JSiI | Jurnal Sistem Informasi
Vol. 13 No. 1 | March 2026, P. 91-97

p-ISSN: 2406-7768
e-ISSN: 2581-2181

IMPLEMENTATION OF XGBOOST AND SUPPORT VECTOR MACHINE FOR BREAST CANCER PREDICTION USING BAYESIAN OPTIMIZATION

Jupron [1], Fajar Agung Nugroho [2]

[1,2] *Department of Informatics Engineering, Faculty of Computer Science, Pamulang University, South Tangerang 15310, Indonesia*

[1]dosen02664@unpam.ac.id
[2]fajar86.unpam@gmail.com

Abstract

Breast cancer remains one of the most prevalent causes of cancer-related mortality among women worldwide, making early and accurate detection critically important. Machine learning techniques have been widely applied for this purpose; however, many existing studies primarily focus on predictive accuracy without providing comprehensive analysis of model optimization and interpretability. This study proposes a comparative framework integrating Support Vector Machine (SVM) and Extreme Gradient Boosting (XGBoost) with Bayesian Optimization to enhance hyperparameter tuning and model performance. The Breast Cancer Wisconsin Dataset, consisting of 569 samples with 30 numerical features, is used for evaluation. The proposed approach includes data preprocessing, dataset splitting, systematic hyperparameter optimization, model training, and performance evaluation. Experimental results show that the XGBoost model achieves superior performance compared to SVM, with an accuracy of 98.24% and an Area Under the Curve (AUC) of 0.994. Further analysis indicates that the model maintains a strong balance between precision and recall, with minimal misclassification. In addition, feature importance analysis reveals that attributes related to tumor size and structural irregularities contribute significantly to the prediction results, supporting the interpretability of the model in a medical context. The main contribution of this study lies in providing a more comprehensive evaluation that combines performance comparison, optimization effectiveness, and feature-level interpretation within a unified framework. The findings demonstrate that the integration of XGBoost and Bayesian Optimization offers a reliable and interpretable approach for breast cancer classification, with strong potential for implementation in machine learning–based clinical decision support systems.

Keywords*: breast cancer, machine learning, XGBoost, Support Vector Machine, Bayesian Optimization.*

I. INTRODUCTION

Breast cancer remains one of the most frequently diagnosed cancers among women worldwide and continues to be a major contributor to cancer-related mortality. Early detection plays a critical role in improving treatment success rates and increasing patient survival. However, the diagnostic process for breast cancer often involves complex clinical assessments and requires expert interpretation, which can delay decision-making. In this context, the integration of machine learning techniques has gained significant attention as a tool to support medical practitioners in achieving faster and more accurate diagnoses [1].

The advancement of machine learning has provided new possibilities in healthcare, particularly in analyzing medical datasets and assisting in predictive modeling. Various classification algorithms have been widely explored to address complex medical problems. Among these, Support Vector Machine (SVM) and Extreme Gradient Boosting (XGBoost) are commonly used due to their strong classification performance. SVM operates by constructing an optimal hyperplane that maximizes class separation, while XGBoost is an ensemble learning approach that improves prediction accuracy through gradient boosting of decision trees in an iterative manner [2].

Recent international studies have demonstrated that advanced machine learning techniques such as XGBoost and Support Vector Machine are highly effective in medical diagnosis, particularly for cancer prediction [3]. Several works report that XGBoost achieves superior performance due to its ensemble learning capability and robustness against overfitting [4], while Bayesian Optimization significantly improves hyperparameter tuning efficiency [5]. However, most of these studies primarily focus on predictive accuracy and lack comprehensive evaluation involving optimization behavior and feature interpretability.

Despite their effectiveness, previous studies have extensively applied SVM, XGBoost, and Bayesian

Received: Jan 13th, 2026 | Published: March 30th, 2026

p-ISSN: 2406-7768
e-ISSN: 2581-2181

Optimization individually or in combination for breast cancer classification. However, many of these studies focus primarily on achieving high accuracy without providing a comprehensive comparison of model behavior, optimization efficiency, and feature interpretability within a unified experimental framework. In addition, limited attention has been given to analyzing how optimized hyperparameters influence model performance and how important features relate to clinical characteristics. This indicates a research gap in providing a more integrated evaluation that combines performance comparison, optimization analysis, and feature-level interpretation.

