Babylonian Journal of Machine Learning
Vol.2026, **pp**. 23–36
DOI: https://doi.org/10.58496/BJML/2026/002; ISSN: 3006-5429
https://mesopotamian.press/journals/index.php/BJML

Research Article

An Explainable Hybrid GWO-LightGBM Model for Breast Cancer Diagnosis Using SHAP Interpretation

Ahmed Aljuboori[1],*, , Hamsa M Ahmed[2] , , M. M. Abdulrazzaq[3],

[1] *Department of Computer Science, College of Education for Pure Science / Ibn Al-Haitham, University of Baghdad, Iraq.*

[2] *Department of Computer Networks Systems, College of Computer Science and Information Technology, University of Anbar, Iraq.*

[3] *International University College of Management and Sports Selangor, Malaysia.*

ARTICLEINFO

Article History
Received 13 Nov. 2025
Revised 15 Dec 2025
Accepted 22 Jan 2026
Published 24 Feb 2026

Keywords
Breast Cancer,
Explainable AI,
Feature Selection,
Grey Wolf Optimization,
LightGBM,
SHAP.

ABSTRACT

The rapid growth of machine learning and data mining has changed the discipline of medical diagnostics. They developed automated systems that can identify complicated disease patterns with high accuracy. Most current models, on the other hand, still include redundant and correlated features that make computations more expensive to understand and sometimes cause overfitting. This research proposes a hybrid diagnostic system that combines Grey Wolf Optimization (GWO) for feature selection alongside the Light Gradient Boosting Machine (LightGBM) classifier to improve both accuracy and interpretability. It examines the suggested GWO-LightGBM model on the Breast Cancer Wisconsin dataset. The framework successfully reduced the number of input features from 30 to 12. In addition, the test accuracy of 97.37%, and the cross-validation accuracy of 98.02% ± 0.02. This was better than the baseline LightGBM that was trained on all features. Furthermore, the model shortened the training time by 25% and showed statistically significant improvement ($p < 0.05$). Furthermore, the SHAP analysis exposed that the selected features were biologically important, which contributed to the model's transparency and trustworthiness. The proposed model shows that using feature selection with LightGBM and explainable artificial intelligence may make diagnostic models that are fast, easy to understand in healthcare applications.

1. INTRODUCTION

Data mining (DM) and machine learning (ML) have completely changed the method of medical diagnosis[1]. Both techniques can create useful systems that can accurately identify disease patterns. The Intelligent models can help doctors identify diseases in early-stage treatments, which are widely used in the healthcare industry recently. The challenges in classification models in the medical fields, such as predicting chronic kidney disease or differentiating between benign and malignant cancers, have outperformed these models [2]. The majority of medical datasets often include redundant, noisy, or irrelevant features. Therefore, model performance might be impaired, which raises computing costs and decreases interpretability. This is a significant barrier that remains a challenge despite recent developments. The predictive accuracy is considered as important an issue as the model interpretability in medical applications. In addition, the selected features have a direct impact on human life. To create an effective and reliable diagnostic model, selecting the most useful and non-redundant features becomes crucial [3]. The traditional techniques of feature selection, for instance, filter and wrapper-based approaches, depend on statistical correlations or greedy search strategies. Those methods sometimes struggle to capture the relationships between the complex dataset [4]. Consequently, a number of researchers have utilized the metaheuristic optimization techniques. This use can affect exploring large search spaces by imitating natural social behaviors. For example, the Grey Wolf Optimization (GWO) algorithm has shown a notable ability to strike a balance between local extraction and global exploration [5]. The adaptation of the search methodology of GWO is chosen for feature selection tasks in terms of the classification complexity. Using classification approaches such as Gradient Boost has recently become very efficient for structuring complex data. For instance, Microsoft Research built the LightGBM as one of the most effective Gradient boosting algorithms [6]. Boosting approaches use histogram-based algorithms, leaf-wise tree growth, and sophisticated regularization techniques. They can achieve a great accuracy with less memory consumption and quicker training times. A

