Cumhuriyet Sci. J., 46(2) (2025) 369-376
DOI: https://doi.org/10.17776/csj.1443598

Cumhuriyet Science Journal

| csj.cumhuriyet.edu.tr | Founded: 2002 ISSN: 2587-2680 e-ISSN: 2587-246X Publisher: Sivas Cumhuriyet University

# Enhanced Breast Cancer Risk Classification Through Genetic Algorithm-Based Feature Selection and Machine Learning Techniques

**Aynur Yonar[1,a,*], Harun Yonar[2,b], Öznur Özaltın[3,c]**

[1] *Department of Statistics, Faculty of Science, Selcuk University, Konya, Türkiye*

[2] *Department of Biostatistics, Faculty of Veterinary, Selcuk University, Konya, Türkiye*

[3] *Department of Actuarial Science, Faculty of Science, Hacettepe University, Ankara, Türkiye*

**Corresponding author*

**Research Article**

*History*
*Received: 27/02/2024*
*Accepted: 03/05/2025*



**ABSTRACT**

Breast cancer remains one of the leading causes of mortality among women worldwide and represents a major global health challenge. Accurate classification of breast tumors as benign or malignant is therefore of critical importance for timely diagnosis and effective treatment. This study aims to enhance breast cancer risk classification by integrating machine learning (ML) techniques with a genetic algorithm-based feature selection method. Initially, multiple ML algorithms are applied to features extracted from digitized images obtained through fine-needle aspiration (FNA) of breast masses. Subsequently, a genetic algorithm-based feature selection approach is employed to identify a subset of the most discriminative features. The results demonstrate that ML models utilizing the feature subsets selected by the genetic algorithm consistently achieve higher classification accuracy compared to their baseline counterparts. This highlights the effectiveness of the proposed feature selection strategy in improving the discriminative capacity of ML models. Beyond the observed improvements in accuracy, the refined ML models developed in this study show potential for more precise and reliable breast cancer diagnoses. By enhancing the performance of ML-based decision support systems, the genetic algorithm-based feature selection approach may contribute to the advancement of personalized treatment strategies in breast cancer care.

***Keywords:*** Breast cancer, Machine learning, Risk classification, Genetic algorithm, Feature selection.

[a] *aynursahin@selcuk.edu.tr* *https://orcid.org/0000-0003-1681-9398*
[b] *hyonar@selcuk.edu.tr* *https://orcid.org/0000-0003-1574-3993*
[c] *oznurozaltin@hacettepe.edu.tr* *https://orcid.org/0000-0001-9841-1702*

## Introduction

Breast cancer stands as one of the most common malignant tumors globally, constituting 10.4% of all cancers. Notably, it holds the unfortunate distinction of being the primary cause of mortality among women aged between 20 and 50 years [1,2]. According to the World Health Organization (WHO), there were 2.3 million cases of breast cancer and 685,000 deaths worldwide in 2020. By the end of 2020, the number of women currently living who have been diagnosed with breast cancer within the previous five years amounted to 7.8 million, establishing it as the most widespread form of cancer globally [3]. Timely detection, prompt diagnosis, and early treatment are imperative in preventing the advancement of the disease and mitigating its mortality rate. Hence, the accurate classification of breast cancers as either benign or malignant is of vital importance [4].

Machine Learning (ML), a subset of artificial intelligence, operates on the principle that machines should be provided with access to data and allowed to autonomously learn and explore. Its focus lies in extracting meaningful patterns from extensive datasets. This field harnesses the power of algorithms and statistical models to enable machines to learn from experience, adapt to changing scenarios, and make informed decisions without explicit programming. The essence of ML lies in its capacity to uncover insights and patterns, enhancing its utility across various domains such as data analysis, pattern recognition, and predictive modeling [5].

Classification serves as a supervised learning approach where a computer program learns from provided data and subsequently generates new observations or classifications in both ML and statistics. The fundamental process entails training an algorithm on a labeled dataset, where the desired output is known, allowing the program to discern patterns and relationships within the data. Following this training phase, the classifier is equipped to predict class labels or categories for new, unseen data based on the patterns it has learned. Leveraging ML algorithms, this process empowers systems to identify intricate patterns within data, finding applications in diverse domains such as image recognition, natural language processing, and fraud detection. Classification plays a pivotal role in enhancing automated decision-making by enabling systems to categorize and interpret data effectively. The accuracy of classification is heavily contingent on the nature of features within a dataset, where the presence of irrelevant or redundant data can impact performance. To enhance classification accuracy, the utilization of feature selection becomes crucial [6]. This process aims to eliminate irrelevant or redundant

features, optimizing the dataset by retaining only the most significant attributes. Through feature selection, the model's efficiency is improved, contributing to more accurate and streamlined classification outcomes [7, 8].

