

Comparative evaluation of supervised machine learning algorithms for breast cancer prediction using the Wisconsin diagnostic dataset

Sunday Akinwamide *, Taiwo Fele and Olufemi Ariyo Ojo

Department of Computer Science, The Federal Polytechnic, Ado Ekiti, Nigeria.

Global Journal of Engineering and Technology Advances, 2025, 24(02), 196-203

Publication history: Received on 10 July 2025; revised on 17 August 2025; accepted on 19 August 2025

Article DOI: <https://doi.org/10.30574/gjeta.2025.24.2.0246>

Abstract

Breast cancer remains a leading cause of cancer-related mortality among women globally. Early diagnosis is critical to improving survival outcomes, and machine learning (ML) models have shown promise in enhancing predictive accuracy in medical diagnostics. This study presents a comprehensive comparative evaluation of six supervised ML algorithms: Logistic Regression (LR), Random Forest (RF), Support Vector Machine (SVM), K-Nearest Neighbours (KNN), Decision Tree (DT), and Gaussian Naive Bayes (GNB), applied to the Breast Cancer Wisconsin (Diagnostic) dataset. Using RandomForestClassifier for feature importance ranking, SMOTE to address class imbalance, and StandardScaler for normalization, the models were trained and evaluated using an 80-20 train-test split. Performance was assessed based on accuracy, precision, recall, and F1-score. Logistic Regression achieved the highest overall performance (97.90% accuracy, 100% precision). Results indicate that linear models can outperform complex classifiers under well-prepared conditions. This study contributes to ML-aided diagnostics by identifying optimal algorithms for breast cancer prediction using clinical imaging-derived features.

Keywords: Breast cancer prediction; Supervised learning; Classification algorithms; SMOTE; Feature selection; Model evaluation; Medical AI

1. Introduction

Breast cancer remains one of the most diagnosed malignancies worldwide and represents a significant contributor to cancer-related mortality among women (Sung et al., 2021). Early detection and accurate diagnosis are critical for improving patient outcomes, reducing treatment complexity, and increasing survival rates (Tabár et al., 2000). In response to the increasing need for reliable and scalable diagnostic solutions, machine learning (ML) has emerged as a powerful approach for predictive analytics in the medical domain. These techniques enable healthcare professionals to process complex and high-dimensional datasets, offering valuable insights that can enhance diagnostic accuracy and inform clinical decisions (Zhu et al., 2019).

A prominent dataset frequently utilized in breast cancer research is the Breast Cancer Wisconsin (Diagnostic) dataset. It comprises features extracted from digitized images of fine needle aspirates (FNA) of breast tissue, capturing structural attributes of cell nuclei that are relevant for malignancy detection. Due to its high reliability and diagnostic clarity, this dataset has served as a benchmark for numerous machine learning classification studies (Mangasarian et al., 1995).

This study conducts a comparative analysis of six widely used supervised ML algorithms: Logistic Regression (LR), Random Forest (RF), Gaussian Naive Bayes (GNB), K-Nearest Neighbours (KNN), Decision Tree (DT), and Support Vector Machine (SVM) to evaluate their effectiveness in predicting breast cancer diagnoses. These models span both

* Corresponding author: Sunday Akinwamide.

linear and non-linear classifiers, offering diverse methodological strengths for binary classification tasks (Sayed et al., 2019).

To enhance the diagnostic performance of the models, the RandomForestClassifier is employed to identify the most important features within the dataset. Feature selection plays a vital role in improving model efficiency by reducing dimensionality and eliminating noise or redundancy. This focused approach allows the classifiers to better distinguish between malignant and benign cases (Zhu et al., 2019).

A major challenge in medical datasets, particularly in breast cancer diagnosis, is class imbalance. Typically, non-cancerous (benign) cases significantly outnumber malignant ones, leading to biased model performance. To address this, the Synthetic Minority Over-sampling Technique (SMOTE) is applied, which generates synthetic examples of the minority class through interpolation, thereby balancing the dataset and enhancing model generalization (Fernández, García, et al., 2018).

