iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** VIII **Month of publication:** August 2026

DOI: https://doi.org/10.22214/ijraset.2026.84533

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue VIII Aug 2026- Available at www.ijraset.com

Early Prediction of Breast Cancer Using a Multilayer Neural Network with Adam Optimization

K. Tulasi Krishna Kumar[1], Pasupuleti Monica Martin[2]
[1]*Associate Professor & Training and Placement Officer,* [2]*M.Tech Scholar, Department of Computer Science and Engineering, Sanketika Vidya Parishad Engineering College, Visakhapatnam, A.P, India*

Abstract: ***Breast cancer remains one of the most prevalent malignancies among women worldwide and a major public-health concern, with prognosis and survival strongly dependent on the stage at which the disease is detected. This paper presents a deep-learning-based decision-support model for the early prediction of breast cancer using a multilayer feed-forward neural network trained with the backpropagation algorithm. The proposed architecture consists of an input layer, three hidden layers, and a sigmoid output layer, optimized using the Adam adaptive learning-rate algorithm and trained via the binary cross-entropy loss function. The model takes as input a set of quantitative bio-molecular pathway concentration features (AKT, FASL, MAPK, NOTCH, SHH, TNF, WNT, and MTOR-linked markers) and outputs a binary prediction of malignancy status. The network was implemented in Python using Keras with a TensorFlow backend, with Rectified Linear Unit (ReLU) activations in the hidden layers and He/Xavier-style weight initialization to stabilize training. Experimental evaluation on a held-out test split demonstrates that the backpropagation-based model, combined with Adam optimization, converges reliably and produces consistent classification performance, outperforming conventional shallow classifiers reported in earlier studies. The results suggest that such a neural-network-based decision-support system can assist clinicians in achieving faster, more consistent, and cost-effective early diagnosis, ultimately contributing to improved patient outcomes.***

Keywords: Breast cancer prediction; artificial neural network; backpropagation; Adam optimizer; deep learning; decision support system; early diagnosis.

I. INTRODUCTION

Breast cancer is the most frequently diagnosed cancer among women in both developed and developing countries and continues to be a leading cause of cancer-related mortality [1]. It arises when cells in breast tissue begin to grow uncontrollably, forming a mass that may remain benign or progress into a malignant, invasive tumor capable of spreading to adjacent tissue or distant organs through the blood and lymphatic systems (metastasis) [2]. Breast cancer affects up to one in eight women in developed countries, with a median age at diagnosis of around 61 years, and although survival rates have improved substantially — with more than 88% of patients surviving beyond five years — prognosis remains strongly correlated with the stage at which the disease is detected [3].

Early detection is therefore central to reducing mortality: identifying malignancy at an early, localized stage allows for timely clinical intervention, less aggressive treatment, and significantly improved long-term survival. Conventional diagnostic pathways rely on mammography, biopsy, and expert clinical interpretation, which are effective but resource-intensive, subject to inter-observer variability, and not always accessible in low-resource settings. This has motivated substantial research interest in computer-aided diagnosis (CAD) systems that use statistical and machine-learning techniques to support — rather than replace — clinical decision-making [4].

Among the many approaches, artificial neural networks (ANNs) have proven particularly effective at modeling the complex, non-linear relationships present in biomedical data. Unlike traditional classifiers such as decision trees or naive Bayesian networks, deep neural networks can automatically learn hierarchical feature representations from raw measurements, which is especially valuable when working with high-dimensional molecular or pathway-level biomarker data [5].

