A Hybrid Fuzzy–Ensemble Machine Learning Framework for Non-Invasive Prediction of HER2 Status in Breast Cancer

Hassan Salarabadi , Dariush Salimi , Seyed Sahand Mohammadi Ziabari * , Mozaffar Aznab

Posted Date: 6 February 2026

doi: 10.20944/preprints202602.0556.v1

Keywords: machine learning; ensemble learning; fuzzy decision systems; HER2 status prediction; breast cancer

 doi:10.20944/preprints202602.0556.v1

Article

A Hybrid Fuzzy–Ensemble Machine Learning Framework for Non-Invasive Prediction of HER2 Status in Breast Cancer

Hassan Salarabadi [1], Dariush Salimi [1], Seyed Sahand Mohammadi Ziabari [2,3,*] and Mozaffar Aznab [4]

[1] Department of Computer Engineering, Faculty of Engineering, University of Zanjan, Zanjan, Iran
[2] Informatics Institute, University of Amsterdam, 1098 XH Amsterdam, The Netherlands
[3] Department of Computer Science and Technology, SUNY Empire State University, Saratoga Springs, NY 12866, USA
[4] Medical Department, Kermanshah University of Medical Sciences and Health Services, Kermanshah, Iran
* Correspondence: sahand.ziabari@sunyempire.edu

Abstract

HER2 status determination is a crucial task in breast cancer prognosis and treatment, yet traditional diagnostic methods such as immunohistochemistry (IHC) and fluorescence in situ hybridization (FISH) are invasive, time-consuming, and costly. Motivated by the need for scalable and data-driven predictive approaches, we propose a hybrid machine learning framework that integrates ensemble learning with fuzzy modeling for HER2 prediction using routinely available clinical and immunohistochemical data. A dataset comprising 624 breast cancer patients from Mahdieh Clinic (Kermanshah, Iran) was analyzed, with extensive feature engineering, scaling, and class balancing applied. We developed an ensemble framework based on tree-based learners (Random Forest, XGBoost, and LightGBM), combined through ensemble strategies and enhanced using fuzzy feature representations and decision threshold optimization. The proposed hybrid model achieved an accuracy of 0.816, an F1-score of 0.814, and an area under the ROC curve (AUC) of 0.862 on the held-out test set, demonstrating strong discriminative capability and balanced classification performance. This work highlights the potential of hybrid fuzzy–ensemble learning for uncertainty-aware predictive analytics in biomedical decision support, aligning with the journal's focus on information processes, intelligent systems, and data mining.

Keywords: machine learning; ensemble learning; fuzzy decision systems; HER2 status prediction; breast cancer

1. Introduction

Breast cancer is one of the most common malignancies affecting both women and, less frequently, men, and it is primarily caused by the uncontrolled growth of breast cells [1,2]. The human epidermal growth factor receptor 2 (HER2) is a transmembrane receptor protein that plays a critical role in regulating cell growth and division [3]. In approximately 15–20% of breast cancer cases, HER2 is overexpressed, leading to aggressive tumor behavior [4]. Accurate identification of HER2 status is therefore crucial for guiding targeted therapies, such as trastuzumab, which have significantly improved clinical outcomes [5]. In spite of its clinical importance, determining HER2 status typically relies on conventional diagnostic techniques, which face several limitations.

Despite its clinical importance, HER2 assessment is predominantly performed using conventional diagnostic techniques, namely immunohistochemistry (IHC) and fluorescence in situ hybridization (FISH) [1]. While these methods are widely accepted as clinical standards, they suffer from several limitations. Both IHC and FISH require specialized laboratory infrastructure, trained personnel, and high-quality reagents, and they are often time-consuming and costly [1,6]. In addition, pre-analytical

 doi:10.20944/preprints202602.0556.v1

factors such as tissue fixation, processing, and staining variability may adversely affect diagnostic accuracy, potentially leading to equivocal or false results [6]. These challenges motivate the exploration of alternative, data-driven approaches that can support or complement existing diagnostic workflows.

Machine learning (ML) techniques have recently emerged as promising alternatives for predicting HER2 status using clinical, histopathological, or genomic data [1,7]. Several studies have shown that ML models, including support vector machines, random forests, and deep learning techniques, can achieve competitive accuracy compared to conventional diagnostic methods [8]. However, most of these studies suffer from important limitations [9,10]. Many rely on relatively small datasets, which restrict their generalizability [11]. Others apply individual algorithms without leveraging ensemble strategies, thereby failing to capture the complementary strengths of different classifiers. Furthermore, relatively little attention has been devoted to feature engineering or uncertainty-aware modeling, both of which are critical for ensuring the reliability of ML-based systems in clinical decision-making.

To address these limitations, this study proposes a hybrid ensemble machine learning framework for HER2 status prediction based on routinely available clinical and immunohistochemical biomarkers. The proposed approach is evaluated on a dataset comprising 624 breast cancer patients collected from Mahdieh Clinic in Kermanshah, Iran. Our framework integrates multiple tree-based classifiers, namely Random Forest, XGBoost, and LightGBM, within ensemble strategies and incorporates fuzzy feature engineering and decision threshold optimization to enhance robustness in borderline and uncertain cases.

2. Related Work

Recent advances in machine learning (ML) and data-driven modeling have demonstrated significant potential in improving diagnostic and prognostic decision-making in Breast Cancer (BC). In particular, the integration of ML with radiomics, multi-omics, and clinical data has enabled the extraction of complex patterns that are difficult to identify using conventional statistical approaches. For instance, Chen et al. conducted a bicentric retrospective study employing ultrasound radiomics and ML techniques to predict pathological prognostic stages in a cohort of 578 BC patients [12]. While promising, the reliance on imaging-derived features and limited external validation may constrain the applicability of such models in routine clinical settings.

