Science World Journal Vol. 21(No 2) 2026 https://dx.doi.org/10.4314/swj.v21i2.12
www.scienceworldjournal.org
ISSN: 1597-6343 (Online), ISSN: 2756-391X (Print)
Published by Faculty of Science, Kaduna State University

A CHAOTIC MAP-DRIVEN JAYA ALGORITHM FOR ROBUST CLINICAL ATTRIBUTES SELECTION IN BREAST CANCER METASTASIS PREDICTION

[1]Bello Abdullahi, [1]Yusuf Sahabi Ali, [1]Mohammed Abdullahi, * [2]Jeremiah Isuwa, [3]Sani Abdullahi, and [4]Tukur Abdullahi

[1]Department of Computer Science, Ahmadu Bello University, Zaria, Nigeria
[2]Department of Computer Science, Federal University of Kashere, Gombe-Nigeria
[3]National Biotechnology Development Agency, Abuja-Nigeria
[4]Department of Computer Science, Kaduna State University, Kaduna, Nigeria

*Corresponding Author Email Address: isuwa.jeremiah@fukashere.edu.ng

ABSTRACT

Breast cancer (BC) remains a leading cause of cancer-related mortality globally, especially due to its metastatic tendencies. Metastatic breast cancer (MBC) occurs when cancer cells spread from the breast to other parts of the body, complicating treatment and reducing survival rates. Predicting MBC is crucial for timely intervention, but challenges persist due to noisy features that limit model accuracy. While traditional machine learning techniques have been applied to predict MBC, they often struggle with identifying key factors for accurate prediction. In this study, we propose enhancing the JAYA algorithm by incorporating a sinusoidal chaotic map for population initialization and to improve exploration during optimization. Specifically, the sinusoidal chaotic initialization is employed to generate a more uniformly distributed and diverse initial population, thereby improving search space coverage, reducing premature convergence, and enhancing the algorithm's ability to identify relevant features for MBC prediction. The enhanced JAYA algorithm was combined with an artificial neural network (ANN) for prediction, and its performance was evaluated using a 5-year MBC dataset obtained from the Kaggle repository, comprising 6,726 instances, 26 features, and a binary class label. The results indicate that the proposed method achieves improved performance across key evaluation metrics, including accuracy, F1 score, sensitivity, and specificity, when compared with the baseline study of Muhammed et al., (2024). Specifically, the proposed method attained an accuracy of 79.76%, an F1 score of 80.13%, a sensitivity of 81.88%, and a specificity of 77.52%, compared to 79.28%, 79.66%, 81.04%, and 77.51%, respectively, reported in the baseline study, thereby demonstrating consistent performance improvements across all evaluation metrics. These improvements were confirmed to be statistically significant using a T-test. Hyperparameter tuning, including adjustments to population size and iteration count, further optimized the method's performance, confirming the benefits of fine-tuning in metaheuristic algorithms for MBC prediction.

Keywords – JAYA Algorithm, Metastatic Breast Cancer (MBC), Feature Selection, Sinusoidal Chaotic Initialization

INTRODUCTION

Cancer is a serious medical condition characterized by the abnormal and uncontrolled growth of cells in the body. (Botlagunta et al., 2023). According to Jiang & Xu, (2022).The disease has a significant global impact, with breast cancer (BC) being the most commonly diagnosed type. Their research further emphasizes that BC is the leading cause of cancer-related mortality among women, with over 2.3 million new cases and more than 685,000 deaths reported in 2020. In Africa, particularly in Nigeria, the rising incidence of BC is becoming a major concern for oncologists (Azubuike et al., 2022). For instance, Azubuike et al., (2022) noted that in 2018, approximately 26,310 cases were reported, with projections indicating an annual increase of around 4,000 cases over the next decade. Additionally, the International Agency for Research on Cancer (IARC) identified 28,380 new BC cases in Nigeria in 2020, representing 22.7% of all cancer cases recorded that year. (Olayide et al., 2023).

Despite its precarious nature, studies have shown that BC confined to the breast has a higher chance of being cured compared to when it metastasizes, a condition known as metastatic breast cancer (MBC) (Abdollahi et al., 2022). Nassar, (2023)noted that when BC spreads to vital organs or tissues, it becomes a leading cause of cancer-related deaths and remains largely incurable. Although not all diagnosed cases of BC progress to metastasis, there is always a risk of future development, making it crucial to both diagnose current cases and predict the likelihood of future recurrences. (Marti et al., 2022).

Traditional methods for detecting BC, such as clinical breast examinations, involve an oncologist manually assessing the breast and surrounding areas for abnormalities like lumps or changes in texture. (Tran et al., 2022). Another common approach is mammography, which uses X-ray imaging to visualize breast tissue. (Tran et al., 2022). When abnormalities are found, a core needle biopsy guided by imaging is often performed to extract a tissue sample for microscopic analysis to confirm the presence of cancer cells. (Abdollahi et al., 2022). However, Jiang & Xu, (2022) highlighted that despite its wide acceptance and use, image-guided core needle biopsies can yield inconclusive results in 5-15% of cases, complicating the decision-making process for oncologists in determining the most effective treatment plan.

However, advances in modern technology, particularly in Artificial Intelligence (AI), have significantly enhanced the accuracy and efficiency of diagnosing and predicting MBC. AI-driven techniques, including Machine Learning (ML) and Deep Learning (DL), leverage large, complex datasets of clinical and diagnostic features to develop predictive models. (Abdollahi et al., 2020). These models not only provide early detection of individuals at risk for BC but also offer critical insights into the likelihood of future metastasis. (Isuwa et al., 2023). By facilitating timely intervention and informed treatment decisions, AI has become an indispensable tool in BC care, surpassing the limitations of traditional diagnostic approaches. (Isuwa et al., 2023).

Although the application of ML and DL techniques has greatly improved MBC prediction, the performance of these predictive models is highly dependent on the quality of the training datasets (Naskar et al., 2023). In many cases, these datasets may include noisy features, which can lead to challenges such as increased computational demands, overfitting, and reduced model interpretability, ultimately affecting overall model accuracy (Abd Elminaam et al., 2022). To mitigate such challenges, selecting only the most relevant features is critical, a process known as Feature Selection (FS) (Isuwa et al., 2022). By focusing on the most influential factors for MBC prediction, FS enhances the precision and dependability of the models, ensuring more effective application in medical diagnostics (Jeremiah et al., 2022). FS techniques have traditionally been grouped into three primary categories. The first is the **filter method**, which evaluates features using statistical measures such as Chi-square and Mutual Information (Hassan et al., 2023). Next is the **wrapper method,** which uses a classifier to assess the quality of randomly selected feature subsets and often relies on heuristic or metaheuristic algorithms for optimization (Hassan et al., 2023). Finally, the **embedded method** integrates the advantages of both the filter and wrapper approaches by evaluating feature relevance during model training. Notable examples of embedded methods include Lasso and Ridge regression, which utilize L1 and L2 regularization, respectively (Isuwa et al., 2021).