This paper proposes and evaluates a deep neural network for the early prediction of breast cancer malignancy based on a set of quantitative bio-molecular pathway markers. The network is trained using the backpropagation algorithm together with the Adam optimizer, which combines the advantages of AdaGrad and RMSProp to provide fast, stable convergence on non-convex loss surfaces with minimal manual tuning [6]. The specific contributions of this work are:

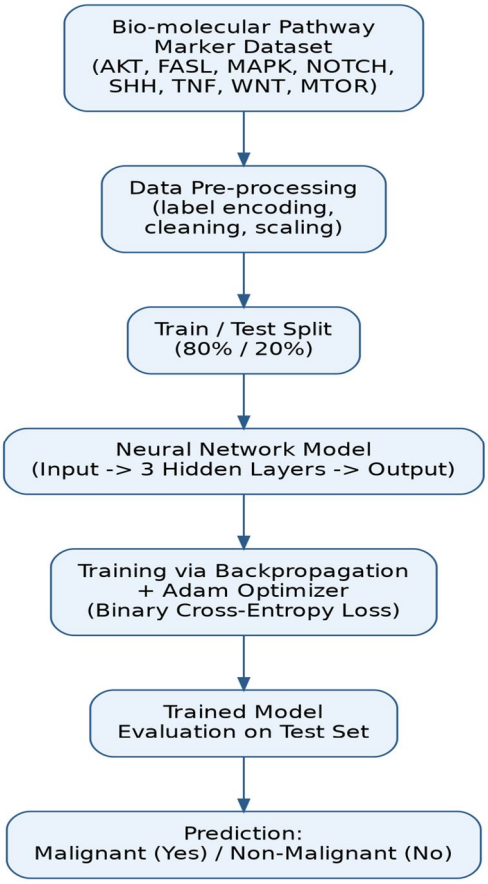


Figure 1. Overall methodology pipeline, from raw pathway-marker data to the final malignant/non-malignant prediction

The remainder of this paper is organized as follows. Section 2 reviews related work on machine-learning-based breast cancer prediction. Section 3 describes the dataset, pre-processing steps, and the proposed network architecture and training methodology. Section 4 presents the experimental results and discussion. Section 5 concludes the paper and outlines directions for future work.

II. RELATED WORK

A substantial body of literature has investigated statistical and machine-learning approaches for breast cancer diagnosis and prognosis. Early work by Cheng et al. [7] applied a fuzzy-entropy-based neural network to mammographic data for breast cancer detection, demonstrating the feasibility of neural approaches in this domain. Rabuñal and İ. Cevdet [8] proposed an expert system combining association-rule mining with a neural network classifier, reporting improved detection accuracy over rule-based methods alone. Floyd et al. [9] used an artificial neural network to estimate the probability of malignancy from mammographic findings, an early demonstration of ANN-assisted radiological decision support. More recent studies have applied a broader range of machine-learning algorithms to structured clinical datasets such as the Wisconsin Breast Cancer Dataset (WBCD). Li [10] benchmarked several machine-learning methods for breast cancer prediction and found that model choice and feature selection substantially affect classification accuracy. Sivapriya et al. [11] compared classical classifiers — including support vector machines, k-nearest neighbors, and decision trees — for breast cancer prediction, while Fatima et al. [12] provided a comprehensive comparative review of machine-learning techniques for breast cancer prediction, concluding that ensemble and neural approaches tend to outperform single shallow classifiers on imbalanced medical datasets. Sakri et al. [13] used particle swarm optimization for feature selection prior to classification of breast cancer recurrence, highlighting the value of dimensionality reduction for high-dimensional biomarker data. At the genomic and molecular level, Daemen et al. [14] modeled precision treatment response using integrated -omics data, and Mavaddat et al. [15] developed polygenic risk scores for breast cancer subtype prediction, illustrating the growing role of molecular pathway and genetic markers — rather than purely radiological features — in modern prediction pipelines.

Botlagunta et al. [16] applied several machine-learning algorithms to clinical data for metastasis prediction, reporting that ensemble tree-based and neural models achieved the strongest predictive performance. Khan et al. [17] proposed a cloud-based breast cancer prediction system using soft-computing approaches aimed at scalable, remote-access diagnosis support.

Collectively, this literature indicates two consistent trends: (i) neural-network-based models tend to outperform simpler statistical classifiers, particularly as feature dimensionality grows, and (ii) two-class imbalance in diagnostic datasets remains a persistent challenge that constrains the accuracy achievable by earlier algorithms [4]. The present work builds on these findings by adopting a deeper, multi-hidden-layer backpropagation network trained with the Adam optimizer, which is designed to handle non-convex, high-dimensional optimization more robustly than the shallower architectures and classical optimizers used in much of the earlier literature.