Complementary efforts have explored the use of multi-omics data for prognosis estimation. Song et al. proposed a prognostic framework for elderly BC patients by integrating mRNA, miRNA, lncRNA, copy number variations (CNVs), and single nucleotide variants (SNVs), highlighting the prognostic relevance of hypoxia-related pathways and immune microenvironment heterogeneity [13]. Although multi-omics models provide rich biological insights, their clinical deployment is often challenged by high costs, data complexity, and limited accessibility. Interpretability and robust feature selection have also gained attention in clinically oriented ML pipelines. For example, Ahmadian et al. proposed an explainable feature selection approach combining particle swarm optimisation with adaptive LASSO for MRI radiogenomics, demonstrating transferable signatures and improved generalisability in a two-center setting [14].

Beyond imaging and omics-based approaches, ML models have also been applied to predict treatment-related outcomes. Lin et al. developed an XGBoost-based model for predicting radiation dermatitis severity in breast cancer patients, incorporating clinical factors, patient-reported outcomes, and cytokine biomarkers [15]. Similarly, Miglietta et al. utilized ML techniques to predict HER2-low phenotype conversion in recurrent breast cancer, emphasizing the role of artificial intelligence in optimizing patient stratification and treatment accessibility [16]. Despite encouraging results, many of these studies are limited by relatively small sample sizes and the use of single-model architectures, which may restrict robustness and generalizability.

The application of ML in oncology extends beyond breast cancer and further underscores its prognostic utility. In bladder cancer, Zhang et al. developed a machine learning-based prognostic signature utilizing proteomics data to predict patient outcomes and treatment response [17]. In prostate

 doi:10.20944/preprints202602.0556.v1

cancer, Gao et al. constructed a programmed cell death-related gene signature using a random forest model, demonstrating that higher risk scores were associated with poorer survival outcomes and diminished immunotherapy benefits [18]. Similarly, Maimaitiyiming et al. proposed a mast cell gene signature that stratified prostate cancer patients into distinct immune-risk groups [19]. In gastric cancer, Liu et al. introduced a deep learning-based pathomics model that achieved high predictive performance for survival outcomes [20].

More recently, ML-based prognostic models have also been investigated in specific breast cancer subpopulations. Wu et al. identified senescence-related molecular subtypes in geriatric breast cancer with distinct prognostic significance [21]. In addition, Emily et al. compared Cox proportional hazards and survival random forest models for breast cancer survival prediction, reporting superior performance of the Cox model in their cohort [22].

Despite these notable advances, several limitations persist across the existing literature. Many studies rely on single-modality data sources, lack ensemble or uncertainty-aware decision mechanisms, or are evaluated on relatively small and population-specific datasets. Moreover, the integration of heterogeneous clinical and immunohistochemical features within robust ensemble frameworks remains underexplored, particularly for HER2 status prediction. Addressing these gaps, the present study proposes a hybrid fuzzy-enhanced ensemble approach that combines multiple tree-based classifiers with fuzzy feature engineering and decision calibration, aiming to improve predictive robustness while maintaining clinical feasibility.

3. Methods and Materials

To ensure reproducibility and methodological transparency, all experiments were implemented in Python using widely adopted machine learning libraries, including Scikit-learn, XGBoost, and LightGBM. Model development and evaluation were conducted in a Jupyter-based environment with fixed random seeds to guarantee consistent results. The proposed framework follows a hybrid pipeline that integrates clinical and immunohistochemical features with fuzzy modeling and ensemble learning. The workflow begins with data preprocessing and feature engineering, followed by fuzzy feature construction to capture uncertainty in key biomarkers. Tree-based classifiers—Random Forest, XGBoost, and LightGBM—are then trained and evaluated both as standalone models and within ensemble strategies. Finally, the model performance is evaluated using standard classification metrics to assess discriminative power and clinical reliability.

3.1. Data Description

This study utilized a retrospective clinical dataset consisting of 624 confirmed breast cancer cases collected from Mahdieh Clinic, Kermanshah, Iran. The dataset includes a combination of clinical and pathological variables routinely used in breast cancer assessment. Specifically, the available features comprise estrogen receptor (ER), progesterone receptor (PR), p53 status, perineural invasion, vascular invasion, metastasis, lymph node involvement, number of involved lymph nodes, tumor size, and the Ki67 proliferation index.

The HER2 status (positive or negative) was defined as the target variable, framing the problem as a binary classification task. All cases were extracted from clinical records and validated by domain experts to ensure data consistency and reliability prior to model development.

3.2. Preprocessing Steps

Several preprocessing steps were applied to enhance data quality and ensure robust model performance. First, missing values in numerical features were handled using mean imputation, which was selected due to the relatively low proportion of missing entries and the approximately symmetric distribution of the variables.

 doi:10.20944/preprints202602.0556.v1

Subsequently, all continuous features were normalized using Min–Max scaling to ensure that variables with different measurement ranges contributed equally to the learning process. The normalization was performed according to:

$$X_{\text{scaled}} = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \tag{1}$$

To address the inherent class imbalance in HER2 status, a Random Over-Sampling strategy was applied to the training set. This approach increases the representation of the minority class by randomly duplicating its samples, thereby mitigating model bias toward the majority class. The class balance can be expressed as:

$$\text{Ratio} = \frac{N_{\text{minority}}}{N_{\text{majority}}} \tag{2}$$

By applying oversampling exclusively to the training data, data leakage was avoided, and the integrity of the test set was preserved for unbiased performance evaluation.

3.3. Fuzzy Feature Engineering

Clinical biomarkers in breast cancer often exhibit uncertainty, imprecision, and gradual transitions rather than crisp boundaries. Conventional machine learning models typically rely on sharp thresholds, which may fail to adequately represent the inherent vagueness of biological processes. To address this limitation, fuzzy feature engineering was employed to transform selected clinical variables into interpretable fuzzy representations.

In this study, fuzzy membership functions were constructed for clinically significant variables, including the Ki67 proliferation index, tumor size, and estrogen receptor (ER) status. These variables were selected based on their established relevance to tumor aggressiveness and HER2 expression. Each variable was mapped into multiple fuzzy sets using triangular membership functions, enabling the model to capture nonlinear relationships and implicit interactions among biomarkers.

