Journal of Applied Data Sciences
Vol. 6, No. 1, January 2025, pp. 71-85
ISSN 2723-6471

The Development of Stacking Techniques in Machine Learning for Breast Cancer Detection

Lucky Lhaura Van FC[1,*], M. Khairul Anam[2], Saiful Bukhori[3], Abd Kadir Mahamad[4], Sharifah Saon[5], Rebecca La Volla Nyoto[6]

[1,6]*Department of Informatics Engineering Universitas Lancang Kuning, Jl. Yos Sudarso, Pekanbaru and 28266, Indonesia*

[2]*Department of Informatics Universitas Samudra, Jl. Prof. Dr. Syarief Thayeb, Meurandeh, Langsa and 24416, Indonesia*

[3]*Departmen of Computer Science, Universitas Jember, Jl. Kalimantan, Jember and 68121, Indonesia*

[4,5]*Fakulti Kejuruteraan Elektrik dan Elektronik, Universiti tun Hussein Onn Malaysia, Parit Raja, Batu Pahat, Johor and 86400, Malaysia*

(Received: September 3, 2024; Revised: October 6, 2024; Accepted: November 21, 2024; Available online: December 27, 2024)

Abstract

This study addresses the challenges of accurately detecting breast cancer using machine learning (ML) models, particularly when handling imbalanced datasets that often cause model bias toward the majority class. To tackle this, the Synthetic Minority Over-sampling Technique (SMOTE) was applied not only to balance the class distribution but also to improve the model's sensitivity in detecting malignant tumors, which are underrepresented in the dataset. SMOTE was effective in generating synthetic samples for the minority class without introducing overfitting, enhancing the model's generalization on unseen data. Additionally, AdaBoost was employed as the meta model in the stacking framework, chosen for its ability to focus on misclassified instances during training, thereby boosting the overall performance of the combined base models. The study evaluates several models and combinations, with K-Nearest Neighbors (KNN) + SMOTE achieving an accuracy of 97%, precision, recall, and F1-score of 97%. Similarly, C4.5 + Hyperparameter Tuning + SMOTE reached 95% in all metrics. The stacking model with Logistic Regression (LR) as the meta model and SMOTE achieved a strong performance with 97% accuracy, precision, recall, and F1-score all at 97%. The best result was obtained using the combination of Stacking AdaBoost + Hyperparameter Tuning + SMOTE, reaching an accuracy of 98%. These findings highlight the effectiveness of combining SMOTE with stacking techniques to develop robust predictive models for medical applications. The novelty of this study lies in the integration of SMOTE and advanced stacking methods, particularly using AdaBoost and Logistic Regression, to address the issue of class imbalance in medical datasets. Future work will explore deploying this model in clinical settings for accurate and timely breast cancer detection.

Keywords: Machine learning, Stacking, Adaboost, Hyperparameter Tuning, SMOTE

1. Introduction

Machine Learning algorithms are commonly used to address problems in various fields such as e-Commerce, education, health, and others [1]. The goal of machine learning is to make life easier and to provide accurate and measurable solutions. To achieve this, machine learning algorithms are utilized to discover new insights and data patterns, or to predict output values from a given set of input variables [2]. However, machine learning algorithms have some limitations, such as incomplete data leading to inaccurate results, which can significantly lengthen the programming process [3]. Furthermore, some studies have shown inconsistency in classification and detection when using basic algorithms [4]. Therefore, previous researchers have attempted to combine algorithms or introduce additional features or methods [5], [6], [7].

Efforts to improve machine learning have been conducted in various ways. For example, [8] enhanced the performance by using the SMOTE method to balance classes. Additionally, ensemble methods with boosting techniques, have been employed. Another study used stacking techniques to improve machine learning model performance [9], while hyperparameter tuning has also been applied to enhance performance [10]. Feature selection has been another approach

***Corresponding author: Lucky Lhaura Van FC (lucky@unilak.ac.id)**
DOI: https://doi.org/10.47738/jads.v6i1.416

to improving performance [11]. Enhancements made by combining different methods or algorithms have generally resulted in better performance than using base algorithms alone. Therefore, this study also employs a combination of methods and algorithms. This research uses a breast cancer dataset processed with three base machine learning algorithms: K-Nearest Neighbors (KNN), Gaussian Naïve Bayes (GNB), and C4.5. KNN is chosen due to its intuitive nature, ease of implementation, and its ability to classify new data based on its proximity to existing classified data without requiring distribution assumptions, making it effective for data with non-linear relationships [12]. GNB, although assuming feature independence, is fast in training and inference, and is suitable for high-dimensional data with good probability estimation [13]. Meanwhile, C4.5, as a decision tree algorithm, can handle both numerical and categorical data, address missing data, and produce models that are easy to interpret with pruning techniques to reduce complexity [14].

Previous studies have shown that using KNN combined with grid search achieved 94.35% accuracy in detecting breast cancer [15]. Another study using GNB for breast cancer prediction achieved 88% accuracy [16]. Additionally, C4.5 combined with Particle Swarm Optimization (PSO) was used for breast cancer classification, achieving 98% accuracy, an improvement from the previous experiment without PSO, which only achieved 93% [17]. Several previous studies have found that using other methods or algorithms can improve accuracy.

This study also employs other methods to improve accuracy, including the SMOTE method for data balancing [18]. Without class balancing in imbalanced data, classification errors are likely. The SMOTE method can be used to enhance algorithm performance. Research conducted by [19] tested financial distress using SVM; without SMOTE, it achieved 95%, and after using SMOTE, it reached 96%. However, SMOTE does not always improve accuracy; it can also reduce accuracy, as shown by [20], where the accuracy achieved with SVM RBF without SMOTE was 83%, but after using SMOTE, it decreased to 81%. Despite the reduction, the performance achieved using SMOTE did not result in overfitting [21].

Additionally, feature selection with Lasso is used to address high-dimensional data issues, minimizing the risk of overfitting during training [22]. The study also employs an ensemble method with stacking techniques, which combine the strengths of base algorithms with an additional model known as a meta-model [23]. The use of stacking is expected to improve model performance in detecting breast cancer. The data used in this research is a public dataset that can be accessed on Kaggle at https://www.kaggle.com/datasets/nancyalaswad90/breast-cancer-dataset/data.

In the stacking phase, this study uses two machine learning algorithms as meta-models: Logistic Regression (LR) and AdaBoost. LR is often chosen as a meta-model in previous studies due to its ability to effectively combine predictions from base models in a simple and reliable manner [24]. AdaBoost is also used as a meta-model because it utilizes a boosting approach that iteratively focuses on correcting errors to improve the overall performance of the stack [25].

The study further employs Optuna for hyperparameter tuning. Optuna provides an efficient, flexible, and user-friendly solution for hyperparameter optimization, making it a powerful choice for automatically improving machine learning model performance [26]. By combining all these methods and algorithms with base algorithms, this study aims to achieve significant improvements in breast cancer detection.

2. Literature Review

Related Research has utilized various methods and datasets, both similar and different, related to diseases. Table 1 shows previous studies.

