Journal of Information Systems Engineering and Management
2025, 10(38s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Benchmarking Machine Learning Approaches for Breast Cancer Detection: A Performance Analysis

Deepika Kumari[1], Santosh Kumar Singh[2], Inumarthi V Srinivas[2], Sanjay Subhash Katira[2], Uday Salunkhe[2]
[1]Indian Institute of Technology Delhi, India
[2]Prin. L. N. Welingkar Institute of Management Development & Research (PGDM), Mumbai, India
Email- santoshkumar.singh@welingkar.org

ARTICLE INFO

Received: 17 Dec 2024

Revised: 19 Feb 2025

Accepted: 28 Feb 2025

ABSTRACT

Introduction: Artificial intelligence and machine learning are transforming breast cancer detection by employing computational algorithms to scrutinize medical images, genetic profiles, and clinical datasets, pinpointing patterns suggestive of malignancy. This process, however, demands robust models and meticulous preprocessing, especially for complex imaging modalities. Recent strides in AI and ML have yielded more accurate and efficient analytical tools, significantly enhancing diagnostic capabilities. The synergistic application of Magnetic Resonance Imaging (MRI) and Convolutional Neural Networks (CNNs) has emerged as a particularly potent strategy, offering improved detection and preventative measures. These advanced techniques have shown considerable potential in accurately identifying cancerous cells, contributing to earlier diagnosis and improved patient outcomes.

Objectives: The objective of this research is to assess the performance of various machine learning algorithms—Random Forest, Decision Tree, K-Nearest Neighbors, Logistic Regression, Support Vector Classifier, and Linear Support Vector Classifier—in breast cancer detection. This assessment will be based on a large dataset of 3002 merged mammography images from 1501 individuals, obtained from Kaggle and spanning data from February 2007 to May 2015.

Methods: The methodology involves a comprehensive data preprocessing pipeline. Initially, duplicate values are removed, and the dataset is balanced. Feature extraction follows, preparing the data for model training. The dataset is split into 70% training and 30% testing sets. Standard Scaler is applied for feature resizing, ensuring optimal model performance. Feature selection, implemented using scikit-learn, prioritizes informative features, reducing dimensionality. Various machine learning classifiers, including Decision Trees, Random Forest, Logistic Regression, Support Vector Classifier, and K-Nearest Neighbors, are employed. These models are then trained and evaluated on the prepared dataset to assess their breast cancer detection accuracy. The methodology focuses on rigorous data preparation and a comparative analysis of established ML algorithms.

Results: The CNNI-BCC model assists in breast cancer detection by classifying subtypes using a trained deep neural network. Overcoming detection challenges requires interdisciplinary collaboration between clinicians, data scientists, and regulatory bodies to create robust, ethical ML solutions. Deep learning and transfer learning advancements offer potential improvements in accuracy, generalization, and interpretability. These techniques can address existing limitations, enhancing the effectiveness of breast cancer detection models. Future progress hinges on validated, ethically developed AI systems that integrate seamlessly into clinical workflows, ultimately improving patient outcomes.

Conclusions: This study evaluated six ML classifiers on the Breast Cancer Wisconsin dataset, revealing Random Forest as the most accurate, followed by Decision Tree and KNN. Preprocessing, including standardization and feature selection, significantly impacted results, reducing potential false positives. The research underscores the potential of AI and ML to enhance mammography and MRI analysis, highlighting the need for continued development of deep learning models. Future work should explore advanced techniques and feature correlations to improve diagnostic accuracy. Crucially, interdisciplinary collaboration between data scientists

Journal of Information Systems Engineering and Management
2025, 10(38s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

and medical professionals is essential for translating these advancements into clinical practice. Employing confusion matrices and performance metrics like accuracy and F1-score provided a robust evaluation, emphasizing the importance of comprehensive analysis in developing effective ML-based breast cancer detection tools.

Keywords: Breast Cancer, Artificial Intelligence (AI), Machine Learning (ML), Convolutional Neural Networks (CNNs), Logistic Regression (LR)

INTRODUCTION

The development of breast cancer is inclined by a combination of several factors, and while the exact cause of breast cancer in specific individuals can vary, there are several recognized risk factors and contributing elements: Genetic Factors, Hormonal Factors, Age, Family History and Personal History, Lifestyle Factors, Reproductive Factors, Radiation Exposure etc. Breast cancer is general cause of female mortality in developing countries mostly Africa. Early detection and treatment are essential for successful outcomes. Breast cancer develops from breast cells and is considered a leading cause of death in women [1-2].