For the Ki67 index, three fuzzy sets—low, medium, and high—were defined to represent varying levels of cellular proliferation. Similarly, tumor size was categorized into small, medium, and large fuzzy sets, reflecting clinically meaningful tumor growth stages. ER status was represented using binary fuzzy memberships corresponding to positive and negative expression. The triangular membership function used for fuzzification is defined as:

$$\mu(x) = \begin{cases} 0, & x \le a \\ \dfrac{x-a}{b-a}, & a < x \le b \\ \dfrac{c-x}{c-b}, & b < x < c \\ 0, & x \ge c \end{cases} \tag{3}$$

where a, b, and c denote the lower bound, peak point, and upper bound of the fuzzy set, respectively. The parameters of the membership functions were defined based on clinical knowledge and empirical data distribution.

The resulting fuzzy membership degrees were incorporated as additional input features and concatenated with the original numerical and categorical variables. This approach allowed the model to encode uncertainty-aware representations while preserving the original feature space.

3.4. Proposed Hybrid Ensemble Model Architecture

This study proposes a hybrid ensemble learning framework that integrates tree-based machine learning models with fuzzy modeling to predict HER2 status in breast cancer patients. The proposed architecture is designed to capture nonlinear relationships among biomarkers, exploit model diversity,

 doi:10.20944/preprints202602.0556.v1

and manage uncertainty inherent in clinical decision-making. An overview of the proposed model is illustrated in Figure 1.

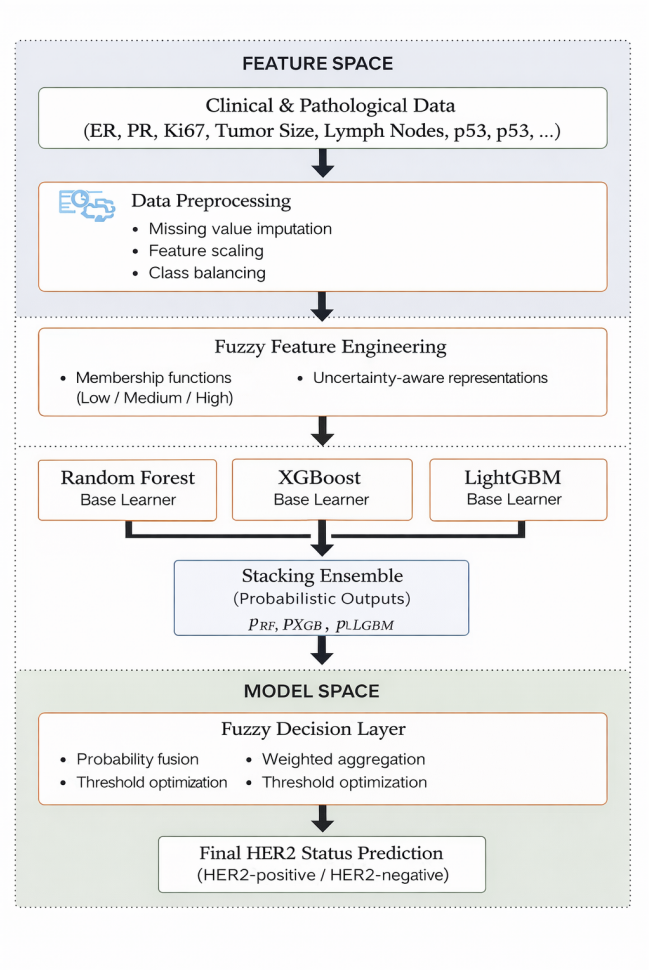

Figure 1. Overview of the proposed hybrid fuzzy ensemble framework for HER2 status prediction.

3.5. Model Training, Validation, and Evaluation Metrics

The proposed framework was trained and evaluated using a structured and reproducible experimental protocol. The dataset was randomly partitioned into training and testing subsets using an 80/20 split with stratified sampling to preserve the original class distribution. All preprocessing steps, including feature scaling, oversampling, and fuzzy feature generation, were applied exclusively to the training set to prevent data leakage.

To enhance decision robustness, threshold optimization was applied to the ensemble probability scores by maximizing the F1-score on the validation data.

Model performance was assessed using accuracy, precision, recall, F1-score, and the area under the receiver operating characteristic curve (AUC):

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{4}$$

$$\text{Precision} = \frac{TP}{TP + FP} \tag{5}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{6}$$

$$\text{F1-score} = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \tag{7}$$

 doi:10.20944/preprints202602.0556.v1

The AUC was computed as:

$$\mathrm{AUC} = \int_0^1 \mathrm{TPR}(\mathrm{FPR})\, d(\mathrm{FPR}) \tag{8}$$

3.6. Experimental Setup

3.6.1. Dataset Overview

All experiments were conducted using Python-based machine learning libraries, including Scikit-learn, XGBoost, and LightGBM, within a Kaggle computational environment. The dataset was partitioned into training (80%) and testing (20%) subsets using stratified sampling to preserve the original class distribution. To prevent information leakage, all preprocessing steps including feature scaling, fuzzy feature construction, and class balancing were applied exclusively to the training data and subsequently transferred to the test set. Model performance was evaluated on the unseen test set using multiple complementary metrics, namely accuracy, F1-score, and the area under the receiver operating characteristic curve (AUC), providing a balanced assessment of both classification correctness and discriminative capability.

3.7. Role of Base Learners in the Ensemble Framework

In the first stage of this study, each base learner was independently employed to predict HER2 status using the same training and testing splits, as well as identical preprocessing procedures. This experiment was conducted to establish baseline predictive performance and to enable a fair comparison with the proposed ensemble framework. Random Forest, XGBoost, and LightGBM models were trained separately on the processed dataset, and their predictive performance was evaluated on the held-out test set. Model performance was assessed using accuracy, F1-score, and area under the ROC curve (AUC), which are commonly adopted metrics in binary medical classification tasks.

Table 1 summarizes the quantitative results obtained by the individual base learners. As shown, all three models achieved competitive performance, with Random Forest yielding the highest AUC, while XGBoost and LightGBM demonstrated comparable accuracy and F1-score values.

Table 1. Quantitative performance metrics of the individual base learners evaluated for HER2 status prediction on the test dataset.

Model	Accuracy	F1-score	AUC
Random Forest	0.816	0.812	0.873
XGBoost	0.800	0.800	0.856
LightGBM	0.784	0.783	0.835

These results indicate that each base learner is capable of capturing relevant patterns in the clinical and pathological data, while exhibiting different strengths in terms of discrimination and classification balance. These complementary behaviors motivate their integration within an ensemble framework, which is investigated in the subsequent subsection.