Table 1. Related Research

No	Researcher	Model	Improvement Technique	Dataset	Accuracy
1	[27]	MLP, DT SVM	Stacking Classifier (LR)	Heart Disease	86%
2	[28]	Naïve Bayes and LR	Stacking Classifier (SVM)	Breast Cancer	87%
3	[29]	RF, SGD, MLP, Adaboost, KNN, XGBoosrt, GBM, SVC, CART, ET	Stacking Classifier (LR)	Heart Disease	91%
4	[30]	Baseline (SGD+KNN), Boosting (GBM+XGB), and Bagging (ETC+RF)	Stacking Classifier and Shaply Explainable AI	Ovarian Cancer	96%

5	[31]	RF, SVC, KNN, LGBM, Bagging, Adaboost	Stacking Classifier (LR)	Heart Failures	80%

Table 1 presents several studies that use various machine learning models, such as MLP (Multilayer Perceptron), Decision Tree (DT), Naïve Bayes, LR, Random Forest (RF), and others, combined with the Stacking Classifier technique. This technique is applied to improve model accuracy on various datasets, including Heart Disease, Breast Cancer, and Ovarian Cancer. For instance, in the first study, the MLP, DT, and SVM models combined with a Stacking Classifier achieved an accuracy of 86% for heart disease prediction. Other studies show that this technique can increase accuracy up to 96% on ovarian cancer datasets, highlighting that the Stacking Classifier method is highly effective in enhancing model performance, especially in the context of disease prediction.

Although these studies exhibit relatively high accuracy, they do not address data balancing. Balancing data, particularly in the context of imbalanced datasets, aims to ensure that each class has a proportional number of samples, preventing the machine learning model from being biased towards the majority class. It can also increase accuracy for minority classes, thereby improving overall model performance [32].

Some drawbacks arise if data balancing is not performed. Overfitting on the majority class may occur, leading to misleading results, such as high accuracy that, when tested on other data, causes classification errors due to the poor performance of the minority class. This issue can be avoided by balancing data using methods such as SMOTE or ADASYN.

This study uses SMOTE to address the problem of data imbalance, as SMOTE is superior to other data balancing methods due to its ability to generate new samples synthetically. Instead of merely duplicating data from the minority class, as in conventional oversampling methods, SMOTE creates new data by interpolating between existing minority class data points. This helps reduce the risk of overfitting, commonly seen in traditional oversampling methods, and provides the model with more diverse patterns from the minority class. As a result, SMOTE is more effective at improving machine learning model performance on imbalanced datasets [33].

Additionally, the study employs the Stacking technique to improve accuracy. Stacking combines predictions from several models and uses another model as a meta-model to generate the final prediction [34]. As shown in table 1, several studies demonstrate that Stacking significantly enhances classification performance, particularly when combined with robust algorithms like AdaBoost. Therefore, this study also compares meta-models between AdaBoost and Logistic Regression, while keeping the base algorithms the same (KNN, GNB, and C4.5). The combination of SMOTE and Stacking techniques has proven effective in improving the performance of heart disease prediction.

3. Methodology

Figure 1 illustrates the research methodology that facilitates the conduct of this study. The research begins with the input of a dataset related to breast cancer, followed by preprocessing the dataset. After that, modeling is performed using various algorithms and their enhancements. The final step is to evaluate the model using a confusion matrix.

ISSN 2723-6471

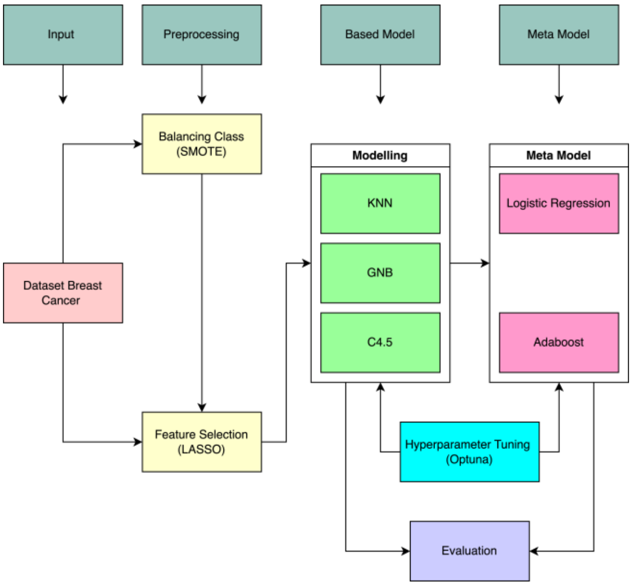

Figure 1. Methodology Flow

3.1.Dataset

The breast cancer dataset used in this study is obtained from Kaggle and consists of 569 records. The dataset comprises numerical data, where each record contains numerical features that describe tumor characteristics such as size, texture, and cell edge clarity, which are then used to predict whether the tumor is benign or malignant. This dataset is well-suited for application in machine learning models for classification tasks.

3.2. Pre-processing

Pre-processing is a crucial step in any machine learning project as it ensures that the data used in the model is clean, structured, and ready for analysis. In this study, pre-processing is carried out to address various challenges associated with the breast cancer dataset obtained from Kaggle, consisting of 569 numerical records. The pre-processing phase of this study involves comprehensive data cleaning on the utilized dataset. Data cleaning includes normalization techniques to ensure all features are on the same scale, thereby minimizing potential bias in the model [35]. Additionally, data transformations, such as log transformation, are employed to reduce skewness in the data distribution. These steps enhance the model's stability and performance during the analysis and prediction phases, ensuring more accurate and reliable results [36], [37].

3.3. Labelling

The initial data has been labelled as 'benign' and 'malignant.' This is the basic stage where each record in the dataset is categorized to facilitate further classification. Figure 2 shows the initial labelling.

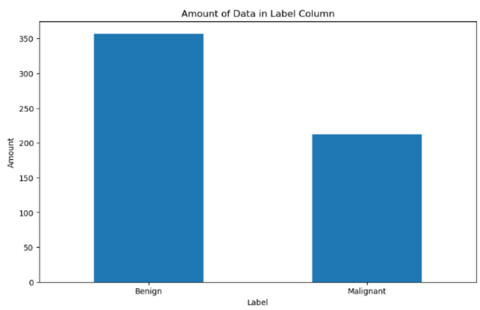

Figure 2. Initial Labelling

3.4. Handling Missing Values

This step ensures that all data in the dataset is complete. However, if there are missing values, steps are taken to fill them using appropriate methods, such as filling with the mean or median values. Mean or median is used to fill missing

ISSN 2723-6471

values in a dataset because both methods help maintain the original distribution of the data without introducing significant bias [38]. Using the mean is appropriate when the data is symmetrically distributed without outliers, as the mean provides a representative central tendency [39]. On the other hand, the median is preferred when the data contains outliers or is skewed, as the median is more robust against outliers and provides a more accurate representation of the central tendency in such distributions [40]. By using mean or median, the dataset becomes more complete, allowing for more accurate and reliable analysis without the interference of missing data.

3.5. Handling Imbalanced Classes

In this breast cancer dataset, the number of 'benign' cases is significantly higher than 'malignant' cases. Such class imbalance can cause machine learning models to predict the majority class (benign) with high accuracy while failing to detect the minority class (malignant), which is more critical. To address this issue, this study employs the SMOTE technique. SMOTE works by synthesizing new examples of the minority class (malignant) based on the existing data, thereby increasing the amount of data in that class and balancing the class distribution [41]. This way, the model can be trained to recognize patterns in both classes more effectively, enhancing its ability to detect malignant tumors. In addition, SMOTE can also work to help overcome the problem of overfitting caused by random oversampling [42]. Figure 3 shows the dataset after balancing with SMOTE.