In recent years, researchers have proposed a wide range of feature selection techniques based on optimization and metaheuristic algorithms. Ye, Xu [9] introduced a feature selection approach utilizing adaptive particle swarm optimization with leadership learning. Ghosh, Datta [10] focused on hyperspectral image data and employed Self-adaptive Differential Evolution (SADE) for feature subset generation. Zhang, Mistry [11] presented a modified version of the Firefly Algorithm (FA) specifically designed to select discriminative features in classification and regression models. Baig, Aslam [12] introduced hybrid approach that combines a Differential Evolution (DE) optimization algorithm to seek out the feature space and yield an optimal subset of features. Sindhu and Ngadiran [13] proposed a Sine-Cosine Algorithm (SCA) with an Elitism strategy and the best new solution update mechanism to select features and enhance classification accuracy. Mafarja and Mirjalili [14] employed a wrapper feature selection model to simultaneously decrease the quantity of features and improve classification accuracy. Abdel-Basset, El-Shahat [15] suggested a novel Grey Wolf Optimizer algorithm via a Two-phase Mutation to select optimal features.

Beyond these techniques, the use of Genetic Algorithm (GA) has garnered considerable attention due to its simplicity, adaptability, and ability to effectively explore complex, high-dimensional search spaces. Recent studies have applied GA in various medical diagnostic tasks, including breast cancer detection. For instance, Sehgal et al. [16] employed GA to optimize hyperparameters in deep learning models, achieving significant improvements in AUC scores for breast cancer classification. Similarly, Yaqoob et al. [17] developed a hybrid GA–deep learning model using RNA-Seq gene expression data, successfully handling high-dimensional biological data. Boumaraf et al. [18] used GA for selecting BI-RADS features from mammogram images and applied a backpropagation neural network for classification. These studies emphasize the importance of integrating metaheuristic-based feature selection with ML to enhance diagnostic accuracy and efficiency in real-world clinical scenarios.

Building on these developments, this study investigates the effect of GA-based feature selection on the performance of various ML classifiers in breast cancer diagnosis. Specifically, five ML algorithms—Decision Tree (DT), K-Nearest Neighbors (KNN), Logistic Regression (LR), Multi-Layer Perceptron (MLP), and Random Forest (RF)—are used to classify samples based on features extracted from digitized images of fine-needle aspiration Fine-Needle Aspiration (FNA) of breast masses. The classification is performed both before and after applying GA for feature selection. The effectiveness of each approach is evaluated using key performance metrics: accuracy, precision, recall, and F1-score.

The aim of this study is to systematically assess how GA-based feature selection influences the performance of different ML classifiers in breast cancer classification. The findings are expected to contribute to the growing body of knowledge on intelligent diagnostic systems and offer practical insights into the design of more accurate and reliable clinical decision support tools.

The remainder of this paper is organized as follows: Section 2 introduces the ML algorithms, GA-based feature selection strategy, and experimental setup. Section 3 presents the application of classifiers to the breast cancer dataset and summarizes the classification results. Section 4 discusses the findings of literature, and Section 5 concludes the study.

## Material And Methods

In this section, a concise overview of the dataset used, ML algorithms, feature selection strategy, and the GA employed in this study is presented.

### *Dataset*

This study utilizes the Breast Cancer Wisconsin (Diagnostic) Dataset obtained from the University of California, Irvine (UCI) Machine Learning Repository [19]. The dataset consists of 30 numerical features that were extracted from digitized images of FNA of breast masses. These features describe characteristics of the cell nuclei present in the images.

Specifically, for each cell nucleus, 10 real-valued attributes are computed:

Radius: Mean of distances from the center to points on the perimeter
Texture: Standard deviation of gray-scale values
Perimeter: Perimeter of the cell nucleus
Area: Area of the cell nucleus
Smoothness: Local variation in radius lengths
Compactness: $(Perimeter^2/Area) - 1$
Concavity: Severity of concave portions of the contour
Concave points: Number of concave portions of the contour
Symmetry: Degree of symmetry of the cell nucleus
Fractal dimension: "Coastline approximation" (fractal measure of complexity)

For each of the above attributes, the mean, standard error, and “worst” (mean of the three largest values) were calculated, resulting in a total of 30 features per instance. The dataset includes a total of 569 samples, comprising 357 benign and 212 malignant tumor instances.

Prior to model training and feature selection, a series of preprocessing steps were applied to the dataset to ensure data quality and improve model performance. First, the dataset was examined for missing values, and no missing observations were detected. All feature values were then normalized using min-max normalization to scale the data to the [0, 1] range, ensuring that features with different scales would not bias the learning algorithms. Since the dataset was relatively balanced, no

resampling methods such as the Synthetic Minority Oversampling Technique (SMOTE) or under sampling were required. Finally, the class labels (benign and malignant) were encoded as binary values (0 and 1) to allow compatibility with the classification algorithms.