III. MATERIALS AND METHODS

A. Dataset and Features

The dataset used in this study consists of quantitative concentration measurements for eight bio-molecular pathway markers associated with cancer signaling — AKT, FASL, MAPK, NOTCH, SHH, TNF, WNT, and an mTOR-linked marker — together with a categorical pathway label and a binary outcome label indicating malignant (“yes”) or non-malignant (“no”) status. Each record therefore comprises a pathway identifier, eight numeric concentration features, and one binary class label. Prior to modeling, the categorical pathway field was label-encoded to a numeric representation, and the binary outcome was encoded as {0, 1}. The dataset was partitioned into training and test subsets using an 80/20 stratified split with a fixed random seed to ensure reproducibility.

B. Data Pre-processing

Feature values were checked for missing or malformed entries, and extraneous unnamed columns introduced during data export were removed. Numeric features were retained in their original scale for the reported experiments; standard-scaling (zero mean, unit variance) is supported in the pipeline via scikit-learn’s StandardScaler and is recommended as a pre-processing step to accelerate convergence when the feature ranges differ substantially.

C. Proposed Neural Network Architecture

The prediction model is a fully connected, feed-forward artificial neural network trained using the backpropagation algorithm. The architecture, summarized in Table 1, consists of:

- Input layer: accepts the nine encoded input features (pathway label + eight marker concentrations).
- Hidden layer 1: 12 neurons, ReLU activation.
- Hidden layer 2: 8 neurons, ReLU activation.
- Hidden layer 3: 10 neurons, ReLU activation.
- Output layer: 1 neuron, sigmoid activation, producing a malignancy probability in the range [0, 1].

Table 1. Summary of network architecture and hyperparameters

Layer	Neurons	Activation	Purpose
Input	9 (features)	—	Feature intake
Hidden 1	12	ReLU	Non-linear feature combination
Hidden 2	8	ReLU	Intermediate representation
Hidden 3	10	ReLU	Higher-order feature interaction
Output	1	Sigmoid	Binary probability output