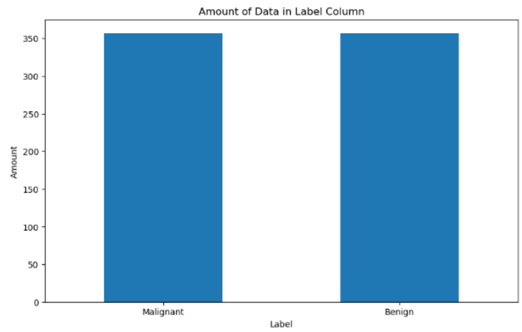

Figure 3. Dataset after SMOTE process

3.6. Feature Scaling

All numerical features in the dataset are scaled to the same range using normalization techniques. This is essential because certain machine learning models, such as KNN, are very sensitive to data scale [43].

3.7. Dimensionality Reduction

Dimensionality Reduction is the process of reducing the number of features (dimensions) in a dataset, aiming to improve the efficiency and accuracy of machine learning models by eliminating less important features [44]. In this study, the Lasso (Least Absolute Shrinkage and Selection Operator) technique is used to select and retain only the most relevant features [22]. Lasso is chosen for breast cancer research due to its ability to automatically perform feature selection, reduce overfitting, and enhance model interpretability. It sets the coefficients of less important features to zero, retaining only the most relevant features for diagnosis. This results in a simpler, more stable, and accurate model, while making it easier to interpret by focusing on the key factors that truly impact the prediction. The process is conducted twice in this study: once before using SMOTE and once after using SMOTE.

Initial Feature Selection, before applying SMOTE, Lasso is used to identify and remove features with little or no contribution to the predictions. This reduces model complexity and prevents overfitting on unbalanced data. After irrelevant features are removed, the resulting dataset has fewer but more informative features. Machine learning models trained on this dataset tend to be faster and more accurate. Figure 4 shows the results of the Lasso process.

Reevaluating Features After Balancing, after balancing the dataset with SMOTE, the class distribution becomes more balanced, which can affect the importance of some features. Therefore, Lasso is reapplied to reassess the relevance of each feature in the context of the balanced dataset. By balancing classes, Lasso can identify new significant features or eliminate features that are no longer relevant after class balancing. This ensures that the model built after SMOTE uses

ISSN 2723-6471

only truly important features, enhancing the model's performance and generalization on more balanced data. Figure 5 shows the Lasso results after balancing.

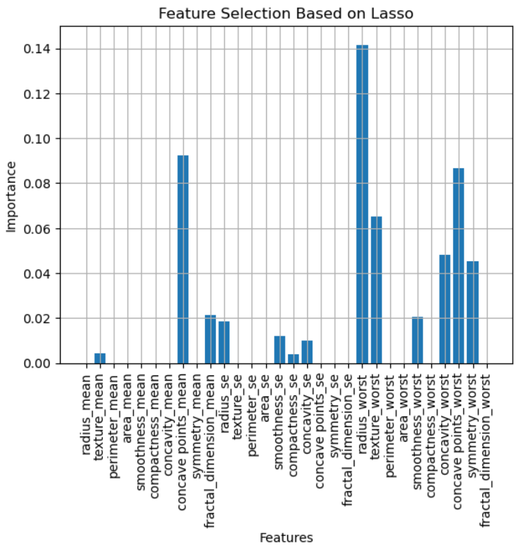

Figure 4. Lasso Process Results

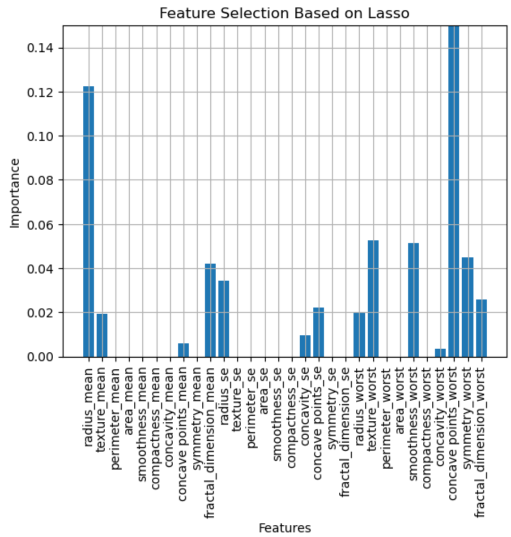

Figure 5. Lasso Process Results after Balancing

3.8.Modeling

After preprocessing, the next step is to build and test different machine learning models to detect breast cancer. The models used in this study include a combination of base algorithms and ensemble techniques to achieve optimal performance. Table 2 shows the models tested in this study. Each model is also tested with the dataset balanced using SMOTE to see how class balancing affects model performance.

Table 2. Model Tested

No	Model	No	Model
1	KNN	11	KNN + Hyperparameter Tuning
2	C4.5	12	C4.5 + Hyperparameter Tuning
3	GNB	13	GNB + Hyperparameter Tuning
4	Stacking LR	14	Stacking LR + Hyperparameter Tuning
5	Stacking Adaboost	15	Stacking Adaboost + Hyperparameter Tuning
6	KNN + SMOTE	16	KNN + SMOTE + Hyperparameter Tuning
7	C4.5 + SMOTE	17	C4.5 + SMOTE + Hyperparameter Tuning
8	GNB + SMOTE	18	GNB + SMOTE + Hyperparameter Tuning
9	Stacking LR + SMOTE	19	Stacking LR + SMOTE + Hyperparameter Tuning
10	Stacking Adaboost + SMOTE	20	Stacking Adaboost + SMOTE + Hyperparameter Tuning

3.8.1. Base Algorithm

KNN: This algorithm classifies data based on its proximity to other data points in the feature space [45]; C4.5 (Decision Tree): This algorithm builds decision trees based on attributes in the dataset to perform classification [46]; GNB: A probabilistic algorithm that uses Bayes' Theorem with an assumption of independence among features [47].

3.8.2. Ensemble Learning

Stacking with Logistic Regression: Logistic regression is used as a meta-model to combine predictions from base models (KNN, C4.5, and GNB) into a more accurate final prediction [48]; Stacking with AdaBoost: In addition to logistic regression, AdaBoost is also used as a meta-model, focusing on improving predictions by placing more weight on errors made by the base models in previous iterations; Combining KNN, C4.5, and GNB offers several advantages. This integration leverages the unique strengths of each algorithm to handle complex data more effectively. KNN, C4.5, and GNB complement each other in handling non-linear data, model interpretability, and inference speed. Additionally, this combination can reduce the risk of overfitting, as the resulting model can better manage diverse data patterns. The outcome is improved accuracy and robustness, making the model more stable and capable of handling various data types with more consistent performance.

3.8.3. Hyperparameter Tuning with Optuna

Hyperparameter tuning was conducted to find the optimal parameter combinations for each algorithm, including maximum depth (max_depth) and minimum samples per split (min_samples_split) for C4.5, as well as other parameters such as the number of neighbors (n_neighbors) for KNN and smoothing variance (var_smoothing) for GNB. By optimizing these hyperparameters, the models can be fine-tuned to achieve a better balance between bias and variance, thereby reducing the risk of overfitting after applying SMOTE. Optuna was chosen for this process due to its capability to efficiently and adaptively search the hyperparameter space, allowing for a broader exploration and ultimately improving overall model performance [49].