3.8. Performance Evaluation of the Proposed Model

The predictive performance of the proposed hybrid fuzzy stacking ensemble model was evaluated on the test set using accuracy, F1-score, and area under the ROC curve (AUC). The model achieved an accuracy of 0.816, an F1-score of 0.814, and an AUC of 0.862, indicating strong discriminative capability in distinguishing HER2-positive from HER2-negative cases.

The ROC curves of the evaluated models are shown in Figure 2. The curves demonstrate the discriminative performance of the base learners and ensemble strategies across thresholds. Overall, ensemble-based approaches achieve competitive and more stable performance relative to individual classifiers.

 doi:10.20944/preprints202602.0556.v1

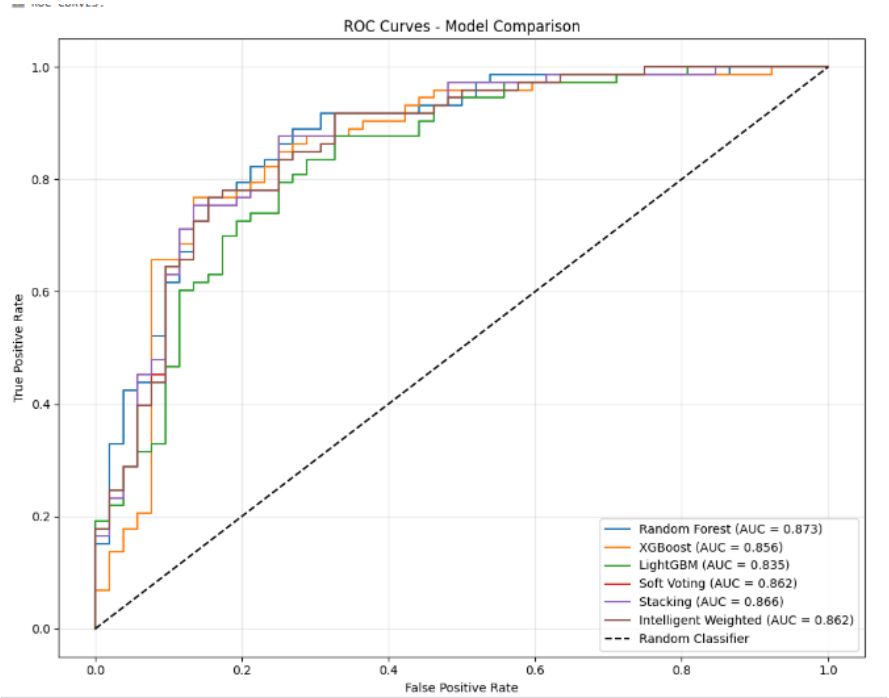

Figure 2. Receiver operating characteristic (ROC) curves of individual base learners and ensemble models for HER2 status prediction. The figure compares Random Forest, XGBoost, LightGBM, and ensemble strategies, including soft voting, hard voting, stacking, and the proposed intelligent weighted ensemble.

To further investigate the effectiveness of ensemble learning, the predictive performance of individual base learners was compared with various ensemble strategies, including hard voting, soft voting, stacking, and an intelligent weighted ensemble. Figure 3 summarizes the performance metrics obtained for all evaluated models. Among individual classifiers, Random Forest achieved the highest discriminative performance (AUC = 0.873). Ensemble-based approaches exhibited competitive performance, with the intelligent weighted ensemble achieving accuracy and F1-score comparable to the best-performing individual model while maintaining robust AUC values. These findings suggest that ensemble strategies can improve stability and balance by integrating complementary decision patterns from multiple classifiers.

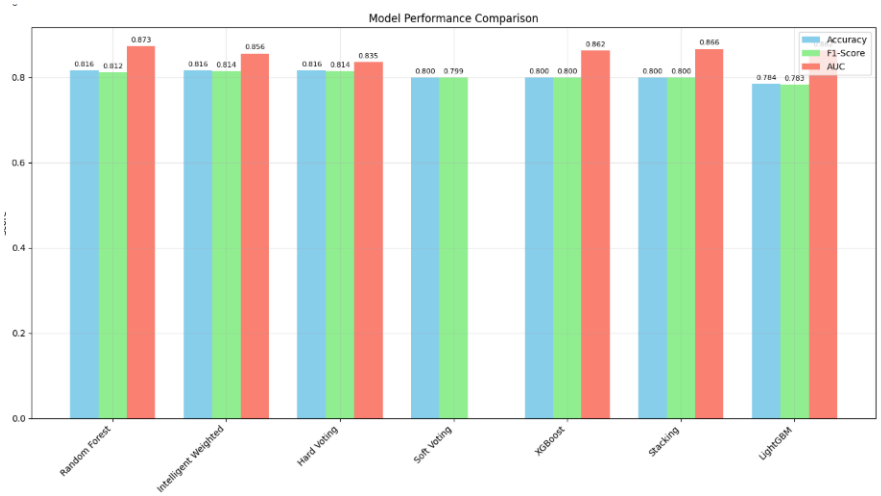

Figure 3. Summary of classification performance metrics for all evaluated base learners and ensemble approaches. The results illustrate the comparative strengths of different modeling strategies and the impact of ensemble integration on predictive accuracy, class balance, and discriminative capability.

 doi:10.20944/preprints202602.0556.v1

3.9. Effect of Threshold Optimization

To further improve classification reliability, threshold optimization was applied to the ensemble probability outputs. Instead of adopting a fixed threshold of 0.5, an optimal threshold was selected by maximizing the F1-score on the validation data. The impact of this optimization on classification accuracy is visualized in Figure 4. As illustrated, threshold tuning leads to a noticeable improvement in balanced performance, particularly by reducing false negative predictions, which is crucial in clinical decision-making contexts.