Although only a single dataset is used in this study, the Breast Cancer Wisconsin (Diagnostic) dataset is widely accepted as a benchmark in the biomedical ML literature. It offers a balanced class distribution, clinically relevant morphological features derived from diagnostic imaging, and enough instances for classification tasks. These characteristics justify its use for evaluating the effectiveness of the proposed feature selection method. Future studies may expand upon this work by applying the same methodology to additional datasets to assess generalizability.

## Machine Learning Algorithms

ML algorithms, which are tasked with classification, are computational techniques designed to categorize data into distinct classes or groups. These algorithms learn patterns from labeled training data and apply these learned patterns to new, unseen data for classification purposes. A brief overview of the algorithms used in this study, including LR, DT, KNN, MLP, and RF, is provided below.

### Logistic Regression (LR)

LR is a linear model specifically developed for classification tasks, rather than regression. This model involves using a logistic function to represent the probabilities linked to various outcomes in a single trial. Despite its linear nature, logistic regression is particularly powerful in binary and multiclass classification problems, making it a widely employed algorithm in various fields such as statistics, machine learning, and epidemiology [20].

### Decision Tree (DT)

DT is a type of non-parametric supervised learning method used for classification and regression applications. A DT algorithm is organized in a hierarchical manner, where each node corresponds to a test of a certain feature, each branch represents the outcome of that test, and each leaf represents either a class label or a distribution of classes [21].

### K-Nearest Neighbors (KNN) Classifier

The KNN method is the predominant Neighbors-based classification method. The ideal selection of the 'k' number relies significantly on the data: typically, a higher 'k' value reduces the impact of noise but can also blur the classification boundaries [22].

### Multi-Layer Perceptron (MLP) Classifier

A supervised learning approach called the MLP classifier is utilized for learning a function. It makes use of stochastic gradient descent, or LBFGS, to optimize the log-loss function. This algorithm can be used for a variety of classification tasks because it excels at identifying complex patterns and correlations in the data [23].

### Random Forest (RF)

RF is one of the ensemble classification approaches that builds multiple DT classifiers on different subsets of the dataset and uses averaging to improve prediction accuracy and reduce overfitting. Each tree in a random forest ensemble is built using a bootstrap sample selected with replacement from the training data. This methodology, known as bagging, contributes to the robustness and generalization capabilities of the overall model [24].

## Feature Selection

Feature selection involves eliminating unnecessary and unimportant features from a dataset to enhance the efficiency of a learning algorithm. However, this task poses significant challenges, primarily due to the vast search space involved. The dataset with n features has a total of $2^n$ potential solutions. The complexity of this task is exacerbated as n increases, which is often the case with advancements in data collection techniques and the growing complexity of the problems being addressed.

Feature selection methods are widely employed in studies to enhance the accuracy of classification. Among the commonly used methods for dimensionality reduction are Principal Component Analysis (PCA) [25], Independent Component Analysis (ICA) [26], and Sequential Forward and Backward Searches [27]. However, these methods often encounter challenges such as local optimal traps or high computational costs.

To address these issues, researchers have turned to evolutionary algorithms like GA [28, 29], Particle Swarm Optimization (PSO) [9, 30], DE [12,31] and Artificial Bee Colony (ABC) [32, 33] optimization. These algorithms have demonstrated success in handling tasks with large feature search spaces. In this study, GA is employed for feature selection, showcasing its effectiveness in optimizing classification outcomes.

## Genetic Algorithm (GA)

GA, conceptualized by Holland [34], is widely recognized as a highly effective search method for approximating solutions in optimization problems. The algorithm kicks off by generating an initial population at random and iterates through a series of steps to facilitate population evolution.

As a population-based method and a prominent class of evolutionary algorithms, GA eliminates concerns related to initial values. Typically, the GA commences with the creation of random initial population and proceeds through an iterative process known as a generation. Each generation encompasses stages such as selection, reproduction involving crossover and mutation operators, evolution, and replacement [35, 36]. The flowchart of GA is shown Figure 1.

The GA was implemented for feature selection using a population size of 30 individuals and a maximum of 50 generations. The algorithm utilized standard binary

tournament selection, single-point crossover with a probability of 0.8, and a mutation rate of 0.02.