3.9. Model Evaluation

The performance of each model is evaluated using metrics such as accuracy, precision, recall, and F1-score, with a confusion matrix used to understand the predictive performance of each model.

4. Result and Discussion

4.1. Algorithm Based

In this study, tests were conducted on three commonly used base algorithms for classification, namely KNN, C4.5, and GNB. The purpose of these tests was to evaluate the performance of each algorithm in detecting breast cancer based on the dataset used. Table 3 shows the results of the tests with the base algorithms.

Table 3. Tests With Base Algorithm

No	Model	Accuracy	Precision	Recall	F1-Score
1	KNN	95%	95%	95%	95%
2	C4.5	96%	96%	96%	96%
3	GNB	93%	93%	92%	92%
4	KNN + SMOTE	97%	97%	97%	97%
5	C4.5 + SMOTE	95%	95%	95%	95%
6	GNB + SMOTE	95%	95%	95%	95%

KNN initially achieved an accuracy of 95%, with precision, recall, and F1-score also at 95%. This indicates that KNN has balanced and fairly reliable performance in detecting tumors. After applying the SMOTE technique, which balances class distribution in the dataset, KNN's performance increased significantly. Accuracy rose to 97%, as did precision, recall, and F1-score, all reaching 97%. This improvement shows that KNN benefited from class balancing, which helped the model better identify malignant tumor cases.

Next, C4.5 initially showed very good results with an accuracy of 96%, as well as precision, recall, and F1-score each at 96%. However, after SMOTE was applied, accuracy and other metrics slightly decreased to 95%. This decrease could be due to the fact that decision tree models like C4.5 might be more prone to overfitting when classes are balanced by synthesizing new data.

Lastly, GNB a probabilistic model, showed initial accuracy of 93% with precision and F1-score both at 93%, but recall was slightly lower at 92%. After applying SMOTE, GNB's performance improved, with accuracy, precision, recall, and F1-score all rising to 95%. This suggests that GNB was able to leverage the more balanced data to enhance its ability to classify malignant tumor cases.

Overall, the test results indicate that applying SMOTE generally improved the performance of the models, particularly with the KNN and GNB algorithms. KNN showed the most significant improvement in accuracy and other metrics after using SMOTE. However, the results for C4.5 suggest that the effect of SMOTE can vary depending on the

ISSN 2723-6471

algorithm used and that in some cases, data balancing can reduce performance due to overfitting or increased model complexity.

4.2. Stacking

Before discussing the test results with the stacking technique, it is important to first understand the basic performance of each algorithm used as a base model. In the initial stage, tests were conducted on the three main algorithms, KNN, C4.5, and GNB. These tests aimed to evaluate the performance of each algorithm individually, both in standard conditions and after applying SMOTE, a technique used to balance an imbalanced dataset [29]. These test results will serve as a basis for comparison to understand how much improvement is provided by the stacking technique, which combines the strengths of several models using LR and AdaBoost as meta models. Table 4 shows the results of testing using the stacking technique and its impact on accuracy, precision, recall, and F1-score.

Table 4. Tests with Stacking Technique

No	Based Model	Meta model	Accuracy	Precision	Recall	F1-Score
1	KNN, C4.5, GNB	LR	96%	96%	96%	96%
2	KNN, C4.5, GNB	Adaboost	96%	95%	96%	96%
3	KNN, C4.5, GNB + SMOTE	LR	97%	97%	97%	97%
4	KNN, C4.5, GNB + SMOTE	Adaboost	95%	95%	95%	95%

Testing with the stacking technique produced interesting results when compared to the tests with the base algorithms. When the three algorithms KNN, C4.5, and GNB were combined using LR as the meta model, the results showed consistent accuracy, precision, recall, and F1-score of 96%. This indicates that Logistic Regression as a meta model was able to effectively combine the strengths of the three base models, resulting in optimal performance.

When the same combination was tested with AdaBoost as the meta model, the results were slightly different. Although accuracy remained the same at 96%, precision slightly decreased to 95%, while recall and F1-score remained at 96%. This small drop in precision suggests that while AdaBoost is known for its strength in boosting model performance, Logistic Regression is more consistent in maintaining precision in this specific combination.

The application of SMOTE to balance the dataset before performing stacking with Logistic Regression as the meta model provided a significant performance boost. Accuracy, precision, recall, and F1-score all increased to 97%, indicating that class balancing through SMOTE helped the model better recognize patterns from the minority class, particularly in the context of breast cancer detection. In contrast, when SMOTE was applied and AdaBoost was used as the meta model, the results showed a slight drop in performance across all metrics (accuracy, precision, recall, and F1-score) to 95%. This may indicate that the combination of AdaBoost with the balanced dataset does not always yield better results, depending on the characteristics of the dataset.

Overall, the stacking technique with Logistic Regression as the meta model showed consistent and strong results, especially after applying SMOTE. In contrast, while AdaBoost also provided good results, certain combinations with SMOTE may require further adjustments to achieve optimal performance. This comparison highlights the importance of choosing the right meta model and considering the effects of techniques like SMOTE in improving overall model performance.

4.3. Hyperparameter Tuning Addition

Before discussing further the results of testing with the addition of hyperparameter tuning, it is important to understand how tuning can enhance the performance of machine learning models. Hyperparameter tuning is the process of adjusting model parameters that are not optimized during training to find the combination of parameters that provides the best performance on the given data [30]. In this study, tuning was performed on several individual algorithms such as KNN, C4.5, and GNB, as well as on stacking techniques using LR and AdaBoost as meta models. Additionally, some tests were also conducted by combining hyperparameter tuning with SMOTE, a technique used to balance an imbalanced dataset. Table 5 shows the results of testing, demonstrating how the combination of various techniques affects the accuracy, precision, recall, and F1-score of each model.

Table 5. Testing With Added Hyperparameter Tuning

No	Model	Accuracy	Precision	Recall	F1-Score
1	KNN + Hyperparameter	96%	97%	96%	96%
2	C4.5 + Hyperparameter	96%	96%	96%	96%
3	GNB + Hyperparameter	94%	94%	93%	93%
4	KNN + Hyperparameter + SMOTE	97%	97%	97%	97%
5	C4.5 + Hyperparameter + SMOTE	95%	95%	95%	95%
6	GNB + Hyperparameter + SMOTE	95%	95%	95%	95%
7	Stacking LR + Hyperparameter	96%	97%	96%	96%
8	Stacking Adaboost + Hyperparameter	95%	95%	95%	95%
9	Stacking LR + Hyperparameter + SMOTE	97%	97%	97%	97%
10	Stacking Adaboost + Hyperparameter + SMOTE	98%	98%	98%	98%

The table presented shows the results of tests conducted after adding hyperparameter tuning to various machine learning models, both individual algorithms and model combinations using stacking techniques. After hyperparameter tuning was performed, KNN showed an increase in accuracy to 96%, with precision at 97%, recall at 96%, and F1-score at 96%, indicating that tuning successfully optimized KNN's performance. C4.5 also experienced an improvement, with all performance metrics at 96%, demonstrating that hyperparameter tuning helped maintain the model's overall performance. For GNB, although there was a slight increase in accuracy and precision after tuning, its performance remained lower than that of KNN and C4.5, with an accuracy of 94% and an F1-score of 93%.

