Tuijin Jishu/Journal of Propulsion Technology
ISSN: 1001-4055
Vol. 47 No. 02 (2026)

An Intelligent Real-Time Breast Cancer Classification in a Cloud Environment Using Optimized ELM

Swarang Raut[1*], Dr. Sunny Sall[2]

*1*M.Tech Student, Department of Computer Engineering, St. John College of Engineering and Management.*
2Head of Department, Department of Computer Engineering., St. John College of Engineering and Management.
.

Abstract:- Breast cancer remains a major global health challenge, particularly for women in remote regions where the medical infrastructure is very poor. AI-powered diagnostic systems supported by cloud computing technology offer a practical way to deliver early detection services to these under-served areas. This work proposes an AI-driven framework that combines Gain Ratio–based feature selection, PCA for reducing overfitting, an optimized Extreme Learning Machine (ELM) for breast cancer classification, and a cloud-enabled platform for remote diagnosis. Using the Wisconsin Diagnostic Breast Cancer (WDBC) dataset, the cloud-based ELM is compared with leading diagnostic methods. Results show that the proposed system outperforms existing approaches, achieving 98.25 % accuracy, 0.9891 recall, 0.9861 precision, an F1-score of 0.9861, and an AUC of 0.9931.

Keywords: *Breast cancer; ELM; Cloud computing; PCA, Real-Time Diagnosis Telehealth etc..*

1. Introduction

Worldwide, malignancies in breast cancer is one of the major causes of mortality in women [1–4]. About 2.3 million new breast cancer cases and roughly 670,000 related deaths occurred worldwide in 2022 (WHO 2025). [5], According to their projections, If current trends continue, global breast cancer incidence and mortality to increase significantly and there could 3.2 million approximately new diagnoses and 1.1 million approximately deaths per year by 2050. Delayed diagnosis continues to be a primary driver of mortality, underscoring the need for advanced, rapid, and accessible diagnostic tools to enhance clinical decision-making.

There are many conventional standard diagnostic practices which includes clinical breast exams, mammography, MRIs and biopsy; however, effectiveness of all these diagnosis systems have limited due to several factors[6, 7]. Some of the major factors affecting the prognosis in BC are availability of experienced medical practitioners, variability in diagnostic interpretation, high costs, and slow processing times especially in rural or under-resourced regions due to lack of modern medical infrastructure [8– 10]. Thus, there is a growing need to enhance the existing diagnostic mechanisms with automated intelligent systems which can offer accurate and scalable real-time solutions. With this intent many machine learning models have been inculcated in the healthcare diagnostic tools to improve the results but many of these approaches heavily rely on local machine processing resulting with limitations in terms of computational power, scalability, and real-time responsiveness [11,57]. Furthermore, the conventional diagnostic models often suffer from overfitting, slow training, or lack of generalization when applied on complex medical datasets. Manual interpretation remains susceptible to human error, and the inefficiency in handling large volumes of data from various healthcare centers further hampers the timely diagnosis [11, 12].

Cloud based paradigms have recently become a scalable and cost-efficient alter- native to locally deployed computing infrastructures, enabling remote data storage, processing, and management through distributed servers [1, 13]. Its browser-based user access mechanisms that allow healthcare professionals to integrate, retrieve, and update clinical records with improved efficiency. Moreover, cloud platforms support large and heterogeneous biomedical datasets—including medical images, EHRs, and also voice-based data streams [14]—while ensuring high service avail- ability and reduced system downtime [15]. In clinical applications, cloud computing facilitates remote diagnostic services for patients with limited mobility or financial constraints. It also enables telehealth

[16] and telemedicine workflows [17], support- ing the transmission of high-resolution medical images, videos, and other patient data from resource-limited regions to fully equipped medical centers. Furthermore, cloud- based systems are increasingly utilized for rapid retrieval of blood and organ donor information during emergency scenarios [18].

To mitigate the challenges associated with the local medical infrastructural services and the significant benefits of the cloud computing; this work proposes a cloud- integrated Optimized Extreme Learning Machine (O-ELM) model. The major objectives of this work include the improved speed and accuracy of diagnosis of medical data classification by enabling the real-time access, processing, and diagnosis through cloud infrastructure, thereby over-coming limitations of localized computing and enhancing accessibility. This optimized ELM is designed by integrating machine learning with cloud computing to handle large datasets, offering faster diagnosis, and be deployed remotely, making it a valuable tool for telemedicine applications. The major contributions after thorough analysis in this paper includes:

- Real-Time Diagnosis via Cloud Integration: The system is deployed on a cloud platform, allowing healthcare providers and diagnostic centers to upload data and receive predictions instantly, without the need for high-end local machines
- Optimization of ELM: With desired feature selection tuning the hyperparameter the basic ELM is enhanced to give the improved accuracy, generalization, and speed. The optimized model reduces training time significantly while maintaining high accuracy on benchmark datasets.
- Performance Improvements: Experimental results of the proposed framework outperform conventional classifiers in terms of precision, recall, F1-score, and efficiency, while exhibiting superior scalability and robustness across heterogeneous medical datasets.

The remainder of this paper is structured as follows: Section 2 provides a review of existing studies on machine learning, deep learning, and cloud-enabled approaches for breast cancer diagnosis. Section 3 outlines the proposed system architecture, covering data pre-processing, optimization techniques, and the designed cloud framework. Section 4 reports the experimental results along with comparative analyses. Section 5 offers a discussion of the results and highlights the study's limitations. Finally, Section 6 concludes the paper and suggests directions for future research.

2. Objectives

The primary objective of this study is to develop an intelligent and efficient breast cancer diagnosis system by integrating machine learning techniques with cloud computing. The study aims to design an optimized Extreme Learning Machine (O-ELM) model to improve classification accuracy, generalization, and computational speed. Another objective is to enable real-time diagnosis through a cloud-based platform, allowing healthcare professionals to access, process, and analyze medical data remotely without relying on high-end local computational resources. The proposed system also focuses on overcoming limitations of conventional diagnostic approaches such as delayed processing, high cost, and dependency on expert interpretation. Furthermore, the study aims to enhance scalability and efficient handling of large and heterogeneous medical datasets using cloud infrastructure. Finally, the performance of the proposed model is evaluated using standard metrics such as accuracy, precision, recall, and F1-score to demonstrate its effectiveness compared to existing method

3. Methods

3.1 Background of Work

To design a robust breast cancer diagnosis framework, in this paper, an extensive survey and analysis is done with an intent to understand previous contributions across three major components: 3.2. Models based of ML and DL BC diagnosis, 3.3. cloud- based healthcare systems, and 3.4. the use and optimization of Extreme Learning Machines (ELM). This section provides an organized literature review based on these areas and highlights gaps that justify the proposed work.