The performance of machine learning models is highly dependent on the selection of appropriate hyperparameters. Manual tuning is often time-consuming and may lead to suboptimal configurations. Therefore, a systematic optimization approach is necessary to efficiently identify optimal parameter settings. Bayesian Optimization has emerged as a powerful method that utilizes probabilistic models to explore the hyperparameter space more effectively compared to conventional techniques such as grid search and random search [6].

This study utilizes the Breast Cancer Wisconsin Dataset, which consists of 569 samples with 30 numerical features representing the morphological characteristics of cell nuclei. The dataset is classified into two categories: benign and malignant. The proposed framework includes data preprocessing, dataset splitting into training and testing sets, and hyperparameter optimization using Bayesian Optimization applied to both SVM and XGBoost models [7].

The experimental results demonstrate that the XGBoost model optimized using Bayesian Optimization achieves the highest performance, with an accuracy of 98.24% on the test data. Furthermore, evaluation using the Receiver Operating Characteristic (ROC) curve yields an Area Under the Curve (AUC) value of 0.994, indicating excellent classification capability. The confusion matrix analysis also shows that most instances are correctly classified, with minimal misclassification [8].

In addition to performance evaluation, this study also analyzes feature importance to provide deeper insights into the model's decision-making process. The results indicate that features such as *radius_worst*, *perimeter_worst*, *area_worst*, and *concave_points_worst* have significant contributions to the classification outcome. These features are associated with tumor size and structural irregularities, which are clinically relevant indicators of malignancy.

Based on the identified gap, this research aims to provide a more comprehensive evaluation by implementing and comparing SVM and XGBoost models optimized using Bayesian Optimization, while also examining optimization behavior and feature relevance. The findings are expected to contribute not only in terms of predictive performance but also in improving the interpretability and practical applicability of machine learning–based decision support systems for breast cancer diagnosis [9].

II. Research Method

This study develops a predictive framework for breast cancer classification by combining Support Vector Machine (SVM) and Extreme Gradient Boosting (XGBoost), where Bayesian Optimization is utilized to systematically search for the most appropriate hyperparameter settings for both models [10], [11]. Within the machine learning paradigm, breast cancer detection is treated as a binary classification task in which each observation is categorized as either benign or malignant based on extracted cellular characteristics. Formally, the dataset can be represented as $D = \{(x_i, y_i)\}_{i=1}^{n}$, where $x_i \in \mathbb{R}^d$denotes the feature vector and $y_i \in \{0,1\}$represents the corresponding label. The main objective is to construct a function $f(x)$: $\mathbb{R}^d \rightarrow \{0,1\}$that is capable of accurately predicting unseen samples [12].

The analysis is conducted using the Breast Cancer Wisconsin Dataset, which contains 569 records described by 30 numerical variables representing the morphological properties of cell nuclei obtained from digitized medical images. These variables include measurements related to size, shape, and surface characteristics such as radius, texture, perimeter, area, smoothness, compactness, concavity, and symmetry. To improve data reliability and ensure stable model performance, several preprocessing steps are carried out. These include removing irrelevant features, converting diagnosis labels into numerical format through encoding, and applying feature scaling using z-score normalization. This normalization step is particularly important for SVM, as it prevents certain variables with larger scales from dominating the learning process [13].

After preprocessing, the dataset is divided into training and testing sets using an 80:20 proportion. The training subset is used to build the predictive models, while the testing subset is reserved for evaluating model generalization. To improve the robustness and reliability of the evaluation, this study also applies k-fold cross-validation (k=5). In this approach, the dataset is divided into five subsets, where the model is trained on four subsets and validated on the remaining subset iteratively. The process is repeated until each subset has been used as validation data. The average performance across all folds is then calculated to provide a more stable and unbiased evaluation.