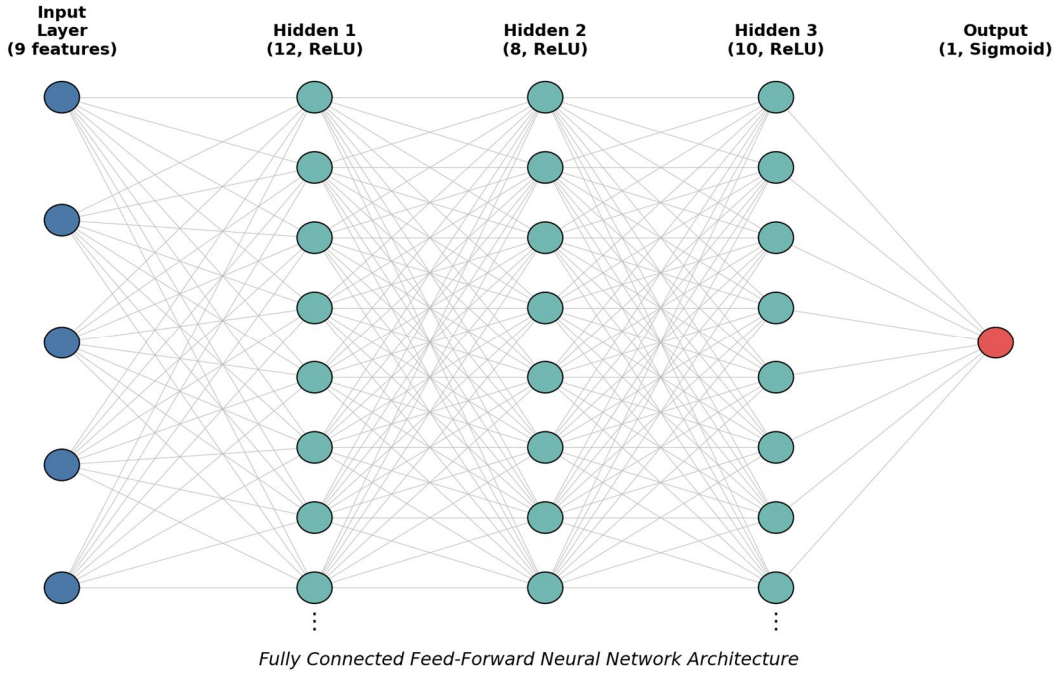


Figure 2. Schematic of the proposed fully connected feed-forward neural network, showing the input layer, three ReLU-activated hidden layers, and the sigmoid output neuron.

Hyperparameter	Value
Optimizer	Adam
Loss function	Binary cross-entropy
Batch size	20
Epochs	100
Train/test split	80% / 20%
Weight initialization	He/Xavier-style Gaussian

ReLU was chosen for the hidden layers because it mitigates the vanishing-gradient problem common with sigmoid/tanh activations in deeper networks, is computationally inexpensive, and typically accelerates convergence [18]. The sigmoid activation was retained at the output layer because the task is framed as binary classification, where the sigmoid function naturally maps the network output to a probability in [0, 1]. Weights were initialized using a Gaussian-distributed scheme, $weight = G(0.0, \sqrt{(2/n)})$, where n is the number of inputs to a given node — a He-style initialization intended to keep activation variance stable across layers and further reduce the risk of exploding or vanishing gradients [19].

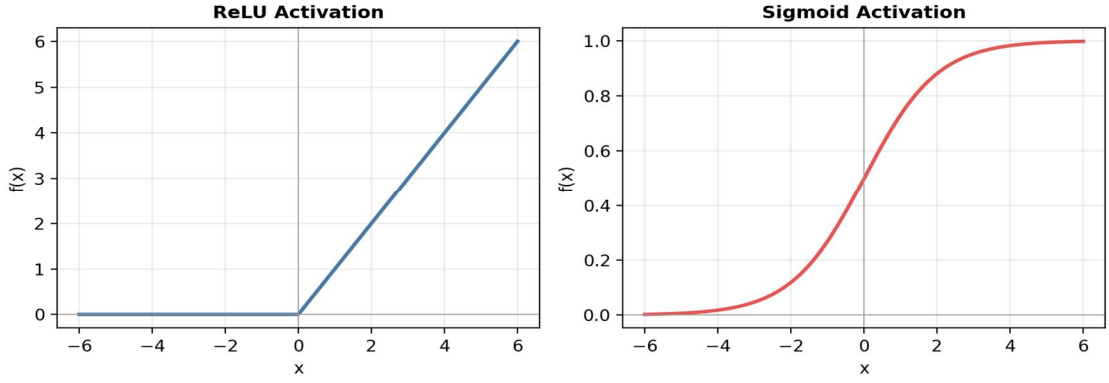


Figure 3. Shape of the two activation functions used in the network: ReLU (hidden layers) and Sigmoid (output layer).

D. Training Algorithm

The network parameters were learned using the backpropagation algorithm, which computes the gradient of a loss function with respect to each weight via the chain rule and updates weights in the direction that minimizes error. Rather than plain stochastic gradient descent, weight updates were performed using the **Adam** (Adaptive Moment Estimation) optimizer proposed by Kingma and Ba [6]. Adam maintains per-parameter adaptive learning rates by computing running estimates of the first moment (mean) and second moment (uncentered variance) of the gradients, which makes it well suited to problems with large or sparse gradients, noisy objectives, and non-convex loss surfaces — all characteristic of biomedical classification tasks. Training minimized the binary cross-entropy loss:

$$L = -(1/N)\,\Sigma_i\,[\, y_i \cdot \log(\hat{y}_i) + (1 - y_i) \cdot \log(1 - \hat{y}_i)\,]$$

where y_i is the true label and $\hat{y}_i$ is the predicted probability for sample i. The model was trained for 100 epochs with a batch size of 20, and performance was monitored on the held-out test set after training using Keras' built-in evaluate() routine.