3.2 Intelligent Breast Cancer Diagnosis Methods

In medical and bio-informatics domain, providing the accurate and efficient diagnostic solution has become a challenging task [19, 20]. In conventional diagnostic process are able to reduce the diagnosing errors, biases and also streamline the procedures to improve the efficiency with support of significant involvement of expert medical expertise. However, this expertise depends on the traditional imaging technologies such as mammography, X-rays, MRI and CT etc. The analysis of large and heterogeneous medical image datasets often limits diagnostic accuracy, increasing the risk of misdiagnosis. Machine Learning and Deep Learning have emerged as robust tools [6]to enhance healthcare decision-making by enabling accurate disease prediction, real-time diagnosis, and cost-effective interventions [19, 21]. ML applications have demonstrated high performance in predicting breast cancer [22], heart disease [23], liver disease [24], lung cancer [25], and prostate cancer [26]. Techniques such as SVM, LR, KNN, ANN, RF, and ensemble models including XGBoost and Auto-ML provide effective solutions for addressing challenges in breast cancer diagnosis.

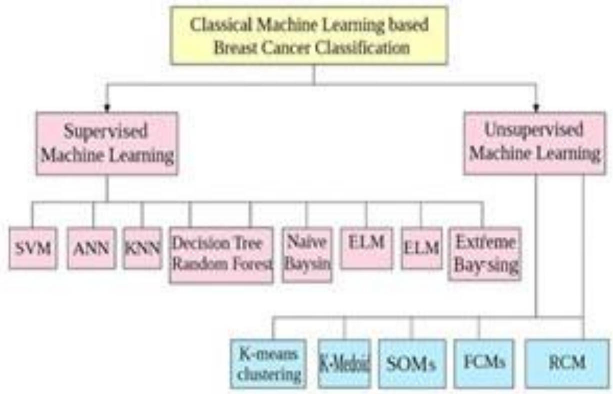


Figure 1: Classical ML based classification of BrC [27]

Deep learning–based CAD is widely used in healthcare—including radiology [4, 28], pathology [29], cardiology [30], pharmacology [31], oncology [32, 33], and genomics [34]—and shows strong breast cancer diagnostic performance, though many studies still rely on traditional ANNs or ML with manual features rather than newer deep architectures like GANs and ELMs [35].

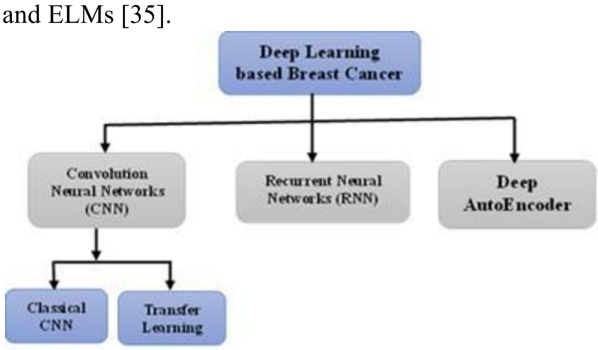


Figure 2: DL based breast cancer classification [27]

3.3 Cloud-Based Healthcare Systems

Model based on ML and DL have shown the significant contribution in early accu- rate detection as well as diagnosis of breast cancer. However, these models are very effective where the advanced medical infrastructure is available. But with increas- ing complexities in medical data processing, the conventional systems required to include more advanced models in real-world healthcare systems for robust automation, reproducibility, and continuous monitoring [36, 37].

The integration of cloud-computing with ML & DL that can support to store large volume of data and to automate the diagnostic system by providing the online services in remote areas where infrastructure is very poor [36]. This section highlights the major research contributions for large-scale healthcare systems breast cancer detection techniques based on machine-learning, cloud-based deployment models with CI/CD practices.

In [37] proposed a intelligent model based on Extreme Gradient Boosting (XGBoost) with an accuracy of 99.12 % for breast-cancer prediction. The study very strongly emphasized the critical role of feature-selection techniques, such as the Bon- ferroni correction, in optimizing model performance. The authors also highlighted the benefits of data pre-processing, including addressing class imbalance so as to improve diagnostic accuracy. [36]. But this study could not inculcate the automated deployment pipeline leading to the limitation in real-life implementation in most circumstances. A hybrid model MoEffNet is described model for breast cancer diagnosis in [38] is a combination of EfficientNet and Mixture of Experts (MoEs). This model according to the data available found to be performed much better than normal CNNs using efficient multi-scale feature extraction using EfficientNets and equaling dynamically modulated gates to yield over 0.99 AUC on multiple mammogram datasets. A deep learning bIn this study it is also found that the deep learning models have been used as support for multiple levels of image abstraction at the cost of heavier computa- tional expense [36]. As such even though this model has outperformed in real-world applications would with heavy cost. This study could have been extended the use of cloud-based edge processing to reduce the diagnosis cost which was not mentioned in this work. In [39] the authors have proposed LightGBM to classify the histopathological breast-cancer images. They enhanced model performance through color histogram analysis, contour and texture feature extraction, and data augmentation, enabling with an 99.8 % accuracy—surpassing other tree-based models such as XGBoost and CatBoost. However, it is very effective for structured image data, but lacked CI/CD automation for continuous model updates. The data presented in [40] is a comprehensive study on MLOps-driven CI/CD pipelines to deploy the machine-learning models. They highlighted that many ML algorithms fail due to insufficient automation mechanism and have suggested improved MLOps model for versioning, monitoring, and scalability. Apart from this, the authors have outlined cloud-deployment best practices, including CI/CD using GitHub Actions and Kubernetes-based delivery for automating a breast-cancer detection system. In [36] the authors have examined the ETL automation processes within ML pipelines using CI/CD workflows. This work demonstrated the effective use of automation tools like Apache Airflow, Jenkins, and Kubernetes to streamline data extraction, transformation, and loading. End-to-End CI/CD pipeline implementation model have reduced the retraining time and manual intervention, thereby improving scalability and operational efficiency. Their results also align with practices using AWS Glue for ETL and AWS SageMaker for automated model deployment, reinforcing the value of cloud-based MLOps.

