Prediction of Axillary Lymph Node Status in Women with Breast Cancer Using Machine Learning Models and Stacking Approach

Sanaz Haghighi[1], *Ali Hariri*[1], *Alireza Shadman*[*], *Soodabeh ShahidSales*[2]
department of Industrial Engineering, Ferdowsi university, Mashhad, Iran[1]
Cancer research center, University of medical sciences, Mashhad, Iran[2]

* Corresponding author: Alireza Shadman , a.shadman@um.ac.ir

Abstract
In breast cancer diagnosis, axillary lymph node (ALN) status plays a critical role in treatment planning. This study aims to enhance ALN involvement prediction using machine learning models. Five models (ANN, SVM, KNN, Random Forest, and Logistic Regression) were evaluated, with two integration approaches using stacking. In the first approach, biopsy and imaging results were excluded, while in the second, they were incorporated along with other demographic and pathological data. Data from 235 breast cancer patients at Omid Hospital in Mashhad (2002-2004) were collected, with 157 patients included in the analysis. RapidMiner software was used for model implementation. The best individual model, ANN, achieved 70.38% accuracy. When biopsy and imaging results were added, KNN and SVM models reached 89.17% accuracy. The optimal performance was obtained from the ANN-SVM stacked model, yielding 95.08% sensitivity and specificity. This study demonstrates that incorporating biopsy and imaging data into the stacked model improves ALN status prediction, offering a more accurate assessment method for breast cancer patients.

Keywords: Machine learning, stacking approach, Lymph node metastasis, Breast cancer

1. Introduction

According to the latest official announcement by the World Health Organization, breast cancer is the second most common cancer worldwide in recent years and ranks fifth among all cancer types in terms of the number of deaths. In Iran, this type of cancer ranks first in terms of incidence and sixth in terms of the death rate. Unfortunately, in breast cancer patients, the disease can spread to other parts of the body through the axillary lymph nodes. Therefore, testing the status of lymph node involvement is a necessary task. Additionally, investigating the involvement of the axillary lymph nodes is a significant factor in determining the most appropriate treatment [1].

In the past, the status of lymph node involvement was determined only through surgery [1]. Later, imaging methods such as mammography and ultrasound, with their minor side effects and non-invasive properties, became more popular for predicting the status of lymph nodes. Magnetic Resonance Imaging (MRI) is a medical imaging technique used in radiology to generate images of the anatomy and physiological processes within the body. Generally considered safe, MRI may pose risks if safety procedures fail or due to human error, leading to potential injuries [2]. Another method is biopsy or sentinel lymph node dissection (SLND). In this method, since samples are taken only from small portions of tissue, a negative biopsy result does not entirely rule out the possibility of cancer cells in the axillary tissue. On the other hand, when the final result indicates involvement of the lymph nodes, it can be concluded that cancer has spread to the axillary region [1].

Based on past studies, the accuracy of each of the mentioned methods is given in Table (1):

Table (1): Analyzing the results of examining surgical, imaging, and biopsy data of patients [1] [2] [3]

4. Method	3. accuracy while reporting involvement	2. accuracy while reporting non-involvement
7. Surgery	6. Between 76.9% to 97.6%	5. Between 76.9% to 97.6%
10. Imaging	9. Between 55% to 78%	8. Between 52% to 70%

13. Biopsy	12. 100 %	11. Between 62% to 85%

According to Table (1), it appears that surgery is generally more reliable than other methods. However, there is a growing trend in utilizing less invasive techniques, such as sentinel lymph node (LN) biopsy or non-invasive methods, for predicting axillary lymph node (ALN) status. This shift is motivated by the complications and morbidity associated with conventional axillary surgery, including issues such as lymphedema, restricted range of motion, and arm paresthesia and pain [4].

The growing popularity of machine learning in forecasting stems from the limitations faced by other methods, as the three mentioned methods struggle to provide reliable and accurate results. Approaches such as artificial neural networks (ANN) and support vector machines (SVM) have demonstrated satisfactory performance in similar studies. Machine learning, trained on datasets that include historical and real-time data, excels at interpreting new data and enhancing predictive accuracy. In summary, the widespread adoption of machine learning is driven by factors such as the availability of significant datasets, advanced computational capabilities, continuous algorithmic developments, the prevalence of open-source frameworks, and its tangible impact on optimizing business processes and decision-making across various industries.

To sum up, this study explores the prediction of lymph node involvement status in breast cancer patients. In addition to traditional methods such as imaging and biopsy, five machine learning models are employed for their increased accuracy and efficiency. The primary objective of this research is to identify the most effective approach for predicting lymph node involvement status in women with breast cancer.