Building on these categories, the metaheuristic-based wrapper methods have gained significant popularity, particularly in recent years. This is due to their ability to efficiently navigate large, complex search spaces and identify optimal or near-optimal feature subsets within a reasonable timeframe. (Agrawal et al., 2021). Unlike filter methods, which can overlook important feature interactions, or embedded methods that are tied to specific models, metaheuristic algorithms offer flexibility and adaptability, making them a preferred choice for FS in high-dimensional data. (Chakravarthy et al., 2024). Their success in balancing exploration and exploitation during the search process ensures more effective and accurate model outcomes. (Chakravarthy et al., 2024). Numerous of these algorithms have been developed and applied to optimize the FS process in medical applications. Notable examples include the Horse Herd Optimization Algorithm (HOA) (Zaimoğlu et al., 2023), Cuckoo Search Algorithm (CS) (Feizi-Derakhsh & Kadhim, 2023), Sand Sat Optimization (SCO) (Alhassan, 2024), Reptile Search Algorithm (RSA) (Z. Elgamal et al., 2022), Mountain Gazelle Optimizer (MGO) (Li et al., 2024), Dwarf Mongoose Optimization (DMO) (Abdelrazek et al., 2024), Gradient-Based Optimizer (GBO) (Abd Elminaam et al., 2022), Harmony Search Algorithm (HSO) (Naskar et al., 2023), and the Jaya Algorithm (JA) (Muhammed et al., 2024) among others.

While these metaheuristic algorithms have demonstrated considerable success in FS, especially in BC prediction, they are not without limitations. A common issue arises during the initialization phase, where the population of potential solutions may be poorly distributed, leading to limited diversity and reducing the algorithm's ability to explore the solution space effectively. (Naskar et al., 2023). Additionally, during the search process, achieving a balance between exploration and exploitation is critical. However, it can be difficult to maintain, often leading the algorithm to get stuck in local optima or converge prematurely. (Abdelrazek et al., 2024). To address these shortcomings, chaotic maps have been integrated into metaheuristic algorithms. Beyond enhancing the initialization phase, chaotic maps are also used to replace random variables in the algorithm's update mechanism. (Abd Elminaam et al., 2022). This substitution allows the algorithms to generate more diverse candidate solutions throughout the optimization process, improving exploration. (Li et al., 2024). By using chaotic sequences to influence key update parameters, the search process becomes more dynamic and unpredictable, significantly reducing the risk of premature convergence and improving overall performance. (Li et al., 2024). Consequently, chaotic maps have proven to be a powerful tool for refining these algorithms, particularly when tackling complex FS tasks in medical applications. (Chakravarthy et al., 2024).

Despite the abundance of studies exploring metaheuristics enhanced with chaotic maps for BC prediction, a gap remains in applying these techniques specifically to MBC prediction. A thorough review of the literature reveals that no study has examined the use of chaotic maps to optimize these algorithms for MBC prediction, despite its critical importance.

Thus, this dissertation aims to address this gap by adapting the JAYA algorithm and enhancing it with the Sinusoidal chaotic map to predict MBC. The Sinusoidal chaotic map will be employed to refine both the initialization and update stages of the JAYA algorithm. This approach introduces greater diversity into the population of solutions during initialization and improves the algorithm's search effectiveness throughout the optimization process. This enhancement is expected to improve the exploration-exploitation balance, reduce the likelihood of premature convergence, and ultimately increase the prediction accuracy for MBC. Specifically, we:

i. Designed an improved initialization strategy for the JAYA algorithm by utilizing a sinusoidal chaotic map to enhance population diversity and search performance.
ii. Developed a predictive framework for 5-year metastatic breast cancer (MBC) using the improved JAYA algorithm and ANN.
iii. Performed a comparative analysis of the performance of the developed framework with existing methods in the literature on a benchmark 5-year MBC dataset, evaluating classification accuracy, F1 score, sensitivity, and specificity.

The remainder of this paper is organized as follows: Section 2 presents a concise review of the foundational concepts and relevant literature. Section 3 describes the proposed framework in detail. Section 4 presents the experimental results along with a comprehensive analysis. Finally, Section 5 concludes the study.

Dimensionality Reduction (DR)

DR involves reducing the number of features in a dataset while preserving its essential information. It is particularly useful for high-dimensional datasets with many features relative to the number of observations. DR helps reduce computational cost, mitigate challenges associated with high-dimensional data, and improve dataset interpretability. The main approaches include feature extraction, such as principal component analysis (PCA), and feature selection methods, such as the Chi-square test, each suited to different data characteristics and analytical goals.

Feature Selection (FS)

FS plays an important role in ML and data analysis by identifying a subset of relevant features from the original dataset. It aims to reduce dimensionality, improve model performance, and enhance interpretability. The FS process typically involves applying a selected method, i.e., filter, wrapper, or embedded, to the full feature set to generate a candidate subset. The subset is then evaluated using a learning algorithm, and the process is repeated iteratively until a stopping criterion is satisfied.

A. Filter Methods of Feature Selection

Filter methods evaluate feature importance independently of any learning algorithm, using statistical measures to rank features by their predictive power. They are computationally efficient and commonly used for high-dimensional datasets, often as a preprocessing step before more complex models. However, they do not capture interactions between features, which can limit their effectiveness in some cases. A common example is the Relief method, which assigns feature relevance scores based on how well features distinguish between neighboring instances from different classes by comparing nearest hits and nearest misses.

The feature weight W_i for feature i is updated using Equation 1:

$$W_i = W_i - diff(i, nearest\ hit) + diff(i, nearest\ miss) \qquad (1)$$

Where $diff(i,j)$ represents the difference between the values of feature i for two instances j. If the feature distinguishes well between different classes, its weight increases; otherwise, it decreases.

B. Wrapper Methods of Feature Selection

Wrapper methods evaluate different subsets of features by training a model and assessing its performance for each subset. This approach directly measures how feature combinations perform under a specific learning algorithm, capturing interactions between features and tailoring the selection process to the model. Popular strategies for wrapper-based feature selection include metaheuristic algorithms (MAs) such as Particle Swarm Optimization (PSO).