Model performance is assessed using four key evaluation metrics: accuracy, precision, recall, and F1 score. Accuracy provides a general measure of classification correctness, while precision and recall focus on the model's capability to identify true positives and minimize false positives. The F1 score, as a harmonic mean of precision and recall, serves as a balanced metric, particularly useful for imbalanced classification problems (Tharwat, 2020).

This study aims to offer a systematic comparison of the six ML algorithms in the context of breast cancer prediction, examining their strengths, limitations, and potential for clinical deployment. By integrating feature selection and data balancing techniques, the study contributes to advancing the reliability and interpretability of ML-based diagnostic systems in oncology.

2. Related Work

The application of ML techniques for breast cancer prediction and diagnosis has gained substantial momentum in recent years. This growing interest is driven by the increasing availability of high-quality medical datasets and the advancement of computational algorithms that enable robust predictive modeling (Almeida et al., 2020). Numerous studies have employed supervised ML algorithms to enhance early detection, a critical component in improving treatment outcomes and survival rates.

Chaurasia and Pal (2020) explored the performance of multiple supervised learning models, including Support Vector Machine (SVM), K-Nearest Neighbours (KNN), Multilayer Perceptron (MLP), and ensemble stacking methods using the Wisconsin Diagnostic Breast Cancer (WDBC) dataset. Among these, SVM demonstrated the highest classification accuracy, surpassing 90%, thereby highlighting its suitability for binary diagnostic tasks.

In a comparative study, Islam et al. (2020) evaluated the effectiveness of SVM, KNN, Decision Tree (DT), Random Forest (RF), and Logistic Regression (LR) models using the Wisconsin Breast Cancer Dataset (WBCD). Their findings showed that Artificial Neural Networks (ANNs) outperformed all traditional classifiers, achieving an exceptional accuracy of 98.57%, precision of 97.82%, and an F1 score of 0.989, underscoring the predictive superiority of deep learning techniques in this domain.

Further emphasizing deep learning, Tiwari et al. (2020) investigated a range of models including KNN, SVM, DT, Naïve Bayes (NB), LR, RF, Convolutional Neural Networks (CNN), and ANN. Using the Breast Cancer Wisconsin (Diagnostic) dataset, their results revealed that ANN and CNN achieved the highest accuracy rates, with CNN reaching 99.3% and ANN 97.3%, affirming the power of feature learning in enhancing diagnostic performance.

In another study, Naji et al. (2021) implemented traditional classifiers such as SVM, RF, LR, DT (C4.5), and KNN on the same diagnostic dataset. Their analysis identified SVM as the most effective algorithm, with a top accuracy of 97.2%, reinforcing the model's consistency across multiple studies.

The work by Nemade and Fegade (2023) investigated a broad range of algorithms including LR, DT, RF, SVM, and KNN while incorporating advanced ensemble techniques such as XGBoost. Their study, based on the WBCD, found that both DT and XGBoost achieved the highest accuracy (97%), with XGBoost also obtaining the highest AUC score of 0.999. Their findings illustrated the performance benefits of boosting algorithms in cancer classification tasks.

Srivenkatesh (2020) also evaluated various ML models including SVM, RF, NB, and LR, applied to a dataset from Kaggle comprising diverse diagnostic features. Notably, RF emerged as the best-performing model, attaining an accuracy of 98.24%, which supports the model's robustness in handling nonlinear data distributions.

Rabiei et al. (2022) focused on a heterogeneous dataset containing demographic, laboratory, and mammographic features. Their study applied RF, Gradient Boosting Trees (GBT), and Multi-Layer Perceptron (MLP), coupled with a genetic algorithm for hyperparameter optimization. Despite the complexity of their feature set, RF achieved the best accuracy at 80%, suggesting potential for ensemble models in multi-modal data settings.

Shravya et al. (2019) employed LR, SVM, and KNN on the WBCD, concluding that SVM again outperformed the other classifiers, achieving an accuracy of 92.7%. This study aligns with others affirming the model's robustness in structured medical datasets.

