International Journal of Scientific Research in Science and Technology
Available online at : www.ijsrst.com
OPEN ACCESS
Print ISSN: 2395-6011 | Online ISSN: 2395-602X
doi : https://doi.org/10.32628/IJSRST

Breast Cancer Screening and Diagnosis using AI : Analyzing the Role of Feature Engineering

Mr. M. Sriivasulu[1], Dr. I. Suneetha[2], Dr. N. Pushpalatha[3]

[1]M.Tech Student, Department of ECE, AITS, Tirupati, Andhra Pradesh, India

[2]Professor, M.Tech, Ph.D., Department of ECE, AITS, Tirupati, Andhra Pradesh, India

[3]Professor and HOD, M.Tech, Ph.D., Department of ECE, AITS, Tirupati, Andhra Pradesh, India

ARTICLEINFO

Article History:

Accepted: 10 Oct 2023
Published: 06 Nov 2023

Publication Issue
Volume 10, Issue 6
November-December-2023
Page Number
29-38

ABSTRACT

Breast cancer, according to the Breast Cancer Institute (BCI), is a highly dangerous disease that affects women worldwide. Medical professionals stress the crucial role of early detection in saving lives. The website cancer.net provides personalized guidance for over 120 cancer types and related hereditary conditions. In the field of breast cancer detection, machine learning techniques are frequently used. In our study, we have introduced an adaptive ensemble voting approach for diagnosing breast cancer using the Wisconsin Breast Cancer database. Our primary aim is to compare and clarify how the logistic algorithm can achieve superior results when combined with ensemble machine learning methods for breast cancer diagnosis, even with a reduced number of variables. Breast tumors fall into two categories: benign tumors, which are non-cancerous, and malignant tumors, which are cancerous.

Keywords : Logistic Regression, SVC, Random Forest, Decision Tree, Cat Boost, K Neighbours, MLP Classifier and GaussianNB.

I. INTRODUCTION

Cancer is widely recognized as one of the most perilous diseases globally, with breast cancer standing out as a particularly formidable threat to women. Tragically, countless women lose their lives to breast cancer each year. The manual detection of breast cancer is not only time-consuming but also poses challenges in terms of accurate classification for physicians. Therefore, it is imperative to explore various automatic diagnostic techniques to enhance cancer detection. Several methods and algorithms have been developed for breast cancer detection, including Support Vector Machine, Naïve Bayes, K-Nearest Neighbors (KNN), and the more recent Convolutional Neural Network (CNN)[11-19], a deep learning algorithm primarily designed for image classification and object detection. In this research,

we utilize the UCI open database for training and testing purposes, which contains two tumor classes: Benign Tumor (non-cancerous) and Malignant Tumor (cancerous). Many researchers are actively engaged in the pursuit of early-stage cancer detection and diagnosis. Detecting cancer in its early stages is less painful and costly to treat, and thus, there is a persistent effort to develop a robust diagnosis system for early tumor detection. This would enable early commencement of treatment and potentially improve the chances of successful resolution. The primary objective of this study is to conduct a comparative analysis of various machine learning algorithms with Artificial Neural Networks (ANN) for breast cancer detection. The subsequent sections outline the structure of this paper. Section 2 offers a literature review of our proposed work. Section 3 elucidates the architecture we have devised. Section 4 provides insights into the methodology employed in our research. Section 5 outlines the feature selection process adopted for our study. In Section 6, we delve into the implementation of the proposed model. Section 7 is dedicated to discussing the results, and Section 8 offers concluding remarks on our research efforts.

II. RELATED WORKS

[1] Utilizing radio frequency-based three-dimensional imaging that capitalizes on differences in the dielectric properties of tissues presents a promising approach for cost-effective, non-invasive, and non-ionizing breast cancer detection. This research paper showcases the application of diverse supervised machine learning techniques to categorize breast tissues into three classes: low-density fatty, high-density fibro-glandular, and malignant. These classifications are made based on analyzing the scattered electric field data collected through strategically positioned antennas encircling the breast tissue. Additionally, the paper includes a comparative analysis of the performance of these machine learning algorithms. Following this classification step, an advanced quantitative non-linear optimization method is employed to enhance the precision of tissue profile reconstruction.