In this study, two classification techniques are applied. The SVM algorithm determines an optimal separating hyperplane by maximizing the margin between classes, with its decision function expressed as $f(x) = w^T x + b$, where $w$and $b$denote the model parameters. On the other hand, XGBoost is an ensemble-based approach that utilizes gradient boosting over decision trees, where multiple weak learners are combined iteratively to produce a more accurate and robust model [14], [15].

To overcome the limitations of manual parameter tuning, Bayesian Optimization is employed as an automated approach for hyperparameter selection. This method constructs a probabilistic surrogate model, such as Gaussian Processes or Tree-structured Parzen Estimators (TPE), to approximate the objective function and guide the search toward optimal regions

Received: Jan 13th, 2026 | Published: March 30th, 2026

p-ISSN: 2406-7768
e-ISSN: 2581-2181

of the parameter space. In this study, several key parameters are optimized. For SVM, these include the regularization parameter C, kernel type, and kernel coefficient γ. For XGBoost, parameters such as the number of trees, learning rate, maximum depth, subsampling ratio, and feature sampling ratio are considered. The optimization is conducted iteratively, typically within 30–50 iterations, to efficiently obtain near-optimal configurations with reduced computational effort [16].

The implementation is performed using Python-based tools, including Scikit-learn for SVM, the XGBoost library for gradient boosting, and optimization frameworks such as Optuna or Hyperopt. These libraries facilitate efficient computation and ensure reproducibility throughout the experimental process.

Once the optimal hyperparameters are identified, both models are trained using the training data and subsequently evaluated on the testing data. Model performance is measured using several evaluation metrics, including accuracy, precision, recall, and F1-score, to provide a comprehensive assessment of classification quality. In addition, further evaluation is carried out through confusion matrix analysis to identify classification errors, ROC curve analysis to assess discriminative performance, and feature importance analysis to examine the contribution of each variable, particularly in the XGBoost model. Finally, a comparative evaluation between SVM and XGBoost is conducted to analyze differences in predictive performance, robustness, and interpretability[17], [18]. The overall workflow of the proposed method is summarized in Figure 1.

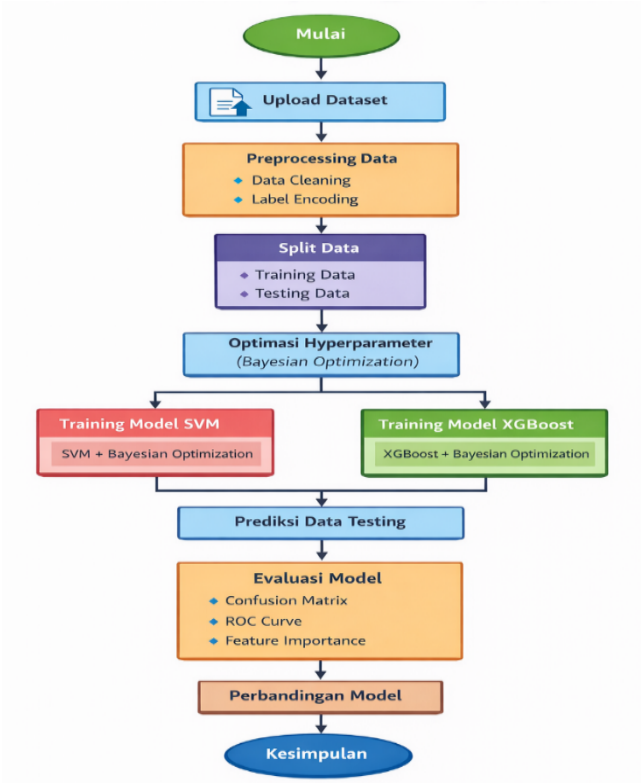


Figure 1. Research workflow for breast cancer prediction using SVM, XGBoost, and Bayesian Optimization.

Figure 1 presents the complete workflow of the proposed study for breast cancer prediction using Support Vector Machine (SVM) and Extreme Gradient Boosting (XGBoost) with hyperparameter optimization based on Bayesian Optimization. The process begins with importing the Breast Cancer Wisconsin Dataset into the computational environment for further analysis.