Ensemble learning was further explored by Nanglia et al. (2022), who implemented both homogeneous (RF) and heterogeneous stacking ensembles combining KNN, SVM, and DT, alongside single models such as LR, NB, ANN, and SGD. Utilizing various K-fold cross-validation strategies, the stacking model outperformed individual classifiers, attaining a validation accuracy of 78%, while base classifiers individually achieved lower accuracy rates.

Lastly, Mahesh et al. (2024) evaluated a blended ensemble model combining SVM, KNN, DT, RF, and LR. Their model, tested on the WBCD, demonstrated an accurate improvement of 98.14%, validating the merit of integrating diverse base learners to maximize diagnostic precision.

Collectively, these studies highlight the versatility and effectiveness of supervised ML algorithms in breast cancer prediction. Models such as LR, RF, GNB, KNN, DT, and particularly SVM, have demonstrated varying strengths depending on dataset characteristics and preprocessing methods. Furthermore, enhancements such as feature selection with RandomForestClassifier and class imbalance correction using SMOTE have consistently been shown to boost model performance. However, a comprehensive evaluation of these algorithms under consistent preprocessing and evaluation criteria remains limited. This study addresses this gap by systematically comparing six ML algorithms using unified preprocessing protocols and a robust evaluation framework.

3. Materials and Methods

3.1. Overview

This section outlines the experimental design and implementation strategy adopted for evaluating the performance of six supervised machine learning algorithms in breast cancer prediction. The Breast Cancer Wisconsin (Diagnostic) dataset served as the foundation for model training and evaluation. To ensure fair and effective comparisons, key preprocessing steps were conducted, including feature selection using the RandomForestClassifier, class imbalance mitigation through the Synthetic Minority Over-sampling Technique (SMOTE), and feature scaling via StandardScaler. Performance evaluation was based on commonly used classification metrics, namely accuracy, precision, recall, and F1 score (Taha & Khedher, 2019; Fang & Yu, 2020).

3.2. Dataset Description

The Breast Cancer Wisconsin (Diagnostic) dataset, retrieved from the UCI Machine Learning Repository, is a widely used benchmark for binary classification tasks in medical machine learning research. It comprises 569 samples, each corresponding to a patient record characterized by 32 numerical attributes derived from digitized images of fine needle aspirates (FNA) of breast masses. These features include metrics such as radius, texture, smoothness, and concavity, which are instrumental in classifying tumors as benign or malignant (Ali & El-Baz, 2017).

- Total samples: 569
- Features: 32 continuous variables
- Target classes: Malignant (1) and Benign (0)
- Class imbalance: Benign cases significantly outnumber malignant ones

This dataset is recognized for its reliability and has served as a standard for both traditional and deep learning-based approaches to cancer prediction.

3.3. Data Preprocessing

Effective preprocessing is critical to the success of ML models. The dataset underwent several preprocessing steps to ensure it was appropriately formatted for training.

3.3.1. Missing Value Handling

The dataset was thoroughly examined for missing or inconsistent entries. No missing values were detected, allowing the dataset to be used without imputation or removal of entries (Bache & Lichman, 2013).

3.3.2. Feature Selection

Feature selection was performed using the RandomForestClassifier to rank input variables by importance. The top-ranking features were retained for training, while features with low predictive value were discarded. This approach enhances model interpretability, reduces overfitting risk, and improves computational efficiency (Fang & Yu, 2020).

3.3.3. Addressing Class Imbalance with SMOTE

To address the inherent class imbalance, SMOTE was employed to synthetically generate new samples from the minority class (malignant). This oversampling technique helps the classifier generalize better and minimizes bias towards the majority class (Fernández, Galar, et al., 2018).

3.3.4. Feature Scaling

StandardScaler was applied to normalize the feature space by ensuring all input variables had a mean of zero and a standard deviation of one. This transformation is essential for ML algorithms that rely on feature magnitude, such as SVM and LR (Jain & Jain, 2015).

3.4. Model Implementation

Each of the six models was implemented using Scikit-learn. A brief overview of each model is provided below:

3.4.1. Random Forest (RF)

An ensemble classifier that aggregates multiple decision trees trained on random subsets of features and data. This method reduces overfitting and enhances generalization (Biau & Scornet, 2016).