[2] Breast cancer is a leading cause of death globally, and its prevalence necessitates effective classification and data mining techniques. These methods are extensively utilized in the medical field for diagnosis and decision-making. This study aims to assess the comparative performance of various machine learning algorithms, including Support Vector Machine (SVM), Decision Tree (C4.5), Naive Bayes (NB), and k Nearest Neighbors (k-NN). The evaluation is conducted using the Wisconsin Breast Cancer (original) dataset, with the primary objective of measuring accuracy, precision, sensitivity, and specificity in their classification capabilities, all while considering computational efficiency and effectiveness. The experimental results reveal that SVM outperforms the other algorithms, achieving the highest accuracy rate at 97.13% and the lowest error rate. These experiments were carried out using the WEKA data mining tool within a simulated environment.

[3] This paper introduces a novel method that combines Weighted-Particle Swarm Optimization (WPSO) for data clustering with Smooth Support Vector Machine (SSVM) for classification purposes. The study evaluates the effectiveness of WPSO clustering by comparing it to conventional methods such as K-means and fuzzy clustering. This comparison is based on criteria like intracluster distance, intracluster variation, and validity indices. The results demonstrate promising performance for the proposed WPSO-SSVM classification approach, achieving an accuracy of 83.76% for diagnosing liver disorders, 98.42% for analyzing WBCD (Wisconsin Breast Cancer Dataset), and 95.21% for detecting mammographic masses. Notably, these accuracy rates outperform those reported in existing literature on the subject.

[4] Scikit-learn offers a wide array of machine learning algorithms, both for supervised and unsupervised tasks, using a consistent and task-

oriented interface. This makes it convenient to compare different methods when tackling a specific problem. It seamlessly integrates with the scientific Python ecosystem, making it versatile for applications beyond traditional statistical data analysis. A noteworthy feature is that these algorithms, written in a high-level language, can serve as fundamental components for custom solutions tailored to specific use cases, such as in the field of medical imaging as discussed in Michel et al. (2011). In the future, there are plans to enhance scikit-learn capabilities with online learning to handle large datasets effectively.

III. Methodology

Proposed system:

Numerous machine learning algorithms exist for predicting and diagnosing breast cancer. Some of these algorithms include DecisionTreeClassifier, Kneighbours Classifier, RandomForestClassifier, Logistic Regression, MLPClassifier, CatBoostClassifier, and GaussianNB. To determine the most effective method for diagnosing breast cancer, we employed an Ensemble Voting approach. Initially, we applied the RandomForestClassifier algorithm to the dataset, followed by individual implementations of the other algorithms. Subsequently, we employed a Voting Ensemble technique to combine these individual results, resulting in the final accuracy assessment.

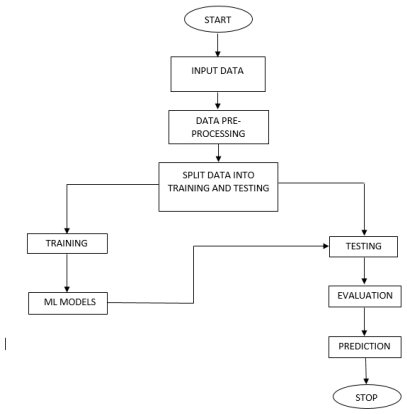


Fig1: Flow of the project

IV. Implementation

Random Forest:

Machine learning techniques for dealing with classification and regression problems include the random forest approach. It employs ensemble learning, a technique for handling challenging issues that combines a number of classifiers. The random forest method offers a wide range of potential decision trees. The random forest approach creates a "forest" that is then the enhancement of machine learning algorithms through a collective meta-algorithm called bagging involves employing techniques like bootstrap aggregation. Bagging contributes to improving the accuracy of these algorithms. Within the realm of machine learning, the random forest method stands out for its ability to accommodate a diverse range of data types and characteristics. Decision trees. By employing the random forest approach with bagging or bootstrap aggregation for training, a "forest" is formed. The bagging collective meta-algorithm improves machine learning algorithms' accuracy. The disadvantages of the choice tree methodology are solved by the random forest approach. Both precision and dataset over fitting are reduced. It creates predictions without the need for numerous package settings, in contrast to Scikit-learn.