In the initial stage, data preprocessing is conducted to ensure data quality and model reliability. This includes data cleaning to remove irrelevant attributes, transformation of categorical diagnosis labels into numerical form using label encoding, and feature scaling through standardization to normalize the input variables. This step is particularly important to improve model stability and ensure that all features contribute proportionally during training.

After preprocessing, the dataset is divided into training and testing subsets using a fixed ratio to enable objective model evaluation. The training data is then used in the hyperparameter optimization phase. In this stage, Bayesian Optimization is employed to explore the hyperparameter search space efficiently. The optimization process iteratively evaluates candidate parameter combinations and selects the most promising configurations based on model performance. For SVM, parameters such as the regularization parameter and kernel-related settings are optimized, while for XGBoost, parameters including learning rate, tree depth, and number of estimators are tuned.

Once the optimal hyperparameters are obtained, both models are trained using the training dataset with the selected configurations. The trained models are subsequently evaluated on the testing dataset to assess their generalization capability.

To provide a comprehensive evaluation, several performance metrics and visualization techniques are utilized. These include the confusion matrix to analyze classification errors, Receiver Operating Characteristic (ROC) curve to evaluate discriminative ability, and feature importance analysis to identify the most influential variables in the prediction process. This analysis also supports a better understanding of how the models make decisions based on input features.

Finally, a comparative analysis between SVM and XGBoost is conducted to determine the relative strengths of each method in terms of predictive accuracy, robustness, and interpretability. The results of this comparison serve as the basis for drawing conclusions and highlighting the contribution of the proposed approach.

III. Results and Discussion

This section presents the experimental results as well as a deeper analysis of the proposed approach based on the research workflow described in the previous section. The process includes dataset preparation, preprocessing, data splitting, hyperparameter optimization using Bayesian Optimization, model training using Support Vector Machine (SVM) and Extreme Gradient Boosting (XGBoost), and performance evaluation using multiple quantitative metrics.

Received: Jan 13th, 2026 | Published: March 30th, 2026

p-ISSN: 2406-7768
e-ISSN: 2581-2181

The dataset used in this study is the Breast Cancer Wisconsin Dataset, which initially consists of 569 instances with 33 attributes representing morphological characteristics of cell nuclei obtained from breast mass images. These features include radius, texture, perimeter, area, smoothness, compactness, concavity, concave points, symmetry, and fractal dimension. During preprocessing, irrelevant attributes such as the ID column and missing or empty fields were removed to improve data consistency. After this step, the dataset contained 31 attributes, including 30 numerical features and one target label. The diagnosis variable was transformed into numerical format using label encoding, where benign tumors were represented as 0 and malignant tumors as 1.

Following preprocessing, the dataset was divided into training and testing subsets using an 80:20 ratio, resulting in 455 samples for training and 114 samples for testing. This separation ensures that model evaluation is conducted on unseen data, allowing for a more reliable assessment of generalization performance.

The next stage involved hyperparameter optimization using Bayesian Optimization to enhance model performance. The optimization process systematically explored the parameter space and identified optimal configurations for each algorithm. For the SVM model, the best parameters were obtained with $C = 34.7376$, $\gamma = 1.0$, and a linear kernel. Meanwhile, the XGBoost model achieved optimal performance with a learning rate of 0.1614, maximum tree depth of 8, 202 estimators, subsample value of 0.7893, and colsample_bytree value of 0.5018. These results indicate that Bayesian Optimization effectively identified parameter combinations that balance model complexity and generalization ability.

After determining the optimal hyperparameter configurations, both models were trained using the training dataset and subsequently assessed on the testing dataset.

In addition to the hold-out evaluation, cross-validation results are also analyzed to assess model stability. The results indicate that XGBoost consistently achieves high performance across different folds, demonstrating strong generalization capability. Meanwhile, SVM shows slightly higher variability, indicating lower robustness compared to XGBoost.

Table 1. Cross-Validation Results

Model	Accuracy (Mean)	Std. Deviation
SVM	96.12%	±0.021
XGBoost	98.36%	±0.015

The evaluation results indicate that XGBoost delivers better overall classification performance compared to SVM. The XGBoost model achieved an accuracy of 98.24%, reflecting its ability to correctly classify nearly all data instances. Based on the classification report, the precision score reached 0.97 for the benign class and 1.00 for the malignant class. In terms of recall, the model achieved 1.00 for benign samples and 0.95 for malignant samples. The F1-score of 0.98 indicates that the model maintains a well-balanced trade-off between precision and recall, highlighting its reliability.