2. Literature Review and Research Background

As previously mentioned, the primary focus of this study is the diagnosis of axillary lymph node involvement or non-involvement in women with breast cancer. In this context, the PubMed database contains 67 articles specifically addressing the prediction of lymph node status in breast cancer patients. Out of these articles, only three studies have used machine learning approaches, all of which applied machine learning for interpreting axillary imaging. Therefore, no study has used demographic and pathological data as inputs for machine learning methods. The oldest paper in this database, titled "Prediction of prognosis in patients with axillary lymph node-positive breast cancer: a statistical study" (1984), identified important variables for short- and long-term prognosis in 97 node-positive breast cancer patients, followed for a minimum of 98 months. The diameter of the primary tumor, categorized as less than or equal to 3 cm or greater than 3 cm, emerged as a crucial prognostic variable. When combined with the presence/absence of tumor cells in the efferent nodal vessels and the mean nuclear area of the tumor cells, it yielded accurate predictions of disease outcomes 60- and 98-months post-operation, achieving success rates of 83% and 80% for the respective time frames. While the number of tumor-bearing nodes remains a significant variable, tumor diameter provided additional valuable information in predicting outcomes [5].

Over the following years, studies in this field have improved and become more sophisticated, aiming to achieve better results with higher accuracy. In this direction, some surveys aim to introduce models to predict the status of axillary lymph nodes.

In 2023, Joachim Diessner and a team of researchers highlighted the correlation between increasing sensitivity and tumor size for various imaging modalities including sonography, mammography, and CT. Notably, MRI demonstrated consistent sensitivity values independent of tumor size, ranging from 41.46% to 52.94%. In contrast, sonography exhibited sensitivity percentages of 18.88% for pT1, 38.86% for pT2, 51.11% for pT3, and 55.88% for pT4. Except for tumor stage pT4, significant differences were observed between conventional imaging methods (sonography and/or mammography) and cross-sectional imaging (MRI and/or CT). For tumor stage pT1, cross-sectional imaging displayed higher sensitivity in detecting positive lymph nodes. The results indicated that the mammography method had the lowest sensitivity at 5.84% for tumor size 1, while MRI or CT achieved the maximum sensitivity of 62.96% for tumor size 4 [6].

In 2023 one study introduces high-definition microvascular imaging (HDMI), a contrast-free ultrasound technique to visualize microvascular structures in lymph nodes. By analyzing morphometric features of tumor microvessels, the researchers achieved high accuracy in identifying metastatic lymph nodes. This noninvasive method offers a promising alternative for breast cancer staging [7].

In the same year, another research combined radiomic features from MRI with deep learning algorithms to create a model for predicting axillary lymph node metastasis in breast cancer. The proposed method demonstrated high accuracy in detecting metastatic lymph nodes, offering significant potential for enhancing clinical decision-making [8].

In 2024, a study developed a predictive model for assessing axillary lymph node metastatic burden in breast cancer patients using radiomic features extracted from 18F-FDG PET/CT scans. The radiomics approach provided a reliable tool for stratifying patients based on metastatic risk, enabling personalized treatment planning [9].
The study by Tuba Laiq in 2025, aimed to predict axillary lymph node metastasis in breast cancer patients using ultrasonographic and clinicopathologic features. In this research, 176 patients were analyzed, revealing a significant correlation between metastasis and cortical thickness greater than 3 mm as well as the absence of a fatty hilum. The overall accuracy of ultrasound in detecting metastasis was 67.6%, with a sensitivity of 84.2% and a specificity of 48.1%. These findings suggest that ultrasound can serve as an effective screening tool; however, due to its low specificity, combining it with other imaging modalities or biopsy is necessary to enhance diagnostic precision [10].
In a general overview, studies on predicting lymph node involvement in women with breast cancer can be broadly categorized into two main groups. These studies primarily focus on either creating models to predict the status of lymph node involvement by considering various prognostic factors or accurately identifying tumors using imaging methods such as mammography, breast ultrasound, magnetic resonance imaging (MRI), or biopsy.
To facilitate a comprehensive comparison, a summarized overview of the previously mentioned papers, along with new additions and the current study, is presented in Table (2). Based on Table (2), several points are noticeable. First, the number of variables in our study is the highest among all studies. It includes both demographic (e.g., age) and pathological (e.g., HER-2) indicators. The highest number of variables in previous studies was 8. Second, in our study, we have used 5 different machine learning models to diagnose lymph node status and also applied the stacking method to evaluate the results of combining all these models. The last and most important point is the accuracy percentage, which is the highest in our study. This is discussed comprehensively in the results section. In addition to the items compared in the table, the following points can be mentioned as other strengths of this research:

- Conducting a distinct analysis of each model using both pathological and demographic variables as inputs.
- Incorporating imaging and biopsy results into the existing input variables and refitting the models.
- using the most robust machine learning models
- Evaluating the combination of results from all models through a stacking approach.
- Employing the Weighted by Relief method for feature selection.

Table (2): Comparison between information and results of similar studies

	Number of patients	**Method of prediction**	**Reported accuracy**	**Number of variables**	**Modeling with clinic pathological variables**	**Modeling with imaging or biopsy**
[5]	200	multivariate logistic regression	70%	8	*	
[6]	382	sonography, mammography, computed tomography [CT] and MRI (imaging methods)	US= 68.89	-		*
[7]	68	a contrast-free ultrasound quantitative microvasculature imaging technique	90%	-		*
[8]	376	MRI-based radiomics model	85%	-		*
[9]	124	The ultrasound, PET/CT, and clinical pathological features	75.1%	-		*
[10]	1055	ANN for interpreting US images	82%	-		*
this study	**235**	**ANN/SVM/KNN/Logistic/Random Forest (and their combination by stacking approach)**	**93.01%**	**15**	*****	*****

Our search revealed no PubMed studies integrating imaging or biopsy results with prognostic factor-based prediction models. This study collected patient data, including biopsy, imaging, and other clinical information, to develop five machine learning models for predicting lymph node involvement. Using a stacking approach, the models were tested in two scenarios: with and without biopsy and imaging results.