The Random Forest Algorithm's Features:

• Comparing it to the decision tree method reveals that it is more accurate.

• It provides a workable strategy for handling missing data.

• It can produce a decent forecast even without hyper-parameter adjusting.

• It resolves decision trees' over fitting issue.

• In each randomly produced tree, a random set of attributes is picked at each node's splitting point.

Decision trees serve as the fundamental components of the random forest algorithm. These decision support structures, resembling trees, are referred to as decision trees. In this discussion, we'll delve into the mechanics of decision trees and explore the functioning of random forest techniques.

The three components of the decision tree are decision nodes, nodes for departs and root nodes. A training dataset is divided into branches via a decision tree approach, and those branches are then further divided. This method is repeated until the area has leaves. No additional separation of the leaf node is possible.

The decision tree's branches represent the traits that are used to predict the result. Links to the branches are provided by decision nodes

Logistic Regression:

Logistic Regression has a rich history in the biological sciences, originating in the early twentieth century. Its applications have extended to various fields, including the social sciences. This statistical technique is particularly useful when dealing with categorical dependent variables.

For instance, Logistic Regression comes into play when we want to make binary predictions, like determining whether an email is spam (1) or not (0) or classifying a tumor as malignant (1) or non-malignant (0). If we were to use linear regression for such tasks, a threshold would be necessary to decide the class. For example, if a tumor is actually malignant, but linear regression predicts a continuous value of 0.4, and we set a threshold of 0.5, it would incorrectly classify the tumor as non-malignant, potentially leading to serious consequences in practical applications. This demonstrates that linear regression, being unbounded, is unsuitable for classification problems. Logistic regression, on the other hand, is designed for such scenarios. Its output strictly ranges from 0 to 1, making it ideal for binary classification tasks. Logistic regression is commonly used in machine learning for precisely these types of problems, where there are two possible outcomes, such as "this or that," "yes or no," or "A or B."

The primary objective of logistic regression is to estimate probabilities of specific events, establishing a connection between input features and the likelihood of particular outcomes. For example, it can predict whether a student will pass or fail an exam based on features like the number of hours spent studying. In this case, the response variable has two values: pass and fail.

SVM:

The primary aim of the support vector machine (SVM) algorithm is to find a hyperplane in an N-dimensional space (where N represents the number of features) that can efficiently separate and classify data points. In this process, several possible hyperplanes are evaluated to distinguish between two classes of data points. The main goal is to identify the hyperplane that provides the largest margin, which is the greatest distance between data points belonging to different classes. By maximizing this margin, we improve the algorithm's capacity to make confident classifications for future data points.

Hyperplanes and Support Vectors:

Hyperplanes serve as crucial decision boundaries for categorizing data points. Depending on the number of features involved, the dimension of the hyperplane varies. For instance, when there are two input features, the hyperplane reduces to a mere line. With three input features, the hyperplane transforms into a two-dimensional plane. However, visualizing hyperplanes becomes progressively challenging as the number of features exceeds three.