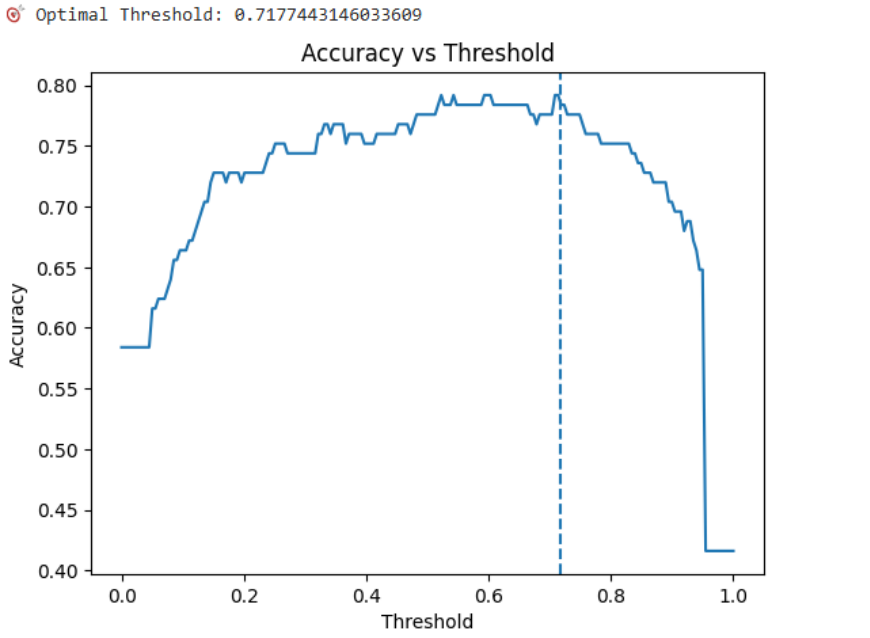

Figure 4. Variation of classification accuracy as a function of decision threshold for the proposed ensemble model. The optimal threshold maximizing performance is highlighted, demonstrating the importance of threshold tuning in clinical decision-making.

The confusion matrix of the proposed model at the optimized threshold is presented in Figure 5. The matrix demonstrates a balanced distribution of true positive and true negative predictions, confirming the effectiveness of the fuzzy stacking framework in minimizing misclassification errors. Notably, the model achieves a high true positive rate for HER2-positive cases, which is clinically significant, as misclassifying HER2-positive patients may delay access to targeted therapies.

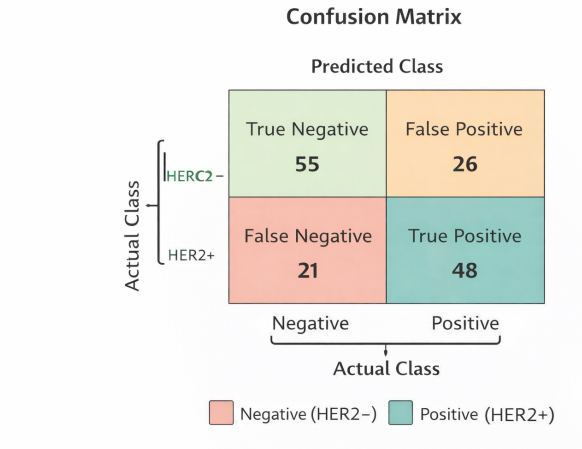

Figure 5. Confusion matrix of the proposed fuzzy ensemble model on the test dataset. The matrix summarizes the classification outcomes, showing true positives (48) and true negatives (55), while maintaining relatively low false positives (26) and false negatives (21).

 doi:10.20944/preprints202602.0556.v1

Overall, the experimental results demonstrate that the proposed hybrid fuzzy stacking ensemble model achieves strong and balanced predictive performance. The integration of fuzzy feature engineering and ensemble learning contributes to improved discrimination stability and decision reliability, supporting the potential of the proposed framework as a decision-support tool for HER2 status prediction.

4. Discussion

Accurate and reliable prediction of HER2 status is a cornerstone of personalized breast cancer management, as it directly influences treatment selection and clinical outcomes. In this study, we investigated a hybrid learning framework that integrates fuzzy feature engineering with multiple ensemble strategies to address key challenges in clinical data, including heterogeneity, uncertainty, and moderate class imbalance.

The experimental results reveal several important insights. First, among individual classifiers, Random Forest demonstrated the strongest overall discriminative performance, achieving an accuracy of 0.816 and an AUC of 0.873, outperforming both XGBoost and LightGBM. This finding suggests that tree-based bagging methods may be particularly well-suited for modeling nonlinear interactions among clinical and pathological biomarkers in HER2 prediction tasks. Second, ensemble strategies did not uniformly outperform the best-performing base learner across all evaluation metrics. While the intelligent weighted ensemble and hard voting approaches achieved accuracy and F1-scores comparable to Random Forest, their AUC values were slightly lower. This observation highlights an important methodological point: ensemble integration does not inherently guarantee superior discrimination, particularly when base learners exhibit correlated decision boundaries or similar error patterns. In such cases, ensemble models may primarily improve prediction stability rather than maximizing separability between classes.

Notably, the stacking ensemble achieved a competitive AUC (0.866), indicating improved ranking capability compared to individual boosting models, even though its overall accuracy remained comparable. This suggests that the stacking mechanism effectively aggregates complementary probabilistic information from base learners, enhancing robustness across decision thresholds rather than optimizing a single operating point.

The role of fuzzy feature engineering is particularly relevant in this context. By transforming continuous biomarkers such as tumor size and Ki67 into fuzzy representations, the model captures gradual transitions between clinical risk states, thereby reducing information loss associated with rigid thresholds. This fuzzy modeling contributes to more stable probability estimates, which is reflected in the relatively consistent AUC values observed across ensemble variants. Importantly, this effect becomes more apparent when threshold optimization is applied, underscoring the interaction between fuzzy representations and decision calibration.

Threshold optimization further enhanced clinical relevance by improving the balance between sensitivity and specificity. Given that false negative predictions in HER2-positive patients may delay access to targeted therapies, prioritizing recall without severely sacrificing precision is critical. The optimized threshold reduced false negatives, as confirmed by the confusion matrix analysis, demonstrating that performance improvements are not solely numerical but clinically meaningful.