MAs are typically population-based methods that operate using a group of candidate solutions rather than a single search agent. In FS, each solution represents a subset of features, and the population is iteratively evolved using algorithm-specific rules to identify the optimal combination. The interaction among multiple solutions enhances exploration and reduces the likelihood of becoming trapped in local minima. However, this population-based search process increases computational cost due to repeated evaluation of multiple candidate solutions.

JAYA Algorithm (JA)

The JAYA Algorithm (JA), proposed by Venkata Rao, (2016). It employs a population-based approach to optimization, drawing inspiration from the principles of continual enhancement and improvement. The term "Jaya," which means "Victory" in Sanskrit, reflects the algorithm's objective of relentlessly seeking superior solutions. (Abed-alguni& AL-Jarah, 2024). Initially intended to solve both constrained and unconstrained mathematical optimization challenges, JA has also been successfully adapted for discrete optimization problems, particularly in the area of FS. (Noshad&Fallahi, 2023).

In contrast to other metaheuristic methods like PSO and Genetic Algorithm (GA), which require adjusting multiple parameters, JA is notable for its simplicity, requiring only tuning of the population size and the number of iterations. This simplicity enhances the optimization process and positions JA as an excellent candidate for FS tasks. (Chaudhuri & Sahu, 2021).

During each iteration, JA evaluates the best and worst solutions in the population, facilitating a steady trajectory toward the optimal solution while concurrently avoiding inferior results. (Chaudhuri & Sahu, 2021). The positional updates for each solution in JA are determined by a specific mathematical equation (2):

$$X_{k,j}^{new} = X_{k,j}^{current} + r_1 * \left(Best_j - \left|X_{k,j}^{current}\right|\right) - r_2 * \left(Worst_j - X_{k,j}^{current}\right|\,) \qquad (2)$$

where k represents the number of solutions ranging from $1\ toN$, and D represents the dimension of the problem ranging from $1\ toD$. $X_{k,j}^{current}$ denotes the value of the k^{th} solution in the j^{th} dimension. $Best_j$ and $Worst_j$ denote the best and worst solutions in terms of fitness for the j^{th} dimension, respectively. r_1 and r_2 are random numbers in the range $[0,1]$, introduced to inject randomness into the search process. Finally, $X_{k,j}^{new}$ represents the updated solution.

C. Embedded Methods of Feature Selection

The embedded method evaluates the significance of features within the model's training framework. Embedded methods integrate FS into the training process, enabling the model to identify the most relevant features while optimizing its performance for the specific task at hand. A widely used embedded method is regularization, i.e., Lasso L1 regularization for regression. This technique penalizes the coefficients of less significant features during training, thereby diminishing their influence on the final predictions.

The regularization term incorporated into the loss function is the sum of the absolute values of the coefficients multiplied by a regularization parameter (lambda). This approach promotes sparsity in the coefficient values, leading to some coefficients being reduced to zero. Consequently, Lasso regression not only aids in FS by discarding less relevant features but also yields a more straightforward and interpretable model. The mathematical representation for Lasso regression can be formulated as shown in Equation 3:

$$minimize\left(\sum_{i=1}^{n}(y_i - \tilde{y}_i)^2 + \lambda\sum_{j=1}^{p}|\beta_j|\right) \qquad (3)$$

where $\sum_{i=1}^{n}(y_i - \tilde{y}_i)^2$ represents the ordinary least squares loss function, where y_i is the observed value, and $\tilde{y}_i$ is the predicted value for the i^{th} observation. It calculates the squared difference between the observed and predicted values, and aggregates it across all observations. The λ represents the regularization parameter, commonly referred to as the penalty term. It governs the degree of regularization applied to the model. An increased value of λ results in more robust regularization, causing more coefficients to be reduced towards zero. Lastly, the $\sum_{j=1}^{p}|\beta_j|$ represents the $L1$ norm of the coefficient vector β, where β_j is the coefficient for the j^{th} feature. The $L1$ norm is the sum of the absolute values of the coefficients. By incorporating this term into the loss function, Lasso enforces a penalty on the absolute magnitude of the coefficients.

Chaotic Maps

Chaotic maps are widely used in MAs to improve the initialization phase of the search process. They introduce controlled randomness into the initial positions of search agents, enhancing

https://dx.doi.org/10.4314/swj.v21i2.12

early-stage exploration. Due to their sensitivity to initial conditions, small changes can produce diverse search trajectories, promoting better coverage of the solution space. This helps MAs avoid local optima and improves the quality of the solutions obtained. As a result, chaotic map-based initialization is commonly used to enhance the performance and robustness of optimization algorithms.

Sinusoidal chaotic map

The sinusoidal chaotic map is a nonlinear mathematical function used to generate chaotic sequences through iterative sine-based operations. It produces highly sensitive and unpredictable dynamics from initial values. Due to these properties, it has been applied in areas such as cryptography and optimization. In metaheuristic algorithms, it is commonly used to improve initialization by introducing diversity and enhancing search efficiency. The mathematical expression of the sinusoidal map is given in Equation 4:

$$x_{k+1} = P.x_k^2 \sin(\pi x_k) \qquad (4)$$

where x_k is the current value, x_{k+1} refers to the next value, and P refers to the controlling parameter for the chaotic map. Lastly, the $\sin(\pi x_k)$ refers to the sine function.

Neural Network (NN) Learning Algorithm

NNs are computational models inspired by the structure of the human brain, consisting of interconnected layers of artificial neurons that process input data to produce outputs. They learn patterns and relationships during training, making them effective for tasks such as classification. A typical NN includes an input layer, one or more hidden layers, and an output layer, as shown in Figure 1. Each layer contains nodes connected by weighted links, which are updated during training using backpropagation to minimize the loss function. This iterative process improves the model's ability to map inputs to accurate outputs.

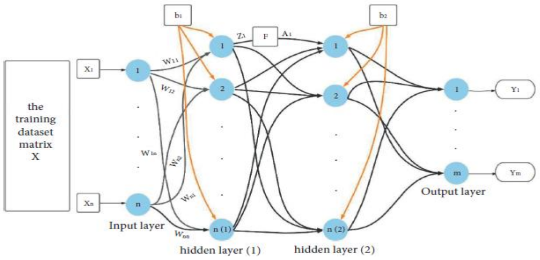

Figure 1: Conventional structure of a NN with two hidden layers (Beghriche et al., 2021)

Table 1 summarizes the reviewed studies, including the specific MAs and chaotic maps used, dataset types, and methodological descriptions.