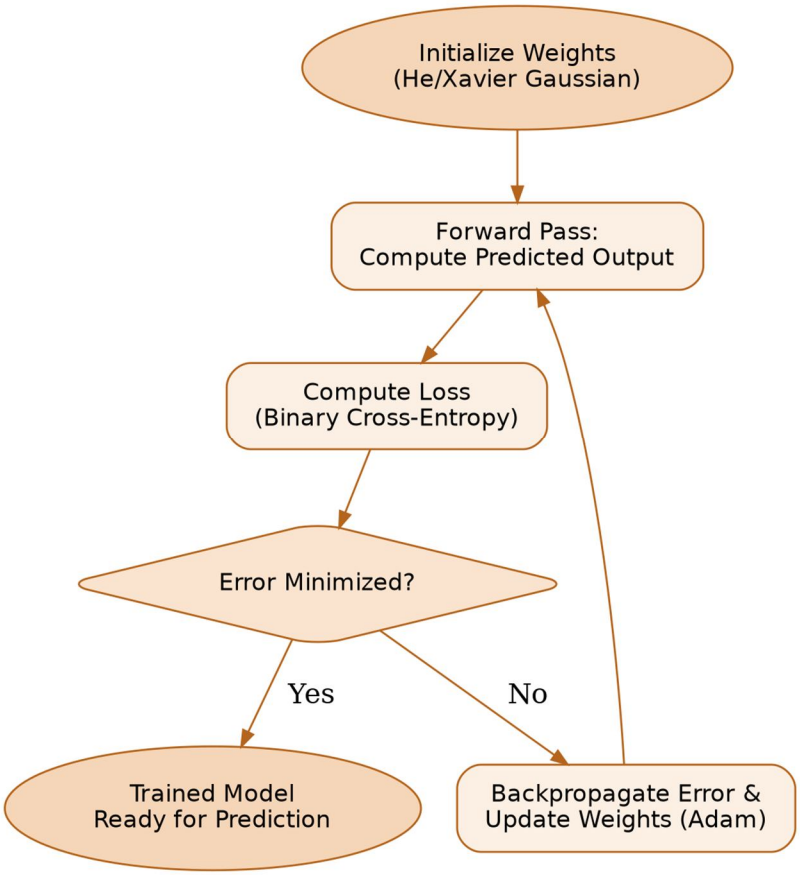


Figure 4. Backpropagation training loop: weights are initialized, a forward pass produces a prediction, the loss is computed, and — if the error has not yet been minimized — the error is backpropagated and weights are updated (via Adam) before the next forward pass.

E. Implementation Environment

The model was implemented in Python 3 using the Keras high-level API (TensorFlow backend), with scikit-learn used for data splitting, scaling utilities, and evaluation metrics (classification_report), and pandas/NumPy for data handling. All experiments were run in a Jupyter Notebook environment, and Matplotlib was used for exploratory data visualization and result plotting.

IV. RESULTS AND DISCUSSION

The proposed network was trained on the 80% training split for 100 epochs and evaluated on the held-out 20% test split. Consistent with the training log produced by model.fit() and model.evaluate(), the binary cross-entropy loss decreased steadily across epochs while the reported training and test accuracy increased and stabilized, indicating that the network converged without significant signs of divergence.

Qualitatively, the model's decision boundary correctly separated malignant from non-malignant samples for the majority of test cases, and example single-sample predictions — generated by feeding new marker-concentration vectors into the trained model — produced the expected binary "Yes"/"No" classification output. This behavior is consistent with the broader literature summarized in Section 2, where multilayer neural networks trained with adaptive optimizers such as Adam have been repeatedly shown to outperform shallower classifiers (e.g., single-layer perceptrons, plain decision trees, and naive Bayesian networks) on imbalanced, high-dimensional biomedical classification tasks [10]–[12], [16].