When SMOTE was applied alongside hyperparameter tuning, KNN achieved a higher accuracy of 97%, with all other metrics also at the same level. This shows that the combination of SMOTE and hyperparameter tuning is very effective in improving KNN's ability to detect patterns from the minority class. On the other hand, with C4.5, the addition of SMOTE after tuning reduced accuracy to 95%, which may be due to the additional complexity from the synthetic data generated by SMOTE. GNB, after the application of SMOTE, showed an improvement in performance, with all metrics increasing to 95%, indicating that GNB was able to effectively utilize the more balanced data, although the improvement was not as significant as with KNN.

In testing with stacking techniques, the use of Logistic Regression as the meta model after tuning resulted in an accuracy of 96%, with precision at 97%, recall at 96%, and F1-score at 96%. This indicates that the stacking model also benefited from tuning, providing very balanced and accurate results. When AdaBoost was used as the meta model, the results were slightly lower, with all metrics at 95%, indicating that Logistic Regression may be more effective in leveraging hyperparameter tuning in this combination.

The best results were obtained when SMOTE was applied together with hyperparameter tuning and stacking. The combination of Logistic Regression as the meta model with SMOTE and tuning resulted in an accuracy of 97%, while the combination with AdaBoost as the meta model provided the greatest increase, with accuracy and all other performance metrics reaching 98%. This shows that the use of SMOTE and hyperparameter tuning, especially when combined with AdaBoost as the meta model, can produce very high performance in detecting patterns from a more balanced dataset. Furthermore, using Receiver Operating Characteristic (ROC) needs to be done to measure the ability of the model to distinguish classes, compare the performance of several models, and determine the optimal threshold [50]. Figure 6 is the ROC result of Stacking Adaboost + Hyperparameter + SMOTE.

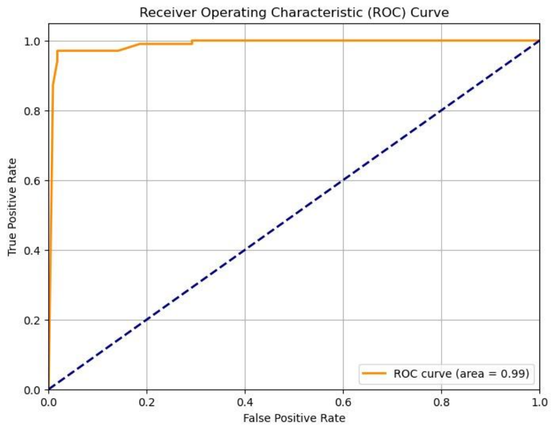

Figure 6. ROC

The Stacking model with AdaBoost, Hyperparameter Tuning, and SMOTE shows excellent performance with an AUC of 0.99. This indicates the model's near-perfect ability to distinguish between positive and negative classes, with a low error rate in classification. The combination of these methods successfully improves sensitivity to the minority class without compromising overall accuracy.

4.4. Overall Discussion of Test Result

In this study, various data processing techniques and machine learning models were tested to detect breast cancer based on the available dataset. The testing involved base algorithms such as KNN, C4.5, and GNB, both individually and in combination with model aggregation techniques (stacking) and hyperparameter tuning. Additionally, the SMOTE was applied to balance class distribution in a dataset that was initially imbalanced between 'benign' and 'malignant' classes.

In the initial tests without applying SMOTE or hyperparameter tuning, KNN, C4.5, and GNB showed good performance with accuracies of 95%, 96%, and 93%, respectively. However, when SMOTE was applied to balance the dataset, there was a significant increase in performance for KNN and GNB, with accuracies rising to 97% and 95%, respectively. This increase indicates that SMOTE is effective in helping models better recognize patterns from the minority class, especially in detecting critical malignant cancer cases, which are fewer in the dataset.

Next, the application of hyperparameter tuning also showed a positive impact on model performance. For KNN, tuning successfully increased accuracy to 96%, while C4.5 showed a similar improvement. GNB, although it improved after tuning, still showed lower performance compared to the other two algorithms. When SMOTE was combined with hyperparameter tuning, KNN showed the highest performance increase, with accuracy reaching 97%. C4.5 and GNB achieved accuracies of 95%. These results confirm that the combination of SMOTE and hyperparameter tuning provides optimal results for KNN. However, for C4.5, the additional complexity from the synthetic data may have reduced model performance.

Further tests were conducted using stacking techniques, where LR and AdaBoost were used as meta models to combine the strengths of several base models (KNN, C4.5, and GNB). The results showed that when LR was used as the meta model, the achieved accuracy was 96%, with balanced precision, recall, and F1-score. Conversely, AdaBoost as the meta model showed a slight decrease in precision, although it still produced strong results. However, the combination of SMOTE with stacking and hyperparameter tuning resulted in the most significant performance increase. When LR was used as the meta model with the combination of SMOTE and tuning, accuracy reached 97%. Meanwhile, AdaBoost in the same combination achieved the highest accuracy of 98%, with all performance metrics at the same level.

Overall, this study shows that techniques such as SMOTE, hyperparameter tuning, and stacking can significantly improve the performance of machine learning models in detecting breast cancer. SMOTE has proven to be very useful in addressing class imbalance issues, while hyperparameter tuning helps optimize model parameters for better performance. The stacking technique, especially when combined with AdaBoost and tuning, provides very satisfactory results, showing great potential in developing more accurate and reliable predictive models for medical applications. This study also improves on previous research, as shown in table 6.

Table 6. Comparison With Previous Research

No	Researcher	Dataset	Model	Accuracy
1	[51]	Heart Disease	Weighted KNN	93.28%
2	[52]	Disease Prediction from Various Symptoms	Naïve Bayes (The proposed optimized model by FCBF, PSO and ACO)	86.10%
3	[53]	Skin Cancer	SVM	96.90%
4	[54]	Breast Cancer	Ensemble Meta-Model	90.00%
5	[55]	Heart Disease	Voting Ensemble (DT, RF, NB, LR, SVM, Gradient Boosting, XGBoost)	96.75%
6	[56]	Heart Disease	KNN	91.99%
7	[57]	Heart Disease	XGBoost with GridSearchCV	87.02%
8	[58]	Breast Cancer	Random Forest	80.00%
9	[59]	Breast Cancer	Random Forest	96.49%
10	This Study	Breast Cancer	Stacking Adaboost + Hyperparameter + SMOTE	98.00%

Table 5 presents a comparison between your study results and previous research that used various models and datasets for disease prediction. In the first study, which used a Weighted KNN model to detect heart disease, the achieved accuracy was 93.28%. This shows that the weighted KNN model performs well in detecting heart disease. The second study used a Naïve Bayes model optimized with techniques such as FCBF, Particle Swarm Optimization (PSO), and Ant Colony Optimization (ACO) to predict various diseases based on symptoms. Despite these optimizations, the achieved accuracy was 86.1%, indicating decent performance, though it was still lower than several other models used in this study.

The third study used a Support Vector Machine (SVM) to detect skin cancer and achieved an accuracy of 96.9%. This demonstrates the effectiveness of SVM in complex classification tasks like skin cancer detection. The fourth study involved the use of a ensemble meta model to breast cancer classification in general, with an accuracy of 90%, indicating that ensemble meta model is a very strong model for disease prediction. In the fifth study, an ensemble voting approach combining several models such as DT, RF, NB, LR, SVM, Gradient Boosting, and XGBoost was used to detect heart disease, achieving an accuracy of 96.75%. This accuracy highlights the advantage of the ensemble approach in enhancing model performance.