Table 1: Summary of the literature highlighting the various MAs and chaotic maps utilized

	Citation	Dataset	Chaotic Map(s)	MA(s).	Description of Study
1	(Muhammed et al., 2024)	5-year MBC dataset	None	JAYA Algorithm	The Jaya algorithm, along with other feature importance measures, including the chi-square and Gini index, was independently combined with a neural network to select top features from a 5-year MBC dataset and predict the subsequent occurrence of metastasis. The study also addressed the challenge of severely imbalanced data by applying SMOTE to the dataset after FS.
2	(Alhassan, 2024)	BreakHis Image dataset from the UCI repository	Logistic	Sand Sat Optimization (SCO) and Remora Optimization Algorithm (ROA)	The study introduces a hybrid chaotic Sand Cat Optimization technique combined with the Remora Optimization Algorithm for feature selection. The primary aim of this research is to leverage the global search capabilities of the Sand Cat Optimization. To enhance the algorithm's exploratory power, the Logistic chaotic sequence is incorporated into the local search's threshold variable.
3	(Chakravarthy et al., 2024)	INbreast and CBIS-DDSM datasets	Tent	Crow Search Optimization algorithm (COA)	The study applied feature fusion using a serial mid-value approach for features extracted via EfficientNet-B4. Subsequently, the Crow Search Optimization Algorithm, enhanced with the Tent chaotic map, was employed for final feature selection. The Tent chaotic map was used to support the search process and facilitate effective exploration of the search space.
4	(Li et al., 2024)	16 medical benchmark datasets from diverse repositories.	Iterative	Mountain Gazelle Optimizer (MGO)	The study introduces an improved Mountain Gazelle Optimizer, building on the newly developed Mountain Gazelle Optimizer to address the feature selection problem in medical datasets. First, the gazelle population is initialized using an iterative chaotic map to enhance population diversity. Second, a nonlinear control factor is incorporated to balance the algorithm's exploration and exploitation phases.

	Citation	Dataset	Chaotic Map(s)	MA(s).	Description of Study
					Lastly, a neighborhood search strategy is applied to the optimal individuals, boosting the algorithm's exploitation and convergence capabilities.
5	(Abdelrazek et al., 2024)	10 datasets from the UCI repository	Chebyshev Circle Gauss Iterative Logistic Piecewise Sine, Singer Sinusoidal Tent	Dwarf Mongoose Optimization (DMO)	A modified version of the Dwarf Mongoose Optimization Algorithm is proposed for feature selection. To enhance DMO's convergence speed and effectiveness, ten chaotic maps were applied to adjust key elements of the Dwarf Mongoose's movement during the optimization process.
6	(Zaimoğlu et al., 2023)	18 datasets from the UCI repository	Logistics Piecewise Singer Sinusoidal Tent	Horse Herd Optimization Algorithm (HOA)	The study proposes a binary chaotic Horse Optimization Algorithm for feature selection. To enhance the diversity of candidate solutions, five different chaotic maps were employed individually to improve the population initialization.
7	(Feizi-Derakhsh & Kadhim, 2023)	8 medical datasets from diverse repositories	Sinusoidal	Cuckoo Search Algorithm (CS)	This approach introduces an enhanced Binary Cuckoo Search Algorithm that incorporates a hybrid Chi-square filter method and a Sinusoidal chaotic map to address feature selection problems. The Chi-square method is used to generate an initial solution, which is then improved to produce a higher-quality final solution. Additionally, the Sinusoidal chaotic map is used to determine the variable values of the step size (α) parameter during the local search process.
8	(Abd Elminaam et al., 2022)	10 datasets from the UCI repository	Chebyshev Circle Gauss Iterative Logistic Piecewise Sine, Singer Sinusoidal Tent	Gradient-Based Optimizer (GBO)	The study proposes a method for selecting the most discriminative features using a novel Chaotic Gradient-Based Optimizer (CGBO) that integrates chaotic maps into the GBO search iterations. Ten chaotic maps were employed to update the parameters, avoid local optima and premature convergence, accelerate convergence, and improve the overall efficiency of GBO.
9	(Z. Elgamal et al., 2022)	20 medical benchmark datasets from the UCI repository.	Circle	ReptileSearch Algorithm (RSA)	The study introduced two key enhancements to the Reptile Search Algorithm for feature selection across 20 medical datasets. First, the Circle chaotic map was utilized during the initialization phase to improve RSA's exploration capabilities in the search space. Then, the enhanced RSA was integrated with the Simulated Annealing local search algorithm to refine the search further and mitigate the risk of becoming trapped in local optima.
10	(Gharehchopogh et al., 2022)	24 UCI standard datasets	Chebyshev Circle Gauss Iterative Logistic Piecewise	Vortex Search Algorithm (VSA)	They proposed a chaotic Vortex Search Algorithm (VSA) for feature selection, incorporating chaos theory to accelerate global convergence and improve overall performance.

	Citation	Dataset	Chaotic Map(s)	MA(s).	Description of Study
			• Sine, • Singer • Sinusoidal • Tent		
11	(Khurma et al., 2020)	23 medical datasets from the UCI repository	• Piecewise • Tent • Circle	Moth Flame Optimization Algorithm (MFO)	In their study, three variants of the Binary Moth Flame Optimization Algorithm, using Piecewise, Tent, and Circle chaotic maps, are introduced and compared as search strategies within a wrapper feature-selection framework. The primary objective of incorporating chaotic maps is to improve the solution initialization process, enabling the optimizer to overcome local minima and achieve global convergence towards the optimal solution.
12	(Agrawal et al., 2021)	21 benchmarks datasets from the UCI repository.	• Chebyshev • Circle • Gauss • Iterative • Logistic Piecewise • Sine, • Singer • Sinusoidal • Tent	Gaining a knowledge-based optimization algorithm (GSK)	Firstly, it represents a binary variant of the GSK algorithm by applying a probability-estimation operator to its two main pillars. And then ten different types of chaotic maps are considered to individually replace the parameters of the GSK algorithm, i.e., the knowledge factor and Knowledge ratio, thereby striking a proper balance between exploration and exploitation and preventing premature convergence, thus enhancing the performance of the GSK.
13	(Z. M. Elgamal et al., 2020)	14 medical benchmark datasets from the UCI repository.	• Chebyshev • Circle • Gauss • Iterative • Logistic Piecewise • Sine, • Singer • Sinusoidal • Tent	Harris Hawks optimization (HHO)	The study introduced two key enhancements to the HHO for feature selection across 14 medical datasets. First, ten chaotic maps were utilized during the initialization phase to improve HHO's exploration capabilities in the search space. Then, the enhanced RSA was integrated with the Simulated Annealing local search algorithm to refine the search further and mitigate the risk of becoming trapped in local optima.
14	(Azar et al., 2020)	5 medical datasets	• Chebyshev • Circle • Gauss • Iterative • Logistic Piecewise • Sine, • Singer • Sinusoidal • Tent	Grey Wolf Optimization (GWO)	This paper presents a robust hybrid dynamic model for feature selection that combines Rough Set (RS) theory, chaos theory, and Binary Grey Wolf Optimization. The GWO parameters are estimated and fine-tuned using ten different chaotic maps. The model is evaluated using five complex medical datasets.