The higher performance of XGBoost relative to SVM can be explained by several underlying factors. As an ensemble learning method, XGBoost integrates multiple decision trees, enabling it to model complex and non-linear patterns more effectively than SVM with a linear kernel. Furthermore, the boosting strategy iteratively refines the model by correcting errors from previous iterations, which improves predictive accuracy and reduces bias. Additionally, the inclusion of regularization mechanisms in XGBoost helps control overfitting, making the model more robust when applied to structured medical data.

Beyond performance evaluation, feature importance analysis was carried out to identify the contribution of each variable to the prediction results. The analysis shows that variables such as *radius_worst*, *perimeter_worst*, *area_worst*, and *concave_points_worst* have the greatest influence on the model output. These features are associated with tumor size, shape irregularity, and structural complexity, which are known indicators of malignancy in clinical practice. This alignment suggests that the model not only performs well but also provides meaningful insights that are consistent with medical knowledge, thereby enhancing its interpretability.

Overall, the findings demonstrate that combining XGBoost with Bayesian Optimization results in a more accurate and dependable classification model compared to SVM. In addition, the inclusion of feature importance analysis and model evaluation contributes to a deeper understanding of how predictions are generated, addressing the need for both high performance and interpretability in medical decision support systems.

To further assess classification performance, a confusion matrix analysis was conducted, as presented in Figure 2. This matrix provides a detailed breakdown of prediction outcomes. The results show that 72 benign samples and 40 malignant samples were correctly classified. However, there were 2 instances where malignant cases were incorrectly predicted as benign, while no benign cases were misclassified as malignant.

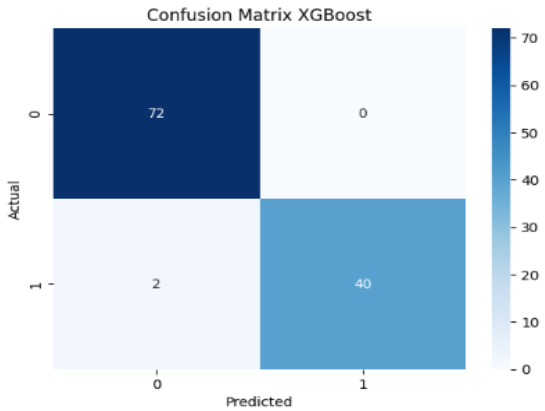


Figure 2. Confusion Matrix of the XGBoost Model

Received: Jan 13th, 2026 | Published: March 30th, 2026

p-ISSN: 2406-7768
e-ISSN: 2581-2181

This result indicates that the model demonstrates very high specificity, as it does not incorrectly label benign cases as malignant, which is important to avoid unnecessary medical interventions. However, the presence of a small number of false negatives (malignant predicted as benign) suggests a potential limitation, as this type of error may delay critical treatment.

This finding also indicates that further validation using cross-validation or additional datasets is necessary to ensure the reliability of the model in real-world clinical applications.

This finding highlights the importance of further optimization or integration with additional clinical validation to minimize such risks in real-world applications.

Further performance evaluation was conducted using the Receiver Operating Characteristic (ROC) curve, as shown in Figure 3. The ROC curve represents the trade-off between the True Positive Rate (TPR) and False Positive Rate (FPR) across different classification thresholds. The XGBoost model achieved an Area Under the Curve (AUC) value of 0.994, which indicates an excellent ability to discriminate between benign and malignant classes.