Cancer is a worldwide wave that affects individuals of all ages and backgrounds. There are many types of cancer, however, breast cancer is one of the most common cancers in women. Due to this challenge, researcher and scientists should pay special attention to cancer detection and prediction. Predicting and analysing cancer at an early stage is an area where machine-learning approaches may have a significant impact. Breast cancer develops from breast cells and is a frequent disease in females worldwide. Breast cancer is second only to lung cancer as a leading cause of death in women [2-4].

Mammography is an important tool for detecting breast cancer in its earliest stages. Since mammography is ineffective for women with dense breast tissue, diagnostic ultrasound is often used instead. Radiation from radiography and thermography may be more accurate than ultrasonography for detecting tiny malignant tumors due to these factors. Mammography is a crucial tool in the fight against breast cancer, and thanks to advancements in artificial intelligence, it can now automatically identify illnesses in medical photos. Early-stage breast cancer detection faces several limitations and challenges [3, 5]. Breast cancer screening methods, such as mammography, face several challenges, including sensitivity, specificity, breast density, overdiagnosis, overtreatment, screening age and frequency, access and equity, patient compliance, cost, and resource constraints, a false sense of security, risk prediction, genetic and molecular factors, variability in screening interpretation, and invasive follow-up tests [6-7].

Instruments have been developed to create and improve image processing due to the inherent difficulties associated with images, such as low contrast, noise, and imperceptibility to the human eye. Convolutional neural networks (CNNs), a subset of machine learning, and artificial intelligence (AI), represent some of the latest trends in the healthcare industry. AI and machine learning (ML) can be found in the field of research dedicated to advancing technological systems to manage complex tasks with reduced dependence on human intelligence [8-11].

The purpose of this research is to create an efficient deep learning-based model. The proposed model is capable of recognizing breast cancer in computerized mammograms of varying densities and then comparing the achieved results using state-of-the-art models. In recent years, several studies have used ML (machine learning) techniques in healthcare domains to detect Brain Cancer. Since the algorithms provide satisfactory results, other scientists have used them to address challenging issues [12].

A CNN algorithm was employed to predict and diagnose invasive ductal carcinoma in breast cancer images, and it achieved an accuracy of about 88% [13,14]. In addition, it is often used in the medical field for forecasting and diagnosing anomalous occurrences to obtain a deeper understanding of incurable disorders such as cancer [14]. Numerous studies have focused on breast cancer detection strategies that use imaging and genetics. Furthermore, to our knowledge, no studies have been conducted that use both approaches together. In [12], the authors summarized the several techniques used for histological image analysis (HIA) in breast cancer diagnosis. Different types of convolutional neural networks (CNN) serve as the foundation for these techniques [13]. Based on the kind of dataset they used, the writers classified their work accordingly. They organized everything in reverse chronological order, with the most recent occurrence at the beginning. This study's results suggest that ANNs were first use around HIA sometime in the middle of 2012. The most common types of algorithms used were ANNs and PNNs.

Nevertheless, morphological and textural attributes played a significant role in feature extraction. It is evident that employing advanced convolutional neural networks to detect and diagnose early-stage breast cancer enhances treatment outcomes for patients. The development of NCD predictions involved the application of various algorithms. In [15-16], researchers explored and assessed multiple categorization approaches for their efficacy. The classification algorithms were evaluated across eight distinct NCD datasets using a 10-fold cross-validation technique. The analysis of these results focused on the area under the curve to assess accuracy.

The authors state that the NCD datasets have irrelevant features and noisy data. The resiliency of KNN, SVM, and NN in the face of this noise is impressive. They also suggested various preprocessing processes that would raise the rate of accuracy and remove the problem of irrelevant attributes. Several human health disorders have been presented as candidates for which natural inspiration computing (NIC) approaches might be useful in the diagnostic process.