From a clinical decision-support perspective, these findings emphasize that model selection should be guided by the intended use case. While Random Forest achieved the highest AUC, ensemble approaches offered comparable accuracy with improved robustness and interpretability through aggregation. Therefore, the proposed fuzzy-enhanced ensemble framework should be viewed not as a replacement for strong base learners, but as a complementary strategy that enhances reliability under varying clinical conditions.

Despite these promising results, several limitations should be acknowledged. The dataset was obtained from a single clinical center, which may limit generalizability. Additionally, the absence of external validation restricts conclusions regarding real-world deployment. Future work should focus

 doi:10.20944/preprints202602.0556.v1

on multi-center validation, incorporation of additional data modalities such as imaging or genomic profiles, and integration of explainability techniques (e.g., SHAP or rule-based fuzzy inference) to strengthen clinician trust.

Table 2. Comparative performance analysis of individual classifiers and ensemble strategies, highlighting their effectiveness and suitability for clinical decision-support in HER2 status prediction.

Model	Accuracy	F1-score	AUC	Key Strength	Clinical Implication
Random Forest	0.816	0.812	0.873	Strong discrimination	Reliable baseline model
XGBoost	0.800	0.800	0.856	Nonlinear boosting	Sensitive to hyperparameters
LightGBM	0.784	0.783	0.835	Computational efficiency	Moderate discrimination
Hard Voting	0.816	0.814	–	Stability	Threshold-independent
Soft Voting	0.800	0.799	0.862	Probabilistic fusion	Improved calibration
Stacking	0.800	0.800	0.866	Robust ranking	Better threshold robustness
Intelligent Weighted Ensemble	0.816	0.814	0.862	Balanced aggregation	Clinically robust decisions

Overall, this study demonstrates that fuzzy-enhanced ensemble learning provides a flexible and clinically meaningful framework for HER2 status prediction. Rather than maximizing a single performance metric, the proposed approach emphasizes robustness, decision reliability, and stable discrimination—key factors for practical adoption in precision oncology.

5. Conclusions

In this study, a hybrid learning framework combining fuzzy feature engineering with ensemble-based classification was proposed for predicting HER2 status in breast cancer patients using routinely available clinical and immunohistochemical data. The primary objective was to develop a robust and clinically meaningful decision-support model capable of handling uncertainty, nonlinear feature interactions, and moderate class imbalance inherent in real-world medical datasets.

Comprehensive experimental evaluations demonstrated that tree-based learning models, particularly Random Forest, achieved strong discriminative performance, with an AUC of 0.873 and an accuracy of 81.6%. Ensemble strategies, including stacking and weighted voting, yielded comparable accuracy and F1-scores, while exhibiting slightly lower but stable AUC values. These findings indicate that, rather than maximizing a single performance metric, the proposed fuzzy-enhanced ensemble framework improves prediction robustness and decision consistency across varying thresholds.

The incorporation of fuzzy feature engineering played a key role in stabilizing probabilistic outputs by modeling gradual transitions in continuous biomarkers such as tumor size and Ki67. This representation reduced information loss caused by hard discretization and contributed to improved calibration and balanced precision–recall behavior. Furthermore, threshold optimization significantly enhanced clinical relevance by reducing false negative predictions for HER2-positive cases, which is critical for timely access to targeted therapies.

Unlike conventional HER2 assessment methods that rely on invasive procedures or costly molecular assays, the proposed framework offers a cost-effective and non-invasive alternative based solely on routinely collected clinical and pathological features. As such, it has the potential to function as a complementary decision-support tool, assisting clinicians in risk stratification and treatment planning rather than replacing standard diagnostic protocols. Nevertheless, several limitations must be acknowledged. The dataset used in this study was derived from a single clinical center, which may limit generalizability. Additionally, the absence of external validation restricts direct clinical deployment.

 doi:10.20944/preprints202602.0556.v1

Future work will focus on validating the proposed framework on multi-center cohorts, incorporating additional data modalities such as imaging radiomics or genomic profiles, and enhancing interpretability through explainable artificial intelligence techniques, including SHAP or fuzzy rule-based explanations. Future work will focus on validating the proposed framework on multi-center cohorts, extending it to additional data modalities (e.g., radiomics and whole-slide pathology) while preserving clinical feasibility, and enhancing interpretability through explainable AI techniques (e.g., SHAP and rule-based fuzzy explanations) [23].

In summary, this study demonstrates that fuzzy-enhanced ensemble learning provides a flexible, reliable, and clinically relevant approach for HER2 status prediction. By prioritizing robustness and decision reliability over isolated performance gains, the proposed framework represents a meaningful step toward intelligent decision-support systems in personalized breast cancer management.

Funding: No external funding was received.

Acknowledgments: The authors gratefully acknowledge the Department of Computer Science and Technology at SUNY Empire State University for providing institutional support.

Institutional Review Board Statement: Not applicable. The work relied exclusively on de-identified, publicly available data and did not involve interaction with human participants or animals.

Informed Consent Statement: This study used retrospective clinical records. All data were de-identified prior to analysis and handled in accordance with relevant institutional guidelines.

Data Availability Statement: The dataset used in this study contains patient-level clinical and immunohistochemical records collected from Mahdieh Clinic (Kermanshah, Iran). Due to privacy and ethical restrictions, the data are not publicly available. De-identified data may be provided by the corresponding author upon reasonable request and subject to institutional approval.