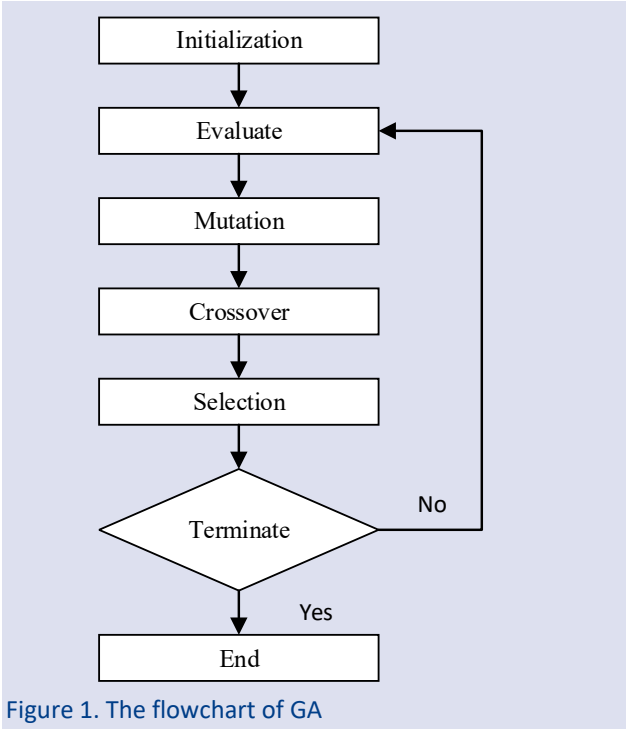


Figure 1. The flowchart of GA

A fitness function based on classification accuracy obtained from a 3-Nearest Neighbor classifier was employed to evaluate feature subsets.

## Application and Experimental Results

In the era of precision medicine, developing accurate and robust models for breast cancer plays a pivotal role in facilitating early diagnosis and improving patient outcomes. To contribute to this goal, the present study examines the impact of GA-based feature selection on the performance of several ML algorithms for classifying breast tumors as benign or malignant.

The Breast Cancer Wisconsin (Diagnostic) dataset from the UCI ML Repository was utilized. A stratified 80:20 train-test split was employed, with 80% of the samples allocated for training and the remaining 20% reserved for testing. Classification models were developed using several well-established ML algorithms, including LR, DT, KNN, MLP, and RF. These models were first evaluated using the full feature set, after which the same classifiers were re-trained on a subset of features selected via a GA-based feature selection method. All implementations were carried out in Python, utilizing the scikit-learn and matplotlib libraries for training, evaluation, and visualization.

All ML models were evaluated using 5-fold cross-validation to ensure robustness and minimize overfitting. The dataset was partitioned into 5 equal subsets, with 4 folds used for training and 1-fold for testing in each iteration. The average performance across all folds was recorded for each metric.

Hyperparameter tuning for classifiers such as DT, KNN, and RF was conducted using grid search within a 5-fold cross-validation framework to determine optimal parameter configurations. For the DT classifier, the maximum tree depth (max_depth) was searched over the range of [3, 5, 10, 15, 20], and the optimal depth was found to be 10. For K-Nearest Neighbors, the number of neighbors (n_neighbors) varied between [1, 3, 5, 7, 9], with 3 neighbors yielding the best performance. In the case of RF, the number of trees (n_estimators) was explored within [50, 100, 200], and the maximum depth was selected from [None, 10, 20, 30]. The best results were obtained using 100 estimators and maximum depth = 20. The final hyperparameter configurations were selected based on the highest average accuracy achieved across five folds.

To assess and compare the classification performance of the baseline and GA-enhanced models, four evaluation metrics were calculated: accuracy, precision, recall, and F1-score. In addition, confusion matrices were generated to provide a visual breakdown of correctly and incorrectly classified instances. The formulas used for these metrics are presented below:

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (1)$$

$$Pr\,e\,cision = \frac{TP}{TP + FP} \quad (2)$$

$$Re\,c\,all = \frac{TP}{TP + FN} \quad (3)$$

$$F1 - score = 2 \times \frac{Pr\,e\,cision \times Re\,c\,all}{Pr\,e\,cision + Re\,c\,al} \quad (4)$$

Here, TP denotes true positives, TN true negatives, FP false positives, and FN false negatives.

Figure 2 presents the confusion matrices of all classifiers, both with and without GA-based feature selection. These matrices reveal the distribution of TP, TN, FP, and FN predictions, allowing for a visual assessment of classification improvements. In most cases, GA leads to a reduction in FP and FN values, indicating enhanced precision and recall.

For example, KNN and MLP show visible improvements after GA, with fewer misclassified instances. DT also benefits from GA by reducing misclassification rates. In contrast, RF displays minimal change in its confusion matrix, consistent with its inherent feature selection capability.

These matrix-based observations align with the metric values in Table 1, confirming that GA contributes to more accurate and reliable classification by eliminating less relevant features and focusing on the most informative ones.