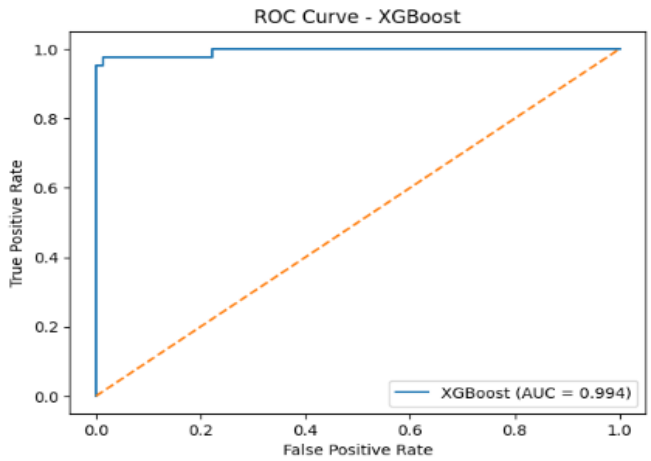


Figure 3. ROC Curve of the XGBoost Model

The high AUC value suggests that the model maintains strong sensitivity and specificity simultaneously, making it highly reliable for classification tasks. This performance also reflects the effectiveness of Bayesian Optimization in identifying optimal hyperparameters that enhance the model's discriminative capability.

In addition to performance metrics, feature importance analysis was conducted to provide interpretability of the model, as presented in Figure 4. The results show that features such as *radius_worst*, *perimeter_worst*, *area_worst*, *concave_points_worst*, and *concave_points_mean* contribute most significantly to the classification process.

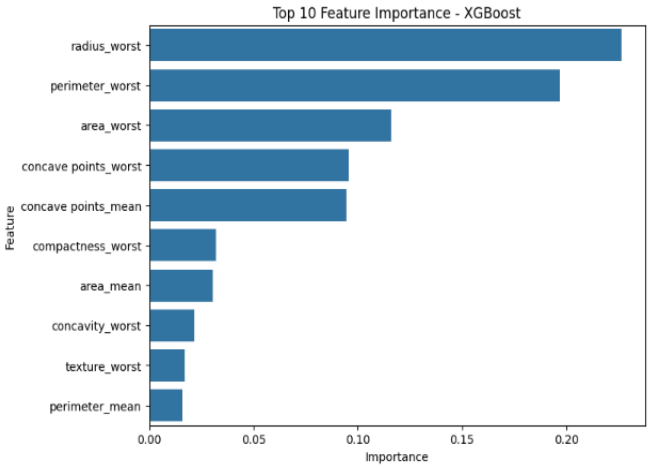


Figure 4. Feature Importance of the XGBoost Model

These variables are closely linked to tumor dimensions, irregular boundaries, and structural complexity, all of which are important clinical indicators of malignancy. For instance, increased values in radius and area are typically associated with abnormal cell proliferation, whereas higher levels of concavity and concave points reflect uneven or distorted tumor edges. The consistency between the model's results and established medical knowledge enhances both the interpretability and reliability of the proposed method.

A comparative evaluation of the SVM and XGBoost models was also performed, as shown in Figure 5. The findings indicate that XGBoost delivers superior accuracy compared to SVM.