References

1. Tanha, J.; Salarabadi, H.; Aznab, M.; Farahi, A.; Zoberi, M. Relationship among prognostic indices of breast cancer using classification techniques. *Informatics in Medicine Unlocked* **2020**, *18*, 100265.
2. Schreurs, M.A.C.; Schmidt, M.K.; Hollestelle, A.; Schaapveld, M.; van Asperen, C.J.; Ausems, M.G.E.M.; van de Beek, I.; Broekema, M.F.; Collée, J.M.; van der Hout, A.H.; et al. Cancer risks for other sites in addition to breast in CHEK2 c.1100delC families. *Genetics in Medicine* **2024**, *26*, 101171.
3. Di Grazia, G.; Conti, C.; Nucera, S.; Stella, S.; Massimino, M.; Giuliano, M.; Schettini, F.; Vigneri, P.; Martorana, F. Bridging the Gap: the role of MDM2 inhibition in overcoming treatment resistance in breast cancer. *Critical Reviews in Oncology/Hematology* **2025**, p. 104834.
4. Hlauschek, D.; Fesl, C.; Gnant, M. The impact of analysis methodology details on invasive breast cancer-free survival in randomized clinical trials. *ESMO Open* **2025**, *10*, 105324.
5. Muñoz, J.; Nucera, S.; Garcia, N.R.; Cebrecos, I.; Oses, G.; Ganau, S.; Sanfeliu, E.; Jares, P.; Marín-Aguilera, M.; Galván, P.; et al. Paclitaxel-related type I Kounis Syndrome in a very young patient with HER2-positive breast cancer and the role of genomics to disentangle a complex therapeutic scenario: a case report and narrative review. *The Breast* **2025**, p. 104465.
6. Jin, Y.; Xue, J.; Li, X.; Zhong, X. Downregulation of PIP4K2C inhibits the breast cancer cell proliferation, migration and invasion. *Translational Oncology* **2025**, *57*, 102420.
7. Borugadda, P.; Kalluri, H.K. A comprehensive analysis of artificial intelligence, machine learning, deep learning and computer vision in food science. *Journal of Future Foods* **2025**.
8. Salarabadi, H.; Iraji, M.S.; Salimi, M.; Zoberi, M. Improved COVID-19 diagnosis using a hybrid transfer learning model with fuzzy edge detection on CT scan images. *Advances in Fuzzy Systems* **2024**, *2024*, 3249929.
9. Rahman, M.A.; Bhuiyan, T.; Ali, M.A. Enhancing aviation safety: Machine learning for real-time ADS-B injection detection through advanced data analysis. *Alexandria Engineering Journal* **2025**, *126*, 262–276.
10. Zhao, Y.; Zhang, W.; Liu, X. Grid search with a weighted error function: Hyper-parameter optimization for financial time series forecasting. *Applied Soft Computing* **2024**, *154*, 111362.
11. Wang, H.; Zhang, J.; Li, Y.; Wang, D.; Zhang, T.; Yang, F.; Zhang, Y.; Yang, L.; Li, P. Deep-learning features based on F18 fluorodeoxyglucose positron emission tomography/computed tomography (18F-FDG PET/CT) to predict preoperative colorectal cancer lymph node metastasis. *Clinical Radiology* **2024**, *79*, e1152–e1158.

 doi:10.20944/preprints202602.0556.v1

12. Chen, L.; Su, Y.; Wang, Y.; Yu, Y.; Huang, D.; Zhang, M.; Chen, X.; Ye, X.; He, Y.; Xue, E.; et al. Ultrasound radiomics-based machine learning and SHapley Additive exPlanations method predicting pathological prognostic stage in breast cancer: a bicentric and validation study. *Clinical Breast Cancer* **2025**.
13. Song, Y.; Wang, C.; Zhou, Y.; Sun, Q.; Lin, Y. A multi-omics-based prognostic model for elderly breast cancer by machine learning: insights from hypoxia and immunity of tumor microenvironment. *Clinical Breast Cancer* **2025**.
14. Ahmadian, M.; Bodalal, Z.; Adib, M.; Mohammadi Ziabari, S.S.; Bos, P.; Martens, R.M.; Agrotis, G.; Vens, C.; Karssemakers, L.; Al-Mamgani, A.; et al. Explainable feature selection combining particle swarm optimisation with adaptive LASSO for MRI radiogenomics: Predicting HPV status in oropharyngeal cancer. *Computer Methods and Programs in Biomedicine* **2026**, *275*, 109204. https://doi.org/10.1016/j.cmpb.2025.109204.
15. Lin, N.; Abbas-Aghababazadeh, F.; Su, J.; Wu, A.J.; Lin, C.; Shi, W.; Xu, W.; Haibe-Kains, B.; Liu, F.F.; Kwan, J.Y.Y. Development of machine learning models for predicting radiation dermatitis in breast cancer patients using clinical risk factors, patient-reported outcomes, and serum cytokine biomarkers. *Clinical Breast Cancer* **2025**.
16. Miglietta, F.; Collesei, A.; Vernieri, C.; Giarratano, T.; Giorgi, C.A.; Girardi, F.; Griguolo, G.; Cacciatore, M.; Botticelli, A.; Vingiani, A.; et al. Development of two machine learning models to predict conversion from primary HER2-0 breast cancer to HER2-low metastases: a proof-of-concept study. *ESMO Open* **2025**, *10*, 104087.
17. Zhang, X.; Li, P.; Ji, L.; Zhang, Y.; Zhang, Z.; Guo, Y.; Zhang, L.; Jing, S.; Dong, Z.; Tian, J.; et al. A machine learning-based prognostic signature utilizing MSC proteomics for predicting bladder cancer prognosis and treatment response. *Translational Oncology* **2025**, *54*, 102349.
18. Gao, F.; Huang, Y.; Yang, M.; He, L.; Yu, Q.; Cai, Y.; Shen, J.; Lu, B. Machine learning-based cell death marker for predicting prognosis and identifying tumor immune microenvironment in prostate cancer. *Heliyon* **2024**, *10*.
19. Maimaitiyiming, A.; An, H.; Xing, C.; Li, X.; Li, Z.; Bai, J.; Luo, C.; Zhuo, T.; Huang, X.; Maimaiti, A.; et al. Machine learning-driven mast cell gene signatures for prognostic and therapeutic prediction in prostate cancer. *Heliyon* **2024**, *10*.
20. Liu, L.; Zhang, Y.; Zhao, X.h.; Zhu, M.; Liang, J.w.; Jiang, Z.y.; Zhang, J.q.; Ma, J.l.; Wang, Z.Z.; Gan, M.f.; et al. Predicting clinical prognosis in gastric cancer using deep learning-based analysis of tissue pathomics images. *Computer Methods and Programs in Biomedicine* **2025**, p. 108895.
21. Wu, X.; Chen, M.; Liu, K.; Wu, Y.; Feng, Y.; Fu, S.; Xu, H.; Zhao, Y.; Lin, F.; Lin, L.; et al. Molecular classification of geriatric breast cancer displays distinct senescent subgroups of prognostic significance. *Molecular Therapy Nucleic Acids* **2024**, *35*.
22. Emily, M.; Meidioktaviana, F.; Nabiilah, G.Z.; Moniaga, J.V. Comparative analysis of machine learning and survival analysis for breast cancer prediction. *Procedia Computer Science* **2024**, *245*, 759–767.
23. Lin, B.Y.; Mohammadi Ziabari, S.S.; Nasser Al Husaini, Y.; Mansoor Alsahag, A.M. SG-MuRCL: Smoothed Graph-Enhanced Multi-Instance Contrastive Learning for Robust Whole-Slide Image Classification. *Information* **2026**, *17*, 37. https://doi.org/10.3390/info17010037.
24. Göransson, S.; Olofsson, H.; Johansson, H.J.; Yan, F.; Vogiatzakis, C.; Liang, S.; Bellato, H.M.; Masvidal, L.; Aksoylu, I.; Hartman, J.; et al. Mechanical control of breast cancer malignancy by promotion of mevalonate pathway enzyme synthesis. *Matrix Biology* **2025**.
25. Zhang, J.; Liu, R.; Gao, S.; Chen, W.; Han, X.; Wang, Z.; Zhou, H.; Wang, Y.; Chen, J.; Ma, Y.; et al. A phase I dose-escalation and dose-expansion study of tibremciclib, a novel CDK4/6 inhibitor, monotherapy and in combination with fulvestrant in HR-positive/HER2-negative advanced breast cancer. *ESMO Open* **2025**, *10*, 105121.
26. Tsutsumi, C.; Ohuchida, K.; Tsutsumi, H.; Shimada, Y.; Yamada, Y.; Son, K.; Hayashida, S.; Katayama, N.; Mochida, Y.; Iwamoto, C.; et al. TIM3 on natural killer cells regulates antibody-dependent cellular cytotoxicity in HER2-positive gastric cancer. *Cancer Letters* **2025**, *611*, 217412.
27. Molinelli, C.; Bruzzone, M.; Blondeaux, E.; Ruelle, T.; Lanzavecchia, C.; De Laurentiis, M.; Russo, S.; Riccardi, F.; Sini, V.; Cognetti, F.; et al. The journey of patients affected by metastatic hormone receptor-positive/HER2-negative breast cancer from CDK4/6 inhibitors to second-line treatment: a real-world analysis of 701 patients enrolled in the GIM14/BIOMETA study. *European Journal of Cancer* **2024**, *213*, 115113.