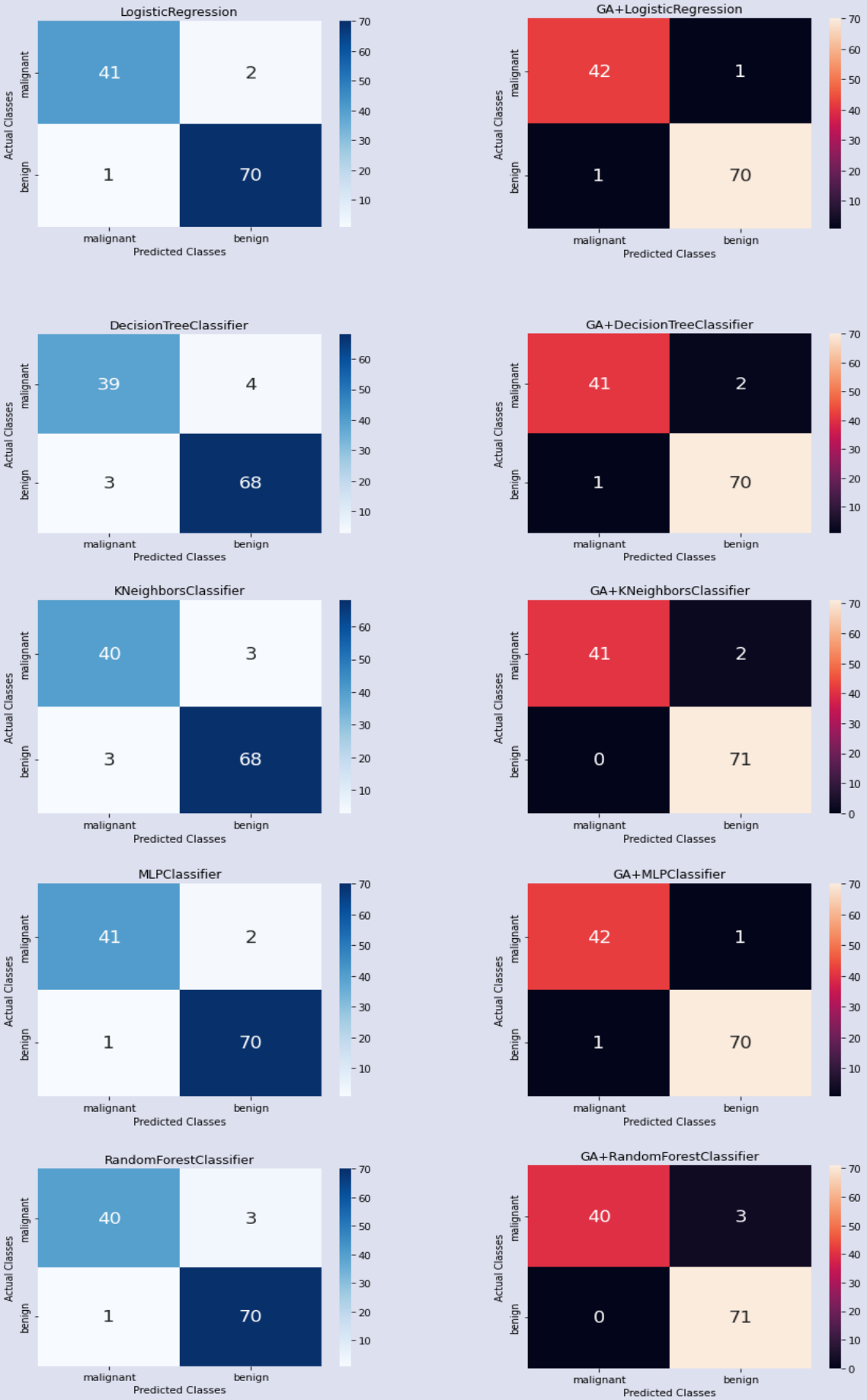


Figure 2. Confusion matrices of the classifiers before and after Genetic Algorithm-based feature selection, showing differences in TP, TN, FP, and FN rates.

The corresponding performance metrics for each model are summarized in Table 1, offering a clear comparison between the baseline ML models and those optimized through GA-based feature selection.

Table 1. Performance comparison of ML algorithms

| Model | Accuracy | Precision | Recall | F1-score |
|---|---|---|---|---|
| LR | 0.9736 | 0.9824 | 0.9859 | 0.9790 |
| GA-LR | 0.9754 | 0.9816 | 0.9830 | 0.9803 |
| DT | 0.9350 | 0.9485 | 0.9492 | 0.9479 |
| GA_DT | 0.9456 | 0.9555 | 0.9549 | 0.9562 |
| KNN | 0.9473 | 0.9577 | 0.9577 | 0.9577 |
| GA_ KNN | 0.9561 | 0.9713 | 0.9774 | 0.9652 |
| MLP | 0.9701 | 0.9781 | 0.9802 | 0.9761 |
| GA-MLP | 0.9719 | 0.9803 | 0.9830 | 0.9776 |
| RF | 0.9631 | 0.9768 | 0.9830 | 0.9707 |
| GA_ RF | 0.9631 | 0.9754 | 0.9802 | 0.9707 |

The numerical results indicate that most classifiers benefit from GA in terms of improved predictive accuracy, recall, and F1-score. Among the classifiers, DT, KNN, and MLP demonstrated the most substantial improvements following feature selection. The accuracy of DT increased from 0.9350 to 0.9456, and its F1-score rose from 0.9479 to 0.9562, indicating a more reliable and generalizable model. KNN also exhibited a consistent enhancement across all metrics, particularly in recall (from 0.9577 to 0.9774), which is essential in medical contexts where false negatives must be minimized. MLP showed increased precision, recall, and F1-score, achieving a final F1-score of 0.9776, reflecting its improved balance between sensitivity and specificity after GA was applied.