3. Methods and Models

3.1. Patients and Datasets

In the initial phase of this study, recorded data was collected from 235 women with breast cancer who were hospitalized at Omid Hospital in Mashhad between 2012 and 2014. Ultimately, 157 patients with breast cancer were included in the study, as the other patients had not undergone lymph node removal surgery, and the status of their lymph node involvement was unknown. Sixty-four patients had lymph node involvement due to metastasis, while no signs of involvement were observed in the remaining 93 patients after lymph node removal surgery. Demographic and clinical information, as well as details about treatment methods and other variables, were extracted from the patients' files. The age range was 25–89 years (average: 50 years), and the tumor size ranged from 0.2 to 14 cm (average: 3.84 cm). Other demographic data are shown in Table (3), and a flowchart describing the research process is shown in Figure (1).

Table (3): some Demographic data for 157 patients

characteristic	**Data set**
No. Of patients	157
involvement	64
Non-involvement	93
Age	
<40 y	47
40_49 y	55
50_59 y	27
60_69 y	22
≥70	6
Involved breast	
Right	64
Left	93
Tumor size	
≤0.2 Cm	62
2.1_4.0 Cm	74
>4 Cm	21

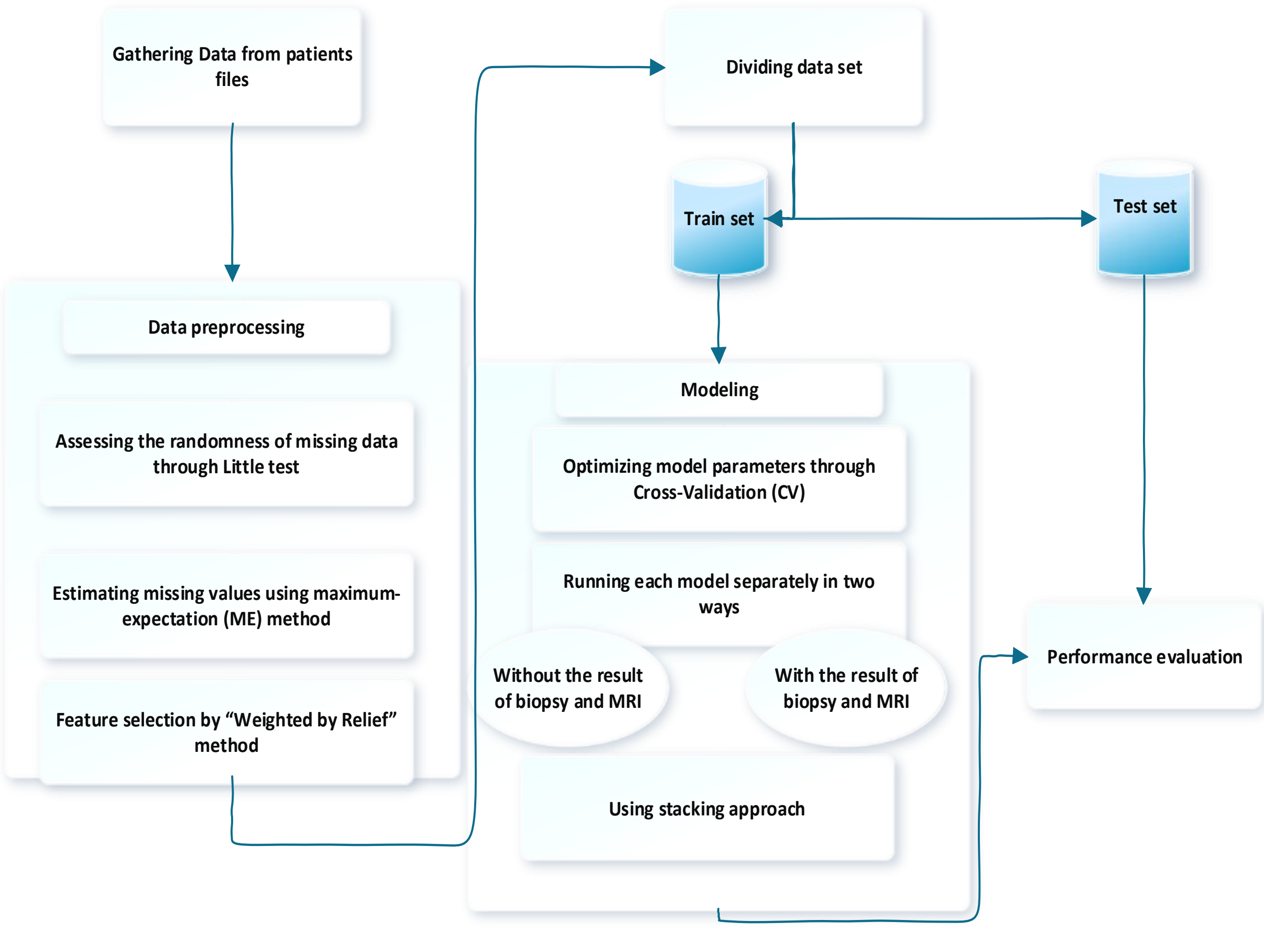

Figure 1: Flowchart of procedures in the data processing and development and evaluation of machine learning models