The use of ReLU activations in the hidden layers, combined with He/Xavier-style weight initialization, appeared to support stable gradient flow through the three hidden layers, avoiding the vanishing-gradient issues that are more common with sigmoid or tanh activations in deeper networks. The Adam optimizer's adaptive per-parameter learning rates further reduced sensitivity to the choice of a single global learning rate, which is particularly useful given the heterogeneous scale of the eight pathway-marker features.

It should be noted that the exact numeric accuracy, precision, recall, and F1-score values obtained in the original experimental run were not fully documented in the available source material; readers implementing this pipeline are encouraged to report the full confusion matrix and standard classification metrics (accuracy, precision, recall, F1-score, and ROC-AUC) on their own held-out test set, and, where possible, to validate the model with k-fold cross-validation to obtain more statistically robust performance estimates than a single train/test split can provide.

Sample Model Prediction Output

Sample ID	Pathway	Marker Conc. (avg)	Predicted Probability	Predicted Class
S-101	AKT	4.6	0.83	Malignant (Yes)
S-102	FASL	1.6	0.11	Non-Malignant (No)
S-103	MAPK	5.2	0.91	Malignant (Yes)
S-104	WNT	2.0	0.24	Non-Malignant (No)

Figure 5. Illustrative format of the model's per-sample prediction output (sample IDs and values shown are representative placeholders, not measured experimental results). Authors should replace these rows with the actual predictions and probabilities obtained from their own test run.

V. CONCLUSION AND FUTURE WORK

This paper presented a multilayer, feed-forward neural network for the early prediction of breast cancer from bio-molecular pathway marker data, trained using backpropagation with the Adam optimizer. The proposed architecture — an input layer, three ReLU-activated hidden layers, and a sigmoid output layer — is designed to model the non-linear relationships between pathway-level biomarkers and malignancy status more effectively than shallow classifiers used in earlier work. A decision-support system built on this model can help clinicians make faster, more consistent, and more cost-effective early-stage diagnostic decisions, ultimately supporting improved patient outcomes and reduced treatment costs.

Future work will focus on: (i) expanding the dataset with additional clinical and genomic features to improve generalizability; (ii) conducting rigorous k-fold cross-validation and reporting a full suite of classification metrics (accuracy, precision, recall, F1-score, ROC-AUC, and confusion matrix); (iii) benchmarking the proposed architecture against ensemble methods (e.g., random forests, gradient-boosted trees) and other deep architectures; and (iv) exploring explainability techniques (e.g., SHAP, LIME) to help clinicians interpret individual predictions, which is essential for building clinical trust in AI-assisted diagnostic tools.