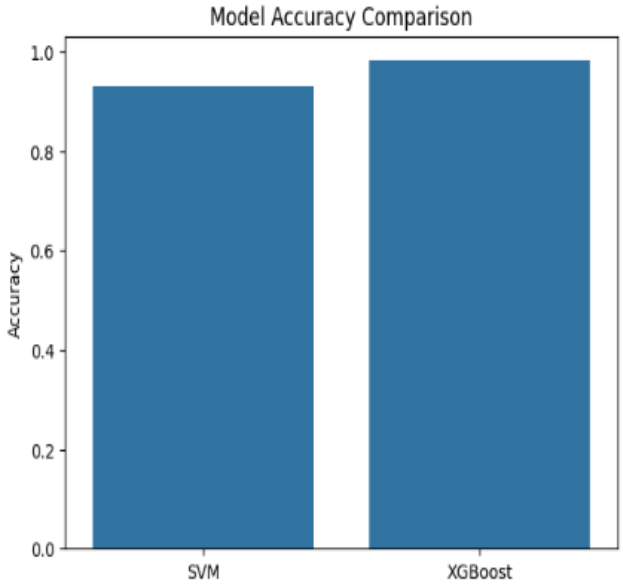


Figure 5. Accuracy Comparison between SVM and XGBoost Models

Received: Jan 13th, 2026 | Published: March 30th, 2026

p-ISSN: 2406-7768
e-ISSN: 2581-2181

The superior performance of XGBoost can be explained by its ensemble learning mechanism, which combines multiple decision trees to model complex and non-linear relationships within the data. In contrast, SVM with a linear kernel has limited capability in capturing such complexity. Additionally, XGBoost incorporates regularization and iterative error correction, which improve generalization performance and reduce overfitting. The integration of Bayesian Optimization further enhances both models by efficiently exploring the hyperparameter space, although its impact is more pronounced in XGBoost due to its larger number of tunable parameters.

Overall, the experimental results demonstrate that the proposed approach, which integrates XGBoost with Bayesian Optimization, provides highly accurate and reliable performance for breast cancer classification. Beyond achieving high accuracy and AUC values, the model also offers interpretability through feature importance analysis, making it suitable as a potential decision support tool for assisting medical professionals in breast cancer diagnosis.

IV. CONCLUSION

This research introduces a comparative and optimized framework for breast cancer classification by combining Support Vector Machine (SVM) and Extreme Gradient Boosting (XGBoost), where Bayesian Optimization is applied to refine hyperparameter selection. The study utilizes the Breast Cancer Wisconsin Dataset, which contains 569 samples characterized by morphological features of cell nuclei. Within this framework, model performance, optimization efficiency, and feature relevance are systematically analyzed in a unified experimental setup.

The results demonstrate that XGBoost consistently outperforms SVM in classification tasks, achieving an accuracy of 98.24% and an AUC score of 0.994. These findings indicate that XGBoost is more capable of modeling complex patterns within the data and producing stable predictions. The confusion matrix further confirms that the model yields only a small number of misclassifications, although a few false negative cases remain, which are important to consider in sensitive medical applications.

In addition to performance evaluation, this study emphasizes the effectiveness of Bayesian Optimization in identifying suitable hyperparameter configurations in an efficient manner, which contributes to improved model accuracy and stability. The feature importance analysis also shows that variables such as *radius_worst*, *perimeter_worst*, *area_worst*, and *concave points worst* have a significant influence on classification outcomes. These features are associated with tumor size and structural irregularities, aligning with clinical indicators of malignancy and enhancing the interpretability of the model.

The key contribution of this work is the provision of a more comprehensive analysis that integrates performance comparison, optimization processes, and feature-level interpretation within a single framework. This approach addresses limitations in prior studies that mainly emphasize accuracy without offering deeper analytical insights. Furthermore, the use of a single data split without cross-validation may limit the robustness of the evaluation results.

However, several limitations should be noted. The use of a single dataset may restrict the generalizability of the findings. Moreover, despite achieving high performance, additional validation using real clinical data and multiple datasets is required to ensure robustness in practical scenarios. Future research should consider exploring other machine learning or deep learning approaches, utilizing more diverse datasets, and incorporating explainable AI methods to further improve transparency.

In practical terms, implementing the proposed framework within clinical decision support systems represents a promising direction. Overall, the combination of XGBoost and Bayesian Optimization offers an effective and dependable solution for breast cancer classification, with strong potential to assist healthcare professionals in enhancing diagnostic accuracy and decision-making processes.

REFERENSI

[1] I. B. Mohammad and M. M. Ahmad, "Bayesian Model Prediction for Breast Cancer Survival: A Retrospective Analysis," *Eur. J. Breast Heal.*, vol. 21, no. 3, pp. 255–264, 2025.

[2] Z. Rais, S. Muhammad Fahmuddin, Saida, and A. T. Utomo, "Implementation of Machine Learning Algorithm with Extreme Gradient Boosting (XGBoost) Method in Hypertension Level Classification," *J. Appl. Sci. Eng. Technol. Educ.*, vol. 7, no. 1, pp. 126–136, 2025.

[3] P. Nalela, D. Rao, and P. Rao, "Predicting Survivability of Cancer Patients with Metastatic Patterns Using Explainable AI," pp. 1–23, 2025.