3.2. Data preprocessing

In data management, the most effective approach is to ensure comprehensive data collection, minimizing or avoiding missing data. However, the presence of missing data in the data set is almost inevitable. In this research, the data related to some variables had missing values [16]. Before selecting the suitable method for handling missing data, it was essential to examine the data for complete randomness. In the context of multivariate quantitative data analysis, little's test was employed to assess this criterion [11]. The results of the test showed that the missing data were entirely random. So, for estimating the missing values, the Expectation-Maximization method was used [12]. In the next step, it was time to choose the suitable variables to enter the model [13]. For this purpose, the Weighted by relief method was used on the training set. Based on this method, among the 15 collected variables, only the following 8 variables applied to the final phase of the research: 1. The involved breast (left, right), 2. the initial condition in the first visit (Primary patient or Recurrence), 3. the degree of tumor malignancy (1,2,3), 4. the status of the estrogen and, 5. progesterone hormone receptors, 6. human growth factor 2 (positive or negative), 7. the period between observing the symptoms by the patient to start the treatment process, 8. the Triple negative of breast cancer (positive or negative). So, it can be concluded that all the demographic variables have been excluded and only pathological variables have been detected effective on the status of lymph nodes.

3.3. Models

In this section, all the models used in this study and Stacking Approach are introduced.

3.3.1. Artificial Neural Network (ANN): ANNs mimic the human brain's structure using layers of interconnected nodes. They are trained by adjusting weights based on input data, enabling them to detect complex patterns. Their adaptability makes them effective for tasks like image recognition and predictive analytics[14].

3.3.2. Support Vector Machine (SVM): SVMs classify data by finding the optimal hyperplane to separate classes, with support vectors defining the boundary. They excel in high-dimensional data, handle non-linear relationships via kernel functions, and are effective with small datasets[15].

3.3.3. K-Nearest Neighbors (KNN): KNN predicts outcomes based on the majority class of neighboring data points. It operates without explicit training, performing calculations during prediction, making it suitable for tasks involving local patterns. It's intuitive but can struggle with large or high-dimensional datasets[16].

3.3.4. Random Forest (RF): RF is an ensemble method that builds multiple decision trees on random subsets of data and combines their outputs for predictions. It is robust, handles high-dimensional and mixed data types, and provides insights into feature importance, reducing overfitting[17].

3.3.5. Logistic Regression (LR): LR is a simple and interpretable algorithm for binary classification. It predicts probabilities using a logistic function and provides insights into feature importance. It is effective when features and outcomes have linear relationships [18].

3.4. Stacking Approach

Stacking combines predictions from multiple base models (level-0) using a meta-model (level-1) for enhanced accuracy. Base models generate predictions, which are aggregated by the meta-model trained on their outputs. The meta-model refines predictions, often using linear regression for regression tasks or logistic regression for classification. This method excels in complex problems but depends on the diversity and effectiveness of base models. Simpler models may be preferred if they offer comparable performance to avoid unnecessary complexity [19].

3.5. Evaluation criteria

The three markers given in Table (4) have been used as the evaluation criteria of the models. It is necessary to explain that the factors of sensitivity, specificity, and accuracy were calculated from the following methods. The TP index denotes the number of accurately predicted positive cases (indicating involvement), while FP represents the count of incorrectly predicted positive cases (erroneous lymph node involvement predictions). Similarly, TN and FN signify the accurate and inaccurate predictions of negative cases (non-involvement), respectively. In medical terminology, sensitivity refers to the percentage of individuals who test positive for a disease among those who actually have the disease. A highly sensitive test is effective in ruling out individuals who do not have the disease. On the other hand, specificity is the percentage of individuals without the disease who test negative for it. A highly specific test aids in correctly identifying individuals who truly have the disease [5] [6].

Table (4): the diagnostic performance indices

sensitivity	$\mathbf{Sensitivity} = \frac{\mathbf{TP}}{\mathbf{TP + FN}}$
specificity	$\mathbf{Specificity} = \frac{\mathbf{TN}}{\mathbf{TN + FP}}$
Accuracy	$\mathbf{Accuracy} = \frac{\mathbf{correct\ predictions}}{\mathbf{all\ predictions}}$

4. Results

Each machine learning model has several hyper parameters that can be adjusted to optimize its performance. Its behavior is controlled by hyper parameters, which have a big effect on how well the model performs. Therefore, prior to presenting the primary results of this research, the adjusted hyper parameters for each model are presented.