Despite recent advances, cloud-based diagnostic systems remain constrained by static processing, limited standardization, insufficient multi-site validation, and minimal edge integration, impeding scalability in low-bandwidth environments. Additionally, secure, privacy-preserving frameworks are critical to safeguard medical data. Deep-learning–based cloud solutions further face challenges of latency, high operational costs, and inadequate real-time pipelines, limiting their utility for time sensitive clinical decisions.

3.4 ELM and Its Variants

Extreme Learning Machine (ELM) also called as single-hidden-layer feedforward neu- ral network (SHLFN), is specifically tailored for rapid training process and strong per- formance compared to traditional intelligent learning approaches. Its computational efficiency in the research aimed at lowering manual involvement, reducing training time, and improving classification accuracy [41]. Moreover, theoretical investigations have been conducted to assess the generalized behavior and global approximation capabilities of ELMs [42].

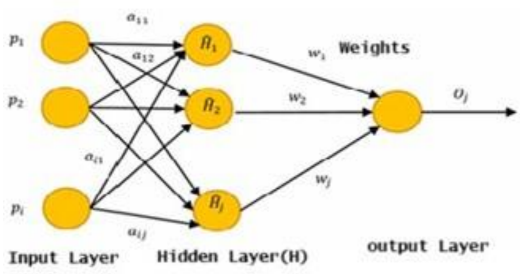


Figure 3: Structure of SLFN

The architecture comprises three main layers: the **Input** P, the **Hidden** H, and the **Output** O. The weight vector **w**, represents the connection from input layer to the hidden layer, **b** denotes the bias terms of the hidden layer and $\boldsymbol{\beta}$ represents the output layer. Training the network involves adjusting these parameters to obtain an optimal solution with high accuracy. The standard ELM framework has been shown to possess universal approximation properties [41, 58] and is typically trained in two stages—

1. **Random initialization of the parameters:** - The parameters w_i and b_i, the parameter of hidden layer and bias of the network respectively are initialized with random values and kept constant throughout the process. Input vector is then mapped to the random feature space using nonlinear, piecewise-continuous activation functions. These randomly assigned parameters have been found to be more efficient than traditionally trained parameters.
2. $\boldsymbol{\beta}_i$ is then obtained by the Moore–Penrose inverse as it is a linear problem given as: $\mathbf{H}\boldsymbol{\beta} = \mathbf{T}$.

Algorithm 1: Basic ELM Model
Require: Training set DS = $\{(\mathbf{x}_i, \mathbf{y}_i)\}^N$
1: Initialize randomly weights $\mathbf{w}_i$ of input layer and biases $\mathbf{b}_i$ of Hidden Layer
2: Compute **H**,output matrix of the hidden-layer
3: Get the output weights with using the Moore–Penrose pseudoinverse:

$$\mathbf{H}\boldsymbol{\beta} = \mathbf{T}$$
$$\boldsymbol{\beta} = \mathbf{H}^{\dagger}\mathbf{T}$$

From the literature and analysis, it is observed that ELM possesses better gen- eralization abilities without iteratively tuned hidden parameters. Also, It is capable of modeling highly complex classification boundaries when provided with a sufficient number of hidden neurons [41]. Building on this capability, several enhanced variants of the ELM framework have been developed to improve its effectiveness in practical scenarios. These include: Random Hidden Nodes enabled ELM [41]; Parame- ter optimization approaches [41, 43, 44], such as the method proposed by Chen and Lv [43]; Structure modification strategies [41, 45, 46]; ELM for online learning [47]; ELM designed for imbalanced data [48]; ELM for big data applications [49, 50]; and ELM for semi-supervised and unsupervised learning [51]. Manifold regularization–based ELM retains the core properties of the traditional ELM while enabling its use on large-scale datasets.

3.5 Identified gaps

From this review, the following research gaps have been identified:

- Limited real-time breast cancer diagnostic systems that integrate machine learning with cloud computing.
- Extreme Learning Machine (ELM), although fast and effective, is rarely used in cloud-based real-time healthcare applications.
- Existing systems often lack end-to-end optimization, including feature selection, dimensionality reduction, and model deployment pipelines.
- Existing systems often lead to overfitting problems due the high volume of data generation and feature extraction.

The proposed system addresses these gaps by implementing a cloud-integrated, real- time diagnostic framework using an optimized ELM model with gain ratio and PCA

3.6 Materials and Methods

The proposed optimized cloud-based breast cancer diagnosis framework which is help- ful for remote monitoring of patient data used for early detection. Its main objective is to classify cases as cancerous or non-cancerous using cloud-stored medical records. The overall architecture is shown in Figure 3.6.1. In this setup, patients visit a local healthcare center where diagnostic information—such as X-rays and clinical parameters are collected and sent to a physician, who then uploads the data to the cloud for automated analysis and diagnosis.

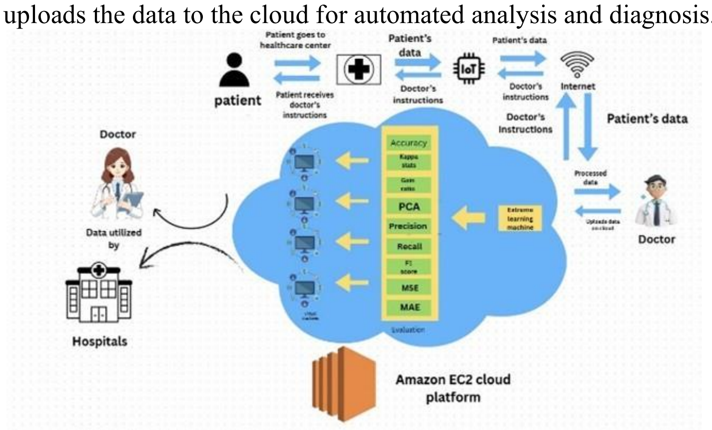


Figure 4: Components of the proposed architecture

In machine learning, model performance relies heavily on selecting attributes that enhance overall accuracy [1, 52–54]. The optimized model operates in three stages: first, the gain ratio method identifies relevant features and discards insignificant ones; second, PCA is applied to address overfitting and reduce computational complexity; finally, ELM is used to classify the selected features.