3.4.2. Decision Tree (DT)

A tree-based model that splits input data based on feature values. Although interpretable and efficient, DTs are prone to overfitting, which can be mitigated through pruning or ensemble methods (Zhang & Singer, 2010).

3.4.3. Logistic Regression (LR)

A probabilistic model used for binary classification. It estimates the probability that a sample belongs to a particular class based on one or more predictor variables (Peng et al., 2002).

3.4.4. K-Nearest Neighbours (KNN)

A non-parametric algorithm that classifies a sample based on the majority vote of its 'k' nearest neighbours. The distance between samples is calculated using metrics such as Euclidean distance (Zhang & Zhao, 2017).

3.4.5. Support Vector Machine (SVM)

A powerful classification model that seeks to find the optimal hyperplane separating two classes with the maximum margin. It supports non-linear decision boundaries using kernel functions (Schölkopf & Smola, 2002).

3.4.6. Gaussian Naive Bayes (GNB)

A probabilistic model based on Bayes' theorem, assuming that features are conditionally independent and normally distributed. Despite its simplicity, it performs competitively on certain datasets (Rish, 2001).

3.5. Evaluation Metrics

Model performance was assessed using four widely adopted metrics derived from the confusion matrix: accuracy, precision, recall, and F1 score (Deng et al., 2016).

- True Positive (TP): Correctly predicted malignant cases
- False Positive (FP): Benign cases wrongly predicted as malignant
- False Negative (FN): Malignant cases wrongly predicted as benign
- True Negative (TN): Correctly predicted benign cases

The formulas for each metric are as follows:

- Accuracy: $(TP + TN) / (TP + FP + TN + FN)$
- Precision: $TP / (TP + FP)$
- Recall (Sensitivity): $TP / (TP + FN)$
- F1 Score: $2 \times (Precision \times Recall) / (Precision + Recall)$

These metrics collectively offer a comprehensive evaluation of model effectiveness, particularly in the context of imbalanced datasets.

3.6. Model Training and Validation

The dataset was split into an 80% training set and a 20% test set to ensure the reliability of the evaluation process. Cross-validation and hyperparameter tuning were conducted during model training to prevent overfitting. Final model performance was reported on the unseen test set, providing an unbiased estimate of generalization capability.

4. Model Evaluation

To assess the predictive capabilities of each machine learning model, a thorough evaluation was conducted using the designated test set. While accuracy served as the primary metric, indicating the proportion of correctly classified samples, it was complemented by three additional performance indicators: precision, recall, and F1-score. Together, these metrics provide a holistic view of model effectiveness, especially in the context of imbalanced datasets commonly encountered in medical classification tasks (Khan & Ahmad, 2016).

The confusion matrix was used to compute these metrics and to visualize classification results in terms of true positives (TP), false positives (FP), false negatives (FN), and true negatives (TN). This approach enables a nuanced interpretation of each model's strengths, particularly its sensitivity to detect malignant cases and its ability to avoid false positives.

4.1. Performance Comparison

Table 1 and Figure 1 present a summary of each model's performance based on the four-evaluation metrics. These results highlight notable differences in predictive accuracy and reliability across the six classifiers.

Table 1 Performance Metrics for Supervised Learning Models

S/N	Model	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)
1	Logistic Regression (LR)	97.90	100.00	95.83	97.87
2	Support Vector Machine (SVM)	96.50	98.55	94.44	96.45
3	K-Nearest Neighbours (KNN)	95.80	100.00	91.67	95.65
4	Random Forest (RF)	95.10	94.52	95.83	95.17
5	Decision Tree (DT)	93.71	92.00	95.83	93.88
6	Gaussian Naive Bayes (GNB)	93.01	95.58	90.28	92.86