4.1. Adjust the hyper parameters of the models

In this study, the trial and error process in conjunction with K-fold CV has been used to adjust the hyper parameters of the models in Rapid Miner software, and their values are given in Table (5). In this process, multiple trials with different parameter values are performed, and for each set of parameters, the model's performance are evaluated using CV. This allows iteratively refine the parameter values based on the model's performance on validation data, ultimately leading to the selection of optimal parameters that generalize well. Moreover, using CV, helps prevent overfitting by providing a more realistic estimate of the model's performance on unseen data. It adds a systematic evaluation component to the trial and error process, ensuring that the selected parameters lead to a model that performs well on a broader range of data.

Table (5): Optimal parameters of the models

ANN

Parameter	**optimal amount**
Training cycle	40
Learning rate	0.2
Momentum	0.9
Decay	No
Shuffle	Yes
normalized	No

SVM

Parameter	**optimal amount**
Kernel Type	multi quadratic
Kernel Cache	٥
C	١٤,٣
Scale	No

Random forest

Parameter	**optimal amount**
Number of Trees	٢١
Apply Pruning	No
Apply pre pruning	No
Max depth	35
Criteria	Gain Ratio

Logistic regression

Parameter	**optimal amount**
Kernel Type	Dot
Kernel Cache	5
C	50
Scale	No

KNN

14. K	**15. 1**

4.2. Models Performance

After determining the optimal values for the model's hyper parameters, initially, the eight selected variables were individually applied to the models. Subsequently, the output of each model was assessed based on the criteria outlined in Table (4). The results of these evaluations are presented in Table (6)-A. Following this, patients' MRI and biopsy results were incorporated as additional variables alongside the previous inputs, and all models were re-executed. The outcomes are provided in Table (6)-B. It is evident that this step has led to improvements in all indicators across all models.

At this stage, an evaluation of the accuracy of the imaging and biopsy methods has been conducted based on the information recorded in patients' files regarding the results of mammography of the axillary lymph nodes at the beginning of the treatment and the outcomes of surgical lymph node removal. The results are presented in Table (7).

Table (6): models Performance

		without mammography and biopsy results			
ANN	**SVM**	**KNN**	**RF**	**LR**	Indicator Approach
٧٠,٣١	٥٦,٣٢	68.75	٦٤,٧١	**71.11**	Sensitivity
٧٠,٣١	٦٣,٤١	68.75	٦٦,٦٧	61.45	Specificity
٧٠,٣٨	٥٨,٤	68.85	٦٥,٥١	64.81	Accuracy
		considering mammography and biopsy results			
83.56	**87.88**	85.71	٨٣,٥٦	79.66	Sensitivity
94.55	90.32	93.10	٩٤,٥٥	75.36	Specificity
88.40	**89.17**	**89.17**	٨٨,١٤	77.5	Accuracy

Table (7): biopsy and imaging results

Method	Sensitivity	Specificity	Accuracy
Imaging	65%	63%	64%
Biopsy	100%	68%	84%

Based on the results shown in Table (6), In the first case, the highest accuracy belongs to the ANN model. But when the model reports the involvement, the LR model is more accurate. the SVM model is the least accurate one, which is only slightly more than 50%. Then after adding biopsy and MRI results to the models, evaluation indicators improved significantly in all models, while the difference between the two indicators Sensitivity and Specificity decreases. The SVM accuracy has improved the most and it reached from 58.4% to 89.17%, which is also the accuracy rate for KNN model. However, if the outcome of the model is involvement, the SVM model with a probability of 87.88% is more reliable. But if the model report is non -involvement, the result of the ANN and RF models are 94.55% which are the best performing. Compared to the information presented in Table (7), the SVM model is the only one exhibiting lower accuracy when contrasted with biopsy and imaging methods. Therefore, separately running all models with pathology variables has better results than running only biopsy or only imaging.

In the case of lymph nodes in breast cancer, as previously stated, if it is not involvement and is diagnosed with error, a vain surgery with complications such as lymphoma and limitation of hand movement is imposed on the patient. On the other hand, if the cancer has reached the lymph nodes, but this is not correctly diagnosed, the cancer may affect the patient's body with wider metastases and threaten her life. Therefore, the correct diagnosis of involvement is more important. Accordingly, the SVM model is recommended.

4.3. Ensemble Method – Stacking

In Table (8), the results of implementing the stacking approach and the different combinations of 5 machine learning models are presented, incorporating the results of biopsy and mammography as additional variables in the models. For example, in composition number 17, In the context of combining Support Vector Machines (SVM) and Artificial Neural Networks (ANN) using stacking, the process typically unfolds as follows:

1. Base models (SVM and ANN): Initially, separate SVM and ANN models are trained on the dataset to make predictions independently. Each model captures different aspects of the underlying patterns in the data.

2. Meta-model formation: The predictions of SVM and ANN models serve as input features for a meta-model. These predictions become new input features for the meta-model, which is often a simpler model such as LR (in this study) or another algorithm capable of combining diverse predictions.
3. Meta-model training: The meta-model is trained on the dataset using SVM and ANN predictions as input features. During this training, the meta- model learns how to weight and combine the predictions of the base models to optimize the overall performance.
4. Final Prediction: After the meta-model is trained, it can be used to make final predictions on new, unseen data.