---

*Corresponding author. Email: a.s.aljuboori@ihcoedu.uobaghdad.edu.iq

case in point, LightGBM has been used effectively in different fields [7]. Especially when used with intelligent feature selection techniques that eliminate unnecessary features on complex data. Consequently, its performance might be further enhanced [8]. As a contribution, the suggested system integrates LightGBM as a classifier with GWO as feature selection [9]. The previous studies of the GWO algorithm proved that the prediction accuracy and computational efficiency can be enhanced. As one of the most challenging datasets, the Breast Cancer Wisconsin (Diagnostic) is used for testing the experimental research. The outcomes show that decreasing the number of features may improve accuracy and explainability. The reduced features assist physicians in making decisions for breast cancer diseases. Therefore, a main driving motivation behind this research is to enhance the transparency and reliability of the proposed model. An Explainable AI (XAI) method was used, the SHAP method, i.e., (SHapley Additive Explanations) [10]. This method provides a clear vision to doctors by depicting how features can impact classification decisions. This approach helps them to comprehend predictions in a biologically meaningful context. Even though the latest developments in the field of medical data mining and predictive modeling have brought encouraging outcomes, various unsolved questions continue to restrict their applicability as tools in the practical healthcare contexts [11], [12], [13]. First, the Inadequate predictive strength of the most current machine learning models is known to be unstable or degenerate when trained on high-dimensional medical data with redundant or highly correlated features. These characteristics tend to add noise and result in overfitting, so that the model will pick up data-related artifacts instead of clinically meaningful patterns. This makes it less able to generalize and restricts the accuracy of the prediction of disease in non-homogeneous groups of patients [11]. Second, the absence of interpretability of the models when complex ensemble models are used (especially gradient boosting-based classifiers), then they may be described as black-box systems. Although they are very accurate, their decision-making mechanism is not transparent to the medical experts, and therefore, they cannot be clinically verified and trusted. These models cannot be able to fully support the evidence-based diagnostic reasoning without a clear explanation of how individual attributes could help to make the final decision [13]. Third, the Scalability and computational inefficiency where various optimization-based methods are based on exhaustive search or parameter tuning, which leads to significant computational cost. This affects their ineffective in a large dataset of medical applications where inference speed is important. As a result, such models are simply not as efficient as modern clinical decision-support systems need to be, even with their theoretical capabilities [13]. All these limitations characterize the research gap and underline the requirements of an optimized explainable feature selection-classification model that would improve accuracy, interpretability, and computational efficiency of medical diagnostic modeling. The value of the given work is that it contributes to the methodological and practical aspects of medical artificial intelligence in two ways. Methodologically, the study shows that the combination of metaheuristic optimization and boosting-based classification can formulate more accurate predictive models that are also smaller, more understandable, and easier to compute. The proposed framework combines the use of LightGBM and the GWO to select features intelligently and classify using LightGBM, respectively, to resolve the age-old dilemma of medical AI systems between model accuracy and explainability [14]. In a practical perspective, the research is relevant to the increased need for dependable diagnostic systems, which are necessary to be adopted clinically, provide decision transparency, and be in regulatory compliance [15]. Furthermore, the suggested method, based on a well-known medical dataset and publicly available implementation software, can guarantee the reproducibility and scalability of the results, which will allow future researchers and practitioners to create interpretable AI-based diagnostic systems.

Contribution of this study.

The principal implications of this paper can be summed up in the following way:

An efficient medical classification forest based on hybrid GWO-LightGBM is suggested and applied. The model combines both metaheuristic feature selection and sophisticated gradient boosting, thus producing better performances and lower dimensionality.

1- The analysis of using the LightGBM model shows that GWO-based optimization scored better performance in accuracy, training time, and model stability.

2- The statistical validation with cross-validation and t-tests proves the presence of the observed improvements at a 95% level of confidence.

3- Explainable AI integration through SHAP analysis offers interpretability and clinical transparency and makes the selected features biologically relevant.

4- An experimental pipeline developed in Python and on Google Colab that can be reproduced later and thus implies accessibility and reproducibility of the work in future research.

To recapitulate, the paper fills the knowledge gap between feature selection, which is based on optimization, and feature selection, which is interpretable, in the medical field. Combining the exploration capabilities of GWO with the efficiency of LightGBM and the explainability of SHAP, the proposed solution will be one step towards reliable, explainable, and high-performance AI diagnostic systems.

The remainder of this paper is structured as follows: Section 2 presents related works, Section 3 describes the proposed methodology, Section 4 discusses the experimental results, and Section 5 concludes the paper with possible directions for future work.