LR, which already exhibited strong baseline performance, experienced marginal gains in accuracy (from 0.9736 to 0.9754) and F1-score (from 0.9790 to 0.9803). However, a slight decline in recall was observed, suggesting that GA may have limited impact on classifiers with inherently stable decision boundaries.

In the case of RF, performance metrics remained unchanged after applying GA. This is consistent with its ensemble nature and built-in feature selection mechanism, where redundant or less informative features are implicitly down weighted during training. As a result, external feature selection techniques like GA may offer minimal added value for RF.

These results suggest that GA can effectively enhance the performance of various ML classifiers, especially those more sensitive to feature dimensionality. By eliminating irrelevant or redundant features, GA helps improve classification robustness and focus on the most discriminative patterns in the data. This is particularly important in medical diagnostics, where high recall and precision are critical to ensuring accurate and safe decision-making.

Therefore, integrating GA as a preprocessing step can be a valuable strategy in the development of more efficient and interpretable ML-based clinical decision support systems. Future research should consider testing the proposed methodology on larger and more diverse datasets, as well as comparing GA with other metaheuristic algorithms to further evaluate its adaptability and generalization capability in complex medical applications.

## Discussion

The results of this study demonstrate that applying GA-based feature selection prior to ML classification significantly enhances diagnostic performance for breast cancer detection. Specifically, classifiers such as DT, KNN, and MLP exhibited notable gains in accuracy, recall, and F1-score after feature reduction. These findings confirm the critical role of eliminating redundant or non-informative attributes in improving classification reliability.

Recent literature also supports the integration of metaheuristic feature selection methods with ML models in cancer classification. For example, Tan et al. [37] applied genetic programming to oral cancer prognosis, reporting an average accuracy of 83.87%, while Sharma et al. [38] achieved 96.66% and 93.06% accuracy using Support Vector Machine (SVM) and Artificial Neural Network (ANN), respectively, on the WBCD dataset. Sidey-Gibbons and Sidey-Gibbons [39] reached up to 96% accuracy using ensemble models. In comparison, our study reported 97.5% accuracy with LR and 97.1% with MLP after GA optimization, demonstrating the competitive performance of our proposed approach.

A key strength of this study lies in its flexibility across multiple classifiers. Unlike many existing studies that evaluate feature selection within a single model, we assessed five classifiers, both linear (LR) and nonlinear (MLP, RF), and observed consistent performance improvements. This adaptability makes the framework broadly applicable in clinical diagnostic modeling. Additionally, reducing the feature set improves interpretability, an essential aspect in medical decision-making.

From a clinical perspective, enhancing classification performance directly supports more reliable decision support systems. Improving recall reduces the risk of false negatives, which is critical for detecting malignant cases early. Simultaneously, improved precision minimizes false positives, preventing unnecessary interventions and patient distress.

Nonetheless, certain limitations must be acknowledged. The current study was conducted using a single, albeit well-established and clinically relevant, dataset (WBCD). While its quality and balance make it ideal for benchmarking, broader validation is required. Future studies should evaluate the proposed framework on additional datasets, including those from real-world clinical environments or public repositories.

Moreover, while GA demonstrated strong performance, it also carries known limitations such as sensitivity to hyperparameters, risk of convergence to local optima, and higher computational cost. Future research may consider comparing GA with alternative metaheuristic or hybrid approaches—such as PSO, ABC,

Grey Wolf Optimizer (GWO) or filter-wrapper ensembles—to enhance both robustness and efficiency.

In summary, the integration of GA with ML classifiers proves to be an effective strategy for breast cancer classification. The proposed method achieves high predictive accuracy, promotes model interpretability, and holds promise for clinical application. However, continued evaluation on diverse data sources and comparative studies with alternative optimization strategies will be essential for building more generalizable and scalable diagnostic systems.

## Conclusions

This study evaluated the effectiveness of GA-based feature selection in enhancing the performance of ML classifiers for breast cancer diagnosis. Using a dataset comprising 30 features derived from fine-needle aspirate images, GA was employed to identify the most informative subset before classification.