Table (8): Stacking implementation results for different modes of combining models (Test set)

NO	ANN	SVM	KNN	R.F	Logistic	Accuracy
1	*	*	*	*	*	**90.70**
2	*	*	*	*		**88.40**
3	*	*	*		*	**88.40**
4	*	*		*	*	**89.39**
5	*		*	*	*	**88.40**
6		*	*	*	*	**88.33**
7	*	*	*			**87.63**
8	*	*		*		**89.74**
9	*	*			*	**90.71**
10	*		*	*		**88.27**
11	*		*		*	**88.40**
12	*			*	*	**91.35**
13		*	*	*		**88.27**
14		*	*		*	**89.87**
15		*		*	*	**88.46**
16			*	*	*	**88.40**
17	*	*				93.01
18	*		*			**87.63**
19	*			*		**90.62**
20	*				*	**90.71**
21		*	*			**89.06**
22		*		*		**89.10**
23		*			*	**88.46**
24			*	*		**88.33**
25			*		*	**89.87**
26				*	*	**88.40**

Based on Table (8), the highest accuracy (93.01 %) is related to the combination of the two models ANN and SVM. Therefore, the results of implementing this combination are in Table (9).

Table (9): Stacking implementation results for the combination of ANN and SVM

Stacking (ANN+SVM)	**Indicator**
٩١,٠٤	Sensitivity
٩٥,٠٨	Specificity
٩٣,٠١	Accuracy

Based on these results, if the disease involves lymph nodes, the model with 91.04% probability predicts this, correctly. On the other hand, if there is no lymph node involvement, the probability of correct prediction increases to 95.31%.

Therefore, it can be concluded that with modeling using machine learning algorithms, the prediction of lymph node involvement in breast cancer patients can be improved to an acceptable level. When, in addition to the demographic and pathological variables, the results of imaging and biopsy of the patient's breast in the initial visit are applied to the models as input variables, the integration of machine learning models with the accumulation approach, increases the accuracy of prediction. So, for example, in the case of non-involvement reporting, the probability of error is only 4.69%. Usually, for more efficient treatment, the physician determines the time intervals for check-ups, according to the conditions of each patient. In Figure (2), In each stage of the study, the prediction accuracy in reporting involvement and non-involvement is given in this study.

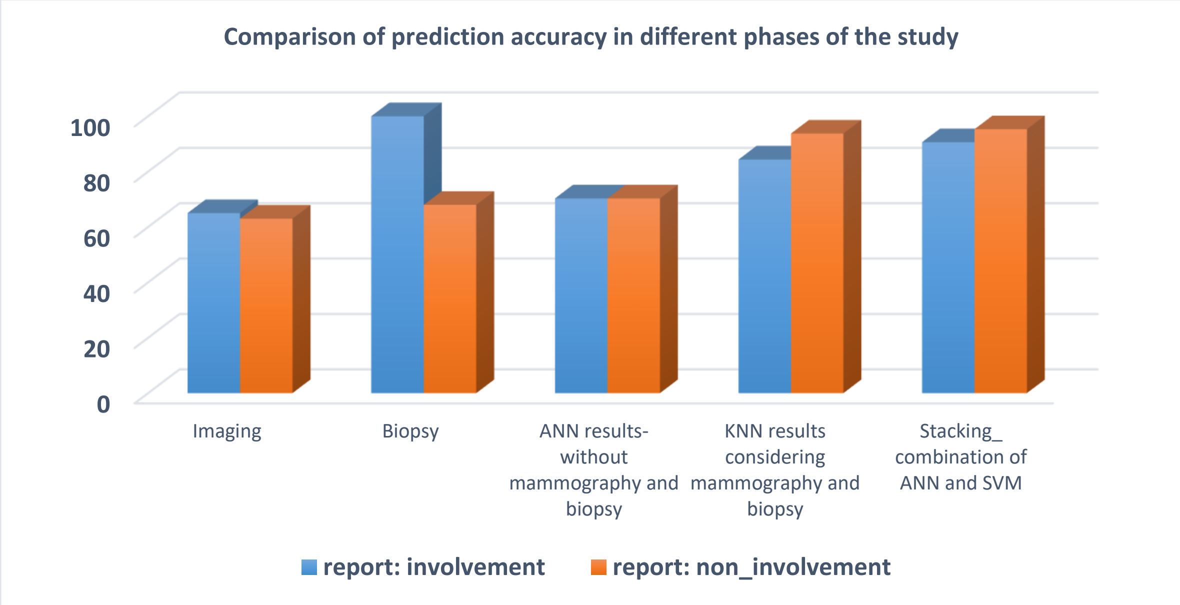

Figure (2): Comparison of prediction accuracy in different phases of the study

Based on these results, it is suggested that if the biopsy results show involvement of the axillary lymph nodes, you should completely trust it. Otherwise, if the biopsy results show no involvement, it is recommended to use the accumulation method to combine the two ANN and SVM models to predict it with the highest accuracy (95.31%). That is, if the model says no involvement, it is correct with a probability of 95.08%. However, since there is still a 4.92% chance of error, regular check-ups are essential for the patient.