Decision Tree:

A tree serves as a powerful metaphor in various aspects of real life and significantly influences a broad spectrum of machine learning techniques, encompassing both classification and regression tasks. In the realm of decision analysis, decision trees emerge as a valuable tool for visually and explicitly depicting decision-making processes. As the name suggests, they adopt a tree-like structure to model decisions. Decision trees find frequent application in data mining to formulate strategies aimed at achieving specific objectives. When represented visually, a decision tree is typically depicted upside down, with the root node positioned at the top. In the graphical representation, bold black text signifies a condition or internal node, serving as the point at which the tree

branches into various edges or paths. As we progress along these branches, we ultimately reach a decision or leaf node. In a classic example like the one illustrated in the left image, these leaf nodes are color-coded, with red indicating death and green signifying survival of passengers. It's important to note that real-world datasets typically incorporate numerous features, rendering the illustrated tree a mere branch within a larger and more complex decision tree. Nevertheless, the elegance of this algorithm lies in its simplicity. It offers clear insight into feature importance and allows for straightforward interpretation of relationships. This approach is commonly referred to as "learning decision trees from data," with the tree depicted above being specifically termed a "Classification tree" due to its purpose of classifying passengers as survivors or casualties. In the case of Regression trees, the representation remains the same, but the objective shifts to predicting continuous values, such as the price of a house. In general, Decision Tree algorithms are denoted as CART, signifying "Classification and Regression Trees."

Cat Boost:

CatBoost, an open-source library for gradient boosting on decision trees, has quickly gained recognition in the machine learning community. Developed by Yandex researchers and engineers, it finds applications in diverse fields like search, recommendation systems, personal assistants, autonomous vehicles, weather prediction, and more, not only at Yandex but also in organizations like CERN, Cloudflare, and Careem taxi. CatBoost's emergence has posed a serious challenge to established players like XGBoost and LightGBM, especially in scenarios where categorical features are prevalent. What sets CatBoost apart is its exceptional performance on benchmarks, particularly when dealing with datasets heavily reliant on categorical features. While its training time may be longer compared to other gradient boosting implementations, its prediction speed is an impressive 13–16 times faster, as demonstrated in Yandex's benchmark tests. Furthermore, CatBoost's default settings serve as an excellent foundation for individuals new to the realm of tree ensembles or those engaging in Kaggle competitions. CatBoost boasts various impressive features, such as improved handling of feature interactions, detailed analysis of feature importance, and built-in support for snapshots. In addition to its effectiveness in classification and regression tasks, CatBoost effortlessly accommodates ranking tasks without any extra configuration. These attributes establish CatBoost as an appealing option for diverse machine learning applications.

K-Nearest Neighbor:

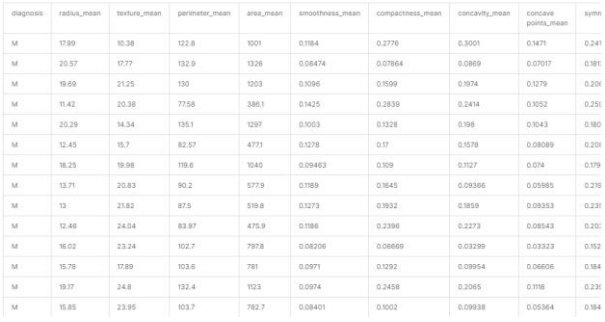
K-Nearest Neighbors (K-NN) stands as one of the simplest machine learning techniques within the realm of supervised learning. This algorithm hinges on the concept of gauging the likeness between new data and existing data points, ultimately assigning the new data to the category that bears the closest resemblance. The essence of K-NN lies in its retention of all the available data, a repository it draws upon when classifying a novel data point. In other words, when new data emerges, the algorithm swiftly assigns it to the most fitting category by seeking similarities with past data. K-NN is a versatile tool, applicable to both regression and classification tasks, although its primary use tends to lean toward classification. One noteworthy characteristic is that K-NN operates as a non-parametric algorithm, devoid of any preconceived assumptions about the underlying data distribution. Often dubbed a "lazy learner" algorithm, K-NN refrains from immediate learning during the training phase. Instead, it preserves the training dataset and takes action only when faced with a new data point. This action involves classifying the new data into the category that exhibits the greatest resemblance to it. Consider an example: Suppose we possess an image of a creature, one that bears resemblance to both cats and dogs. Our objective is to determine whether it's a cat or a dog. In this scenario, the K-NN algorithm proves invaluable. It operates

based on a similarity metric, effectively identifying similar features between the new data and known images of cats and dogs. Ultimately, it assigns the creature to either the cat or dog category based on the most closely aligned features.