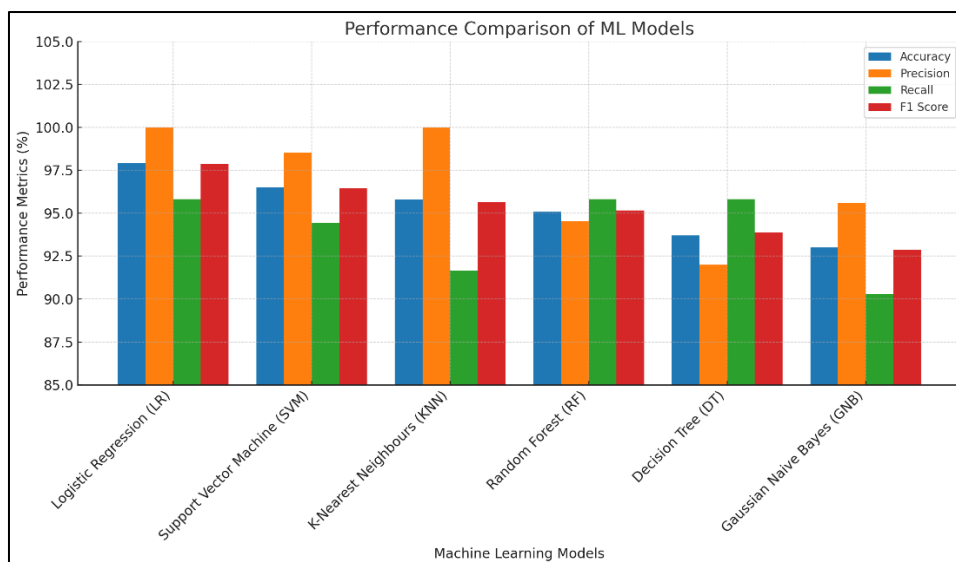


Figure 1 Performance metrics (Accuracy, Precision, Recall, F1 Score) for six ML models on the Breast Cancer Wisconsin (Diagnostic) dataset

5. Discussion

The experimental results revealed that Logistic Regression (LR) demonstrated the most effective performance among the six evaluated models, achieving the highest accuracy (97.90%) and perfect precision (100%). These results indicate LR's robust ability to correctly identify malignant cases while minimizing false positives. The combination of high recall (95.83%) and flawless precision positions LR as a highly dependable diagnostic tool, particularly in clinical scenarios where false alarms can lead to unnecessary stress and procedures for patients.

Support Vector Machine (SVM) and K-Nearest Neighbours (KNN) also delivered strong classification performance, achieving high precision and F1 scores. SVM's precision of 98.55% and recall of 94.44% indicate a well-balanced capability in differentiating between benign and malignant samples. KNN, while also attaining perfect precision (100.00%), recorded a slightly lower recall (91.67%), suggesting that it may occasionally fail to detect some malignant cases. In clinical practice, such false negatives are particularly undesirable, as they can delay critical treatment for affected individuals.

Both Random Forest (RF) and Decision Tree (DT) exhibited comparable recall values (95.83%), highlighting their sensitivity to malignant instances. However, these models showed moderate declines in precision, 94.52% for RF and 92.00% for DT, resulting in slightly lower overall accuracy and F1 scores. These results suggest that while tree-based models can identify true positives effectively, they may be more prone to false positives, potentially over-predicting malignancy.

Gaussian Naive Bayes (GNB), despite being computationally efficient and conceptually simple, underperformed in comparison to the other classifiers. It achieved the lowest accuracy (93.01%) and recall (90.28%), which can be attributed to its assumption of feature independence, a limitation that is often unrealistic in medical datasets where features are highly correlated. These deficiencies reduce GNB's reliability in complex diagnostic applications.

Overall, the results underscore the importance of selecting a model not solely based on accuracy but also on a balanced trade-off between precision and recall. In clinical contexts, especially for life-threatening conditions like breast cancer, a model's ability to minimize false negatives is often more critical than achieving perfect precision alone. Therefore, models like Logistic Regression and SVM, which maintain high levels of both precision and recall, may offer the most clinically valuable solutions for early breast cancer detection.