5. Discussion

This study investigated the stacking approach combining five machine learning models to predict lymph node involvement in breast cancer. Two modes were analyzed: one using only demographic and pathology data and another incorporating biopsy and imaging results. The best-performing combination, ANN and SVM, achieved a prediction accuracy of 93.01%, significantly improving early lymph node evaluation compared to traditional methods.

This is the first study applying stacking to deep learning for lymph node metastasis prediction, highlighting its importance for guiding surgical decisions and therapies. Traditional non-invasive methods like CT, MRI, ultrasound, and PET scans provide valuable insights but have limitations. Physical exams and biopsies, while common, struggle to detect micro-metastases early. Advances in deep learning offer new opportunities to overcome these challenges.

In 2023, Joachim Diessner and colleagues examined the sensitivity of imaging methods based on tumor size. MRI showed consistent sensitivity (41.46%-52.94%) regardless of size, while sonography's sensitivity increased with tumor stage, ranging from 18.88% (pT1) to 55.88% (pT4). Cross-sectional imaging (MRI/CT) outperformed conventional methods (sonography/mammography) for smaller tumors, achieving up to 62.96% sensitivity for large tumors, compared to mammography's lowest at 5.84%. These results highlight the advantages of MRI and CT for detecting lymph nodes, particularly in early stages [20].

In 2024, a group of researchers, aimed to enhance the diagnosis of axillary lymph node metastasis in breast cancer patients using a radiomics-based nomogram derived from photoacoustic imaging. Photoacoustic imaging data from peritumoral regions were collected and analyzed using radiomics techniques. The predictive accuracy of the novel model was compared with traditional methods, demonstrating that incorporating radiomics features from peritumoral areas significantly improved sensitivity and overall diagnostic performance. The proposed model achieved a predictive accuracy of 90.5%, highlighting its potential as a non-invasive and reliable tool for assessing axillary lymph node metastasis in breast cancer patients [21].

Another study in the same year, explores the evolution of axillary lymph node metastasis management in breast cancer, from traditional axillary surgery to recent advancements in preoperative diagnosis and axillary management. It reviews historical surgical approaches, highlighting their limitations and the shift toward less invasive techniques. The study emphasizes the role of modern imaging modalities, including ultrasound, MRI, and novel molecular imaging, in improving preoperative lymph node assessment. Additionally, it discusses updated strategies such as sentinel lymph node biopsy and targeted axillary dissection, which aim to reduce overtreatment while maintaining oncological safety. These advancements contribute to a more personalized and less invasive approach to axillary management in breast cancer patients [22].

Studies indicate that PET/MRI has the highest accuracy among diagnostic methods (excluding modeling) for assessing lymph node involvement in breast cancer patients. This imaging method is safer due to reduced radiation exposure and offers improved predictive accuracy. However, its widespread use is limited by several disadvantages:

- High initial costs and space requirements.
- Limited availability compared to standalone PET or MRI.
- Lack of standardized protocols and combined PET/MRI reporting.
- Long acquisition times (up to 60 minutes).
- Gamma radiation exposure from injected radiotracers.

In contrast, predictive modeling methods provide notable advantages that address these limitations while offering higher accuracy:

- No need for costly infrastructure or radiation exposure.
- Ability to uncover complex data patterns and integrate diverse information (clinical, genetic, lifestyle).
- Flexibility to test unexperienced scenarios and adapt to new technologies or biomarkers.
- Continuous learning for improved accuracy with new data.
- High processing speed, remote accessibility, and better handling of incomplete data.
- Ensemble methods for combining models, enabling robust predictions.

These benefits make modeling a powerful, accessible alternative for predicting lymph node involvement.

Our study has notable limitations and benefits that require attention:

1. Retrospective Nature: The study relied on a limited dataset, and its findings are influenced by the composition of this data. Larger and prospective studies are necessary to validate these results for clinical application.

2. Diagnosis Stability: Lymph node metastasis and non-metastasis diagnoses are inherently unstable and dependent on the timing of breast surgery. Some patients initially diagnosed with negative lymph nodes might develop positive lymph nodes over time if followed up longer.

3. Single-Center Study: As the research was conducted in a single center, further studies involving multiple centers or regions are needed to generalize the findings more confidently.

On the other hand, there are also some benefits:

1. Novel Modeling Approach: Unlike most studies focusing on improving imaging methods or comparing them with biopsy results, this study integrated biopsy and MRI results as input variables into the models alongside demographic and pathological data, enhancing prediction accuracy.

2. Advanced Machine Learning Models: The use of five combined machine learning models, including ANN and SVM, and the comparison of their outputs provided a unique and innovative strategy for predicting lymph node status.

3. Enhanced Accuracy: The direct use of biopsy and MRI results—biopsy capturing cellular-level information and MRI imaging anatomical and physiological processes—explains the improvement in accuracy.

Conclusion

The stacking approach of ANN and SVM, incorporating biopsy and MRI results alongside other patient data, provides a reliable prediction of lymph node involvement in early breast cancer. If biopsy results confirm involvement, medical professionals can directly plan treatments to prevent metastasis without needing further model predictions.