MULTILAYER PERCEPTRON:

In the realm of deep learning, TensorFlow, Keras, Microsoft Cognitive Toolkit (CNTK), and PyTorch are widely recognized as popular libraries. However, it's worth noting that the well-known machine learning library, Scikit-learn, is also capable of basic deep learning modeling. Specifically, Scikit-learn offers a Multilayer Perceptron (MLP) implementation. This MLP has some distinct characteristics. Firstly, it lacks an activation function in the output layer. For regression tasks, it employs the square error as the loss function, while for classification, it uses cross-entropy. Notably, it can handle both single and multiple target values for regression. Unlike other prominent deep learning packages like Keras, Scikit-learn's MLP implementation does not support GPU acceleration. Additionally, it has limitations when it comes to fine-tuning parameters such as activation functions and weight initializers for individual layers. A Multilayer Perceptron (MLP) is a type of feedforward artificial neural network (ANN). The term "MLP" can be somewhat ambiguous, sometimes referring broadly to any feedforward ANN and at other times specifically to networks comprising multiple layers of perceptrons with threshold activation. An MLP typically consists of a minimum of three layers: an input layer, a hidden layer, and an output layer. Each node, except for those in the input layer, represents a neuron utilizing a nonlinear activation function. MLP leverages a supervised learning technique known as backpropagation for training, distinguishing itself from linear perceptrons through its multiple layers and non-linear activation functions. This characteristic enables MLP to discern data that lacks linear separability.

Dataset:



diagnosis	radius_mean	texture_mean	perimeter_mean	area_mean	smoothness_mean	compactness_mean	concavity_mean	concave points_mean	symm
M	17.99	10.38	122.8	1001	0.1184	0.2776	0.3001	0.1471	0.241
M	20.57	17.77	132.9	1326	0.08474	0.07864	0.0869	0.07017	0.181
M	19.69	21.25	130	1203	0.1096	0.1599	0.1974	0.1279	0.206
M	11.42	20.38	77.58	386.1	0.1425	0.2839	0.2414	0.1052	0.259
M	20.29	14.34	135.1	1297	0.1003	0.1328	0.198	0.1043	0.180
M	12.45	15.7	82.57	477.1	0.1278	0.17	0.1578	0.08089	0.208
M	18.25	19.98	119.6	1040	0.09463	0.109	0.1127	0.074	0.179
M	13.71	20.83	90.2	577.9	0.1189	0.1645	0.09366	0.05985	0.219
M	13	21.82	87.5	519.8	0.1273	0.1932	0.1859	0.09353	0.235
M	12.46	24.04	83.97	475.9	0.1186	0.2396	0.2273	0.08543	0.203
M	16.02	23.24	102.7	797.8	0.08206	0.06669	0.03299	0.03323	0.152
M	15.78	17.89	103.6	781	0.0971	0.1292	0.09954	0.06606	0.184
M	19.17	24.8	132.4	1123	0.0974	0.2458	0.2065	0.1118	0.236
M	15.85	23.95	103.7	782.7	0.08401	0.1002	0.09938	0.05364	0.184

Dataset description:

- **Radius (mean, se, worst):** The average size of the tumor's radius or distance from the center to the outer edge.
- **Texture (mean, se, worst):** A measure of variation in color or intensity of the pixel values in the image.
- **Perimeter (mean, se, worst):** The total length of the tumor's boundary.
- **Area (mean, se, worst):** The total area of the tumor.
- **Smoothness (mean, se):** A measure of the variation in radius lengths within the tumor.
- **Compactness (mean, se):** A measure of how compact the shape of the tumor is, calculated as perimeter^2 / area - 1.0.
- **Concavity (mean, se):** A measure of the severity of concave portions of the contour.
- **Concave Points (mean, se):** The number of concave portions of the contour.
- **Symmetry (mean, se, worst):** A measure of symmetry in the tumor's shape.
- **Fractal Dimension (mean, se, worst):** A measure of the complexity of the tumor's shape.

