hai-Ling Zha, Hui Wang, Xin-Pei Liu, Jia-Zhen Pan, Li-Wen Du, Meng-Jun Cai, Min Zong, and Cui-Ying Li. 2022. "Ultrasound-Based Radiomics Nomogram for Differentiation of Triple-Negative Breast Cancer from Fibroadenoma." *The British Journal of Radiology*, February, 20210598.

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.

Lalami, Imane, Julie Labrosse, Isabelle Cedrin-Durnerin, Marjorie Comtet, Claire Vinolas, Fabien Krief, Christophe Sifer, Maeliss Peigne, and Michael Grynberg. 2022. "Is Letrozole during Ovarian Stimulation Useful in Breast Cancer Patients Undergoing Fertility Preservation to Reduce Early Luteal Progesterone Levels Following GnRH-Agonist Trigger?" *Reproductive Biology and Endocrinology: RB&E* 20 (1): 87.

Lee, Inki, Ilhan Lim, Byung Hyun Byun, Byung Il Kim, Chang Woon Choi, Kyo Chul Lee, Choong Mo Kang, et al. 2022. "The Prediction of HER2-Targeted Treatment Response Using Cu-Tetra-Azacyclododecanetetra-Acetic Acid (DOTA)-Trastuzumab PET/CT in Metastatic Breast Cancer: A Case Report." *Journal of Breast Cancer*, January. https://doi.org/10.4048/jbc.2022.25.e5.

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.

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

Tschodu, Dimitrij, Bernhard Ulm, Klaus Bendrat, Jürgen Lippoldt, Pablo Gottheil, Josef A. Käs, and Axel Niendorf. 2022. "Comparative Analysis of Molecular Signatures Reveals a Hybrid Approach in Breast Cancer: Combining the Nottingham Prognostic Index with Gene Expressions into a Hybrid Signature." *PloS One* 17 (2): e0261035.

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.

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.

Yao, Kevin, Chun-Yip Tong, and Chao Cheng. 2022. "A Framework to Predict the Applicability of Oncotype DX, MammaPrint, and E2F4 Gene Signatures for Improving Breast Cancer Prognostic Prediction." *Scientific Reports* 12 (1): 2211.

Yu, Mengjie. 2017. *Breast Cancer Prediction Using Machine Learning Algorithm*.

Tables and Figures

Table 1. Predicted Accuracy of Breast Cancer Prediction for 10 different sample sizes.

GROUP	Algorithms	Average accuracy
1	Novel Multi Layer Perceptron	90.40
2	Support vector machine (SVM)	86.45

Table 2. Statistical analysis of Novel Multi Layer Perceptron and Randomized Based Algorithm. Mean accuracy, Standard deviation and standard error values are obtained for 10 sample datasets.

	Algorithms	N	Mean	Std.Deviation	Std.Error Mean
Accuracy	Novel Multi Layer Perceptron	10	90.45	.23871	.07549

	Support Vector Machine (SVM)	10	86.45	.15728	.07549

Table 3. Independent sample T-test with 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 Differences**	**Std.Error Differences**	**95% Confidence Interval of the Differences**	
								Lower	**Upper**
Equal Variances assumed	1.822	.194	44.779	18	.000	4.0480	.09040	3.858	4.23792
Equal Variances not assumed	1.822	.194	44.77	18	.000	4.0480	.09040	4.240	3.85594

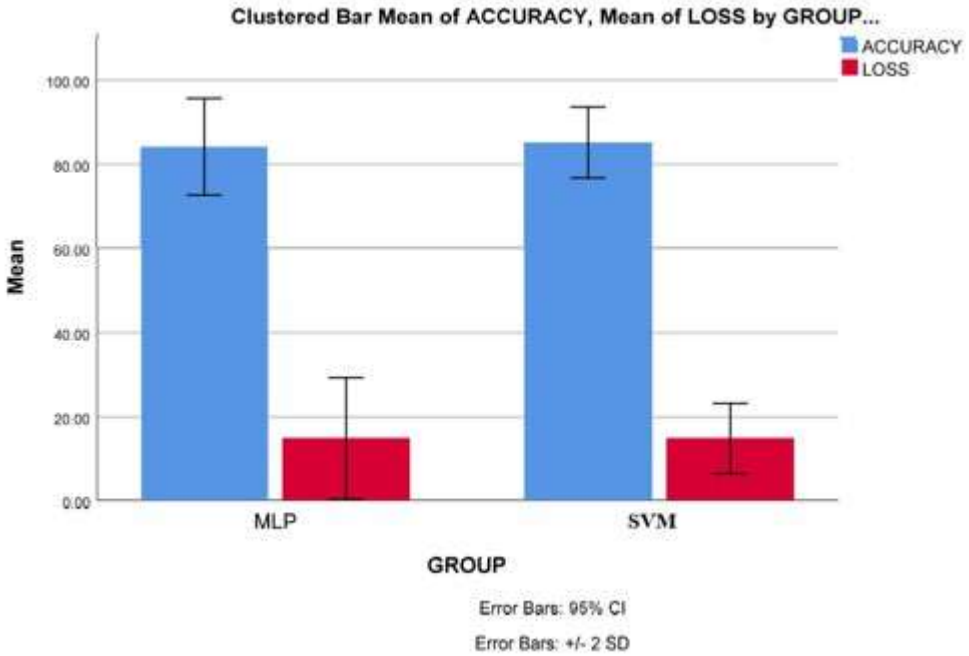

Fig 1. Comparison of Novel Multi Layer Perceptron and Support Vector Machine (SVM) in terms of means and accuracy. The mean accuracy of the Novel Multi Layer Perceptron is better than Support Vector Machine (SVM). X-axis: Novel Multi Layer Perceptron vs Support Vector Machine (SVM), Y-axis: Mean accuracy. Error Bar +/-1 SD with 95% Cl.