References

[1]. Lazar, A. M., et al. (2023). Feasibility of sentinel lymph node biopsy in breast cancer patients with axillary conversion after neoadjuvant chemotherapy—A single-tertiary centre experience and review of the literature. Diagnostics, 13(18), 3000. https://doi.org/10.3390/diagnostics13183000
[2]. Tafreshi, N. K., et al. (2010). Molecular and functional imaging of breast cancer. Cancer Control, 17(3), 143–155. https://doi.org/10.1177/107327481001700302
[3]. de Freitas, R., Jr., et al. (1991). Accuracy of ultrasound and clinical examination in the diagnosis of axillary lymph node metastases in breast cancer. European Journal of Surgical Oncology, 17(3), 240–244. PMID: 2044777
[4]. Duan, Y., et al. (2023). Multimodal radiomics and nomogram-based prediction of axillary lymph node metastasis in breast cancer: An analysis considering optimal peritumoral region. Journal of Clinical Ultrasound, 51(7), 1231–1241. https://doi.org/10.1002/jcu.23520
[5]. Fehm, T., et al. (2005). Prediction of axillary lymph node status of breast cancer patients by tumorbiological factors of the primary tumor. Strahlentherapie und Onkologie, 181(9), 580–586. https://doi.org/10.1007/s00066-005-1374-y
[6]. Diessner, J., et al. (2023). Evaluation of different imaging modalities for axillary lymph node staging in breast cancer patients to provide a personalized and optimized therapy algorithm. Journal of Cancer Research and Clinical Oncology. https://doi.org/10.1007/s00432-022-04221-9
[7]. Ferroni, G., et al. (2023). Noninvasive prediction of axillary lymph node breast cancer metastasis using morphometric analysis of nodal tumor microvessels in a contrast-free ultrasound approach. Breast Cancer Research. https://doi.org/10.1186/s13058-023-01670-z
[8]. Liu, Z., et al. (2023). Predicting axillary lymph node metastasis in breast cancer patients: A radiomics-based multicenter approach with interpretability analysis. European Journal of Radiology. https://doi.org/10.1016/j.ejrad.2024.111522
[9]. Li, Y., et al. (2024). Prediction of the axillary lymph-node metastatic burden of breast cancer by 18F-FDG PET/CT-based radiomics. BMC Cancer, 24, Article 12476. https://doi.org/10.1186/s12885-024-12476-3
[10]. Laiq, T., et al. (2025). Prediction of axillary lymph node metastasis in breast cancer patients based on ultrasonographic-clinicopathologic features. Pakistan Journal of Medical Sciences, 41(1). https://doi.org/10.12669/pjms.41.1.10384
[11]. Li, C. (2013). Little's test of missing completely at random. The Stata Journal, 13(4), 795–809. https://doi.org/10.1177/1536867x1301300407
[12]. Molenberghs, G., & Verbeke, G. (2005). Multiple imputation and the expectation-maximization algorithm. In Models for Discrete Longitudinal Data (pp. 311–331). https://doi.org/10.1007/0-387-28980-1_28
[13]. Chandrashekar, G., & Sahin, F. (2014). A survey on feature selection methods. Computers & Electrical Engineering, 40(1), 16–28. https://doi.org/10.1016/j.compeleceng.2013.11.024
[14]. Grossi, E., & Buscema, M. (2007). Introduction to artificial neural networks. European Journal of Gastroenterology & Hepatology, 19(12), 1046–1054. https://doi.org/10.1097/MEG.0b013e3282f198a0
[15]. Evgeniou, T., & Pontil, M. (2001). Support vector machines: Theory and applications. In Machine learning and its applications: Advanced lectures (pp. 249–257). Springer-Verlag. https://doi.org/10.1007/3-540-44673-7_12
[16]. Guo, G., et al. (2003). KNN model-based approach in classification. In OTM Conferences/Workshops (pp. 986–996). https://doi.org/10.1007/978-3-540-39964-3_62
[17]. Ali, J., et al. (2012). Random forests and decision trees. [Conference paper].
[18]. Peng, J., Lee, K., & Ingersoll, G. (2002). An introduction to logistic regression analysis and reporting. The Journal of Educational Research. https://doi.org/10.1080/00220670209598786
[19]. Kwon, H., Park, J., & Lee, Y. (2019). Stacking ensemble technique for classifying breast cancer. Healthcare Informatics Research, 25(4), 283–289. https://doi.org/10.4258/hir.2019.25.4.283
[20]. Diessner, J., et al. (2023). Evaluation of different imaging modalities for axillary lymph node staging in breast cancer patients to provide a personalized and optimized therapy algorithm. Journal of Cancer Research and Clinical Oncology. https://doi.org/10.1007/s00432-022-04221-9
[21]. Huang, Z., et al. (2025). Enhancing axillary lymph node diagnosis in breast cancer with a novel photoacoustic imaging-based radiomics nomogram: A comparative study of peritumoral regions. Academic Radiology. https://doi.org/10.1016/j.acra.2024.10.018
[22]. Wu, T., et al. (2025). Axillary lymph node metastasis in breast cancer: From historical axillary surgery to updated advances in the preoperative diagnosis and axillary management. BMC Surgery, 25, Article 2802. https://doi.org/10.1186/s12893-025-02802-2