4. Results and Discussion:

The following images will visually depict the process of our project.

Home page: The user can view the web application's home page here.

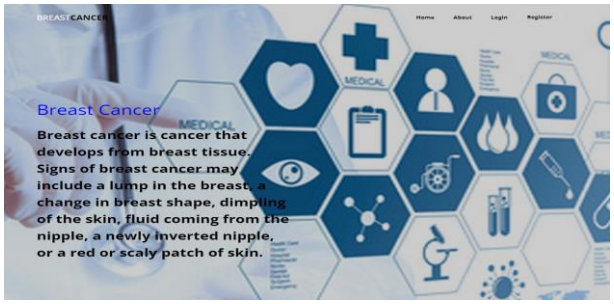


Fig2: Home Page

About:

Fig3: About Page

Registration:

The user can register attending this page by providing all the necessary details.

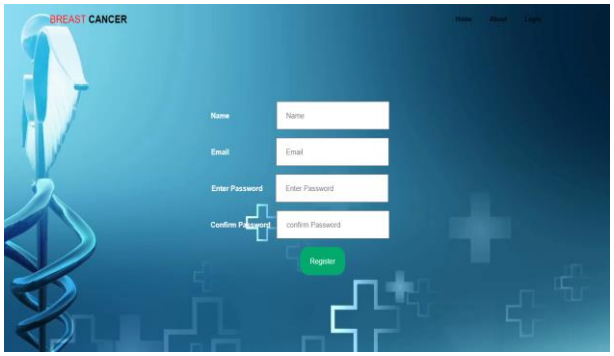


Fig4: Registration Page

Login:

The user can sign in using their authorized credentials here.

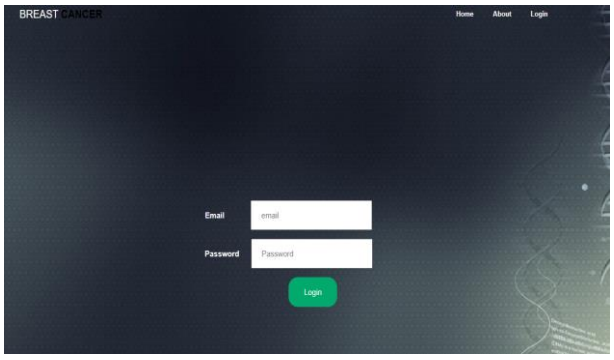


Fig5: Login Page

User Home Page:

The user's home page will be shown after a successful login.

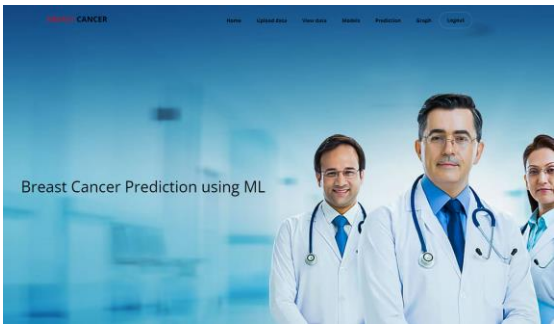


Fig6: User home Page

Upload data:

Upload relevant data for model training. Preprocess and clean data. Split into training and testing sets. Train model using algorithms.

Fig7: Upload data Page

View Data:

View Data enables analysis, visualization, and understanding of information, turning raw numbers into insights for informed decision-making and strategy planning.

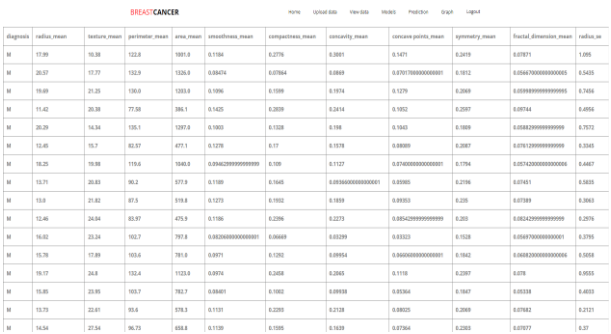


Fig8: View Data Page.