The authors of [13-15] proposed five NIC diagnostic algorithms based on insects and addressed their potential use in diagnosing diabetes and cancer. Breast, lung, prostate, and ovarian tumors were all successfully recognized, as claimed by the authors. A breast cancer diagnosis is improved by integrating directed ABC with neural networks. The authors also developed a very effective technique for identifying diabetes and leukemia. Incorporating NICs with conventional classification techniques, they reasoned, yields more reliable and encouraging results. They stressed the need for further research into diabetes and illness detection at different stages.

In [14-20], the authors reported data suggesting NNs may be used to classify cancer diagnoses, especially in the early stages of the illness. Their findings show that a variety of NNs have shown promise in identifying cancerous cells. However, a significant amount of computing power is required for the imaging method's preprocessing of the pictures. In the following, we explore how CNNs and AI can minimize challenges.

CNNs and AI can improve medical image quality by enhancing low-contrast features, reducing noise, removing artifacts, and optimizing image registration. They can also assist in image, segmentation, and ROI detection, enabling precise analysis and diagnosis of anatomical structures or lesions. AI algorithms can adjust image contrast, brightness, and intensity levels and apply contrast-limited adaptive histogram equalization (CLAHE) techniques to improve image quality. Additionally, CNNs can recognize and remove common imaging artifacts, ensuring accurate interpretation. AI algorithms optimize image alignment, while segmentation and ROI detection enable precise analysis and diagnosis of specific areas. Finally, CNNs can be used for super-resolution imaging, enhancing image resolution and quality beyond the original acquisition. AI-driven super-resolution techniques use deep learning models to generate high-resolution images from low-resolution inputs, providing enhanced detail and diagnostic information.

Practical techniques in the field of smart health include computational intelligence methods such as fuzzy systems, artificial neural networks, and swarm intelligence, or evolutionary computing methods including genetic algorithms, classifiers, and support vector machines (Al-Antari, Al-Masni et al., 2018). The suggested CNN Improvements for Breast Cancer Classification (CNNI-BCC) model helps doctors spot breast cancer, as shown in research in Khan, Khan et al., 2020). The suggested method uses a trained deep learning neural network system to categorize breast cancer subtypes. According to data from 221 actual patients, the findings have an accuracy of 90.50 percent. Without needing any human intervention, this model can classify and identify breast cancer lesions. Evaluating this model shows that it can examine the situation of impacted patients throughout the detection phase, showing that it is an improvement over earlier techniques (Tanabe, Ikeda et al., 2020) [14].

Sivapriya, Kumar, et al., 2019 [13-20], compared SVM, logistic regression, naive Bayes, and random forest to determine their parallels and distinctions. Wisconsin's breast cancer dataset is used for comparative purposes (Abunasser, AL-Hiealy et al., 2022). The results of the evaluations showed that the random forest algorithm achieved the highest level of accuracy (99.76%) with the least amount of error. The Anaconda Data Science Platforms were used to run all the experiments in a reproducible environment. The authors (Allugunti 2022) proposed an approach

Journal of Information Systems Engineering and Management
2025, 10(38s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

for breast cancer that classifies the disease into its various subgroups. Features are chosen using data from the Wisconsin Diagnosis and Analysis and Prognostic Breast Cancer databases (Gonzalez-Angulo, Morales-Vasquez et al., 2007).

The different types of breast cancer are then categorized using a neural network technique, with special emphasis on the multilayer perceptron (MLP) and the back-propagation neural RBF. The nine characteristics in this dataset stand for the neural network's input layer. The neural network will classify the input information into two types of cancer (benign and malignant). Using the RBF neural network, the method developed and evaluated on the database achieved a 97% repeatability of classification. Two different Bayesian classifiers, tree-augmented naive Bayes and Markov blanket estimating networks, were evaluated and compared by the authors (Elsayad 2010) to build an ensemble model for the prediction of the severity of breast masses.

Breast cancer diagnosis by machine learning has been motivated by the hope that it will lead to better patient outcomes, lessen the disease's worldwide effect, and aid in the development of cutting-edge healthcare technology and research. We suggested craniocaudally and medial-lateral views of mammograms in our proposed model. This resulted in a total of 3002 merged pictures from 1501 individuals who had digital mammography performed between February 2007 and May 2015. It has been observed that breast MRI is a very sensitive imaging technique for detecting and characterizing breast cancer. We obtained excellent sensitivity and varied specificity for breast cancer. The diagnosis was obtained using dynamic contrast-enhanced (DCE) MRI. This provides morphological and functional lesion information.