 doi:10.20944/preprints202602.0556.v1

28. Zhang, Q.; Guo, D.; Xu, G.; Xie, R.; Deng, Y.; Fu, P.; Wan, J. Tumor-associated macrophages suppress estrogen receptor-β expression in triple-negative breast cancer through the PI3K/AKT pathway. *Experimental and Therapeutic Medicine* **2026**, *31*, 1–10.
29. Ye, G.; Zhang, C.; Zhuang, Y.; Liu, H.; Song, E.; Li, K.; Liao, Y. An advanced nomogram model using deep learning radiomics and clinical data for predicting occult lymph node metastasis in lung adenocarcinoma. *Translational Oncology* **2024**, *44*, 101922.
30. Bressane, A.; Toda, D.H.R.; Negri, R.G.; Formiga, J.K.S.; Bankole, A.O.; Bankole, A.R.; Sharifi, S.; Moruzzi, R. Physics-informed feature engineering with fuzzy symbolic regression for predicting settling velocity in water treatment. *Journal of Water Process Engineering* **2025**, *78*, 108749.
31. Siqueira, V.S.d.M.; Cuadros, M.A.S.L.; Munaro, C.J.; de Almeida, G.M. Expert system for early sign stuck pipe detection: Feature engineering and fuzzy logic approach. *Engineering Applications of Artificial Intelligence* **2024**, *127*, 107229.
32. Liu, J.; Wu, J.; Pan, J.; Zhao, P. Fuzzification-back propagation neural network-based model prediction for robotic arm positioning error reduction. *Engineering Applications of Artificial Intelligence* **2025**, *151*, 110639.
33. Wang, J.; Feng, X.; Yu, Y.; Wang, X.; Werghi, N.; Han, X.; Zhou, H.; Shi, K.; Zhong, S.; Cai, J. Fuzzy actor–critic learning-based interpretable control and stability-informed guarantee with error mapping for discrete-time nonlinear system. *Chaos, Solitons & Fractals* **2025**, *199*, 116878.
34. Li, P.; Cheng, G.; Zhao, Y.; Liu, L.; Chen, M.; Ma, Z.; Chen, X. An interpretable reasoning schema for fuzzy knowledge graph under interval type-2 fuzzy model. *Applied Soft Computing* **2025**, p. 113534.
35. Ni, C.; Huang, H.; Cui, P.; Ke, Q.; Tan, S.; Ooi, K.T.; Liu, Z. Light Gradient Boosting Machine (LightGBM) to forecasting data and assisting the defrosting strategy design of refrigerators. *International Journal of Refrigeration* **2024**, *160*, 182–196.
36. Makwana, S.; Vaghela, D.; Chan, C.K.; Naik, N. Novel hybrid approach of random forest and stacking ensemble to improve fresh water yield prediction in mobile wick solar still. *Desalination and Water Treatment* **2025**, p. 101490.
37. Koohmishi, M.; Connolly, D.P. Railway ballast fouling detection using thermal imaging: Integration of LSTM and XGBoost. *Transportation Geotechnics* **2025**, p. 101889.
38. Gao, J.; Zhang, J.; Gong, Y.; Yang, K.; Quan, W.; Li, L.; Wu, Y.; Wang, R.; Cheng, H. A robust soil quality index for mountain ecosystems driven by XGBoost and network analysis. *Geoderma* **2025**, *463*, 117558.
39. Hendiani, S.; Walther, G. Double-layer multi-criteria group decision-making approach using neutralized possibility degree-based decision matrix with fuzzy information. *Engineering Applications of Artificial Intelligence* **2024**, *133*, 108276.