The sixth and seventh studies also focused on heart disease prediction, with the KNN model achieving an accuracy of 91.99%, and XGBoost with GridSearchCV achieving an accuracy of 87.02%. These studies show that while KNN provides good performance, optimizing the model with GridSearchCV can result in lower accuracy, suggesting that dataset complexity and characteristics have a significant impact. The eighth and ninth studies focused on detecting breast cancer using a Random Forest model. The results showed significant accuracy variation, with one study achieving only 80% accuracy, while another achieved 96.49%, indicating that model performance can vary greatly depending on specific implementation and data quality.

This study used a Stacking approach with AdaBoost as the meta model, combined with hyperparameter tuning and SMOTE, yielding the highest accuracy of 98%. This places the current study above the others in Table 6 showing that this combination of optimization techniques is highly effective in detecting breast cancer, providing a significant improvement over other models tested on the same or similar datasets.

5. Conclusion

This study successfully demonstrated that the combination of Stacking technique, AdaBoost as a meta model, hyperparameter tuning, and SMOTE can significantly enhance the performance of machine learning models in detecting breast cancer. Using an imbalanced dataset, the application of SMOTE proved highly effective in balancing class distribution, thereby improving the model's sensitivity in detecting malignant tumors. Hyperparameter tuning also played a crucial role in optimizing each model's performance, leading to higher accuracy and balanced performance metrics. The best results were achieved with the combination of Stacking with AdaBoost and SMOTE, which reached

ISSN 2723-6471

an accuracy of 98%, surpassing other models in previous studies. In conclusion, this comprehensive approach not only improves model accuracy but also enhances model reliability in clinical applications, making it a potential tool for early breast cancer diagnosis. Further research could explore applying this technique to larger and more diverse datasets to ensure broader generalization of the proposed model.

6. Declarations

6.1. Author Contributions

Conceptualization: L.L.V.F.C., M.K.A., S.B., A.K.M., S.S., R.L.V.N.; Methodology: S.S.; Software: L.L.V.F.C.; Validation: L.L.V.F.C., S.S., dan R.L.V.N.; Formal Analysis: L.L.V.F.C., S.S., dan R.L.V.N.; Investigation: L.L.V.F.C.; Resources: S.S.; Data Curation: S.S.; Writing Original Draft Preparation: L.L.V.F.C., S.S., dan R.L.V.N.; Writing Review and Editing: S.S., L.L.V.F.C., dan R.L.V.N.; Visualization: L.L.V.F.C.; All authors have read and agreed to the published version of the manuscript.

6.2. Data Availability Statement

The data presented in this study are available on request from the corresponding author.

6.3. Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

6.4. Institutional Review Board Statement

Not applicable.

6.5. Informed Consent Statement

Not applicable.

6.6. Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] J. Kufel, K. Bargieł-Łączek, S. Kocot, M. Koźlik, W. Bartnikowska, M. Janik, Ł. Czogalik, P. Dudek, M. Magiera, A. Lis, et al. "What is machine learning, artificial neural networks and deep learning?—Examples of practical applications in medicine," *Diagnostics*, vol. 13, no. 15, pp. 1–22, 2023, doi: 10.3390/diagnostics13152582.

[2] N. Seman and N. A. Razmi, "Machine learning-based technique for big data sentiments extraction," *IAES International Journal of Artificial Intelligence*, vol. 9, no. 3, pp. 473–479, Sep. 2020, doi: 10.11591/ijai.v9.i3.pp473-479.

[3] B. van Giffen, D. Herhausen, and T. Fahse, "Overcoming the pitfalls and perils of algorithms: A classification of machine learning biases and mitigation methods," *Journal of Business Research*, vol. 144, no. 1, pp. 93–106, May 2022, doi: 10.1016/j.jbusres.2022.01.076.

[4] M. M. Taye, "Understanding of machine learning with deep learning: Architectures, workflow, applications and future directions," *Computers*, vol. 12, no. 5, pp. 1–26, May 2023, doi: 10.3390/computers12050091.

[5] Y. Yunitasari, A. Musdholifah, and A. K. Sari, "Sarcasm detection for sentiment analysis in Indonesian tweets," *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, vol. 13, no. 1, pp. 53–62, 2019, doi: 10.22146/ijccs.41136.

[6] C. J. Varshney, A. Sharma, and D. P. Yadav, "Sentiment analysis using ensemble classification technique," *in 2020 IEEE Students' Conference on Engineering and Systems (SCES), Institute of Electrical and Electronics Engineers Inc.*, Jul. 2020, pp. 1–20, doi: 10.1109/SCES50439.2020.9236754.

[7] B. Brahimi, M. Touahria, and A. Tari, "Improving sentiment analysis in Arabic: A combined approach," *Journal of King Saud University - Computer and Information Sciences*, vol. 33, no. 10, pp. 1242–1250, 2019, doi: 10.1016/j.jksuci.2019.07.011.

[8] P. A. R. Mukhlashin, A. Fitrianto, A. M. Soleh, and W. Z. A. Wan Muhamad, “Ensemble learning with imbalanced data handling in the early detection of capital markets,” *Journal of Accounting and Investment*, vol. 24, no. 2, pp. 600–617, May 2023, doi: 10.18196/jai.v24i2.17970.

[9] T. Kawichai, "Stacking ensemble learning with regression models for predicting damage from terrorist attacks," *ECTI Transactions on Computer and Information Technology*, vol. 18, no. 3, pp. 250–259, 2024, doi: 10.37936/ecti-cit.2024183.255276.

[10] M. Arifin, W. Widowati, and F. Farikhin, "Optimization of hyperparameters in machine learning for enhancing predictions of student academic performance," *Ingenierie des Systemes d'Information*, vol. 28, no. 3, pp. 575–582, Jun. 2023, doi: 10.18280/isi.280305.

[11] T. Zhao, Y. Zheng, and Z. Wu, "Feature selection-based machine learning modeling for distributed model predictive control of nonlinear processes," *Computers and Chemical Engineering*, vol. 169, no. 1, pp. 1–17, Jan. 2023, doi: 10.1016/j.compchemeng.2022.108074.

[12] X. T. Dang and T. T. Le, "KNN-SMOTE: An innovative resampling technique enhancing the efficacy of imbalanced biomedical classification," *in Machine Learning and Other Soft Computing Techniques: Biomedical and Related Applications, N. Hoang Phuong, N. T. Huyen Chau, and V. Kreinovich, Eds., Cham: Springer Nature Switzerland*, vol. 2024, no.1, pp. 111–121, 2024, doi: 10.1007/978-3-031-63929-6_11.

[13] T. P. Handayani, W. Hasyim, and Nursetia, “Preliminary Evaluation of Gaussian Naive Bayes for Multi-Label Hate Speech and Abusive Language Detection on Indonesian Twitter,” *Journal of International Multidisciplinary Research*, vol. 1, no. 1, pp. 159–165, Nov. 2023, doi: 10.62504/jimr532.

[14] I. D. Mienye and N. Jere, “A Survey of Decision Trees: Concepts, Algorithms, and Applications,” *IEEE Access*, vol. 12, no. 1, pp. 86716–86727, Jun. 2024, doi: 10.1109/ACCESS.2024.3416838.