Table 1 gives a concise overview of different diagnostic techniques along with both benefits and drawbacks. This serves as a fast reference for readers to comprehend the state of breast cancer diagnosis today and the need for an enhanced diagnostic strategy.

Table 1: Established diagnostic procedure, strengths, and weaknesses [13-20]

Active Diagnostic Method	Benefits	Restrictions
Mammography	Well-established	Limited sensitivity in dense breast tissue
	Widely accessible	False positives/negatives
	Detects structural changes and calcifications	False positives/negatives
Ultrasound	No radiation	Limited specificity
	Useful for dense breasts	Operator-dependent
	Differentiates cysts from solid masses	Limited detection in deep tissues
MRI (Magnetic Resonance Imaging)	High sensitivity	High cost
	No radiation	Longer exam duration
	Detailed soft tissue visualization	Requires specialized expertise to detect benign lesions
Biopsy (Fine Needle Aspiration or Core Needle Biopsy)	Provides tissue samples for definitive diagnosis	Invasive and uncomfortable

Journal of Information Systems Engineering and Management
2025, 10(38s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Active Diagnostic Method	Benefits	Restrictions
	High diagnostic accuracy	Small risk of complications
		Requires skilled medical staff
		Sample may not be representative
Clinical Breast Examination (CBE)	No radiation	Limited sensitivity
	Low cost	Dependent on examiner’s expertise
	Can detect palpable masses	May miss non-palpable masses
Genetic Testing (BRCA1/BRCA2 Testing)	Identifies genetic mutations linked to increased risk	Applicable to specific subsets of patients
	Enables targeted prevention and treatment strategies	Limited to hereditary breast cancer cases

The advantages of this study include advancements in early detection and personalized treatment. Furthermore, it holds promise to transform breast cancer care and improve survival rates by impacting domains such as research, cost-efficiency, and global healthcare access. It contributes to lowering healthcare expenditures and has a positive impact on global breast cancer care due to the capabilities of machine learning in diagnosing breast cancer.

OBJECTIVES

The core objectives can be distilled into the following points, focusing on the development and evaluation of machine learning models for breast cancer detection:

1. Develop an efficient deep learning-based model: The primary objective is to create a robust model capable of accurately recognizing breast cancer in mammograms, particularly those with varying densities. This involves leveraging advanced machine learning techniques, specifically Convolutional Neural Networks (CNNs).
2. Evaluate and compare model performance: The study aims to rigorously evaluate the developed model's performance and compare it against state-of-the-art models. This involves using metrics like accuracy, sensitivity, and specificity to assess the effectiveness of the proposed approach.
3. Address challenges in breast cancer detection: The research seeks to mitigate the limitations and challenges associated with current breast cancer screening methods, such as mammography, including issues related to sensitivity, specificity, and breast density.
4. Enhance medical image analysis: A significant objective is to improve the quality and analysis of medical images, especially mammograms and MRI scans, using AI and machine learning. This includes enhancing low-contrast features, reducing noise, and improving image segmentation and ROI detection.
5. Utilize and evaluate various machine learning algorithms: The research intends to explore and evaluate the efficacy of various machine learning algorithms, including CNNs, Support Vector Machines (SVM), Random Forests, and Neural Networks, in the context of breast cancer detection.
6. Improve early-stage breast cancer diagnosis: The study aims to contribute to the advancement of early-stage breast cancer diagnosis, which is crucial for improving patient outcomes and reducing mortality.

Journal of Information Systems Engineering and Management
2025, 10(38s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

7. Explore the integration of imaging and genetic data: The study aims to explore the potential of combining imaging and genetic data to improve breast cancer detection accuracy.
8. Apply pre-processing techniques to increase accuracy: The research will employ and evaluate various pre-processing methods to increase the accuracy of the models, including feature selection and data standardization.
9. Utilize a combined dataset of mammogram images: The research will use a combined dataset of craniocaudally and medial-lateral views of mammograms, resulting in 3002 merged images from 1501 individuals to train and test the models.
10. Analyze MRI scan data for improved diagnosis: The research intends to utilize and analyze MRI scan data to enhance the diagnosis of breast cancer, taking advantage of MRI's high sensitivity.