3.6.1 Gain Ration Method for Feature Selection

For the feature selection, Gain Ratio a filter-based approach is used [1, 55] where the ranking of the attributes is done according to their relevance. This method helps in reducing bias there by normalizing the score using the attribute's split information Unlike Information Gain method.

First calculate the information gain of the feature F, which help in finding the possible decrease in entropy of the dataset DS when it is partitioned based on the values of, when F

$$Gain(DS, F_g) = Entropy(DS) - \sum_{v \in Values(Fg)} \frac{|DS_v|}{|DS|} \cdot Entropy(DS_v) \qquad (3.1)$$

Here, DS_v denotes the subset of samples where attribute $F_g = v$ and $Entropy(DS)$ quantifies the impurity in dataset DS, calculated as:

$$Entropy(DS) = -\sum_{i=1}^{C} p_c \log_2(p_c) \qquad (3.2)$$

Where p_c denotes the probability of class C in the dataset DS, i.e., the proportion of samples belonging to class C. Now calculate the Now calculate Split information as

$$Split(D, F) = -\sum_{v \in Values(F)} \frac{|D_v|}{|D|} \log_2\left(\frac{|D_v|}{|D|}\right) \qquad (3.3)$$

Thus, the Gain Ratio is given as the ratio of Information Gain to Split Information and represented as:

$$GainRatio(F) = \frac{Gain(D,F)}{SplitInfo(F)} \qquad (3.4)$$

Where

- $Gain(DS, F_g) = Entropy(DS) - \sum_{v \in Values(F_g)} \frac{|DS_v|}{|DS|} \cdot Entropy(DS_v)$
- $SplitInfo(F_g) = -\sum_{v \in Values(F_g)} \frac{|DS_v|}{|DS|} \cdot \log_2\left(\frac{|DS_v|}{|S_i|}\right)$

3.6.2 PCA for Overfitting

Overfitting is a well-known challenge in medical diagnostic models, where classifiers may capture patterns specific to the training data, thereby limiting their generalization capability. In this proposed cloud-integrated diagnosis system, possibility of overfitting may arise due to high-dimensional noisy and highly correlated input features. Principal Component Analysis (PCA), a dimensionality reduction algorithm is employed to alleviate overfitting challenges by transforming the original feature space into a reduced set of orthogonal principal components. This set of data is ordered according to the proportion of variance preserved from the original data.
Assume your dataset after Gain Ratio feature selection is represented as:

$$F_g = \{f_1, f_2, f_3, \dots, f_n\}, \quad f_i \in R^d \tag{3.5}$$

Where

- $F_g \in R^{n\times d}$ is the matrix of n samples with d features (e.g., clinical measurements)
- Each row f_i represents a feature vector for one patient

Step 1: Find the data center} by subtracting the mean of each feature

$$\bar{f} = \frac{1}{n}\sum_{i=1}^{n} f_i \tag{3.6}$$

$$F_{centred} = F - \bar{f} \tag{3.7}$$

Step 2: Compute the covariance matrix Σ

$$\Sigma = \frac{1}{n-1} F_{centred}^{\top} F_{centred} \tag{3.8}$$

This $d \times d$ matrix captures the relationships (covariances) between all pairs of features.

Step 3: Eigen decomposition} of the covariance matrix:

$$\Sigma \upsilon_i = \lambda_i \upsilon_i \tag{3.9}$$

Where

- λ_i is the eigenvalue representing the variance captured by component
- υ_i is the corresponding eigenvector indicating the direction of the principal component

Step 4: Arrange the eigenvalues λ_i in descending order to Select top-κ principal components

$$\boldsymbol{\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \cdots \geq \lambda_d} \tag{3.10}$$

now choose the top- κ eigenvectors

$$V_\kappa = [\upsilon_1, \upsilon_2, \dots, \upsilon_\kappa], \quad V_\kappa \in R^{d\times\kappa}, \quad \kappa < d \tag{3.11}$$

The accumulated co-variance retained to decide the appropriate no of principal components

$$ACV = \frac{\sum_{i=1}^{\kappa}\lambda_i}{\sum_{i=1}^{d}\lambda_i} \tag{3.12}$$

Step 5: Project the data onto the new subspace

$$Z = X_{centred} \cdot V_\kappa \tag{3.13}$$

Where $Z \in R^{n\times\kappa}$ is the transformed dataset with reduced dimensions.

Step 6: Now use this reduced dataset Z to the ELM classifier for classification.

3.6.3 Classification using ELM

Extreme Learning Machine (ELM) [1]is a single-layer feedforward neural network (SLFN) that achieves fast learning by randomly assigning input weights and hidden- layer biases, while analytically computing the output weights. By avoiding iterative backpropagation, ELM significantly reduces training time and computational complexity.
Figure 3.3.1 shows the working of the ELM model for breast cancer detection and Figure 3.3.2 shows the integration of Cloud with modified ELM Mode

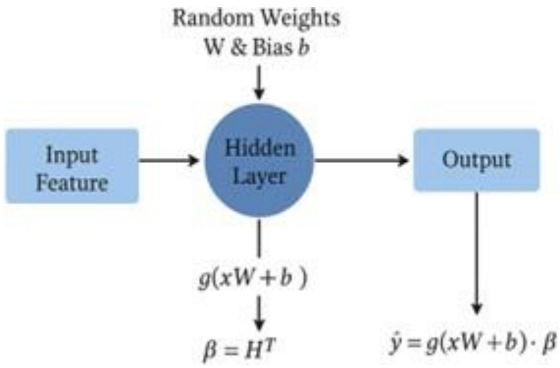


Figure 5: Working of ELM for breast cancer detection

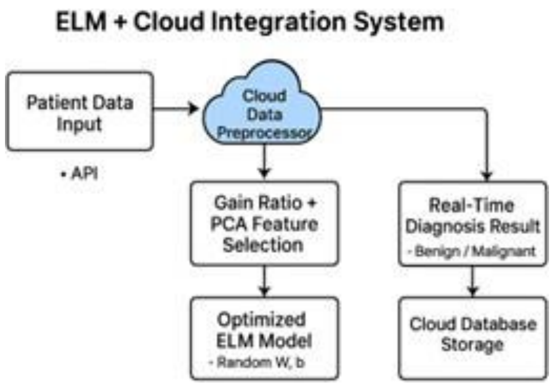


Figure 6: ELM Model Integrated with cloud

Algorithm 2: ELM-Based Classification
Input:
Dataset $DS \in R^{M\times N}$, Labels $T \in R^{M\times 1}$, Hidden neurons L
Output:
Output weights β, Prediction $\hat{y}$
begin
1. Randomly initialize weights $\boldsymbol{W} \in R^{N\times L}$ and $\boldsymbol{b} \in R^{L}$ weights and biases respectively
2. Find hidden layer output: $\boldsymbol{H} = g(DS \cdot \boldsymbol{W} + \boldsymbol{b})$
3. Get the output weights: $\boldsymbol{\beta} = \boldsymbol{H}^{\dagger}T$
4. *for* each new feature vector F_g *do*
5. Compute prediction: $\hat{y} = g\left(F_g \cdot \boldsymbol{W} + \boldsymbol{b}\right) \cdot \boldsymbol{\beta}$
end