6. Conclusion and Future Work

This study presented a comparative analysis of six supervised machine learning algorithms for breast cancer prediction using the Breast Cancer Wisconsin (Diagnostic) dataset. Among the evaluated models, Logistic Regression emerged as

the most effective, achieving the highest accuracy and perfect precision, while Support Vector Machine and K-Nearest Neighbours also demonstrated strong performance. Tree-based models like Random Forest and Decision Tree offered high sensitivity but at the cost of slightly reduced precision. Gaussian Naive Bayes, although computationally efficient, showed limitations due to its simplifying assumptions.

The findings highlight the critical role of feature selection, data preprocessing, and model choice in developing reliable diagnostic systems. In clinical applications, models with balanced precision and recall, such as Logistic Regression and SVM are especially valuable for minimizing both false positives and false negatives.

Future work should explore the integration of ensemble methods (e.g., XGBoost, LightGBM), deep learning architectures, and real-world patient data to enhance generalizability. Additionally, implementing explainable AI (XAI) techniques could improve model interpretability, fostering greater trust and adoption in clinical settings.

Compliance with ethical standards

Disclosure of conflict of interest

No conflict of interest to be disclosed.

References

- [1] Ali, A. M., & El-Baz, A. (2017). Early detection of breast cancer using machine learning techniques. *Procedia Computer Science*, 105, 182–189. <https://doi.org/10.1016/j.procs.2017.11.026>
- [2] Almeida, S. M., de Melo, C. B., & Alves, A. C. (2020). Machine learning approaches in breast cancer: A review. *International Journal of Medical Informatics*, 140, 104164. <https://doi.org/10.1016/j.ijmedinf.2020.104164>
- [3] Bache, K., & Lichman, M. (2013). UCI Machine Learning Repository. University of California, Irvine. <http://archive.ics.uci.edu/ml>
- [4] Biau, G., & Scornet, E. (2016). A Random Forest Guided Tour. *Test*, 25(2), 197–227.
- [5] Chaurasia, V., & Pal, S. (2020). Applications of machine learning techniques to predict diagnostic breast cancer. *SN Computer Science*, 1(5), 270.
- [6] Deng, X., Liu, Q., Deng, Y., & Mahadevan, S. (2016). An improved method to construct basic probability assignment based on the confusion matrix. *Information Sciences*, 340, 250–261. <https://doi.org/10.1016/j.ins.2016.01.033>
- [7] Fang, Y., & Yu, L. (2020). A novel method for breast cancer diagnosis based on SMOTE and random forest. *IEEE Access*, 8, 203066–203078. <https://doi.org/10.1109/ACCESS.2020.3031631>
- [8] Fernández, A., Galar, M., Barrenechea, E., Derrac, J., & Herrera, F. (2018). An overview of ensemble methods for dealing with imbalanced data. *IEEE Transactions on Neural Networks and Learning Systems*, 29(6), 1992–2006.
- [9] Fernández, A., García, S., Galar, M., Prati, R. C., Krawczyk, B., & Herrera, F. (2018). *Learning from imbalanced data sets*. Springer International Publishing. <https://doi.org/10.1007/978-3-319-98074-4>
- [10] Islam, M. M., Haque, M. R., Iqbal, H., Hasan, M. M., Hasan, M., & Kabir, M. N. (2020). Breast cancer prediction: A comparative study using machine learning techniques. *SN Computer Science*, 1, 1–14.
- [11] Jain, A. K., & Jain, A. (2015). A survey of feature selection techniques in machine learning. *International Journal of Computer Applications*, 111(12), 29–34.
- [12] Khan, S., & Ahmad, S. (2016). A systematic review on performance metrics for classification algorithms. *International Journal of Computer Applications*, 140(1), 1–7. <https://doi.org/10.5120/19794-3755>
- [13] Mahesh, T. R., Vinoth, K. V., Vivek, V., Karthick, R. K. M., & Sindhu, M. G. (2024). Early predictive model for breast cancer classification using blended ensemble learning. *International Journal of System Assurance Engineering and Management*, 15(1), 188–197.
- [14] Mangasarian, O. L., Street, W. N., & Wolberg, W. H. (1995). Breast cancer diagnosis and prognosis via linear programming. *Operations Research*, 43(4), 570–577. <https://doi.org/10.1287/opre.43.4.570>
- [15] Naji, M. A., El Filali, S., Aarika, K., Benlahmar, E. H., Abdelouahid, R. A., & Debauche, O. (2021). Machine learning algorithms for breast cancer prediction and diagnosis. *Procedia Computer Science*, 191, 487–492.