The experimental findings confirmed that applying GA led to improvements in classification accuracy, recall, and F1-score, particularly for DT, KNN, and MLP. These results underscore the importance of dimensionality reduction and the role of relevant feature selection in improving diagnostic accuracy.

While GA offers valuable advantages—such as adaptability and effective search capability—it also presents certain limitations. These include sensitivity to parameter tuning, potential convergence to local optima, and high computational cost. To overcome these challenges, future work should explore and compare alternative metaheuristic approaches in similar classification tasks.

Additionally, validating the proposed approach on more diverse and larger datasets, beyond the WBCD, such as those publicly available, will be crucial in assessing generalizability and scalability. Overall, this study supports GA as a promising preprocessing technique for improving ML-based breast cancer diagnostic systems, while also highlighting the importance of comparative evaluations and broader dataset applications in future research.

## Conflicts of interest

There are no conflicts of interest in this work.

## References


[1] Lacoviello, L., Bonaccio, M., de Gaetano, G., and Donati, M. B. Epidemiology Of Breast Cancer, A Paradigm Of The "Common Soil" Hypothesis, *Seminars In Cancer Biology*, (2021) 4-10.

[2] Siegel, R. L., Miller, K. D., and Jemal, A., Cancer Statistics, *CA: A Cancer Journal For Clinicians*, 69 (1) (2019), 7-34.

[3] WHO. Available at https://www.who.int/news-room/fact-sheets/detail/breast-cancer. Retrieved March 10,2023.

[4] Mridha, M. F., Hamid, M. A., Monowar, M. M., Keya, A. J., Ohi, A. Q., Islam, M. R., and Kim, J.-M., A Comprehensive Survey On Deep-Learning-Based Breast Cancer Diagnosis, *Cancers*, 13 (23) (2021) 6116.

[5] Mehrotra, D., Basics Of Artificial Intelligence & Machine Learning, *Notion Press*, (2019).

[6] Segal, M. R., Machine Learning Benchmarks And Random Forest Regression, (2004).

[7] Blum, A. L. and Langley, P., Selection Of Relevant Features And Examples In Machine Learning, *Artificial Intelligence*, 97 (1-2) (1997) 245-271.

[8] Chen, R.-C., Dewi, C., Huang, S.-W., and Caraka, R. E., Selecting Critical Features For Data Classification Based On Machine Learning Methods, *Journal of Big Data*, 7 (1) (2020) 52.

[9] Ye, Z., Xu, Y., He, Q., Wang, M., Bai, W., and Xiao, H., Feature Selection Based on Adaptive Particle Swarm Optimization With Leadership Learning, *Computational Intelligence Neuroscience*, (2022).

[10] Ghosh, A., Datta, A., and Ghosh, S., Self-Adaptive Differential Evolution For Feature Selection In Hyperspectral Image Data, *Applied Soft Computing*, 13 (4) (2013) 1969-1977.

[11] Zhang, L., Mistry, K., Lim, C. P., and Neoh, S. C., Feature Selection Using Firefly Optimization For Classification And Regression Models, *Decision Support Systems*, 106 (2018) 64-85.

[12] Baig, M. Z., Aslam, N., Shum, H. P., and Zhang, L., Differential Evolution Algorithm As A Tool For Optimal Feature Subset Selection In Motor Imagery EEG, *Expert Systems With Applications*, 90 (2017) 184-195.

[13] Sindhu, R., Ngadiran, R., Yacob, Y. M., Zahri, N. A. H., and Hariharan, M., Sine–Cosine Algorithm For Feature Selection With Elitism Strategy And New Updating Mechanism, *Neural Computing Applications*, 28 (2017) 2947-2958.

[14] Mafarja, M. and Mirjalili, S., Whale Optimization Approaches for Wrapper Feature Selection, *Applied Soft Computing*, 62 (2018) 441-453.

[15] Abdel-Basset, M., El-Shahat, D., El-Henawy, I., De Albuquerque, V. H. C., and Mirjalili, S., A New Fusion Of Grey Wolf Optimizer Algorithm With A Two-Phase Mutation For Feature Selection, *Expert Systems with Applications*, 139 (2020) 112824.

[16] Sehgal, A., Sehgal, M., La, H. M., and Bebis, G., Deep Learning Hyperparameter Optimization for Breast Mass Detection in Mammograms, International Symposium on Visual Computing, Cham: Springer Nature Switzerland, (2022) 270–283.

[17] Yaqoob, A., Verma, N. K., Aziz, R. M., and Shah, M. A., RNA-Seq Analysis for Breast Cancer Detection: A Study on Paired Tissue Samples Using Hybrid Optimization and Deep Learning Techniques, *Journal of Cancer Research and Clinical Oncology*, 150(10) (2024) 455.

[18] Boumaraf, S., Liu, X., Ferkous, C., and Ma, X., A New Computer-Aided Diagnosis System with Modified Genetic Feature Selection for BI-RADS Classification of Breast Masses in Mammograms, *BioMed Research International*, 2020(1) (2020) 7695207.