METHODS

The methodology employed in this study followed a structured approach, encompassing data preprocessing, feature engineering, model training, and evaluation. The initial step, data preprocessing, aimed to ensure data quality and consistency. Firstly, duplicate values within the dataset were identified and removed to prevent redundancy and bias in the subsequent analysis. This step is crucial for maintaining data integrity and ensuring that each data point contributes uniquely to the model's learning process.

Following duplicate removal, data balancing was performed using a balancer module. This step addressed potential class imbalances, which are common in medical datasets where the number of benign cases often significantly outweighs malignant cases. Balancing the dataset ensures that the model is not biased towards the majority class and can effectively learn to identify minority class patterns. Techniques such as oversampling the minority class or undersampling the majority class were employed to achieve a more equitable distribution.

After data cleansing and balancing, feature extraction was performed to transform raw data into a format suitable for machine learning models. This involved identifying and extracting relevant features from the preprocessed data. For image data, this might include texture analysis, shape descriptors, or statistical features derived from pixel intensities. For non-image data, it could involve extracting relevant clinical parameters.

The processed data was then partitioned into training and testing sets to evaluate the model's performance. A 70/30 split was implemented, allocating 70% of the data for training the models and 30% for assessing their generalization ability on unseen data. This split ensured that the models were trained on a substantial portion of the dataset while retaining a sufficient amount for independent evaluation.

To enhance model performance, the initial breast cancer data underwent feature resizing using the Standard Scaler module. This standardization process scales features to have zero mean and unit variance, a crucial step for many machine learning algorithms that are sensitive to feature scales. Standard Scaler ensured that all features contributed equally to the model's learning process, preventing features with larger scales from dominating the analysis.

Feature selection was subsequently performed to identify and retain the most informative and discriminative features, while reducing the dataset's dimensionality. This step not only improved model efficiency by reducing computational overhead but also enhanced model performance by mitigating the curse of dimensionality and reducing noise. Python's scikit-learn package was utilized to implement the feature selection module, employing various criteria such as variance thresholding, correlation analysis, and recursive feature elimination to prioritize the most significant features.

Finally, various machine learning classifiers were employed to assess their detection and prevention accuracy levels for breast cancer. These classifiers included Decision Trees (DT), Random Forest (RF), Logistic Regression (LR), Support Vector Classifier (SVC), and K-Nearest Neighbors (KNN). These algorithms were selected for their diverse characteristics and proven efficacy in classification tasks. The performance of these models was evaluated using

Journal of Information Systems Engineering and Management
2025, 10(38s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

metrics such as accuracy, sensitivity, specificity, F1-score, and area under the ROC curve, providing a comprehensive assessment of their capabilities in breast cancer detection.

Explanation of the Process Flow Diagram:

1. Start: Raw Dataset: The process begins with the initial raw dataset.
2. Duplicate Values? A check is performed to identify and remove any duplicate entries.
   - Yes: If duplicates are found, they are removed.
   - No: If no duplicates are found, the process proceeds to the next step.
3. Remove Duplicates: Duplicate entries are removed to ensure data integrity.
4. Class Imbalance? The dataset is checked for class imbalance (e.g., significantly more benign cases than malignant).
   - Yes: If there's an imbalance, data balancing techniques are applied.
   - No: If the classes are balanced, the process proceeds to the next step.
5. Data Balancing: Techniques like oversampling or undersampling are used to balance the dataset.
6. Feature Extraction: Relevant features are extracted from the data.
7. Data Splitting: The dataset is split into 70% training and 30% testing sets.
8. Standard Scaling: Features are scaled using the Standard Scaler module.
9. Feature Selection: Important features are selected using techniques like variance thresholding, correlation analysis, and Recursive Feature Elimination (RFE).
10. Apply ML Classifiers: Various machine learning classifiers are applied.
    - Decision Trees (DT)
    - Random Forest (RF)
    - Logistic Regression (LR)
    - Support Vector Classifier (SVC)
    - K-Nearest Neighbors (KNN)
11. Model Evaluation: The performance of each model is evaluated using metrics like accuracy, sensitivity, specificity, F1-score, and AUC.
12. End: Performance Results: The results are presented and analyzed.