[15] T. A. Assegie, "An optimized K-nearest neighbor based breast cancer detection," *Journal of Robotics and Control (JRC)*, vol. 2, no. 3, pp. 115–118, May 2021, doi: 10.18196/jrc.2363.

[16] D. Zuo, L. Yang, Y. Jin, H. Qi, Y. Liu, and L. Ren, "Machine learning-based models for the prediction of breast cancer recurrence risk," *BMC Medical Informatics and Decision Making*, vol. 23, no. 1, pp. 1–14, Dec. 2023, doi: 10.1186/s12911-023-02377-z.

[17] R. H. Saputra and B. Prasetyo, “Improve the Accuracy of C4.5 Algorithm Using Particle Swarm Optimization (PSO) Feature Selection and Bagging Technique in Breast Cancer Diagnosis,” *Journal of Soft Computing Exploration*, vol. 1, no. 1, pp. 47–55, Oct. 2020, doi: 10.52465/joscex.v1i1.9.

[18] J. H. Joloudari, A. Marefat, M. A. Nematollahi, S. S. Oyelere, and S. Hussain, “Effective Class-Imbalance Learning Based on SMOTE and Convolutional Neural Networks,” *Applied Sciences (Switzerland)*, vol. 13, no. 6, Mar. 2023, doi: 10.3390/app13064006.

[19] M. M. R. Mubarak, Y. H. Chrisnanto, and P. N. Sabrina, “Implementation of Random Forest Using Smote and Smoteenn in Customer Churn Classification in E-Commerce,” *Enrichment: Journal of Multidisciplinary Research and Development*, vol. 1, no. 8, pp. 463–477, Nov. 2023, doi: 10.55324/enrichment.v1i8.69.

[20] L. L. Van Fc, M. K. Anam, M. B. Firdaus, Y. Yunefri, and N. A. Rahmi, "Enhancing machine learning model performance in addressing class imbalance," *COGITO Smart Journal*, vol. 10, no. 1, pp. 478–490, 2024.

[21] Hartono and E. Ongko, "Avoiding overfitting and overlapping in handling class imbalanced using hybrid approach with smoothed bootstrap resampling and feature selection," *International Journal on Informatics Visualization*, vol. 2, no. 6, pp. 343–348, 2022, doi: 10.30630/joiv.6.2.985.

[22] N. R. Octavianto and A. Wibowo, "Stacking classifier method for prediction of human body performance," Indonesian Journal of Electrical Engineering and Computer Science, vol. 34, no. 3, pp. 1832–1839, Jun. 2024, doi: 10.11591/ijeecs.v34.i3.pp1832-1839.

[23] J. Liu, Y. Ma, W. Xie, X. Li, Y. Wang, Z. Xu, Y. Bai, P. Yin, & Q. Wu, "Lasso-based machine learning algorithm for predicting postoperative lung complications in elderly: A single-center retrospective study from China," *Clinical Interventions in Aging*, vol. 18, no.1, pp. 597–606, 2023, doi: 10.2147/CIA.S406735.

[24] A. Ghasemieh, A. Lloyed, P. Bahrami, P. Vajar, and R. Kashef, "A novel machine learning model with stacking ensemble learner for predicting emergency readmission of heart-disease patients," *Decision Analytics Journal*, vol. 7, no. 1, pp. 1–13, Jun. 2023, doi: 10.1016/j.dajour.2023.100242.

[25] Q. Sui and S. K. Ghosh, "Active learning for stacking and AdaBoost-related models," *Stats (Basel)*, vol. 7, no. 1, pp. 110–137, Mar. 2024, doi: 10.3390/stats7010008.

[26] S. Hadianti and W. A. G. Kodri, "Optimization of the machine learning approach using Optuna in heart disease prediction," *Journal Medical Informatics Technology*, vol. 1, no. 3, pp. 59–64, Sep. 2023, doi: 10.37034/medinftech.v1i3.15.

[27] R. R. Sarra, I. I. Gorial, R. R. Manea, A. E. Korial, M. Mohammed, and Y. Ahmed, "Enhanced stacked ensemble-based heart disease prediction with Chi-square feature selection method," *Journal of Robotics and Control (JRC)*, vol. 5, no. 6, pp. 1–8, 2024, doi: 10.18196/jrc.v5i6.23191.

[28] K. Natarajan, V. V. Kumar, T. R. Mahesh, M.Abbas, N. Kathamuthu, E. Mohan & J. R. Annand, "Efficient heart disease classification through stacked ensemble with optimized firefly feature selection," *International Journal of Computational Intelligence Systems*, vol. 17, no. 1, pp. 1–14, Dec. 2024, doi: 10.1007/s44196-024-00538-0.

[29] S. Mohapatra, S. Maneesha, S. Mohanty, P. K. Patra, S. K. Bhoi, K. S. Sahoo, A. H. Gandomi, "A stacking classifiers model for detecting heart irregularities and predicting cardiovascular disease," *Healthcare Analytics*, vol. 3, no. 1, pp. 1–10, Nov. 2023, doi: 10.1016/j.health.2022.100133.

[30] N. Abuzinadah, P. S. Kumar, A. A. Alarfaj, E. A. Alabdulqader, M. Umer, T.H. Kim, S. Alsubai, Ashraf, "Improved prediction of ovarian cancer using ensemble classifier and SHAP explainable AI," *Cancers (Basel)*, vol. 15, no. 24, pp. 1–21, Dec. 2023, doi: 10.3390/cancers15245793.

[31] C. C. Chiu, C. M. Wu, T. N. Chien, L. J. Kao, C. Li, and H. L. Jiang, "Applying an improved stacking ensemble model to predict the mortality of ICU patients with heart failure," *Journal of Clinical Medicine*, vol. 11, no. 21, pp. 1–10, Nov. 2022, doi: 10.3390/jcm11216460.

[32] Herianto, B. Kurniawan, Z. H. Hartomi, Y. Irawan, and M. K. Anam, "Machine learning algorithm optimization using stacking technique for graduation prediction," *Journal of Applied Data Sciences*, vol. 5, no. 3, pp. 1272–1285, Sep. 2024, doi: 10.47738/jads.v5i3.316.

[33] H. Hairani, A. Anggrawan, and D. Priyanto, "Improvement performance of the random forest method on unbalanced diabetes data classification using SMOTE-Tomek link," *International Journal on Informatics Visualization*, vol. 1, no. 7, pp. 258–264, Mar. 2023, doi: 10.30630/joiv.7.1.1069.

[34] B. L. V. S. R. Krishna, V. Mahalakshmi, and G. K. M. Nukala, "A stacking model for outlier prediction using learning approaches," *International Journal of Intelligent Systems and Applications in Engineering*, vol. 12, no. 2S, pp. 629–638, 2023.

[35] A. P. Joshi and B. V. Patel, “Data Preprocessing: The Techniques for Preparing Clean and Quality Data for Data Analytics Process,” *Oriental journal of computer science and technology*, vol. 13, no. 0203, pp. 78–81, Jan. 2021, doi: 10.13005/ojcst13.0203.03.

[36] Y. Zhao, Z. Huang, L. Gong, Y. Zhu, Q. Yu, and Y. Gao, "Evaluating the impact of data transformation techniques on the performance and interpretability of software defect prediction models," *IET Software*, vol. 2023, no. 1, pp. 1–30, Nov. 2023, doi: 10.1049/2023/6293074.