The Proposed Method

This section provides a comprehensive discussion of the proposed method, detailing the techniques and enhancements introduced. Additionally, it covers the benchmark datasets used, the parameter configurations, and the experimental design.

Chaotic-based initialization strategy for JAYA

A chaotic-based initialization strategy leverages the properties of chaotic systems, nonlinear dynamics, ergodicity, and deterministic randomness to introduce diversity and distribution in MAs. Unlike conventional random initialization, chaotic sequences provide a more uniform and comprehensive coverage of the search space, minimizing the risk of premature convergence and improving global search capabilities. Additionally, chaotic systems offer an adaptive mechanism for initializing control parameters, helping to balance exploration and exploitation more effectively throughout the optimization process. This ensures a more stable and efficient convergence, particularly in complex or high-dimensional problems.

In this study, the sinusoidal chaotic map is employed for both the initialization of the population and the control parameter in the JAYA algorithm. Its simplicity, deterministic yet diverse nature, and smooth iterative transitions make it an ideal choice for improving search diversity and parameter tuning. Compared with other chaotic maps, the sinusoidal chaotic map offers greater stability and avoids extreme values, enabling more effective exploration of the solution space and improving overall algorithm performance.

Experimental Configuration

To improve the performance and stability of the JAYA algorithm's optimization process, the Sigmoid transfer function is employed as shown in Equations 5 and 6. It maps fitness values to a binary space, where 0 indicates a feature is not selected and 1 indicates it is.

$$S\left(x_i^d\right) = \frac{1}{1+e^{-x_i^d}} \qquad (5)$$

$$x_i^d = \begin{cases} 1, if S\left(x_i^d\right) > \alpha, \\ 0, \ otherwise \end{cases} \qquad (6)$$

where $\alpha \sim U(0,1)$, $x \epsilon \mathbb{R}$ represents the likely solution and d the feature's continuous value at a time. To evaluate solution quality, a fitness function is used, as in Equation 7. By assigning a fitness value to each solution, the algorithm can distinguish between better and worse solutions, facilitating the selection and evolution of optimal candidates.

$$fitness\ function = \alpha \Delta_R(D) + \beta \frac{|Y|}{|T|} \qquad (7)$$

where $\Delta_R(D)$ stands for the classifier's error rate. $|Y|$ stands for the size of the chosen feature subset a,nd $|T|$ stands for the feature count in the dataset. The $'\alpha' \in [0,1]$ represents the classifier's error rate's weight. $'\beta' = (1-\alpha)$ is the significance of the features reduction. To assess the effectiveness and statistical significance of the proposed method relative to other approaches, classification accuracy, F1 score, sensitivity, specificity, standard deviation (SD), and the t-test are used with the ANN classifier.

Datasets Description

The study utilized a 5-year MBC dataset, as referenced by Marti et al., (2022) and Xu et al., (2023), consisting of 6,726 samples from both benign and malignant cases. The dataset includes 26 clinical features and a binary target class indicating whether or not a patient developed metastasis within 5 years of initial treatment. Notably, 92% of the samples are from unlikely (benign) cases, while only 8% are from likely (malignant) cases, making the dataset highly imbalanced.

The Proposed Chaotic Map-Driven JAYA for MBC Prediction

Figure 2 illustrates the system architecture of the proposed method. The architecture is structured into three successive stages, each designed to facilitate the overall optimization process. These stages include: Data Preprocessing, Handling Data Imbalance, and Optimization Using the improved JAYA Algorithm.

Data Preprocessing

The original 5-year MBC dataset contains three clearly irrelevant columns/features: a patient ID column and two unnamed columns with over 70% missing values. These columns are excluded during the initial data-cleaning phase to ensure data quality and relevance. The dataset predominantly consists of categorical data, with all values being non-numeric. To facilitate analysis, a LabelEncoder is used to encode these categorical variables as numerical values. Further preprocessing steps include normalizing the numerical features with MinMaxScaler, which scales them to a specified range, typically between 0 and 1. This step ensures consistency and enhances the algorithm's performance during analysis.

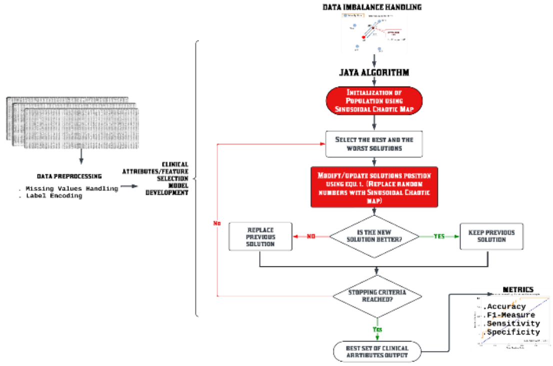

Figure 2: The Proposed System Architecture

Handling Data Imbalance

During this stage, the Synthetic Minority Oversampling Technique (SMOTE) is applied to address data imbalance in the preprocessed MBC dataset. As discussed in Section 3.3, the dataset contains 6,726 instances, of which only 533 (8%) are labeled malignant. This represents a clear case of severe class imbalance, which poses significant challenges for learning algorithms. Such imbalances can lead to biased model training, in which the algorithm becomes overly focused on the majority class, resulting in poor predictive performance on the minority class. Challenges include skewed decision boundaries, reduced sensitivity to malignant cases, and lower overall model reliability.

By using SMOTE, synthetic instances of the minority class are generated, effectively increasing its representation in the dataset. This approach not only balances the class distribution but also helps create a more robust model that effectively learns patterns from both majority and minority classes, thereby improving classification accuracy and recall for malignant cases.