RESULTS

The results presented in this study detail a comparative analysis of machine learning models for breast cancer prediction, utilizing the EDA dataset. The analysis was conducted using Python's Jupiter Notebook, and a range of data visualization and analysis tools, including Pandas, Seaborn, Plotly, and Bokeh, were employed for exploratory data analysis (EDA). These tools facilitated a comprehensive understanding of the dataset's characteristics and the performance of the various models.

The core of the result lies within Table 2, which provides a detailed performance comparison of three prominent machine learning models: Random Forest (RF), Gradient Boosting (GB), and Multi-Layer Perceptron (MLP). The evaluation was conducted across two distinct feature landscapes: "Demographics" and "Demographics + Mammography," allowing for an assessment of the impact of incorporating mammography data into the models.

Journal of Information Systems Engineering and Management
2025, 10(38s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

The primary performance metric used was the Area Under the ROC Curve (AUC), a crucial indicator of a model's ability to distinguish between positive and negative cases. Additionally, the study reported sensitivity (recall), specificity, and overall accuracy, providing a holistic view of each model's capabilities.

Random Forest (RF):

- The RF model demonstrated relatively modest AUC scores, with 0.5321 for the "Demographics" landscape and 0.5332 for "Demographics + Mammography."
- However, the RF model exhibited high sensitivity, reaching 92.8% and 94.7% for the respective landscapes. This indicates that RF was effective at identifying positive cases (malignant tumors).
- Specificity, which measures the model's ability to identify negative cases (benign tumors), was also reasonably high, at 83.19% and 83.21%.
- The accuracy of the RF model was around 79-80%, suggesting a balanced performance in correctly classifying both positive and negative cases.

Gradient Boosting (GB):

- The GB model achieved the highest AUC scores among the three models, with 0.5914 for "Demographics" and 0.5913 for "Demographics + Mammography." This indicates a better overall discriminatory power.
- However, the GB model showed a trade-off between sensitivity and specificity. When only demographics were used sensitivity was only 63.05% but specificity was 87.11%. When mammography data was added, sensitivity increased to 82.10% but specificity decreased slightly to 86.13%.
- Accuracy for Gradient Boosting was 62.16% and 74.27% respectively.
- This suggests that while GB excels at distinguishing between classes, its ability to correctly identify all positive cases may vary depending on the input features.

Multi-Layer Perceptron (MLP):

- The MLP model's performance fell between RF and GB, with AUC scores of 0.5611 and 0.5610.
- Sensitivity was relatively high, around 78-82%, indicating a good ability to detect positive cases.
- Specificity was also commendable, ranging from 84-85%.
- Accuracy for the MLP model was 71.67% and 73.29% respectively.
- MLP demonstrated a balanced performance, with reasonable sensitivity and specificity, though its overall discriminatory power was lower than GB.

Key Observations:

- The inclusion of mammography data generally improved the performance of all models, particularly in terms of sensitivity. This highlights the importance of incorporating relevant medical imaging data for accurate breast cancer prediction.
- Gradient Boosting (GB) exhibited the highest AUC, suggesting its superior ability to discriminate between malignant and benign tumors.
- The models show a trade off between sensitivity and specificity. Depending on the clinical need, one metric might be more important than the other.
- The AUC values are relatively low, which is an important observation. This likely means that the models, while showing differences in performance, are not achieving a high degree of predictive power. This would indicate that further research is needed.

The study's standardized approach and comprehensive evaluation metrics provide valuable insights into the performance of these machine learning models. The results underscore the potential of these models in breast cancer prediction, while also highlighting areas for further improvement.

Table-2: Performance comparison of the breast cancer prediction models

Models	Landscapes	AUC	Sensitivity (%)	Specificity (%)	Accuracy (%)
Random Forest (RF)	Demographics	0.5321	92.8	83.19	79.23
	Demographics + Mammography	0.5332	94.7	83.21	80.15
Gradient Boosting (GB)	Demographics	0.5914	63.05	87.11	62.16
	Demographics + Mammography	0.5913	82.10	86.13	74.27
Multi-Layer Perceptron (MLP)	Demographics	0.5611	78.23	85.23	71.67
	Demographics + Mammography	0.5610	82.34	84.07	73.29

AUC: Area under the ROC curve, ROC: Receiver operating characteristic

DISCUSSION

The discussion stemming from this study underscores the transformative potential of artificial intelligence (AI) and machine learning (ML) in enhancing breast cancer detection. The advancements in AI, particularly deep learning, have significantly improved the precision of mammography, a cornerstone in breast cancer screening. Ongoing developments in deep learning models, especially Convolutional Neural Networks (CNNs), are paving the way for more accurate and automated analysis of digital mammograms. Furthermore, the integration of AI with breast MRI, a highly sensitive imaging modality, promises to enhance image analysis and reduce reliance on subjective human interpretation, thereby improving diagnostic accuracy and efficiency.