[19] UCI, Available at https://archive.ics.uci.edu/dataset/17/breast+cancer+wisconsin+diagnostic Retrieved January 28, 2024.

[20] Maalouf, M. and Strategies, Logistic Regression in Data Analysis: An Overview, *International Journal of Data Analysis Techniques*, 3 (3) (2011) 281-299.

[21] ZhuParris, A., de Goede, A. A., Yocarini, I. E., Kraaij, W., Groeneveld, G. J., and Doll, R. J., Machine Learning Techniques for Developing Remotely Monitored Central Nervous System Biomarkers Using Wearable Sensors: A Narrative Literature Review, *Sensors*, 23 (11) (2023) 5243.

[22] Dillen, A., Lathouwers, E., Miladinović, A., Marusic, U., Ghaffari, F., Romain, O., Meeusen, R., and De Pauw, K., A Data-Driven Machine Learning Approach For Brain-Computer Interfaces Targeting Lower Limb Neuroprosthetics, *Frontiers In Human Neuroscience*, 16 (2022) 949224.

[23] Awadallah, M. A., Abu-Doush, I., Al-Betar, M. A., and Braik, M. S. Metaheuristics For Optimizing Weights, *Neural Networks In Comprehensive Metaheuristics, Elsevier,* (2023) 359-377.

[24] Fawagreh, K., Gaber, M. M., and Elyan, E., Random Forests: From Early Developments To Recent Advancements, *Systems Science Control Engineering: An Open Access Journal*, 2 (1) (2014) 602-609.

[25] Yu, X., Chum, P., and Sim, K.-B., Analysis The Effect Of PCA For Feature Reduction In Non-Stationary EEG Based Motor Imagery Of BCI System, *Optik* 125 (3) (2014) 1498-1502.

[26] Guo, X., Wu, X., Gong, X., and Zhang, L. (Year) Envelope Detection Based On Online ICA Algorithm And Its Application To Motor Imagery Classification. in:*2013 6th International IEEE/EMBS Conference on Neural Engineering (NER),* (2013) 1058-1061 IEEE.

[27] Chandrashekar, G. and Sahin, F., A Survey On Feature Selection Methods. Computers Electrical Engineering, 40 (1) (2014) 16-28.

[28] Aalaei, S., Shahraki, H., Rowhanimanesh, A., and Eslami, S., Feature Selection Using Genetic Algorithm For Breast Cancer Diagnosis: Experiment On Three Different Datasets, Iranian journal of basic medical sciences, 19 (5) (2016) 476.

[29] Leardi, R., Boggia, R., and Terrile, M., Genetic Algorithms As A Strategy For Feature Selection, Journal of chemometrics, 6 (5) (1992) 267-281.

[30] Sakri, S. B., Rashid, N. B. A., and Zain, Z. M., Particle Swarm Optimization Feature Selection For Breast Cancer Recurrence Prediction, IEEE Access, 6 (2018) 29637-29647.

[31] Wang, P., Xue, B., Liang, J., and Zhang, M., Multiobjective Differential Evolution For Feature Selection In Classification, IEEE Transactions on Cybernetics, (2021)

[32] Hancer, E., Xue, B., Karaboga, D., and Zhang, M., A Binary ABC Algorithm Based On Advanced Similarity Scheme For Feature Selection, Applied Soft Computing, 36 (2015) 334-348.

[33] Rostami, O. and Kaveh, M., Optimal Feature Selection For SAR Image Classification Using Biogeography-Based Optimization (BBO), Artificial Bee Colony (ABC) And Support Vector Machine (SVM): A Combined Approach Of Optimization And Machine Learning. Computational Geosciences, 25 (2021) 911-930.

[34] Holland, J. H. (1975) Adaptation In Natural And Artificial Systems The University of Michigan Press.

[35] Talbi, E.-G., Metaheuristics: From Design To Implementation, John Wiley & Sons, 2009.

[36] Yang, X.-S., Engineering Optimization: An Introduction with Metaheuristic Applications, John Wiley & Sons,(2010).

[37] Tan, M. S., Tan, J. W., Chang, S.-W., Yap, H. J., Kareem, S. A., and Zain, R. B., A Genetic Programming Approach to Oral Cancer Prognosis, *PeerJ*, 4 (2016).

[38] Sharma, A., Kulshrestha, S., and Daniel, S. B., Machine Learning Approaches for Cancer Detection, *International Journal of Engineering and Manufacturing (IJEM)*, 8(2) (2018) 45–55.

[39] Sidey-Gibbons, J. A. M., and Sidey-Gibbons, C. J., Machine Learning in Medicine: A Practical Introduction, *BMC Medical Research Methodology*, 19(1) (2019) 64.