Optimization Using the Enhanced JAYA Algorithm

During this stage, the balanced dataset is passed to the sinusoidal enhanced JAYA algorithm, which is tasked with intelligently selecting the most relevant features for MBC prediction. The sinusoidal chaotic map is used to replace pseudorandom numbers in two critical areas: population initialization and the update-stage initialization. These enhancements ensure a more diverse and effective search of the solution space, enabling the algorithm to identify optimal feature subsets with higher precision.

During each iteration of the FS process, the ANN classifier is used to evaluate the predictive quality of the selected features. By computing classification performance metrics (e.g., accuracy) based on the current feature subset, the ANN algorithm provides a feedback loop that guides the selection process toward refinement. This integration of JAYA with ANN ensures the selection of an informative feature set, ultimately enhancing the overall predictive capability and computational efficiency of the MBC classification model.

Parameters, Values, and Settings

The parameters and configurations for the JAYA algorithm, sinusoidal chaotic map, and ANN are summarized in Table 2. Initial settings were adopted from existing studies to ensure a fair comparison, and key parameters, particularly population size and the number of candidate solutions, were fine-tuned to optimize performance for the dataset and problem context.

Table 2: Parameter values and settings used in the proposed method

JAYA Algorithm	population size = 20, number of iterations = 20
Neural Network	Number of neurons (input layer) = number of dataset features, activation function (2 hidden layers) = Relu, activation function (output layer) = sigmoid, number of neurons (first hidden layer) = 64, number of neurons (second hidden layer) = 32, number of neurons (output layer) = 1, optimizer = adam, loss = binary_crossentropy, epoch = 10, batch_size = 32, validation_split = 0.1, dropout (2 layers) = 0.2
Sinusoidal CM	$x_0 = 0.7, P = 2.3$
Other Settings	train_test_spilt = 70:30, random_state = 42, number of algorithms run = 10

Result Presentation and Analysis

Experimental setup

All experiments were conducted in Python using a Jupyter Notebook environment. Computations were performed on a system with an Intel(R) Core(TM) i7-6600U CPU at 2.80 GHz and 8 GB of RAM. The presentation of results is carried out in two phases. First, the proposed method is compared with the findings of Muhammed et al. (2024) to assess its performance and effectiveness. This comparison is limited to a single study due to the scarcity of literature specifically focused on breast cancer metastasis. Subsequently, key parameters, particularly population size and the number of iterations, are fine-tuned to further enhance performance. The experimental results are presented in both tabular and graphical formats, providing a clear view of the proposed method's performance. The tables highlight key evaluation metrics, while the figures illustrate performance trends and comparative analyses. This combined presentation ensures an intuitive and comprehensive evaluation of the proposed approach.

Comparison with Existing Study

Table 3 summarizes the results obtained from the experiments conducted. In presenting these results, the notation $x \pm y$ is employed, where x represents the mean value of the metric under consideration, and $\pm y$ denotes the corresponding SD, reflecting the stability of the method across multiple runs. From Table 3, it is evident that the proposed method outperforms the existing study across all evaluated metrics, demonstrating its effectiveness in handling MBC prediction. Regarding **accuracy,** the proposed method achieved 79.76%, compared to 79.28% in the existing study, representing an improvement of 0.48 percentage points. For the **F1-score**, the proposed method attained 80.13%, surpassing the existing study's 79.66% by 0.47 percentage points. In terms of **sensitivity,** the proposed method achieved a remarkable 81.88%, outperforming the existing study's 81.04% by 0.84 percentage points. This improvement highlights the proposed method's enhanced ability to identify malignant cases correctly. For **specificity**, the proposed method achieved 77.52%, slightly surpassing the existing study's 77.51% by 0.01 percentage points. While the improvement in specificity is marginal, it reinforces the method's robustness in correctly classifying benign cases.

Table 3: Results from experiments with existing work in the literature

Methods	Accuracy	F1-Score	Sensitivity	Specificity
Proposed Method	**0.7976±0.0065**	**0.8013±0.0089**	**0.8188±0.0275**	**0.7752±0.0253**
Muhammed et al., (2024)	0.7928±0.0074	0.7966±0.0129	0.8104±0.0377	0.7751±0.0274

Figure 3 provides a graphical representation of these results, enabling a more visual explanatory comparison between the proposed and existing methods.

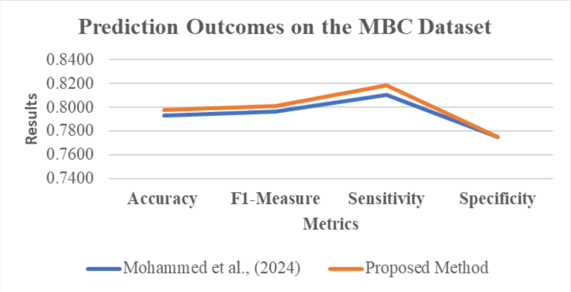

Figure 3: Performance comparison of the proposed method with the existing study

The experimental results presented highlight the effectiveness of the proposed methodology across multiple evaluation metrics. The improvements observed across metrics validate the enhancements introduced at each stage of the methodology. The data pre-processing steps, including data cleaning and normalization, streamlined the dataset by ensuring quality and uniformity. The use of the MinMaxScaler enhanced feature compatibility with the JAYA algorithm, while the LabelEncoder facilitated the seamless numerical representation of categorical data. These foundational steps minimized noise and redundancy, contributing to a more effective optimization and classification process.

Addressing the severe class imbalance in the MBC dataset through SMOTE proved instrumental in improving the classification performance. As shown in the results, the balanced dataset achieved higher sensitivity (81.88%), highlighting the model's ability to correctly identify malignant cases, which is critical for MBC prediction. Without SMOTE, the classifier would likely have been biased toward the majority class, resulting in reduced performance for the minority class. By replacing pseudorandom numbers with sinusoidal chaotic sequences during population initialization and parameter updates, the method ensured a more diverse and representative exploration of the solution space. This improvement directly addresses challenges such as premature convergence and insufficient search diversity that are common in the traditional JAYA algorithm. The iterative FS process, guided by feedback from the ANN classifier, yielded a highly relevant feature subset. This dynamic interaction between the JAYA algorithm and the ANN classifier contributed to superior classification metrics. For instance, the marginal yet consistent improvements in accuracy (79.76% vs. 79.28%) and F1-Score (80.13% vs. 79.66%) indicate that the proposed method is superior in balancing precision and recall. Furthermore, although the improvements across metrics were marginal, the T-test results shown in Table 4 confirm that these differences are statistically significant. A T-test evaluates whether the means of the two groups are statistically different, and in this study, it was used to compare the proposed method's performance with that of the existing study.

Table 4: T-test results between the proposed and existing works

Metrics	P_Values
Accuracy	0.0353
F1-Measure	0.0175
Sensitivity	0.0113
Specificity	0.0013

With a significance level of 0.05, all p-values were below this threshold, indicating that the improvements achieved by the proposed method are not due to chance but represent genuine advancements. This statistical significance reinforces the reliability and validity of the proposed methodology, even in the context of seemingly marginal metric gains.

Further Hyperparameter Tuning of the Proposed Method

Optimizing hyperparameters in MAs is essential for enhancing their performance in specific problem domains. These parameters play a pivotal role in determining the algorithm's convergence rate, maintaining a balance between exploration and exploitation, and ensuring overall effectiveness. This experiment further evaluates the impact of parameter tuning on the algorithm's overall performance. Using all four evaluation metrics, a comprehensive analysis was conducted to examine how variations in key parameters, such as population size and the number of iterations, influence the algorithm's effectiveness. These parameters were fine-tuned with a population size of 30 and an iteration count of 70. These values were selected after rigorous experimentation across multiple configurations to achieve optimal performance. Subsequently, a statistical test was performed on the results to assess their statistical significance and validate the observed improvements.

From Table 5, it is clear that the fine-tuned model slightly outperformed the original method, particularly in the F1 score and sensitivity metrics.

Table 5: Results from hyperparameter tuning

Method	Accuracy	F1-Score	Sensitivity	Specificity (Spec)
Finetuned model	0.7971±0.0071	**0.8020±0.0113**	**0.8218±0.0399**	0.7722±0.0373
Original model	**0.7976±0.0065**	0.8013±0.0089	0.8188±0.0275	**0.7752±0.0253**

While these improvements are marginal, the T-test results presented in Table 6 confirm that the observed differences are statistically significant. Notably, bolded values in the table correspond to p-values greater than the significance threshold of 0.05, indicating that these differences are not statistically significant. Among the evaluated metrics, only specificity exhibited a p-value exceeding the significance level. This outcome suggests that the fine-tuning process had a limited impact on specificity, potentially due to the model's inherent stability or the dataset's minimal variability. This insight highlights the complex influence of parameter adjustments on individual performance metrics.

Table 6: T-test results between the original and finetuned methods

Metrics	P_Values
Accuracy	0.0456
F1-Measure	0.0510
Sensitivity	0.0111
Specificity	**0.0611**

In conclusion, the proposed method demonstrated competitive performance across multiple metrics, with fine-tuning further enhancing its effectiveness in certain areas such as F1 score and sensitivity. The T-test results confirm the statistical significance of these improvements, reinforcing their robustness. However, the limited impact on specificity highlights opportunities for further refinement.

CONCLUSION AND FUTURE RESEARCH DIRECTIONS

This research focused on developing an enhanced feature selection method for predicting breast cancer metastasis using an improved JAYA algorithm. The study began with data preprocessing, in which irrelevant columns and missing values were addressed, and categorical data were numerically encoded. Further preprocessing involved normalizing the dataset using MinMaxScaler to standardize feature ranges. Given the severe class imbalance, SMOTE was applied to balance the dataset, thereby improving model training and performance. The JAYA algorithm, enhanced with a sinusoidal chaotic map, was employed for intelligent feature selection. This enhancement addressed issues such as limited diversity and premature convergence by improving the initialization of the population and control parameters. An artificial neural network was integrated to iteratively evaluate the quality of the selected feature subsets, guiding the optimization process. Extensive experiments were conducted, and results demonstrated that the proposed method outperformed the existing approach across all metrics, including accuracy, F1-score, sensitivity, and specificity. Fine-tuning of hyperparameters, such as population size and iteration count, further enhanced the algorithm's effectiveness, with statistical tests confirming the significance of the improvements. This work advances predictive models for metastatic breast cancer by proposing an optimized, data-driven approach that balances exploration and exploitation, addresses class imbalance, and achieves superior predictive performance.

Several areas exist for future enhancement of the proposed method. Incorporating advanced machine learning models, such as ensemble techniques, could significantly improve classification accuracy and robustness. Extending the enhanced JAYA algorithm to other medical or high-dimensional datasets would help validate its generalizability and effectiveness across diverse domains. Additionally, exploring alternative chaotic maps may uncover optimal configurations for population initialization and parameter tuning, further improving performance.

REFERENCES

Abd Elminaam, D. S., Ibrahim, S. A., Houssein, E. H., & Elsayed, S. M. (2022). An Efficient Chaotic Gradient-Based Optimizer for Feature Selection. *IEEE Access*, *10*, 9271–9286.

Abdelrazek, M., Abd Elaziz, M., & El-Baz, A. H. (2024). CDMO: Chaotic Dwarf Mongoose Optimization Algorithm for feature selection. *Scientific Reports*, *14*(1), 1–18.

Abdollahi, J., Davari, N., Panahi, Y., &Gardaneh, M. (2022). Detection of Metastatic Breast Cancer from Whole-Slide Pathology Images Using an Ensemble Deep-Learning Method. *Archives of Breast Cancer*, *9*(3), 364–376.

Abdollahi, J., Keshandehghan, A., Gardaneh, M., Panahi, Y., &Gardaneh, M. (2020). Accurate Detection of Breast Cancer Metastasis Using a Hybrid Model of Artificial Intelligence Algorithm. *Archives of Breast Cancer*, *7*(1), 22–28.

Abed-alguni, B. H., & AL-Jarah, S. H. (2024). IBJA: An improved binary DJaya algorithm for feature selection—Journal *of Computational Science*, *75*, 102201.

Agrawal, P., Ganesh, T., & Mohamed, A. W. (2021). Chaotic gaining sharing knowledge-based optimization algorithm: an improved metaheuristic algorithm for feature selection. *Soft Computing*, *25*(14), 9505–9528.

Alhassan, A. M. (2024). An improved breast cancer classification with a hybrid chaotic sand cat and Remora Optimization feature selection algorithm. *PLoS ONE*, *19*(4 April), 1–30.

Azar, A. T., Anter, A. M., & Fouad, K. M. (2020). Intelligent system for feature selection based on rough set and chaotic binary grey Wolf optimisation. *International Journal of Computer Applications in Technology*, *63*(1–2), 4–24.

Azubuike, S. O., Hayes, L., Sharp, L., Alabi, A., Oyesegun, R., Azubuike, S. O., Sharp, L., & Alabi, A. (2022). *Socio-economic status and the risk of breast cancer among Nigerian women : a case control study*.

Beghriche, T., Djerioui, M., Brik, Y., Attallah, B., &Belhaouari, S. B. (2021). An Efficient Prediction System for Diabetes Disease Based on a Deep Neural Network. *Complexity*, *2021*.

Botlagunta, M., Botlagunta, M. D., Myneni, M. B., Lakshmi, D., Nayyar, A., Gullapalli, J. S., & Shah, M. A. (2023). Classification and diagnostic prediction of breast cancer metastasis on clinical data using machine learning algorithms. *Scientific Reports*, *13*(1), 1–17.

Chakravarthy, S., Nagarajan, B., Kumar, V. V., Mahesh, T. R., Sivakami, R., & Annand, J. R. (2024). Breast Tumor Classification with Enhanced Transfer Learning Features and Selection Using Chaotic Map-Based Optimization. *International Journal of Computational Intelligence Systems*, *17*(1).

Chaudhuri, A., & Sahu, T. P. (2021). A hybrid feature selection method based on the Binary Jaya algorithm for micro-array data classification. *Computers and Electrical Engineering*, *90*(November 2020), 106963.

Elgamal, Z. M., Binti, N., Yasin, M., Tubishat, M., Alswaitti, M., &Mirjalili, S. (2020). *An Improved Harris Hawks Optimization Algorithm With Simulated Annealing for Feature Selection in the Medical Field*. *8*.

Elgamal, Z., Sabri, A. Q. M., Tubishat, M., Tbaishat, D., Makhadmeh, S. N., & Alomari, O. A. (2022). Improved Reptile Search Optimization Algorithm Using Chaotic Map and Simulated Annealing for Feature Selection in the Medical Field. *IEEE Access*, *10*, 51428–51446.

Feizi-Derakhsh, M. R., & Kadhim, E. A. (2023). An Improved Binary Cuckoo Search Algorithm For Feature Selection Using the Filter Method And Chaotic Map. *Journal of Applied Science and Engineering*, *26*(6), 897–903.

Gharehchopogh, F. S., Maleki, I., &Dizaji, Z. A. (2022). Chaotic vortex search algorithm: metaheuristic algorithm for feature selection. In *Evolutionary Intelligence* (Vol. 15, Number 3). Springer Berlin Heidelberg.

Hassan, I. H., Abdullahi, M., Isuwa, J., Isah, A., & Kim, J. (2023). *Microarray Cancer Gene Selection using Pelican Optimization Algorithm*. *5*(1), 609–620.

Isuwa, J., Abdullahi, M., & Abdulrahim, A. (2022). *Hybrid particle swarm optimization with sequential one-point flipping algorithm for feature selection*. (July), 1–18.

Isuwa, J., Abdullahi, M., Ali, Y. S., Kim, J., Hassan, I. H., & Buba, J. R. (2023). Optimizing Microarray Cancer Gene Selection using Swarm Intelligence : Recent Developments and An Exploratory Study Optimizing Microarray Cancer Gene Selection using Swarm Intelligence : Recent. *Egyptian Informatics Journal*, *24*(4), 100416.

Isuwa, J., Mohammed, A., Yusuf, S. A., Idris, M. N., & Garko, B. S. (2021). *Hybridization of Metaheuristic Algorithms with Local Search Methods for Optimizing the Feature Selection Process : A Survey*. (Iccait).

Jeremiah, I., Abdullahi, M., Yusuf, S. A., & Idris, M. N. (2022). *Integration of Specific Local Search Methods in Metaheuristic Algorithms for Optimizing the Feature Selection Process : A Survey*. *4*(1), 34–48.

Jiang, X., & Xu, C. (2022). Deep Learning and Machine Learning with Grid Search to Predict Later Occurrence of Breast Cancer Metastasis Using Clinical Data. *Journal of Clinical Medicine*, *11*(19).

Khurma, R. A., Aljarah, I., &Sharieh, A. (2020). An efficient moth flame optimization algorithm using chaotic maps for feature selection in medical applications. *ICPRAM 2020 - Proceedings of the 9th International Conference on Pattern Recognition Applications and Methods*, (January), 175–182.

Li, Y., Geng, Y., & Sheng, H. (2024). An improved mountain gazelle optimizer based on a chaotic map and spiral disturbance for medical feature selection, in *PLoS ONE* (Vol. 19, Number 7, July).

Marti, J. L. G., Brufsky, A., Wells, A., & Jiang, X. (2022). Machine Learning to Discern Interactive Clusters of Risk Factors for Late Recurrence of Metastatic Breast Cancer. *Cancers*, *14*(1).

Muhammed, B. J., Yusuf Musa Malgwi, J. I., Mohammed Abdullahi, & Zuriani Mustaffa. (2024). A Neural Network-based Prediction of Later Occurrence of Breast Cancer Metastasis in an Imbalanced Data Setting. 1–8.

Naskar, A., Pramanik, R., Hossain, S. K. S., Mirjalili, S., & Sarkar, R. (2023). Late-acceptance hill-climbing-aided chaotic harmony search for feature selection: An empirical analysis on medical data—Expert *Systems with Applications*, *221*(February), 119745.

Nassar, H. B. (2023). Classification for Imbalanced Breast Cancer Dataset Using Resampling Methods. *Ijcsns*, (February).

Noshad, A., & Fallahi, S. (2023). A new hybrid framework based on deep neural networks and JAYA optimization algorithm for feature selection using SVM applied to classification of acute lymphoblastic Leukaemia. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, *11*(4), 1549–1566.

Olayide, A., Isiaka, A., Ganiyu, R., Samuel, O., Olalekan, O., Julius, O., & Anya, R. (2023). *Breast Cancer Treatment and Outcomes in Nigeria : A Systematic Review and Meta-analysis*. *8*(3), 591–598.

Tran, T., Le, U., & Shi, Y. (2022). An effective up-sampling approach for breast cancer prediction with imbalanced data: A machine learning model-based comparative analysis. *PLoS ONE*, *17*(5 May).

Venkata Rao, R. (2016). Jaya: A simple and new optimization algorithm for solving constrained and unconstrained optimization problems. *International Journal of Industrial Engineering Computations*, *7*(1), 19–34.

Xu, C., Coen-Pirani, P., & Jiang, X. (2023). Empirical Study of Overfitting in Deep Learning for Predicting Breast Cancer Metastasis. *Cancers*, *15*(7).

Zaimoğlu, E. A., Yurtay, N., Demirci, H., & Yurtay, Y. (2023). A binary chaotic horse herd optimization algorithm for feature selection: Engineering *Science and Technology, an International Journal*, *44*.