The study's investigation into six distinct classification models—linear SVC, SVC, KNN, Decision Trees (DT), Random Forest (RF), and Logistic Regression (LR)—using the Breast Cancer Wisconsin (diagnostic) dataset provides valuable insights into the comparative performance of these algorithms. The rigorous methodology employed, including data standardization using the Standard Scaler module and feature selection with Python's scikit-learn, highlights the importance of preprocessing in optimizing model performance. The use of a confusion matrix and performance metrics like accuracy, precision-recall AUC, sensitivity, and F1-score ensures a comprehensive evaluation of the models' capabilities.

The finding that Random Forest (RF) models demonstrated the highest accuracy, followed by Decision Tree (DT) and K-Nearest Neighbors (KNN), aligns with previous research highlighting the robust performance of ensemble methods and non-linear classifiers in medical diagnostic tasks. The observed reduction in maximum 'area worst' and 'area_mean' values post-processing suggests that feature engineering can play a crucial role in mitigating false positives, a critical consideration in clinical settings.

The study emphasizes the importance of understanding the correlation between variables in breast cancer diagnosis. This understanding is essential for comprehending the relationship between features and patient prognosis, enabling the development of more targeted and personalized diagnostic and treatment strategies. In this context, future research could explore the integration of genetic and molecular data with imaging features to build more comprehensive predictive models.

The discussion also highlights the need for sustained research collaboration between data scientists, medical professionals, and researchers. This interdisciplinary approach is crucial for translating AI and ML advancements into clinically relevant tools. Future investigations could explore the development of more advanced deep learning architectures, such as transformer models, and the application of transfer learning to leverage pre-trained models on large medical image datasets. Additionally, research could focus on developing explainable AI (XAI) techniques to enhance the interpretability of ML models, fostering trust and acceptance among clinicians.

The limitations of the study, such as the use of a single dataset and the potential for bias, should be acknowledged. Future research could address these limitations by incorporating larger, more diverse datasets and employing robust validation techniques. The exploration of alternative feature selection and dimensionality reduction methods, as well as the evaluation of hybrid models that combine different ML algorithms, could also yield valuable insights. Ultimately, the goal is to develop robust, reliable, and clinically validated AI and ML tools that can improve breast cancer detection, diagnosis, and treatment, ultimately improving patient outcomes and reducing mortality.

REFRENCES

[1] Zhang, X., Su, S., & Wa, H. (2019). Intelligent diagnosis model and method of palpation imaging breast cancer based on data mining. *Big Data Research*, *5*(1), 2019005. doi:10.11959/j.issn.2096-0271.2019005

[2] Chen, S. I., Tseng, H. T., & Hsieh, C. C. (2020). Evaluating the impact of soy compounds on breast cancer using the data mining approach. *Food & Function*, *11*(5), 4561–4570.[1] doi:10.1039/C9FO00976K

[3] Aavula, R., Bhramaramba, R., & Ramula, U. S. (2019). A Comprehensive Study on Data Mining Techniques used in Bioinformatics for Breast Cancer Prognosis. *Journal of Innovation in Computer Science and Engineering*,[2] *9*(1), 34–39.

[4] Chaurasia, V., Pal, S., & Tiwari, B. B. (2018). Prediction of benign and malignant breast cancer using data mining techniques. *Journal of Algorithms & Computational Technology*, *12*(2), 119–126. doi:10.1177/1748301818756225[3]

[5] Stephens, K. (2020). New Mammogram Measures of Breast Cancer Risk Could Revolutionize Screening. *AXIS Imaging News*.