Model:

Model training involves feeding data to a machine learning algorithm, allowing it to learn patterns and make predictions or decisions. The process is iterative,

often optimizing a loss function to improve the model's accuracy.

Fig9: Model Page.

Prediction:

Whatever information we provide the model after successful training will help it predict hate speech.

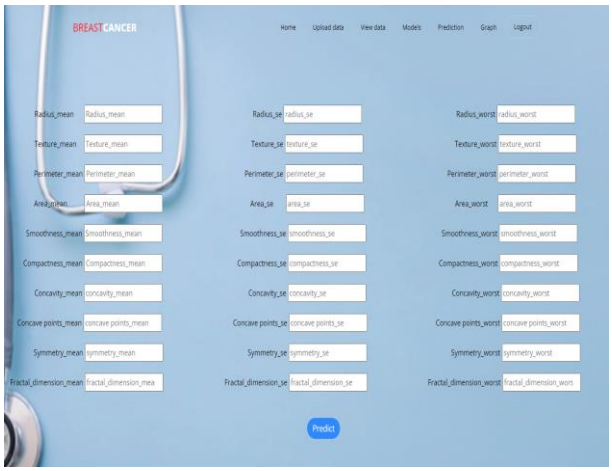


Fig10: Prediction Page

Graph:

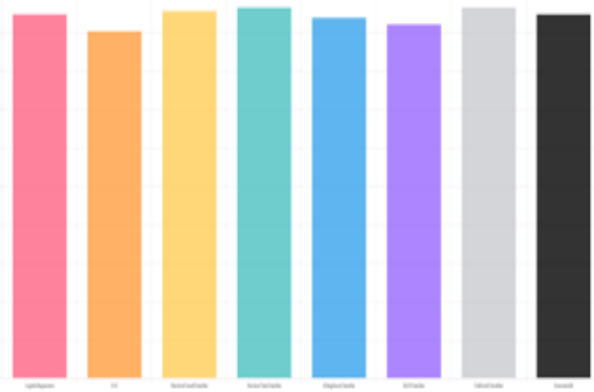


Fig11: Prediction Page

V. Conclusion

This study introduces an ensemble machine learning approach for breast cancer diagnosis, as depicted in the provided table and graph. The paper demonstrates the effectiveness of this proposed method, which utilizes only 16 features for cancer diagnosis. The future research direction aims to extend the analysis to encompass all the features available in the UCI dataset in pursuit of even higher accuracy. The research findings highlight the potential of neural networks in analyzing human vital data, enabling pre-diagnosis without requiring specialized medical knowledge. The methodology involves an Ensemble Machine Learning algorithm incorporating both Logistic Regression and Neural Networks for breast cancer detection and diagnosis. The dataset is pre-processed using standardization, followed by the application of the Univariate Feature Selection algorithm, which employs the chi-squared method to identify the top 16 features from the UCI dataset. Logistic and neural network algorithms are then applied to these 16 selected features, and a voting algorithm is used to combine their results, resulting in a notably improved accuracy compared to individual machine learning algorithms. This approach proves particularly valuable, given that a large number of features can increase the implementation cost of the model.

VI. REFERENCES

[1]. 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.

[2]. C. Deng and M. Perkowski, "A Novel Weighted Hierarchical Adaptive Voting Ensemble Machine Learning Method for Breast Cancer Detection," 2015 IEEE International Symposium on Multiple-Valued Logic, Waterloo, ON, 2015, pp. 115-120.

[3]. A. Qasem et al., "Breast cancer mass localization based on machine learning," 2014 IEEE 10th International Colloquium on Signal Processing and its Applications, Kuala Lumpur, 2014, pp. 31-36.

[4]. 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.

[5]. J. A. Bhat, V. George and B. Malik, "Cloud Computing with Machine Learning Could Help Us in the Early Diagnosis of Breast Cancer," 2015 Second International Conference on Advances in Computing and Communication Engineering, Dehradun, 2015, pp. 644-648.