3.7 Dataset and Evaluation Metrics

3.7.1. Dataset Used for Experimentation

The Wisconsin Breast Cancer Diagnosis (WBCD) dataset [1]is a widely used bench- mark for assessing machine learning–based breast cancer classification techniques. Provided by the University of Wisconsin, the dataset comprises 569 samples, with 357 benign and 212 malignant cases. Each instance is described by 30 real-valued features derived from digitized fine-needle aspirate (FNA) images of breast masses. These features are grouped into three categories—mean, standard error, and worst values—computed for ten morphological characteristics, including radius, texture, perimeter, area, smoothness, compactness, concavity, concave points, symmetry, and fractal dimension.

3.7.2. Evaluation Criteria

The evaluation model produces four possible predictive outcomes, indicating whether a sample is correctly or incorrectly classified as positive or negative. These outcomes include True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN). Such classifications are essential for quantifying the model's performance metrics and diagnostic accuracy.

$$\text{Accuracy} = \frac{TP + TN}{\text{Total Number of Samples}}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

$$\kappa = \frac{p_o - p_e}{1 - p_e}$$

where p_o represents the observed agreement, p_e denotes the expected agreement, y_j is the true label, and $\widehat{y_j}$ is the predicted label.

$$\text{F1-Score} = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$

4. Results

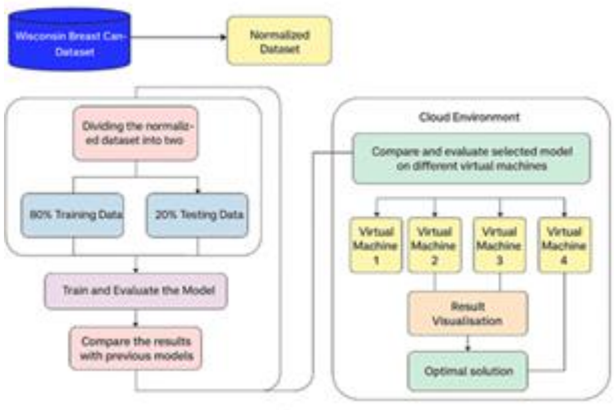


Figure 7: Experimental Setup in Cloud Environment for BC Detection with ELM Model

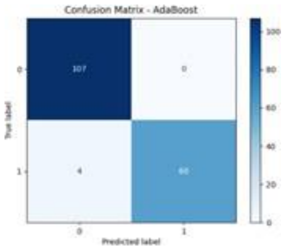
4.1 Standalone Envionment:

The Experiment is carried out on standalone machine with advanced configuration of 16 GB RAM, an Intel i5-10th generation of processor (2.71 GHz), and 1 TB HDD. For implementation of Classification models Python with PyCharm IDE [56] is used and performance is tested using WBCD dataset using multiple performance metrics.

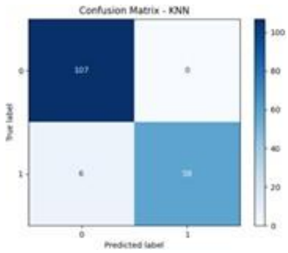
Figure 8: User Interface for Cloud Based ELM Model for BC Detection

Table 1: Performance metrics on standalone machine using different machine learning

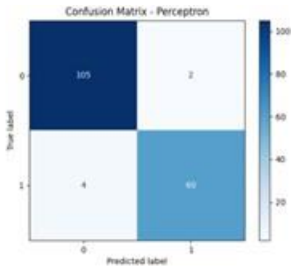
Model	Accuracy	Kappa	Precision	Recall
AdaBoost	0.9298	0.8460	0.9375	0.9545
KNN	0.9064	0.7913	0.9211	0.9375
Perceptron	0.8304	0.6614	0.9765	0.7545
SVM	0.9298	0.8447	0.9464	0.9464
ELM	0.9692	0.6046	0.7912	1.0000

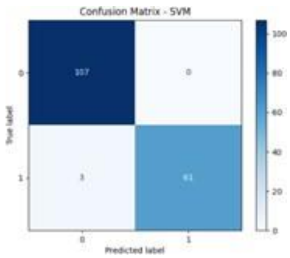


(a)

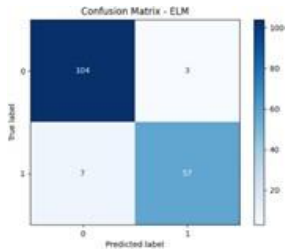


(b)





(d)



(e)

Figure 9: Confusion Matrices: a. AdaBoost b. KNN c. Perceptron d. SVM e. ELM

Table 2: Performance Metrics for Different Sample Sizes

Sample Size	Accuracy	Kappa	Precision	Recall	F1-Score
50	0.9340	0.8312	0.8950	0.9861	0.9382
100	0.9461	0.7921	0.8926	0.9869	0.9381
150	0.9561	0.7851	0.8859	1.0000	0.9410
200	0.9691	0.6052	0.7918	1.0000	0.8840
250	0.9651	0.4383	0.7301	0.9119	0.8114