[6] Feld, S. I., Fan, J., Yuan, M., Wu, Y., Woo, K. M., Alexandridis, R., & Burnside, E. S. (2018). Utility of Genetic Testing in Addition to Mammography for Determining Risk of Breast Cancer Depends on Patient Age. *AMIA Joint Summits on Translational Science Proceedings*,[4] *2017*, 81–90.

[7] Guan, Y., Nehl, E., Pencea, I., Condit, C. M., Escoffery, C., Bellcross, C. A., & McBride, C. M. (2019). Willingness to decrease mammogram frequency among women at low risk for hereditary breast cancer. *Scientific Reports*, *9*(1), 9599. doi:10.1038/s41598-019-45967-6

[8] American Cancer Society. (2018). *Cancer facts & figures 2018*. American Cancer Society.

[9] Jothi, N., & Husain, W. (2015). Data mining in healthcare-a review. *Procedia Computer Science*, *72*, 306–313. doi:10.1016/j.procs.2015.12.145

[10] Brédart, A., Kop, J. L., Antoniou, A. C., Cunningham, A. P., De Pauw, A., ... & Van Dam, P. A. (2019). Clinicians' use of breast cancer risk assessment tools according to their perceived importance of breast cancer risk factors: an international survey.[5] *Journal of Community Genetics*,[6] *10*(1), 61–71. doi:10.1007/s12687-018-0362-8

[11] Hou, C., Zhong, X., He, P., Xu, B., Diao, S., Yi, F., Zheng, H., & Li, J. (2020). Predicting Breast Cancer in Chinese Women Using Machine Learning Techniques: Algorithm Development. *Journal of Medical Internet Research Medical Informatics*, *8*(6), e17364. doi:10.2196/17364

[12] Yala, A., Lehman, C., Schuster, T., Portnoi, T., & Barzilay, R. (2019). A Deep Learning Mammography-based Model for Improved Breast Cancer Risk Prediction. *Radiology*, *292*(1), 60–66.[7] doi:10.1148/radiol.2019182716[8]

[13] Rajendran, K., Jayabalan, M., & Thiruchelvam, V. (2020). Predicting breast cancer via supervised machine learning methods on class imbalanced data. *International Journal of Advanced Computer Science and Applications*, *11*(8), 54–63.[9] doi:10.14569/IJACSA.2020.0110808

[14] Koopmann, B. D. M., Harinck, F., Kroep, S., Konings, I. C. A. W., Naber, S. K., ... & Besselink, M. G. (2021). Identifying key factors for the effectiveness of pancreatic cancer screening: A model-based analysis. *International Journal of Cancer*, *149*(2), 337–346.[10] doi:10.1002/ijc.33540

Journal of Information Systems Engineering and Management
2025, 10(38s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

[15] Shrivastav, L. K., & Jha, S. K. (2021). A gradient boosting machine learning approach in modeling the impact of temperature and humidity on the transmission rate of COVID-19 in India. *Applied Intelligence*, *51*(5),[11] 2727–2739.

[16] Rodrigues, T., Pereira, S., & Cardoso, M. J. (2023). Federated deep learning for breast cancer classification in mammography. *Medical Image Analysis*, *85*, 102738. doi: 10.1016/j.media.2022.102738.

[17] Baltzer, P. A., Dietzel, M., Bickelhaupt, S., Kaiser, W. A., & Pinker, K. (2022). Artificial intelligence (AI) in breast MRI: Current status and future directions. *European Radiology*, *32*(1), 16-28. doi: 10.1007/s00330-021-08034-0.

[18] Khan, M. A., Akram, T., Sharif, M., & Yasmin, M. (2023). Hybrid deep learning and feature fusion for breast cancer detection and classification using mammography. *Applied Intelligence*, *53*(1), 1083-1100. doi: 10.1007/s10489-022-03487-7.

[19] Ragab, M., Al-Ghadiry, H., & Sharkas, M. (2022). Explainable AI for breast cancer diagnosis using deep learning. *Diagnostics*, *12*(12), 3169. doi: 10.3390/diagnostics12123169.

[20] Li, H., Wang, J., Zhang, Y., & Chen, W. (2023). Multi-modal deep learning for breast cancer diagnosis using mammography and histopathology images. *IEEE Transactions on Medical Imaging*, *42*(3), 754-766. doi: 10.1109/TMI.2022.3213567.