[6]. 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 International Conference on Computational Intelligence and Computing Research (ICCIC), Chennai, 2016, pp. 1-5.

[7]. H. R. Mhaske and D. A. Phalke, "Melanoma skin cancer detection and classification based on supervised and unsupervised learning," 2013 International conference on Circuits, Controls and Communications (CCUBE), Bengaluru, 2013, pp. 1-5.

[8]. S. Aruna, S. P. Rajagopalan and L. V. Nandakishore, "An algorithm proposed for Semi-Supervised learning in cancer detection," International Conference on Sustainable Energy and Intelligent Systems (SEISCON 2011), Chennai, 2011, pp. 860-864.

[9]. Y. Tsehay et al., "Biopsy-guided learning with deep convolutional neural networks for Prostate Cancer detection on multiparametric MRI," 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017), Melbourne, VIC, 2017, pp. 642-645.

[10]. S. Nayak and D. Gope, "Comparison of supervised learning algorithms for RF-based breast cancer detection," 2017 Computing and Electromagnetics International Workshop (CEM), Barcelona, 2017, pp. 13-14. doi: 10.1109/CEM.2017.7991863

[11]. P. Harish, S. Baskar, " A survey on detection, recognition, segmentation and classification of brain tumor," Int. J. Eng. Adv. Technol. 8 (2) (2019) 249–254.

[12]. P. Harish, S. Baskar, " Safe engineering applications for diagnosis of brain abnormalities using adaptive local binary pattern based texture analysis," J.Green Eng. 10 (9) (2020) 4727–4743

[13]. Harish Pasupulati and Baskar S (2020) "Safe engineering methods for detecting brain tumor using heat distribution merged saliency detection," Journal of Green Engineering, 10(10), pp. 7946–7961. http://www.jgenng.com/volume10-issue 10.php

[14]. Chandrakala Kuruba, N. Pushpalatha, Gandikota Ramu, I. Suneetha, M. Rudra Kumar, P. Harish "Data mining and deep learning-based hybrid health care application," Applied Nanoscience (Switzerland), 2023, 13(3), pp. 2431-2437.

[15]. Ramani K, Suneetha I, Pushpalatha N, Harish P, Yugundhar P, "Chronic Disease Prediction through Supervised Learning Techniques," Journal of Pharmaceutical Negative Results, 2022, 13(4), pp. 990-998.

[16]. Ramani K, Suneetha I, Pushpalatha N, Harish P, Yugundhar P, "Efficient Multi-Object Identification and Guiding System for Visually Impaired," Journal of Pharmaceutical Negative Results, 2022, 13(4), pp. 1019-1023.

[17]. Silpa C, Suneetha I, Reddy Hemantha G, Arava R.P.R, Bhumika Y, "Medication Alarm: A Proficient IoT -Enabled Medication Alarm for Age Old People to the Betterment of their Medication Practice," Journal of Pharmaceutical Negative Results, 2022, 13(4), pp. 1041-1046.

[18]. Suneetha I, Venkateswarlu T, "Noise suppression using non-linear PAR model for binary images," 2013 15th International Conference on Advanced Computing

Technologies (ICACT), Rajampet, India 2013 https://ieeexplore.ieee.org/document/6710525.

[19]. Suneetha I, Venkateswarlu T, "Image enhancement through contrast improvement using linear parameterized gradient intercept model," ARPN Journal of Engineering and Applied Sciences, 2012, 7(8), pp. 945-949.

Cite this article as :

Mr. M. Sriivasulu, Dr. I. Suneetha, Dr. N. Pushpalatha, "Breast Cancer Screening and Diagnosis using AI : Analyzing the Role of Feature Engineering", International Journal of Scientific Research in Science and Technology (IJSRST), Online ISSN : 2395-602X, Print ISSN : 2395-6011, Volume 10 Issue 6, pp. 29-38, November-December 2023.
Journal URL : https://ijsrst.com/IJSRST523105107