[4] R. Oktafiani, A. Hermawan, and D. Avianto, "The Effect of Data Split Composition on Breast Cancer Classification Performance Using Machine Learning Algorithms," Journal of Science and Information, vol. 9, no. April, pp. 19–28, 2023.

[5] I. Solikin, "Data Mining Techniques for Efficient Breast Cancer Prediction," Fidelity: Journal of Electrical Engineering, vol. 3, no. 3, pp. 63–67, 2021.

[6] L. Widya Astuti, Faradilla, I. Saluza, and M. Fadhiel Alie, "Optimization of Breast Cancer Classification Using Forward Selection in Naive Bayes," Global Informatics Scientific Journal, vol. 11, no. 2, pp. 63–67, 2020.

[7] M. R. Andryan, M. Fajri, and N. Sulistyowati, "Performance Comparison of XGBoost and Support Vector Machine (SVM) Algorithms for Breast Cancer Diagnosis," JIKO (Journal of Informatics and Computer Science), vol. 6, no. 1, p. 1, 2022.

Received: Jan 13th, 2026 | Published: March 30th, 2026

p-ISSN: 2406-7768
e-ISSN: 2581-2181

[8] N. Nurzilla, "Prediction of Breast Cancer Tumor Growth Using a Machine Learning-Based Linear Regression Model," *Journal of Artificial Intelligence Applications*, vol. 1, no. 1, pp. 28–35, 2025.

[9] M. Abdillah et al., "Implementation of XGBoost for Chronic Kidney Disease Classification Using the Chronic Kidney Disease Dataset," *JATISI (Journal of Informatics Engineering and Information Systems)*, vol. 12, no. 3, 2025.

[10] T. Tinaliah and T. Elizabeth, "Prediction of Breast Cancer Types Using Support Vector Machine Based on Recursive Feature Elimination," *JATISI*, vol. 11, no. 3, pp. 1–9, 2024.

[11] Y. N. Kholisho and Z. Amri, "Village Administration System," *JITE (Journal of Informatics and Telecommunication Engineering)*, vol. 9, no. July, pp. 281–291, 2025.

[12] J. Kusuma, B. H. Hayadi, Wanayumini, and R. Rosnelly, "Comparison of Multi-Layer Perceptron (MLP) and Support Vector Machine (SVM) Methods for Breast Cancer Classification," *MIND (Multimedia Artificial Intelligence Network Database) Journal*, vol. 7, no. 1, pp. 51–60, 2022.

[13] R. Resmiati and T. Arifin, "Breast Cancer Patient Classification Using the Support Vector Machine Method with Backward Elimination," *SISTEMASI: Journal of Information Systems*, vol. 10, no. 2, pp. 381–393, 2021.

[14] A. Desiani, D. A. Zayanti, I. Ramayanti, F. F. Ramadhan, and G. Giovillando, "Comparison of Support Vector Machine (SVM) and Logistic Regression Algorithms," *Journal of Artificial Intelligence and Information Technology*, vol. 4, no. 1, pp. 33–42, 2025.

[15] H. Imaduddin, B. A. Hermansyah, and F. A. Salsabilla B, "Comparison of Support Vector Machine and Decision Tree Methods in Breast Cancer Classification," *Cyberspace: Journal of Information Technology Education*, vol. 5, no. 1, p. 22, 2021.

[16] N. Tri, R. Adiningrum, R. Rianti, C. Prianto, T. Informatics, and J. Barat, "Design and Development of a Machine Learning-Based Cancer Prediction Application," vol. 11, no. 3, 2023.

[17] E. S. Septiany, H. H. Handayani, T. Al Mudzakir, and A. F. N. Masruriyah, "Optimization of the Support Vector Machine Method Using Recursive Feature Elimination and Forward Selection for Breast Cancer Classification," *TIN: Applied Informatics of Nusantara*, vol. 5, no. 2, pp. 144–154, 2024.

[18] Z. Syahputra and R. Kurniawan, "High Performance Computing Architecture for Computer Network Systems," *Journal of Computer Networks, Architecture and High Performance Computing*, vol. 7, no. 1, pp. 341–352, 2025.

Received: Jan 13th, 2026 | Published: March 30th, 2026