- [16] Nanglia, S., Ahmad, M., Khan, F. A., & Jhanjhi, N. Z. (2022). An enhanced predictive heterogeneous ensemble model for breast cancer prediction. *Biomedical Signal Processing and Control*, 72, 103279.
- [17] Nemade, V., & Fegade, V. (2023). Machine learning techniques for breast cancer prediction. *Procedia Computer Science*, 218, 1314–1320.
- [18] Peng, C. Y. J., Lee, K. L., & Entinger, T. (2002). An Introduction to Logistic Regression Analysis and Reporting. *The Journal of Educational Research*, 96(1), 3–14.
- [19] Rabiei, R., Ayyoubzadeh, S. M., Sohrabei, S., Esmaeili, M., & Atashi, A. (2022). Prediction of breast cancer using machine learning approaches. *Journal of Biomedical Physics & Engineering*, 12(3), 297.
- [20] Rish, I. (2001). An empirical study of the naive Bayes classifier. In *IJCAI 2001 Workshop on Empirical Methods in Artificial Intelligence*, 41–46.
- [21] Sayed, G. I., Hassanien, A. E., & Abraham, A. (2019). Breast cancer detection using swarm optimization-based feature selection with k-nearest neighbor classifier. In *Proceedings of the 2019 9th International Conference on Computer Engineering and Bioinformatics* (pp. 77–81). <https://doi.org/10.1145/3365992.3366017>
- [22] Schölkopf, B., & Smola, A. (2002). *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*. MIT Press.
- [23] Shravya, C., Pravalika, K., & Subhani, S. (2019). Prediction of breast cancer using supervised machine learning techniques. *International Journal of Innovative Technology and Exploring Engineering (IJITEE)*, 8(6), 1106–1110.
- [24] Srivenkatesh, M. (2020). Prediction of breast cancer disease using machine learning algorithms. *International Journal of Innovative Technology and Exploring Engineering (IJITEE)*, 9(4), 2868–2878.
- [25] Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., & Bray, F. (2021). Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians*, 71(3), 209–249. <https://doi.org/10.3322/caac.21660>
- [26] Tabár, L., Vitak, B., Chen, H. H., Yen, M. F., Cohen, A., Tot, T., Chiu, S. Y., & Duffy, S. W. (2000). The Swedish two-county trial: Impact of mammographic screening on breast cancer mortality during 3 decades. *Radiology*, 260(3), 658–663. <https://doi.org/10.1148/radiol.11110469>
- [27] Taha, M. H., & Khedher, L. (2019). Comparative analysis of machine learning algorithms for breast cancer diagnosis. *Procedia Computer Science*, 159, 1663–1672.
- [28] Tharwat, A. (2020). Classification assessment methods. *Applied Computing and Informatics*, 17(1), 168–192. <https://doi.org/10.1016/j.aci.2018.08.003>
- [29] Tiwari, M., Bharuka, R., Shah, P., & Lokare, R. (2020). Breast cancer prediction using deep learning and machine learning techniques. *SSRN*. <https://doi.org/10.2139/ssrn.3558786>
- [30] Zhang, H., & Singer, Y. (2010). Generalization Bounds for Decision Trees. *Machine Learning*, 80(3), 277–299.
- [31] Zhang, H., & Zhao, J. (2017). A survey of K-nearest neighbour classification algorithm and its applications. *International Journal of Computer Applications*, 170(7), 22–28.
- [32] Zhu, W., Xie, L., Han, J., & Guo, X. (2019). The application of deep learning in cancer prognosis prediction. *Cancers*, 11(4), 603. <https://doi.org/10.3390/cancers11040603>