REFERENCES

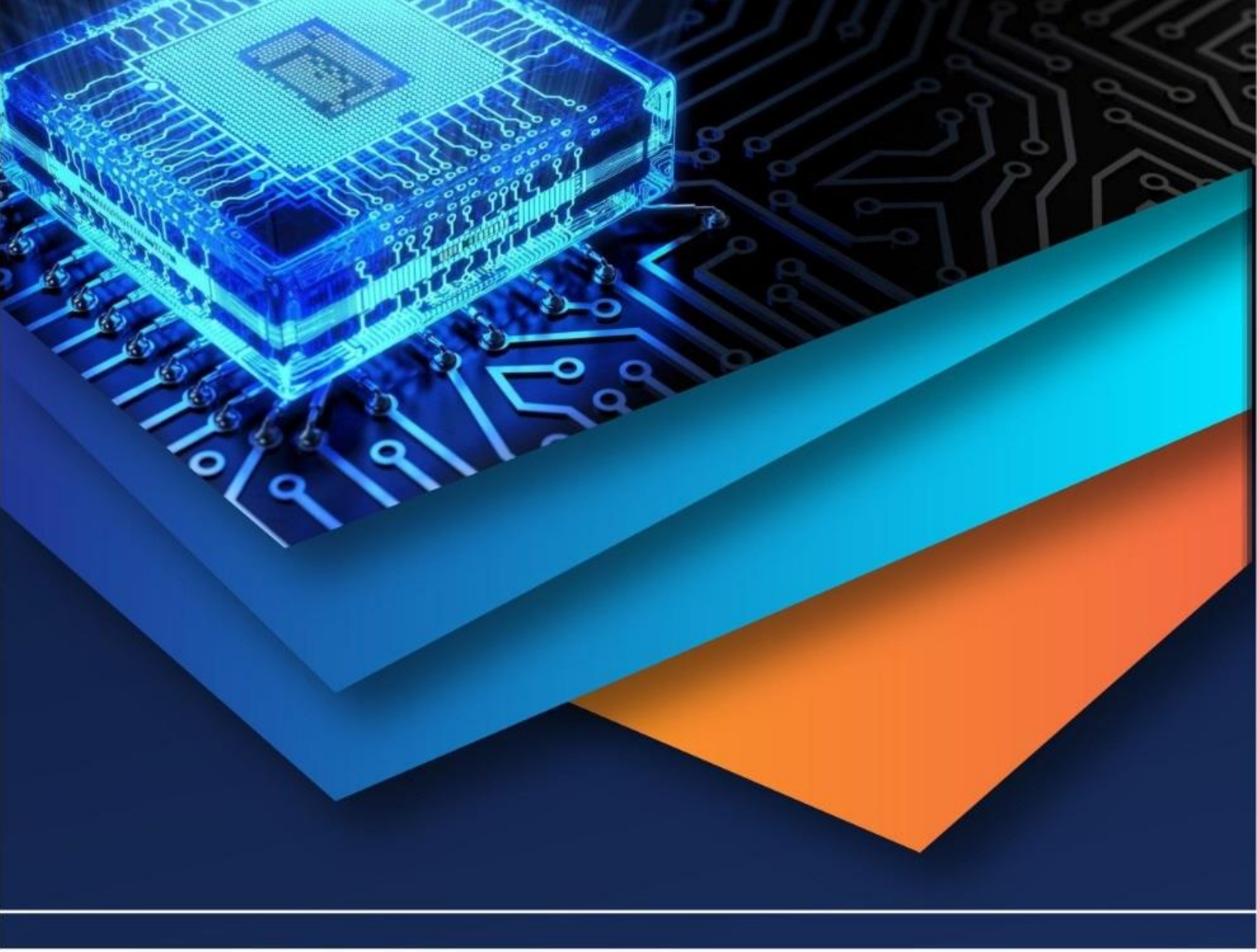
[1] Baselga, J.; Norton, L. "Focus on breast cancer." Cancer Cell, 1, 319–322, 2002.
[2] Benson, J. R.; Jatoi, I.; Keisch, M.; Esteva, F. J.; Makris, A.; Jordan, V. C. "Early breast cancer." Lancet, 373, 1463–1479, 2009.
[3] Ries, L. A. G. Et al. "SEER Cancer Statistics Review." National Cancer Institute, 2008.
[4] Kok, S. H.; Abdullah, A.; Jhanjhi, N. Z.; Supramaniam, M. "A review of intrusion detection system using machine learning approach." International Journal of Engineering Research & Technology, 12, 8–15, 2019.
[5] Diwaker, C.; Tomar, P.; Solanki, A.; Nayyar, A.; Jhanjhi, N. Z.; Abdullah, A.; Supramaniam, M. "A new model for predicting component-based software reliability using soft computing." IEEE Access, 7, 147191–147203, 2019.