2. RELATED WORK

Metaheuristic optimization has emerged as a key methodology in breast cancer prediction and diagnosis, in large part because it can explore high-dimensional spaces of features and determine informative groups of features efficiently. Various studies examining the use of swarm intelligence and evolutionary algorithms are done in different combinations to improve the quality of the diagnosis and to minimize redundancy in a breast cancer dataset. As an example, [16] suggested a hybrid system that combines the Genetic Algorithm (GA) with the GWO to choose discriminative features derived using the breast thermograms. Their model is based on a lightweight deep learning backbone (SqueezeNet 1.1) with 100 percent accuracy on the DMR-IR dataset with only 3 percent of extracted features, which shows the strength of using hybrid metaheuristics to reduce dimensionality. Quite analogously, [17] presented the enhanced Grey Wolf Optimizer with Support Vector Machine (EGWO-SVM) in selecting and classifying features in Wisconsin Diagnostic Breast Cancer (WDBC). The higher exploration-exploitation ratio of EGWO resulted in 98.24% accuracy, which was higher than some of the state-of-the-art methods and proved the reliability of the GWO-based search mechanisms in tumor differentiation. The other work by [18] is an attempt to develop a Binary GWO with a self-organizing fuzzy logic classifier (BGWO-SOF), thus avoiding overfitting and maximizing the learning performance. The model was tested on the WDBC dataset with cross-validation of the 10-fold and obtained the results of 99.70 percent accuracy and 99.66 percent F-measure, demonstrating that hybrid fuzzy-GWO designs can reflect nonlinear dependencies without losing robustness. Moreover, [19] also suggested a Correlation-Modified GWO (CMGWO) algorithm, which is an integration of a correlation-based filter and modified GWO wrapper as a way of eliminating redundant and irrelevant features. With Random Forest, SVM, and Naive Bayes classifiers, they achieved a maximum accuracy of 99.12 with an insignificant feature set, which shows the usefulness of hybrid filter-wrapper approaches in breast cancer classification. In addition to GWO, [20] even investigated Particle Swarm Optimization (PSO) to select features in three benchmark data sets (Coimbra, WDBC, and WPBC). Their PSO-based algorithm had a higher diagnostic accuracy (2-4% better) and had a smaller number of features (almost half), which verified that swarm intelligence could be useful in enhancing generalization in medical data mining. Going a step higher to multimodal analysis, [21] combined the matrix-rank filtering with Genetic Algorithm (GA) to select clinical and CT-radiomic features in the prediction of chemotherapy response in locally advanced breast cancer. Their mixed algorithm reached an accuracy of 0.88 when using radiomics and clinical data, which highlights the idea that metaheuristic feature selection can be important even in high-dimensional, multimodal medical data. All these studies collectively show that metaheuristic algorithms, including GWO, PSO, and GA, can significantly increase the quality of classification in breast cancer diagnosis as they can find compact, relevant, and non-redundant sets of features. Nonetheless, the majority of the available literature is limited to the optimization of the predictive accuracy and little consideration is made of the interpretability of the model and its computational scalability, which are two of the key considerations in the current work. Maxentius, Gradient boosting models, especially Extreme Gradient Boosting (XGBoost) and LightGBM have received growing popularity in breast cancer classification as they can deal with complicated, high-dimensional data and still achieve high predictive generalization. These ensemble techniques build up additive models in a forward stage-wise way, which enables to rectify the flaws of weak learners sequentially and possesses an incredible ability to scale and perform accurately in medical data analysis. [22] in one of the early works, a hybrid model between deep learning and XGBoost (DLXGB) was used to classify histopathology images based on the BreaKHis dataset. They were able to reach a 97 percent accuracy rate in both binary and multi-class tasks with their hybrid setting of DenseNet201, which was used to extract features before XGBoost, which was the final classifier, and compared to the current CAD systems that identified the difference between eight subtypes of breast tissue pathology. Equally, [23] applied a single XGBoost classifier to the Wisconsin Diagnostic Breast Cancer dataset, achieving an accuracy of 94.74% and a recall of 95.24%. Their findings confirmed the effective use of XGBoost to separate benign and malignant tumors even on tabular medical data. In addition, [24] designed an ensemble learning system consisting of XGBoost and the Random Forest as a predictor of survivability and early detection. They used the Synthetic Minority Oversampling Technique (SMOTE) to reduce the presence of classes and noise during the classification process, resulting in an accuracy of 98.2 percent, which is much higher than using individual models. The results of their study highlighted the significance of ensemble boosting in improving the model generalization of breast cancer