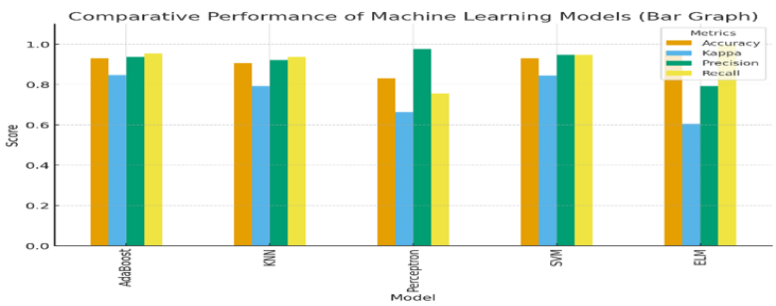
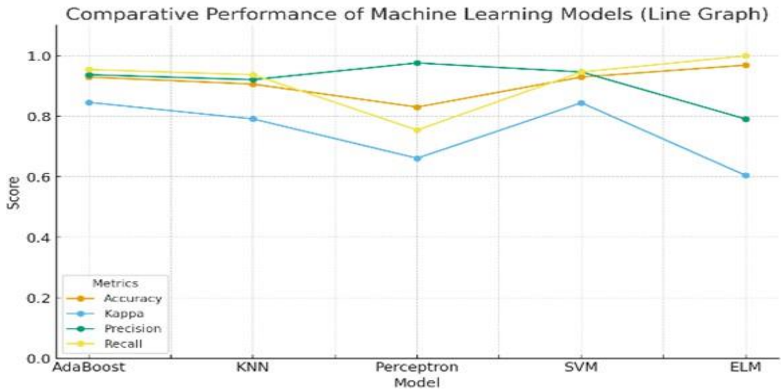


Figure 10: Comparative Analysis of Machine Learning Models

Figure 11: Comparison of Machine Learning Algorithms line Graph



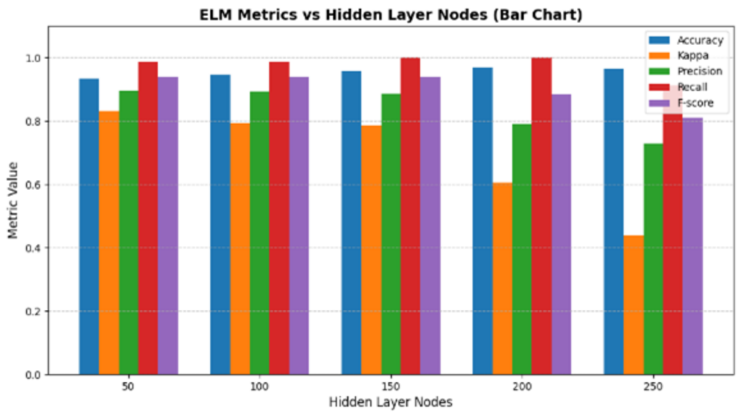


Figure 12: ML Algorithm Performance vs. Hidden Neurons

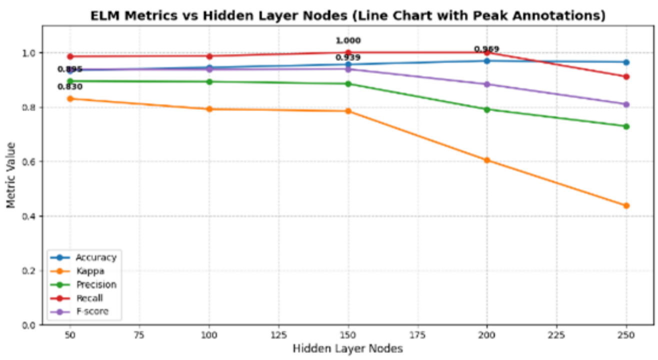


Figure 13: ML Algorithm Performance vs. Hidden Neurons

4.2 Cloud Envionment:

The ELM model was deployed on Amazon EC2 (PaaS) to reduce execution time and improve accuracy, with results compared to a standalone system. Table 3 shows

Table 3: Performance Analysis with Varying Hidden Layer Nodesin Cloud Environment

No. of HLN	Accuracy	Kappa	Precision	Recall	F-score

50	0.9358	0.7918	0.8927	0.9869	0.9376
100	0.9472	0.7425	0.8753	0.9587	0.9153
150	0.9521	0.7274	0.8708	0.9739	0.9195
200	0.9690	0.5174	0.7954	0.9043	0.8464
250	0.9649	0.5474	0.8127	0.9874	0.8916
300	0.9889	0.5469	0.8254	0.9886	0.8927

the how ELM out-perform to meet the highest accuracy of 0.9892 at HLN=300 hidden nodes on a system configured with 4 vCPUs and 16 GB RAM, and its perfor- mance tends to improve as CPU capacity, memory, and hidden node counts increase. Figure 15 represents the comparison of outcomes between cloud and the standalone environment.

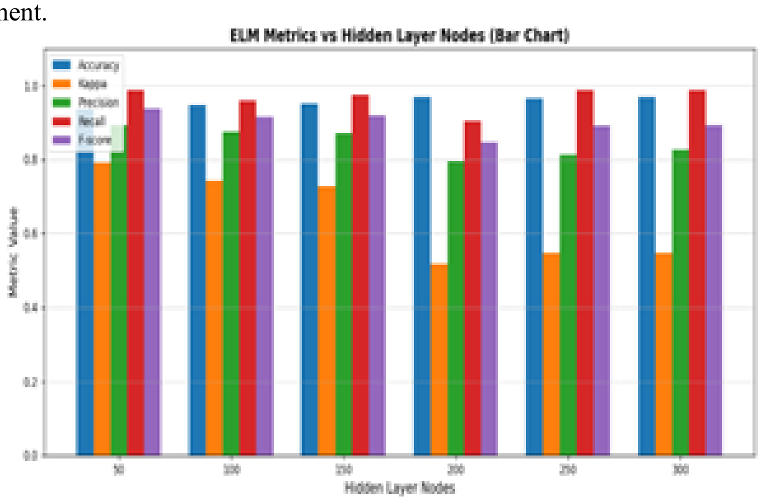


Figure 14: Analysis of ELM in cloud Environment with varying hidden layer nodes.

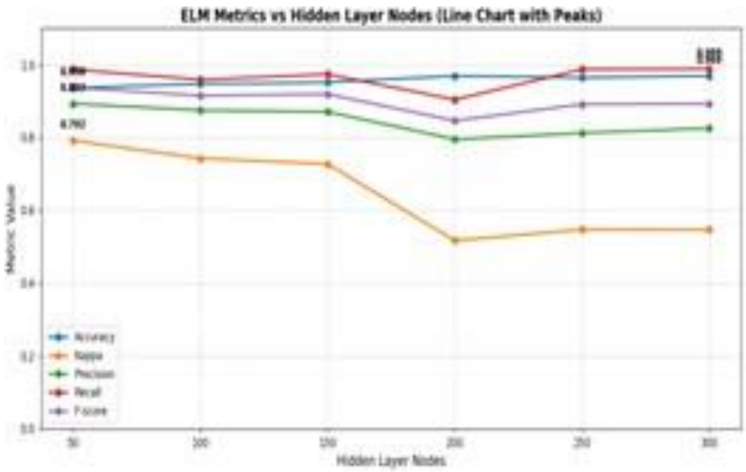


Figure 15: Analysis of ELM in cloud Environment with varying hidden layer nodes.

5. Discussion

The experimental results demonstrate that the proposed cloud-integrated optimized Extreme Learning Machine model achieves superior performance compared to conventional classification techniques. The integration of feature selection and dimensionality reduction techniques such as Gain Ratio and PCA significantly enhances the model's accuracy and generalization capability while reducing computational complexity. Furthermore, the deployment of the system on a cloud platform enables real-time data processing and remote accessibility, making it highly suitable for telemedicine applications. The proposed framework effectively addresses the limitations of traditional diagnostic methods, including dependency on local computational resources and delayed processing.

However, the model performance may vary depending on dataset quality and size, and further validation on large-scale real-world datasets is required to improve its robustness and reliability.

5.1.1 Conclusion:

This study presented an optimized cloud-based breast cancer diagnostic model using the Extreme Learning Machine (ELM) and the effective analysis on the WBCD dataset. By integrating gain ratio–based feature selection and PCA-driven dimensionality reduction, the proposed system achieved fast and accurate classification while mitigating overfitting. Comparative analysis showed that ELM consistently outperformed several conventional classifiers on a standalone system. When deployed in a cloud environment using Amazon EC2, the ELM model exhibited further improvements in accuracy, execution time, and overall stability due to scalable computational resources. The cloud-based implementation achieved notable performance, including an accuracy of 98.91 %, recall of 0.9825, precision of 0.9861, F1-score of 98.61 %, and an AUC of 0.9931. These results indicate that cloud-enabled ELM is a promising approach for reliable and efficient breast cancer diagnosis, particularly for remote or resource-constrained healthcare settings.

But still this cloud-based framework further can be improved the classification accuracy on larger real-time datasets which includes medical imaging, EHR and also can be integrated with IoT computing with improved security features

Refrences

[1] Lahoura, V., et al. Cloud computing-based framework for breast cancer diagnosis using extreme learning machine. Diagnostics 11, 241 (2021).

[2] Momenimovahed, Z. & Salehiniya, H. Epidemiological characteristics of and risk factors for breast cancer in the world. Breast Cancer 11, 151–164 (2019).

[3] Mohammed, M., et al. Neural network and multi-fractal dimension features for breast cancer classification from ultrasound images. Computers & Electrical Engineering 70, 871–882 (2018).

[4] Munir, K., et al. Detection and screening of COVID-19 through chest computed tomography radiographs using deep neural networks. In Data Science for COVID-19, Elsevier, 63–73 (2021).

[5] World Health Organization. Breast cancer—WHO.int (2024). URL: https://www.who.int/news-room/fact-sheets/detail/breast-cancer . Accessed 23 November 2024.

[6] Al-Balas, M., et al. Clinical outcomes of screening and diagnostic mammography in a limited resource healthcare system. BMC Women's Health 24, 191 (2024).

[7] Subasi, A., Kandpal, A. D., Raj, K. A. & Bagci, U. Breast cancer detection from mammograms using artificial intelligence. In Applications of Artificial Intelligence in Medical Imaging, Academic Press, 109–136 (2023).

[8] Mohammadiaghdam, N., Doshmangir, L., Babaie, J., Khabiri, R. & Ponnet, K. Determining factors in the retention of physicians in rural and underdeveloped areas: a systematic review. BMC Family Practice 21, 216 (2020).

[9] Gizaw, Z., Astale, T. & Kassie, G. M. What improves access to primary healthcare services in rural communities? A systematic review. BMC Primary Care 23, 313 (2022).

[10] Aborode, A. T., et al. Bringing lab to the field: Innovations in point-of-care diagnostics for tropical diseases in resource-limited settings. Advances in Biomarker Sciences and Technology 7, 28–43 (2025).

[11] Malherbe, K. Revolutionizing breast cancer screening: Integrating AI with clinical examination for targeted care in South Africa. Journal of Radiology Nursing 44, 195–202 (2025).

[12] Goh, S. S. N. & Hartman, M. From promise to practice: Harnessing AI for breast cancer screening. Intelligent Oncology 1, 4–6 (2025).

[13] Al-Fuqaha, A., Guizani, M., Mohammadi, M., Aledhari, M. & Ayyash, M. Internet of things: A survey on enabling technologies, protocols, and applications. IEEE Communications Surveys & Tutorials 17, 2347–2376 (2015).

[14] Mohammed, M. A., et al. Voice pathology detection and classification using convolutional neural network model. Applied Sciences 10, 3723 (2020).

[15] Botta, A., De Donato, W., Persico, V. & Pescapé, A. Integration of cloud computing and IoT: A survey. Future Generation Computer Systems 56, 684–700 (2016).

[16] Dorsey, E. R. & Topol, E. J. State of telehealth. The New England Journal of Medicine 375, 154–161 (2016).
[17] Wilson, L. S. & Maeder, A. J. Recent directions in telemedicine: Review of trends in research and practice. Healthcare Informatics Research 21, 213–222 (2015).
[18] Ayeni, B., et al. A web-based system for the discovery of blood banks and donors in emergencies. In Advances in Intelligent Systems and Computing, Vol. 1181, 592–600 (Springer, 2019).
[19] Ahmed, K. A., et al. Advancing breast cancer prediction: Comparative analysis of ML models and deep learning-based multi-model ensembles on original and synthetic datasets. PLOS ONE 20, e0326221 (2025).
[20] Park, S.-H. & Han, K. Methodologic guide for evaluating clinical performance and effect of AI technology for medical diagnosis and prediction. Radiology 286, 800–809 (2018).
[21] Islam, M. M., et al. Breast cancer prediction: A comparative study using machine learning techniques. SN Computer Science 1, 290 (2020).
[22] Naji, M. A., et al. Machine learning algorithms for breast cancer prediction and diagnosis. Procedia Computer Science 191, 487–492 (2021).
[23] Shah, D., Patel, S. & Bharti, S. K. Heart disease prediction using machine learning techniques. SN Computer Science 1, 345 (2020).
[24] Ma, H., Xu, C.-F., Shen, Z., Yu, C.-H. & Li, Y.-M. Application of machine learning techniques for clinical predictive modeling: A cross-sectional study on nonalcoholic fatty liver disease in China. BioMed Research International 2018, 4304376 (2018).
[25] Raoof, S. S., Jabbar, M. A. & Fathima, S. A. Lung cancer prediction using machine learning: A comprehensive approach. IEEE, 108–115 (2020).
[26] Akinnuwesi, B. A., et al. Application of support vector machine algorithm for early differential diagnosis of prostate cancer. Data Science and Management 6, 1–2 (2023).
[27] Sahu, A., Das, P. K. & Meher, S. Recent advancements in machine learning and deep learning-based breast cancer detection using mammograms. Physica Medica 103, 103138 (2023).
[28] Huang, D., et al. The application and development of deep learning in radiotherapy: A systematic review. Technology in Cancer Research & Treatment 20, 15330338211016386 (2021).
[29] Janowczyk, A. & Madabhushi, A. Deep learning for digital pathology image analysis: A comprehensive tutorial with selected use cases. Journal of Pathology Informatics 7, 29 (2016).
[30] Ghorbani, A., et al. Deep learning interpretation of echocardiograms. NPJ Digital Medicine 3, 10 (2020).
[31] Hutchinson, L., et al. Models and machines: How deep learning will take clinical pharmacology to the next level. CPT: Pharmacometrics & Systems Pharmacology 8, 131 (2019).
[32] Munir, K., Frezza, F. & Rizzi, A. Deep learning for brain tumor segmentation. Springer, 189–201 (2021).
[33] Munir, K., Frezza, F. & Rizzi, A. Brain tumor segmentation using 2D-UNet convolutional neural network. Springer, 239–248 (2021).
[34] Zou, J., et al. A primer on deep learning in genomics. Nature Genetics 51, 12–18 (2019).
[35] Mridha, M. F., et al. A comprehensive survey on deep learning-based breast cancer diagnosis. Cancers 13, 6116 (2021).
[36] Munagala, S. V., et al. A cloud-based CI/CD pipeline for machine learning in healthcare: Breast cancer prediction. SSRN (2025). URL: https://ssrn.com/abstract=5241747
[37] Rahman, M. A., et al. Enhancing early breast cancer detection through advanced data analysis. IEEE Access 12, 161941–161953 (2024).
[38] Ahmed, H. O. A. & Nandi, A. K. High performance breast cancer diagnosis from mammograms using mixture of experts with EfficientNet features (MOEFFNet). IEEE Access 12, 133703–133725 (2024).
[39] Kanber, B. M., et al. LightGBM: A leading force in breast cancer diagnosis through machine learning and image processing. IEEE Access 12, 39811–39832 (2024).
[40] Kreuzberger, D., Kühl, N. & Hirschl, S. Machine learning operations (MLOps): Overview, definition, and architecture. IEEE Access 11, 31866–31879 (2023).
[41] Wang, J., Lu, S., Wang, S.-H. & Zhang, Y. A review on extreme learning machine. Multimedia Tools and Applications 81 (2022).
[42] Huang, G.-B., Ding, X. & Zhou, H. Optimization method based extreme learning machine for classification. Neurocomputing 74, 155–163 (2010).
[43] Chen, K., Lv, Q., Lu, Y. & Dou, Y. Robust regularized extreme learning machine for regression

using iteratively reweighted least squares. Neurocomputing 230, 345–358 (2017).
[44] Zhang, P. & Yang, Z. A robust AdaBoost.RT based ensemble extreme learning machine. Mathematical Problems in Engineering 2015, 1–12 (2015).
[45] Bai, Z., Kasun, L. & Huang, G.-B. Generic object recognition with local receptive fields based extreme learning machine. Procedia Computer Science 53, 391–399 (2015).
[46] Deng, W., Bai, Z., Huang, G.-B. & Zheng, Q. A fast SVD-hidden-nodes based extreme learning machine for large-scale data analytics. Neural Networks 77, 14–28 (2016).
[47] Huang, G.-B., Liang, N.-Y., Rong, H.-J., Saratchandran, P. & Sundararajan, N. On-line sequential extreme learning machine. IASTED, 123–130 (2005).
[48] Zong, W., Huang, G.-B. & Chen, Y. Weighted extreme learning machine for imbalance learning. Neurocomputing 101, 229–242 (2013).
[49] Xin, J., Wang, Z., Qu, L. & Wang, G. Elastic extreme learning machine for big data classification. Neurocomputing 149, 464–471 (2015).
[50] Yang, R., Xu, S. & Feng, L. An ensemble extreme learning machine for data stream classification. Algorithms 11, 107 (2018).
[51] Liu, B., Xia, S.-X., Meng, F.-R. & Zhou, Y. Manifold regularized extreme learning machine. Neural Computing and Applications 27, 255–269 (2015).
[52] Khalid, S., Khalil, T. & Nasreen, S. A survey of feature selection and feature extraction techniques in machine learning, 372–378 (2014).
[53] Spencer, R., Thabtah, F., Abdelhamid, N. & Thompson, M. Exploring feature selection and classification methods for predicting heart disease. Digital Health 6, 205520762091477 (2020).
[54] Gabryel, M. & Damaševičius, R. The image classification with different types of image features. In Lecture Notes in Computer Science, Vol. 10245, 497–506 (Springer, 2017).
[55] Jia, P., Dai, J., Pan, Y. & Zhu, M. Novel algorithm for attribute reduction based on mutual-information gain ratio. Journal of Zhejiang University 40, 1041–1044 (2005).
[56] JetBrains. PyCharm: The Python IDE for professional developers. https://www.jetbrains.com/pycharm/ (2020).
[57] Shabina Sayed Ashwini, M. Save, Tatwadarshi P. Nagarhalli, Kirtida Naik, Manish Rana, Sunny Sall, Anas Dange.Leveraging Machine Learning to Enhance Public Health Outcomes: A Comprehensive Approach to Disease Prediction, Prevention and Management. SEEJPH Volume XXVI, 2025, ISSN: 2197-5248 (2025).