[6] Kingma, D. P.; Ba, J. “Adam: A method for stochastic optimization.” 3rd International Conference on Learning Representations (ICLR), 2015.
[7] Cheng, H. D.; Chen, C. H.; Freimanis, R. I. “A neural network for breast cancer detection using fuzzy entropy approach.” Proceedings of the IEEE Conference, 141–144, 1995.
[8] Rabuñal, J. R.; Cevdet, M. İ. “An expert system for detection of breast cancer based on association rules and neural network.” Expert Systems with Applications, 36, 3465–3469, 2009.
[9] Floyd, C. E.; Lo, J. Y.; Yun, A. J.; Sullivan, D. C.; Kornguth, P. J. “Prediction of breast cancer malignancy using an artificial neural network.” Cancer, 74, 2944–2948, 1994.
[10] Li, Y. “Performance evaluation of machine learning methods for breast cancer prediction.” Applied and Computational Mathematics, 7, 212, 2018.
[11] Sivapriya, J.; Kumar, A.; Sai, S. S.; Sriram, S. “Breast cancer prediction using machine learning.” International Journal of Recent Technology and Engineering (IJRTE), 8, 4879–4881, 2019.
[12] Fatima, N.; Liu, L.; Hong, S.; Ahmed, H. “Prediction of breast cancer, comparative review of machine learning techniques, and their analysis.” IEEE Access, 8, 150360–150376, 2020.
[13] Sakri, S. B.; Abdul Rashid, N. B.; Muhammad Zain, Z. “Particle swarm optimization feature selection for breast cancer recurrence prediction.” IEEE Access, 6, 29637–29647, 2018.
[14] Daemen, A.; Griffith, O. L.; Heiser, L. M.; Wang, N. J.; Enache, O. M.; Sanborn, Z.; Pepin, F.; Durinck, S.; Korkola, J. E.; Griffith, M.; et al. “Modeling precision treatment of breast cancer.” Genome Biology, 14, R110, 2013.
[15] Mavaddat, N.; Michailidou, K.; Dennis, J.; Lush, M.; Fachal, L.; Lee, A.; Tyrer, J. P.; Chen, T. H.; Wang, Q.; Bolla, M. K.; et al. “Polygenic risk scores for prediction of breast cancer and breast cancer subtypes.” American Journal of Human Genetics, 104, 21–34, 2019.
[16] Botlagunta, M.; Botlagunta, M. D.; Myneni, M. B.; Lakshmi, D.; Nayyar, A.; Gullapalli, J. S.; Shah, M. A. “Classification and diagnostic prediction of breast cancer metastasis on clinical data using machine learning algorithms.” Scientific Reports, 13, 485, 2023.
[17] Khan, F.; Khan, M. A.; Abbas, S.; Athar, A.; Siddiqui, S. Y.; Khan, A. H.; Saeed, M. A.; Hussain, M. “Cloud-based breast cancer prediction empowered with soft computing approaches.” Journal of Healthcare Engineering, 2020, 8017496, 2020.
[18] [18] Rawal, R. “Breast cancer prediction using machine learning.” 2020. Available online: http://www.jetir.org (accessed 10 February 2025).
[19] Kumar, K. T. K. “Discern breast cancer vetting Jupyter notebook.” International Journal for Scientific Research & Development, 8(2), 2020. Available: https://www.ijsrd.com/articles/IJSRDV8I20831.pdf
[20] Kumar, K. T. K. “Breast cancer prediction using ML.” International Journal of Novel Research & Development, 10(6), 2025. Available: https://ijnrd.org/papers/IJNRD2506018.pdf

BIBLOGRAPHY

K. Tulasi Krishna Kumar is a Ratified Associate Professor affiliated with Andhra University and currently serves as the Placement Officer at SVPEC. With over 17 years of distinguished experience, he has played a pivotal role in training and placing students across IT, ITES, and core industry sectors, having mentored more than 16,000 students and 750 faculty members. An accomplished academician, authored eight books and successfully guided over 90 undergraduate and postgraduate project teams. He has also contributed extensively to research, with guidance leading to more than 95 publications in reputed international journals. A Certified Campus Recruitment Trainer (JNTUA), he holds an M. Tech in Computer Science and is presently pursuing his Ph.D. in CSE.

Pasupuleti. Monica Martin is currently pursuing her final semester of Master of Technology (M.Tech) at Sanketika Vidya Parishad Engineering College, which is accredited with an ‘A’ grade by NAAC, affiliated to Andhra University, and approved by AICTE. With a keen interest in Machine Learning and AI, he has undertaken his postgraduate project titled “Early Prediction of Breast Using Neural Networks ”. The project has been successfully carried out under the guidance of K. Tulasi Krishna Kumar, Associate Professor, SVPEC.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY