

# Dual Metaheuristic Optimization Algorithms for Breast Cancer Diagnosis Model

Musa Bala Usman<sup>1</sup>, Jimoh-Mahmud Aishat Oladayo<sup>2</sup>, Buraimoh Olawale Henry<sup>3</sup>, Lawal Hussein Taiye<sup>4</sup>,  
Oladejo Qazeem Ayokunle<sup>5</sup>

<sup>1</sup>Department of Software Development, Directorate of ICT, Ibrahim Badamasi Babagida University, Lapie,  
Minna, Niger State, Nigeria

<sup>2,3,5</sup>Department of Computer Science, Alhikmah University, Ilorin, Kwara State, Nigeria

<sup>4</sup>National Open University of Nigeria, Ilorin, Kwara State, Nigeria

Hlawal@noun.edu.ng<sup>4</sup>

## Abstract

*Data mining techniques are increasingly applied in the medical field for disease diagnosis and classification, including breast cancer, due to their high diagnostic accuracy. A major challenge in breast cancer diagnosis lies in managing numerous features, which makes efficient feature extraction and selection essential for improving classification performance. This study aims to conduct a comparative analysis of machine learning algorithms using dual meta-heuristic optimization techniques—Firefly and Cuckoo algorithms for effective breast cancer diagnosis. The Wisconsin Breast Cancer Dataset was used as the benchmark dataset. Two feature selection methods, Firefly and Cuckoo optimization algorithms, were employed to identify the most relevant features. These selected features were then classified using six machine learning algorithms: Support Vector Machine (SVM), Decision Tree (C4.5), K-Nearest Neighbors (KNN), Neural Network (NN), Logistic Regression (LR), and Random Forest (RF) with stochastic gradient enhancement. Model performance was evaluated using accuracy, precision, recall, F1-measure, kappa statistics, and model-building time. The experimental results revealed that the highest classification accuracy of **97.74%** was achieved with the Cuckoo optimization algorithm combined with the RF classifier. Meanwhile, the best precision, recall, and F1-measure values (all **0.9780**) were obtained with the Firefly optimization algorithm coupled with the RF classifier. These findings highlight the effectiveness of meta-heuristic feature selection in improving breast cancer diagnosis.*

**Keywords:** Breast cancer, machine learning algorithm, data mining, meta-heuristic optimization

## 1. Introduction

The healthcare industry has changed over the ages due to the introduction of new medical techniques (Ostańska, Aebisher, & Bartusik-Aebisher, 2021). To serve individuals in need, the medical community has historically brought together the best and brightest members of society (Nunzio et al., 2022). Doctors have created new technologies and better treatment techniques for everything from cancer therapy and childbirth to heart attacks. Early detection of any disease is crucial for treatment or cure in its early stages and for

extending life expectancy, making it the most significant aspect of the medical field or health care (Ginsburg et al., 2021). One such illness where early diagnosis can lower cancer patient mortality is cancer. There is a very high risk of breast cancer in developing nations, which is viewed as a serious threat (Sohail, Jiadong, Irshad, Uba, & Abir, 2018). Breast cancer is a prevalent form of cancer that arises from abnormal cell development, resulting in tumors that alter normal cells in the breast area of the body (Dharambir et al., 2023). The specific location where cell or tissue cancer forms determines the cancer's classification as a disease in the medical field. Prompt detection is the best way to reduce breast cancer (Rahman & Siti Nurzaimah Nazhirah Zaim, 2024). A timely diagnosis involves a proper and trustworthy diagnostic method that enables medical professionals to distinguish between benign and malignant breast tumors without performing a surgical biopsy.

In the medical field, distinguishing tumors can be challenging, even for experienced professionals and specialists. As a result, automating diagnostic models is essential for tumor diagnosis. Numerous researchers have utilized data mining algorithms to detect cancer survivability in patients, and these algorithms have proven to be highly effective in cancer diagnosis (Kaur et al., 2022). Data mining techniques have significantly contributed to converting vast amounts of data into specific and pertinent information for knowledge discovery and predictive purposes (Kaderye, Arif, & Kundu, 2024).

In the fields of pattern recognition, statistics, and data mining, feature selection is recognized as an active research subject. By removing features that have little to no predictive value, feature selection is used to produce a subset of input variables (Shardlow, McIntyre, Fluck, McIntyre, & Taal, 2016). This method can greatly increase the comprehensibility of classifier models and frequently create a model that generalizes more effectively to unknown points (Shardlow et al., 2016). A key pre-processing step in the data mining technique is feature selection, which is used to reduce large data dimensions (Saini & Sharma, 2018), eliminate unsuitable attributes, improve learning accuracy, and make results easier to understand (Hayashi & Takano, 2020; Awotunde et al., 2022)

Metaheuristic optimization algorithms as feature selection approaches encompass a variety of nature-inspired techniques, such as the Cuckoo Algorithm, Genetic Algorithm (GA), Firefly Algorithm, Tabu Search (TS), Bat Optimization Algorithm (BOA), Ant Colony Optimization (ACO), Genetic Programming (GP), and Particle Swarm Optimization (PSO), among others (Kaveh & Mesgari, 2023; Navarro & Navarro, 2024). These techniques are particularly effective for feature selection due to their straightforward representation of feature subsets and ease of evaluation. Additionally, the capacity of these heuristic approaches to navigate large search spaces efficiently is highly beneficial for feature selection problems. Understanding the impact of these metaheuristic algorithms on breast cancer prediction is crucial. Among the aforementioned metaheuristic optimization algorithms, firefly and cuckoo have been considered to perform optimally in the selection of relevant features before the classification phase. The two algorithms' efficiencies as optimization approaches are due to the fact that cuckoo offers a simple but still effective approach to solving complex optimization problems, outperforming other algorithms in many cases due to their robust exploration and exploitation of solution spaces. The firefly algorithm is important for its ability to solve complex optimization problems effectively and efficiently, particularly in areas like engineering, networking, and image processing, by mimicking the social behaviour of fireflies (Alloush, Rebdawi, & Trab, 2024; Panliang et al., 2025).

## 2. Related Works

Sarker et al. (2025) handled the huge data by reducing the features while maintaining performance. A Wrapper Model with Metaheuristic Algorithms: Whale Optimization, Bald Eagle Search, and Sea Lion Optimization and Extreme Gradient Boost Classifier. SHAP explained feature importance for both overall and individual predictions. Our study achieved F-scores of 97.43%, 95%, and 94.74% for SLOA\_XGB, BESA\_XGB, and WOA\_XGB, respectively, on the Breast Cancer Coimbra dataset. Each method reduced the features from 9 to 4. SHAP analysis identified Glucose as having the highest impact on model predictions. Also, the model identified a link between the mean values of certain features and misclassification likelihood. This study analyzed data from 116 subjects, with the SLOA\_XGB classifier having the best performance: 97.43% of F-score, 97.14% of accuracy, 97.14% of precision, and 100% of recall using only Glucose, Age, Resistin, and Adiponectin. Bilal et al. (2024) developed the improved quantum-inspired binary Grey Wolf Optimizer with the Support Vector Machines Radial Basis Function Kernel. This hybrid approach was to enhance the accuracy of breast cancer classification by determining the optimal Support Vector Machine parameters. The motivation for this hybridization lies in the need for improved classification performance compared to existing optimizers such as Particle Swarm Optimization and Genetic Algorithm. The evaluation of the proposed IQI-BGWO-SVM approach on the MIAS dataset based on accuracy, sensitivity, and specificity. Furthermore, the application of IQI-BGWO-SVM for feature selection was explored, and the results were compared. Experimental findings demonstrate that the suggested QI-BGWO-SVM technique outperforms state-of-the-art classification methods on the MIAS dataset with a resulting mean accuracy, sensitivity, and specificity of 99.25%, 98.96%, and 100%, respectively, using a tenfold cross-validation dataset partition. EL Kati, Wang, Taresh, and Ali (2024) built a hybrid optimization algorithm using a simple gradient search mechanism into a metaheuristic technique, the multi-verse optimizer (MVO), to facilitate the search for a global optimal solution in the high-dimensional search space of DL models. A DL model for BC diagnosis is developed based on a three-hidden-layer DNN whose parameters are optimized using the proposed hybrid optimizer. Experimental analysis was carried out using the Wisconsin Breast Cancer dataset (WBCD) and the Wisconsin Diagnosis Breast Cancer (WDBC) dataset, each of which is divided into 70% for training and 30% for testing. For comparison reasons, similar DL models trained using various optimizers, including gradient-based, metaheuristic, and recently-proposed hybrid optimization algorithms, are also analyzed. The results revealed the effective performance of the proposed optimizer in terms of attaining the most accurate DL model in the fastest convergence rate. The proposed model gave accuracy of 93.5%, precision of 88.06%, specificity of 93.06%, sensitivity of 95.64%, F1 score of 91.67%, and Matthew's correlation coefficient (MCC) of 87.14% on WBCD, and accuracy of 96.73%, precision at 93.38%, specificity at 95.83%, sensitivity at 98.25%, F1 score at 95.75%, and MCC at 93.18% on WDBC, in just six epochs. Thirumalaisamy et al (2023) developed a comprehensive classification technique for identifying breast cancer by applying a synthesized CNN, an enhanced optimization algorithm, and transfer learning. The primary goal was to assist radiologists in rapidly identifying anomalies. To overcome inherent limitations, ACO was modified with opposition-based learning (OBL). The Enhanced Ant Colony Optimization (EACO) methodology was then employed to find the optimal hyperparameter values for the CNN architecture. The proposed framework combined the Residual Network-101 (ResNet101) CNN architecture with the EACO algorithm, resulting in a new model dubbed EACO-ResNet101. Experimental analysis was conducted on the MIAS and DDSM (CBIS-DDSM) mammographic datasets. Compared to conventional methods, our proposed model achieved an impressive accuracy of 98.63%, sensitivity of 98.76%, and specificity of 98.89% on the CBIS-DDSM dataset. On the MIAS dataset, the proposed model achieved a classification accuracy of 99.15%,

a sensitivity of 97.86%, and a specificity of 98.88%. These results demonstrate the superiority of the proposed EACO–ResNet101 over current methodologies.

### 3. Methodology

The study simulated a breast cancer model in a Weka 3.9.4 data mining environment. The raw dataset was acquired from the publicly available Wisconsin breast cancer dataset. The dataset was preprocessed to normalize the data before the simulation procedures. Two optimization metaheuristic methods, the Cuckoo Algorithm and Firefly Algorithm, were used to rank the features based on their relevance to choose discriminant features, or, most importantly for breast cancer diagnosis, features from the Breast Cancer dataset. Some selected: C4.5, SVM, KNN, NN, LR, and RF were used to classify whether the cancer is malignant or benign. The framework of the breast cancer diagnostic model is represented in Figure 1

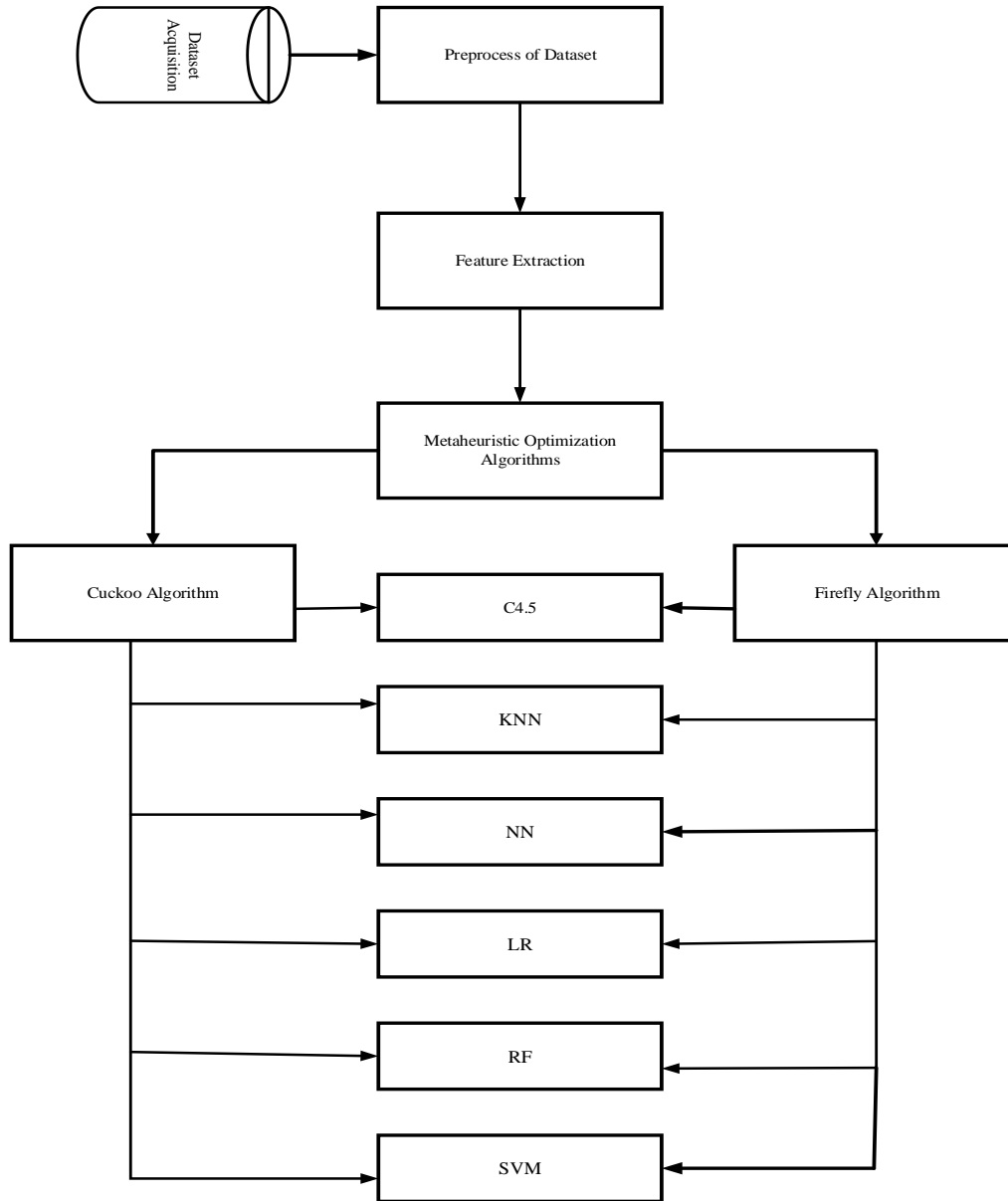


Figure 1: Breast Cancer Diagnostic Model

### 3.1 Data Acquisition and Preprocessing

The constructed breast cancer diagnosis model was evaluated in this work using the Wisconsin Breast Cancer Dataset (WDBC). The following are specifics of the attributes contained in the WDBC dataset: For every cell nucleus, the following 10 real-valued properties are calculated: radius, texture, perimeter, area, smoothness, compactness, concavity, concave points, symmetry, and fractal dimension. The ID number and diagnosis (M = malignant, B = benign) are also recorded. A digital image of a fine needle aspiration (FNA) of a breast mass is used to calculate these features. They describe the characteristics of the cell nuclei observed in the image. When the length of the radial line segments formed by the snake's

centroid and individual snake points is averaged, it gives the radius of a single nucleus. The total distance between consecutive snake points constitutes the nucleus's perimeter. By adding half of the pixels on the snake's perimeter and calculating the number of pixels on its interior, the area is computed. The formula provides a measure of the compactness of the cell nuclei by combining the area and perimeter. The difference between a radial line's length and the mean length of the lines surrounding it is used to calculate smoothness. The main axis, or the longest chord across the center, is used to quantify symmetry. Next, the length difference between lines in both directions that are perpendicular to the nuclear boundary and the major axis is measured.

Mandelbrot's "coastline approximation" is used to estimate the fractal dimension of a nuclear barrier. Increasingly larger "rulers" are used to measure the nucleus's circumference. The observed perimeter decreases as the size of the ruler grows, reducing the measurement's accuracy. An approximation of the fractal dimension can be obtained by measuring the downward slope and graphing the log of the observed perimeter against the log of ruler size. A greater number, across all shape criteria, indicates a less regular contour and, thus, a higher likelihood of malignancy. The variation of the grayscale intensities in the component pixels is used to quantify the texture of the cell nucleus.

### 3.2 Features Selection

Three meta-heuristic optimization techniques: cuckoo search algorithm, particle swarm optimization, and firefly were employed in this work. The following subsections give the algorithms of metaheuristic optimization algorithms used:

#### 3.2.1 Cuckoo Search Algorithm

The obligatory brood parasitism of certain cuckoo species, which involves laying their eggs in other bird species' nests, serves as the inspiration for the cuckoo search algorithm (CSA). As seen in Figure 2. Wei & Yu (2018) outline the steps of the cuckoo algorithm that were implemented.

```

Objective function:  $f(\mathbf{x})$ ,  $\mathbf{x} = (x_1, x_2, \dots, x_d)$ ;
Generate an initial population of  $n$  host nests;
While (t < MaxGeneration) or (stopping criterion)
  Get a cuckoo randomly (say,  $i$ ) and replace its solution by performing Lévy flights;
  Evaluate its quality/fitness  $F_i$ 
  [For maximization,  $F_i \propto f(\mathbf{x}_i)$  ];
  Choose a nest among  $n$  (say,  $j$ ) randomly;
  if ( $F_i > F_j$ ),
    Replace  $j$  by the new solution;
  end if
  A fraction ( $p_a$ ) of the worse nests are abandoned and new ones are built;
  Keep the best solutions/nests;
  Rank the solutions/nests and find the current best;
  Pass the current best solutions to the next generation;
end while

```

Figure 2: Cuckoo Algorithm

### 3.2.2 Firefly Algorithm

This algorithm is a computationally efficient, nature-inspired metaheuristic population-based optimization algorithm that mimics a firefly's attraction to a flashing light (Ali et al., 2014). The firefly algorithm stepwise approach is described in Figure 2

```
Begin
1) Objective function:  $f(\mathbf{x})$ ,  $\mathbf{x} = (x_1, x_2, \dots, x_d)$ ;
2) Generate an initial population of fireflies  $x_i$  ( $i = 1, 2, \dots, n$ );
3) Formulate light intensity  $I$  so that it is associated with  $f(\mathbf{x})$ 
   (for example, for maximization problems,  $I \propto f(\mathbf{x})$  or simply  $I = f(\mathbf{x})$ ;)
4) Define absorption coefficient  $\gamma$ 

while (t < MaxGeneration)
  for i = 1 : n (all n fireflies)
    for j = 1 : i (n fireflies)
      if ( $I_j > I_i$ ),
        Vary attractiveness with distance  $r$  via  $\exp(-\gamma r)$ ;
        move firefly  $i$  towards  $j$ ;
        Evaluate new solutions and update light intensity;
      end if
    end for j
  end for i
  Rank fireflies and find the current best;
end while
end
```

Figure 3: Algorithm for Firefly

## 4. Results and Discussions

The experimental analysis was conducted in WEKA data mining, which consists of a Graphical User Interface (GUI) Chooser, which enables users to navigate to different interfaces for diverse operations. The applications are with five buttons: Explorer, Experimenter, Knowledge Flow, Workbench, and Simple CLI, as illustrated in Figure 3.



Figure 4: WEKA Tool Applications Interface (GUI Chooser)

From the interface, the Explorer button is clicked to navigate to the Explorer Interface. This interface is required for this study to achieve its purpose. The pictorial view is shown in Figure 3

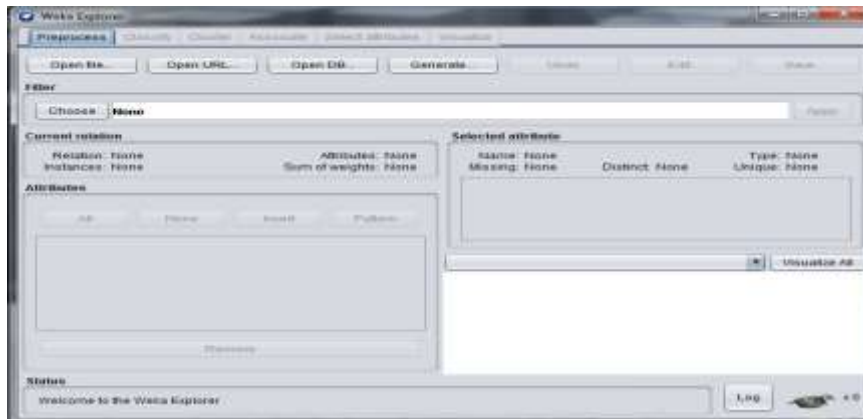


Figure 5: The WEKA Explorer Interface.

The interface shown in Figure 3 supports operations such as data pre-processing, classification, Clustering, Association, Selection of attributes, and Visualization of data. This study employed the WEKA Explorer Interface for data pre-processing and classification operations.

#### 4.1 Dataset Loading

The breast cancer Wisconsin (original) dataset is loaded into the WEKA data mining tool as shown in Figure 6.

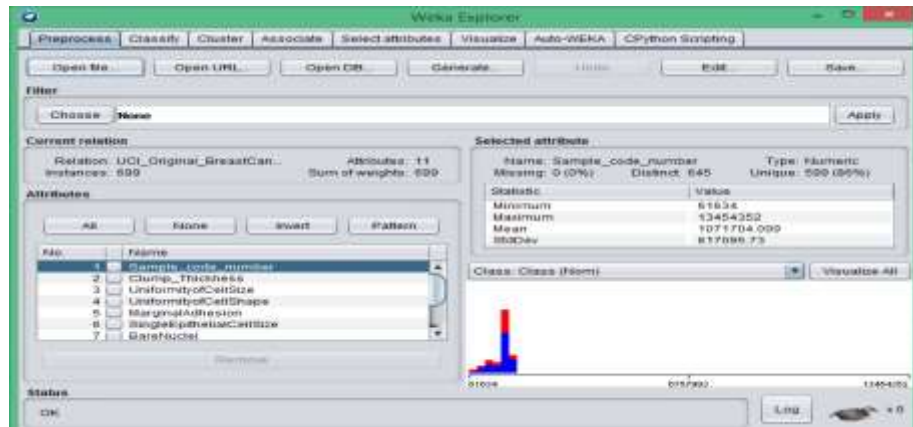


Figure 6: Loading of the Dataset

#### 4.2 Data Pre-processing

The breast cancer Wisconsin (original) dataset, after loading, was pre-processed with the WEKA data mining tool. The initial class value type was transformed from numerical to a categorical variable by using the textual word for the numbers as described in Figure 7.



Figure 7: Data Pre-processed Stage

#### 4.3 Feature / Attribute Selection Algorithms

The breast cancer Wisconsin (original) dataset after the pre-processed stage was passed to four different swarm intelligence algorithms: the firefly algorithm and cuckoo algorithm to obtain the most relevant attributes from the dataset. The nine most relevant features, which are Clump Thickness, Uniformity of Cell Size, Uniformity of Cell Shape, Marginal Adhesion, Single Epithelial Cell Size, Bare Nuclei, Bland Chromatin, Normal Nucleoli, and Mitoses, were selected for breast cancer classification as shown in the following Figures 8 and 9.



From Table 1, the best accuracy value of 97.7090% was obtained in the RF classifier, the best precision value of 0.9770 was obtained in the RF classifier, the best recall value of 0.9770 was recorded in the RF classifier, and the best kappa statistics value of 0.9376 was obtained in the RF classifier. The lowest time of 0.6s taken to build a model was achieved in the KNN classifier.

Table 2: Feature Selection Using Firefly Algorithm

Classifier	Accuracy (%)	Precision	Recall	F-Measure	Kappa-Statistic	Time-Taken (s)
C4.5	95.6580	0.9560	0.9560	0.9660	0.8899	0.04
KNN	95.4590	0.9540	0.9540	0.9540	0.8975	0
SVM	96.9950	0.9700	0.9700	0.9700	0.9437	0.08
NN	96.2790	0.9630	0.9630	0.9650	0.8975	1.18
LR	96.7670	0.9670	0.9670	0.9670	0.9340	0.4
RF	97.8040	0.9730	0.9780	0.9780	0.9477	0.14

From Table 2, the best accuracy value of 97.8040% was obtained in RF, the best precision value of 0.9730 was obtained in RF, the best recall value of 0.9780 was recorded in RF, the best F1-measure value of 0.9780 was recorded in RF, and the best kappa statistics value of 0.9477 was obtained in RF. The lowest time value of 0s taken to build a model was achieved in KNN.

Table 3: Feature Selection Using Cuckoo Algorithm

Classifier	Accuracy (%)	Precision	Recall	F-Measure	Kappa-Statistic	Time-Taken (s)
C4.5	94.7630	0.9470	0.9470	0.9470	0.8898	0.03
KNN	95.3059	0.9530	0.9530	0.9530	0.8920	0
SVM	96.9980	0.9700	0.9700	0.9700	0.9337	0.38
NN	95.3690	0.9540	0.9540	0.9540	0.8978	0.84
LR	96.7660	0.9660	0.9660	0.9660	0.9240	0.03
RF	97.9350	0.9730	0.9730	0.9730	0.9390	0.11

From Table 3, the highest accuracy of 97.9350% was obtained in the RF classifier, the highest precision value of 0.9730 was obtained in the RF classifier, the highest recall value of 0.9730 was recorded in the RF classifier, the highest F1-measure value of 0.9370 was recorded in the RF classifier, and the highest kappa statistics of

0.9390 was obtained in the RF classifier. The lowest time value of 0s taken to build a model was recorded in KNN.

#### 4.5 Results of Comparative Analysis for Breast Cancer Diagnostic Model.

The comparative analysis was conducted for the different evaluation metrics used in this study. The different results are compared and represented in the following graphical representation.

##### 4.5.1 Comparative Analysis of Breast Cancer Diagnostic Model (Accuracy %)

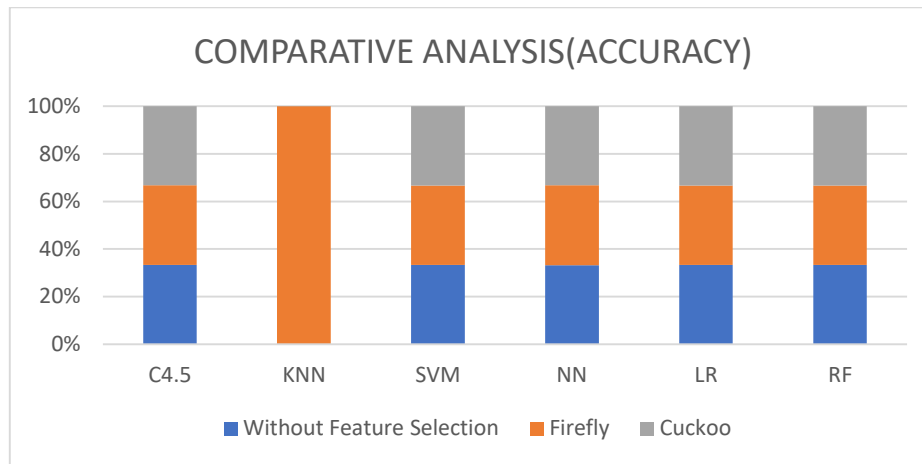


Figure 10: Comparative Analysis (Accuracy)

From Figure 10, the best accuracy value of 97.9350 % was obtained in feature selection using the Cuckoo algorithm applying the RF classifier, while the lowest accuracy of 95.6580% was in feature selection, with feature selection using the firefly algorithm applying the C4.5 classifier.

##### 4.5.2 Comparative Analysis of Breast Cancer Diagnostic Model (Precision)

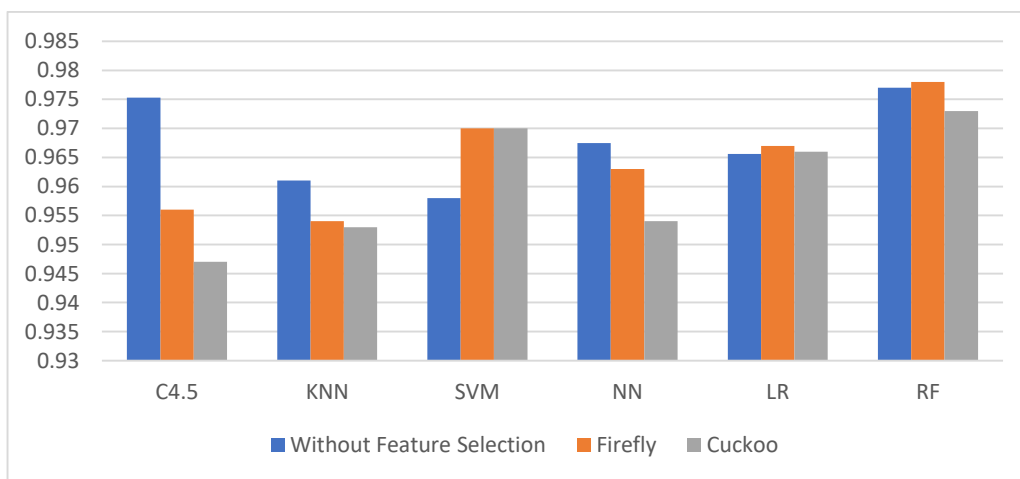


Figure 11: Comparative Analysis (Precision)

From Figure 11, the best precision value of 0.9780 was obtained in feature selection using the Firefly algorithm applying the RF classifier, while the lowest precision value of 0.9470 was recorded in feature selection using the Cuckoo algorithm applying the C4.5 classifiers.

#### 4.5.3 Comparative Analysis of Breast Cancer Diagnostic Model (Recall)

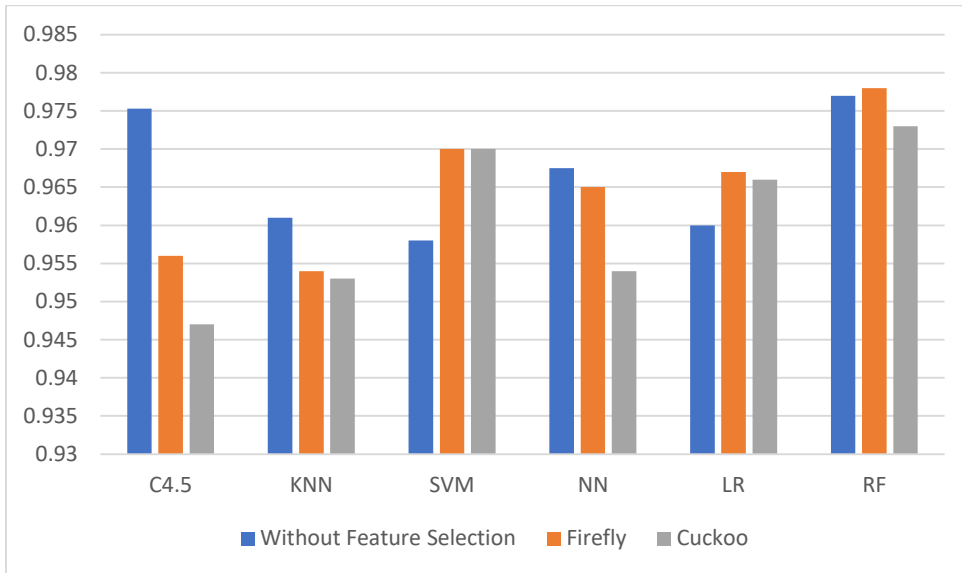


Figure 12: Comparative Analysis (Recall)

From Figure 12, the best recall value of 0.9780 was obtained in feature selection using the Firefly algorithm applying the RF classifier, while the lowest recall value of 0.9470 was recorded in the Cuckoo feature selection.

#### 4.5.4 Comparative Analysis of Breast Cancer Diagnostic Model (F1-Measure)

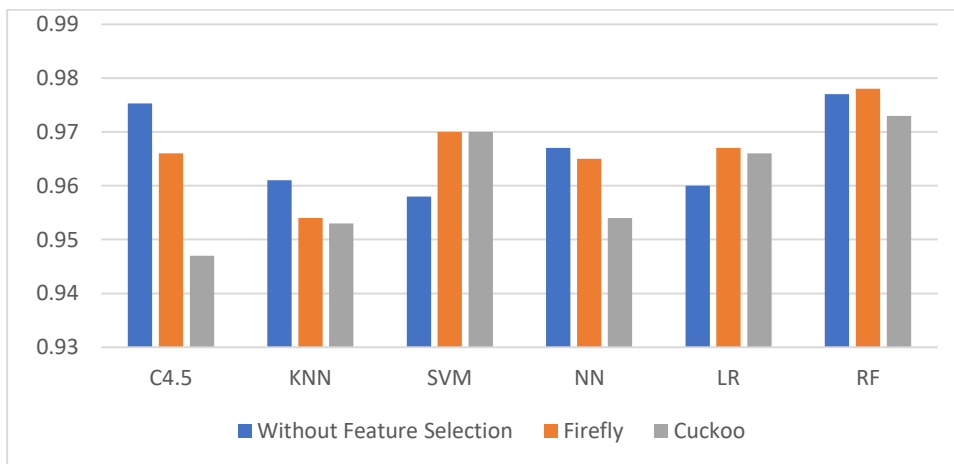


Figure 13: Comparative Analysis (F1-measure)

From Figure 13, the best F1-measure value of 0.9780 was obtained in feature selection using the Firefly algorithm applying the RF classifier, while the lowest F1-measure value of 0.9470 was recorded in feature selection using Cuckoo.

#### 4.5.5 Comparative Analysis of Breast Cancer Diagnostic (Kappa Statistic)

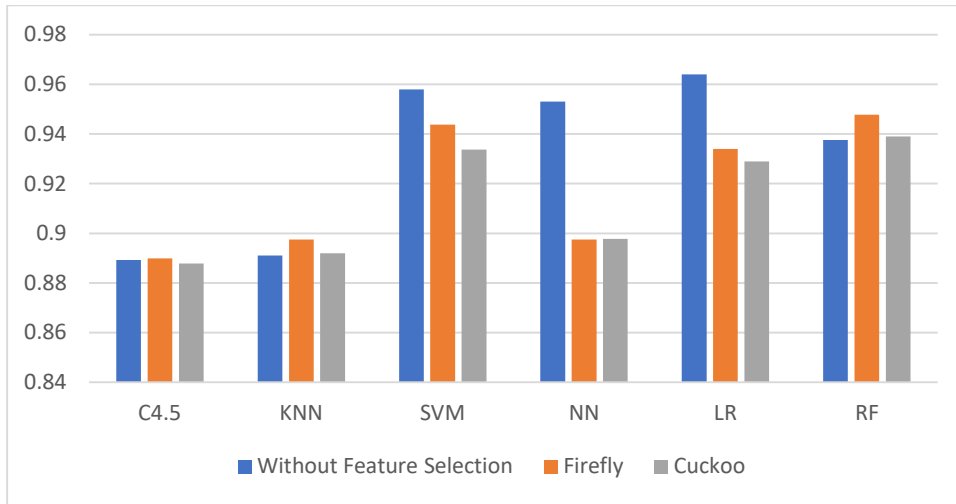


Figure 14: Comparative Analysis (F1-measure)

From Figure 14, the best kappa statistic value of 0.9640 was obtained when no feature selection was applied to the LR classifier, while the lowest kappa statistic value of 0.8878 was recorded in feature selection using Cuckoo applying C4.5.

#### 4.5.6 Comparative Analysis of Time Taken to Build Breast Cancer Diagnostic (Time-Taken)

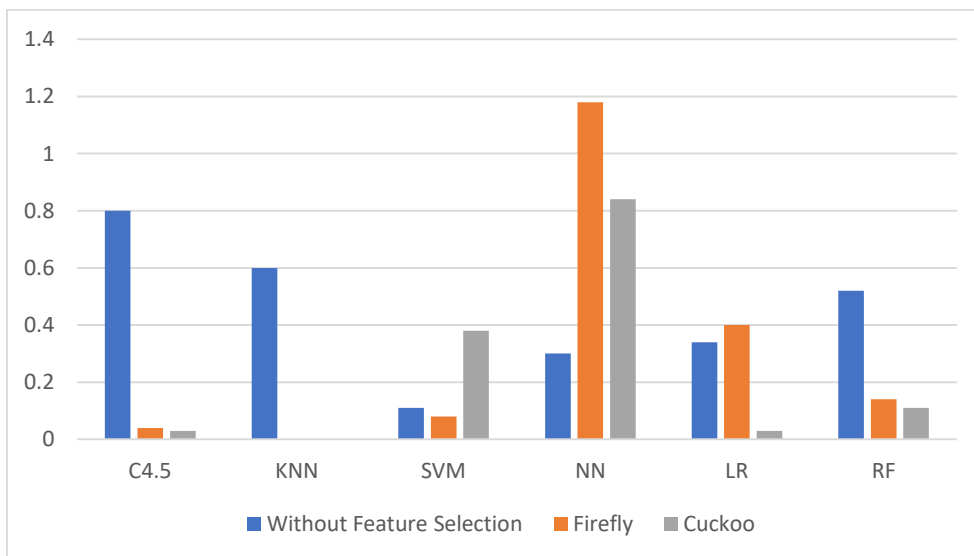


Figure 15: Comparative Analysis (Time-Taken)

From Figure 15, the lowest time value of 0s taken to build a model was recorded in feature selection using Firefly and Cuckoo, applying the KNN classifier. The breast cancer diagnostic model recorded the highest time taken value of 1.18s in NN taken to build the breast cancer diagnostic model.

## 5. Conclusion

In the development of a predictive or diagnostic model employing the data mining approach, some of the data may sometimes contain more features needed or wrong features (irrelevant features). The use of unneeded or irrelevant features in building a model may result in more memory space being used during the training process and classification phase. This paper developed a breast cancer diagnostic model to examine the effect of two metaheuristic-based attribute selection techniques on breast cancer disease diagnosis. The outputs of the comparative study showed the effectiveness of the attribute selection in terms of accuracy, recall, precision, kappa statistic, and time taken to build a model. Further research should consider more nature-inspired metaheuristic optimizations algorithms for robust and comprehensive comparative study.

## References

- Alloush, A., Rebdawi, G., & Trab, M. S. A. (2024). Modeling Information Diffusion in Online Social Networks Using a Modified Firefly Algorithm. *Journal of Information and Organizational Sciences*, 48(2), 311–322. <https://doi.org/10.31341/jios.48.2.5>
- Awotunde, J. B., Misra, S., Adeniyi, A. E., Abiodun, M. K., Kaushik, M., & Lawrence, M. O. (2022). A Feature Selection-Based K-NN Model for Fast Software Defect Prediction. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. [https://doi.org/10.1007/978-3-031-10542-5\\_4](https://doi.org/10.1007/978-3-031-10542-5_4)
- Bilal, A., Imran, A., Baig, T. I., Liu, X., Abouel Nasr, E., & Long, H. (2024). Breast cancer diagnosis using a support vector machine optimized by improved quantum-inspired grey wolf optimization. *Scientific Reports*, 14(1), 1–25. <https://doi.org/10.1038/s41598-024-61322-w>
- Dharambir Kashyap, Deeksha Pal, Riya Sharma, Vivek Kumar Garg, Neelam Goel, Deepika Koundal, Atef Zaguia, Shubham Koundal, and A. B. (2023). Retracted: Global Increase in Breast Cancer Incidence: Risk Factors and Preventive Measures. *BioMed Research International*, 9872034. <https://doi.org/10.1155/2023/9872034>
- EL kati, Y., Wang, S. L., Taresh, M. M., & Ali, T. A. A. (2024). Enhancing breast cancer diagnosis using deep learning and gradient multi-verse optimizer: a robust biomedical data analysis approach. *PeerJ Computer Science*, 10, 1–23. <https://doi.org/10.7717/PEERJ-CS.2578>
- Ginsburg, O., Yip, C., Brooks, A., Cabanes, A., Caleffi, M., Y, J. D., ... Neoplasticas, E. (2021). HHS Public Access. *HHS Public Access*, 126(Suppl 10), 2379–2393. <https://doi.org/10.1002/cncr.32887>.Breast
- Hayashi, Y., & Takano, N. (2020). One-dimensional convolutional neural networks with feature selection for highly concise rule extraction from credit scoring datasets with heterogeneous attributes. *Electronics (Switzerland)*, 9(8), 1–15. <https://doi.org/10.3390/electronics9081318>
- Kaderye, G., Arif, A., & Kundu, R. (2024). Data Mining in Different Fields: A Study. *International Journal of Innovative Science and Research Technology (IJISRT)*, 9(3), 1828–1838. <https://doi.org/10.38124/ijisrt/ijisrt24mar1384>

- Kaveh, M., & Mesgari, M. S. (2023). Application of Meta-Heuristic Algorithms for Training Neural Networks and Deep Learning Architectures: A Comprehensive Review. *Neural Processing Letters*, 55(4), 4519–4622. <https://doi.org/10.1007/s11063-022-11055-6>
- Navarro, M. M., & Navarro, B. B. (2024). Hybrid Cuckoo Search and Genetic Algorithm for Optimizing Electricity Forecast Hybrid Cuckoo Search and Genetic Algorithm for Optimizing Electricity Forecast. *2023 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*, 602–606. <https://doi.org/10.1109/IEEM58616.2023.10406830>
- Nunzio, A. Di, Giarra, A., Toscanesi, M., Amoresano, A., Piscopo, M., Ceretti, E., ... Montano, L. (2022). Comparison between Macro and Trace Element Concentrations in Human Semen and Blood Serum in Highly Polluted Areas in Italy. *International Journal of Environmental Research and Public Health*, 19(18). <https://doi.org/10.3390/ijerph191811635>
- Ostańska, E., Aebisher, D., & Bartusik-Aebisher, D. (2021). The potential of photodynamic therapy in current breast cancer treatment methodologies. *Biomedicine and Pharmacotherapy*, 137, 1–10. <https://doi.org/10.1016/j.biopha.2021.111302>
- Panliang, M., Madaan, S., Babikir Ali, S. A., Gowrishankar, J., Khatibi, A., Alsoud, A. R., ... Santhosh, A. J. (2025). Enhancing feature selection for multi-pose facial expression recognition using a hybrid of quantum-inspired firefly algorithm and artificial bee colony algorithm. *Scientific Reports*, 15(1), 1–23. <https://doi.org/10.1038/s41598-025-85206-9>
- Rahman, H. A., & Siti Nurzaimah Nazhirah Zaim, U.S.S. S. and A. A. J. (2024). Prognostic Factors Associated with Breast Cancer-Specific Survival from 1995 to 2022: A Systematic Review and Diseases, 1–20.
- Saini, O., & Sharma, S. (2018). A review of dimension reduction techniques in data mining. *Comput. Eng. Intell. Syst*, 9(1), 7–14.
- Sarker, P., Ksibi, A., Jamjoom, M. M., Choi, K., Nahid, A. Al, & Samad, M. A. (2025). Breast cancer prediction with feature-selected XGB classifier, optimized by metaheuristic algorithms. *Journal of Big Data*, 12(1), 1–26. <https://doi.org/10.1186/s40537-025-01132-7>
- Shardlow, A., McIntyre, N. J., Fluck, R. J., McIntyre, C. W., & Taal, M. W. (2016). Chronic Kidney Disease in Primary Care: Outcomes after Five Years in a Prospective Cohort Study. *PLoS Medicine*, 13(9), e1002128. <https://doi.org/10.1371/journal.pmed.1002128>
- Sohail, N., Jiadong, R., Irshad, M., Uba, M. M., & Abir, S. I. (2018). Data mining techniques for Medical Growth: A Contribution of Researcher reviews. *IJCSNS International Journal of Computer Science and Network Security*, 18(10), 5–10. [https://doi.org/http://paper.ijcsns.org/07\\_book/201810/20181002.pdf](https://doi.org/http://paper.ijcsns.org/07_book/201810/20181002.pdf)
- Thirumalaisamy, S., Thangavilou, K., Rajadurai, H., & Saidani, O. (2023). Breast Cancer Classification Using a Synthesized Deep Learning Model with Metaheuristic Optimization Algorithm. *Diagnostics*, 13, 1–21.