[37] R. M. West, "Best practice in statistics: The use of log transformation," *Annals of Clinical Biochemistry*, vol. 59, no. 3, pp. 162–165, May 2022, doi: 10.1177/00045632211050531.

[38] S. Alam, M. S. Ayub, S. Arora, and M. A. Khan, "An investigation of the imputation techniques for missing values in ordinal data enhancing clustering and classification analysis validity," *Decision Analytics Journal*, vol. 9, no. 1, pp. 1–4, Dec. 2023, doi: 10.1016/j.dajour.2023.100341.

[39] P. Mishra, C. M. Pandey, U. Singh, A. Gupta, C. Sahu, and A. Keshri, "Descriptive statistics and normality tests for statistical data," *Annals of Cardiac Anaesthesia*, vol. 22, no. 1, pp. 67–72, Jan. 2019, doi: 10.4103/aca.ACA_157_18.

[40] J. M. de la Rubia, "Standardized distance from the mean to the median as a measure of skewness," *Open Journal of Statistics*, vol. 13, no. 3, pp. 359–378, 2023, doi: 10.4236/ojs.2023.133018.

[41] N. G. Ramadhan, "Comparative analysis of ADASYN-SVM and SMOTE-SVM methods on the detection of type 2 diabetes mellitus," *Scientific Journal of Informatics*, vol. 8, no. 2, pp. 276–282, Nov. 2021, doi: 10.15294/sji.v8i2.32484.

[42] M. Khairy, T. M. Mahmoud, and T. Abd-El-Hafeez, "The effect of rebalancing techniques on the classification performance in cyberbullying datasets," *Neural Computing and Applications*, vol. 36, no. 3, pp. 1049–1065, Jan. 2024, doi: 10.1007/s00521-023-09084-w.

[43] M. Pagan, M. Zarlis, and A. Candra, “Investigating the impact of data scaling on the k-nearest neighbor algorithm,” *Computer Science and Information Technologies*, vol. 4, no. 2, pp. 135–142, Jul. 2023, doi: 10.11591/csit.v4i2.pp135-142.

[44] H. T. Pham, J. Awange, and M. Kuhn, “Evaluation of Three Feature Dimension Reduction Techniques for Machine Learning-Based Crop Yield Prediction Models,” *Sensors*, vol. 22, no. 17, pp. 1–18, Sep. 2022, doi: 10.3390/s22176609.

[45] K. Satyanarayana, D. Shankar, and D. Raju, "An approach for finding emotions using seed dataset with KNN classifier," *Turkish Journal of Computer and Mathematics Education*, vol. 12, no. 10, pp. 2838–2846, 2021, doi: 10.17762/turcomat.v12i10.4930.

[46] Y. Q. Song, X. Yao, Z. Liu, X. Shen, and J. Mao, "An improved C4.5 algorithm in bagging integration model," *IEEE Access*, vol. 8, no.1 , pp. 206866–206875, 2020, doi: 10.1109/ACCESS.2020.3032291.

[47] V. Vichianchai and S. Kasemvilas, "A New Term Frequency with Gaussian Technique for Text Classification and Sentiment Analysis," *Journal of ICT Research and Applications*, vol. 15, no. 2, pp. 152–168, Oct. 2021, doi: 10.5614/itbj.ict.res.appl.2021.15.2.4.

[48] S. Mohapatra, I. Mishra, and S. Mohanty, "Stacking model for heart stroke prediction using machine learning techniques," *EAI Endorsed Transactions on Pervasive Health and Technology*, vol. 9, no. 1, pp. 1–5, May 2023, doi: 10.4108/eetpht.9.4057.

[49] P. Srinivas and R. Katarya, "hyOPTXg: OPTUNA hyper-parameter optimization framework for predicting cardiovascular disease using XGBoost," *Biomed Signal Process Control*, vol. 73, no. 1, pp. 1–10, Mar. 2022, doi: 10.1016/j.bspc.2021.103456.

[50] S. G. Begum and P. K. Sree, "Drug recommendation using a 'reviews and sentiment analysis' by a recurrent neural network," *Indonesian Journal of Multidisciplinary Science*, vol. 2, no. 9, pp. 3085–3094, Jun. 2023, doi: 10.55324/ijoms.v2i9.530.

[51] K. A. Barry, Y. Manzali, M. Lamrini, F. Rachid, and M. Elfar, "Heart Disease Prediction Using Weighted K -Nearest Neighbor Algorithm," *Operations Research Forum*, vol. 76, no. 5, pp. 1–16, Aug. 2024, doi: 10.1007/s43069-024-00356-2.

[52] Y. Khourdifi and M. Bahaj, "Heart disease prediction and classification using machine learning algorithms optimized by particle swarm optimization and ant colony optimization," *International Journal of Intelligent Engineering and Systems*, vol. 12, no. 1, pp. 242–252, 2019, doi: 10.22266/ijies2019.0228.24.

[53] P. Banasode, M. Patil, and N. Ammanagi, "A melanoma skin cancer detection using machine learning technique: Support vector machine," in *IOP Conference Series: Materials Science and Engineering*, IOP Publishing Ltd, Feb. 2021, pp. 1–5, doi: 10.1088/1757-899X/1065/1/012039.

[54] M.D. Ali, A. Saleem, H. Elahi, M.A. Khan, M. I. Khan, M. M. Yaqoob, K. U. Farooq, A. Al-Rasheed, "Breast Cancer Classification through Meta-Learning Ensemble Technique Using Convolution Neural Networks," *Diagnostics*, vol. 13, no. 13, pp. 1–19, Jul. 2023, doi: 10.3390/diagnostics13132242.

[55] B. P. Doppala, D. Bhattacharyya, M. Janarthanan, and N. Baik, "A reliable machine intelligence model for accurate identification of cardiovascular diseases using ensemble techniques," *Journal of Healthcare Engineering*, vol. 2022, no. 1, pp. 1–13, 2022, doi: 10.1155/2022/2585235.

[56] T. A. Assegie, "Heart disease prediction model with k-nearest neighbor algorithm," *International Journal of Informatics and Communication Technology (IJ-ICT)*, vol. 10, no. 3, pp. 225–230, Dec. 2021, doi: 10.11591/ijict.v10i3.pp225-230.

[57] C. M. Bhatt, P. Patel, T. Ghetia, and P. L. Mazzeo, "Effective heart disease prediction using machine learning techniques," *Algorithms*, vol. 16, no. 2, pp. 1–14, Feb. 2023, doi: 10.3390/a16020088.

[58] R. Rabiei, S. M. Ayyoubzadeh, S. Sohrabei, M. Esmaeili, and A. Atashi, "Prediction of breast cancer using machine learning approaches," *Journal of Biomedical Physics and Engineering*, vol. 12, no. 3, pp. 297–308, Jun. 2022, doi: 10.31661/jbpe.v0i0.2109-1403.

[59] A. Khalid, A. Mehmood, A. Alabrah, B. F. Alkhamees, F. Amin, H. AlSalman, G. S. Choi, "Breast cancer detection and prevention using machine learning," *Diagnostics*, vol. 13, no. 19, pp. 1–21, Oct. 2023, doi: 10.3390/diagnostics13193113.