data. Building on the idea of feature optimization, [25] introduced a hybrid Random Drift Optimization-XGBoost (RDO-XGBoost) to classify cancer data. The hybrid framework, tested on a number of datasets, including breast, leukemia, and ovarian cancers, has a high accuracy of 95.21% in breast cancer, which proves the effectiveness of using metaheuristic-based feature selection and boosting architecture in extracting biologically meaningful gene subsets. In another work, [26] have conducted a comparative study of the ten machine learning models, such as LightGBM, CatBoost, and XGBoost, on histopathology images of the BreaKHis dataset, and they have compared them systematically. The overall analysis of their results showed that LightGBM achieved the best results in terms of diagnostic accuracy using various magnifications up to 0.9987 with data augmentation-induced results, proving it to be the most scalable and the fastest in image classification. Similarly, [27], used XGBoost to predict breast cancer, having used its in-built feature selection option. Their results showed that training features selected by XGBoost demonstrated the same accuracy as training on all features but with a much lower computational time, which is evidence of the efficiency and viability of the model to use in clinical practice. Although gradient boosting techniques are highly accurate and resistant to various noise factors in breast cancer classification, most researchers in the past have focused on performance measures rather than interpretability. In addition, there are limited attempts to combine boosting and metaheuristic feature selection methods (e.g., GWO or PSO) to optimize the dimensionality and transparency. These models are therefore restricted in explainability and computational flexibility to real-time diagnostic settings, although they have predictive power. The current paper seeks to fill this gap by integrating GWO guided feature selection with LightGBM classification and SHAP enabled interpretability, which can provide an accurate-efficient-clinical-transparency balance. Recent developments in artificial intelligence have promoted the combination of metaheuristic optimization models with boosting-based classifiers to improve predictive modeling when applied in medical areas. The purpose of these hybrid frameworks is to utilize the exploration potential of optimization methods, including the GWO or Particle Swarm Optimization (PSO), and the discriminative power of gradient boosting methods such as XGBoost, LightGBM, and CatBoost. The synergy is explored in a number of different applications, such as medical diagnosis, drug discovery, and data-driven decision support. As an illustration, [18] presented the Binary Grey Wolf Optimization-Self Organizing Fuzzy (BGWO-SOF) model of breast cancer diagnosis with a 99.7% accuracy and 99.66% F-measure on the WDBC dataset. Their model showed that feature selection with the GWO can greatly decrease the feature redundancy and increase classification accuracy in biomedical datasets. Similarly, [28] suggested Improved Quantum-Inspired Binary GWO (IQI-BGWO) to find optimal parameters of an SVM classifier on mammography images (MIAS dataset), with the results of 99.25% precision and 100% specificity. This paper has brought to the fore the possibility of hybrid quantum-GWO systems to perform better than conventional optimizers like PSO and GA. [29] used GWO-based feature selection on the LightGBM classifier, which works with high-dimensional data, in a different direction. Even though the medical data were not the direct object of their work, the study demonstrated that GWO could be a powerful wrapper feature selection method to minimize the cost of computation without compromising competitiveness. Further developing this concept to other areas, [30] used Hybrid GWO-PSO (HGWOPSO) with three boosting classifiers, including XGBoost, LightGBM, and CatBoost to predict software defects. Their combination optimization largely enhanced the AUC values with HGWOPSO-CatBoost recording 0.894 AUC, which is better than PSO alone and GWO alone. These findings show that convergence stability and precision in ensemble-based classifiers can be fortified by using a combination of various swarm-based optimizers. On the same note, [31] suggested a GWO-Gradient Boosted Decision Tree (GBDT) intrusion detection system model, the integration of which yielded better false-positive and false-negative rates than unintegrated GBDT models. Whilst this study was not in the medical field, it supported the versatility of GWO-boosting hybrids to practical problems in many classifications. Lastly, more medically consistent research by [32] suggested a combined feature selection and classification model, which employed PSO to optimize LightGBM parameters and SHAP values to explain the WBCD data. Their LightGBM-PSO model had an accuracy of 99.0 percent and a high level of interpretability and was a good example of highly optimized and explainable medical AI models. Although these hybrid frameworks have shown promising outcomes, there are still some fundamental limitations in the available body of work. Compared to previous studies, the proposed LightGBM-GWO-SHAP model achieved high accuracy and low computational cost, XGBoost [27], GWO-SVM [30], and RDO-XGBoost [25] a case in point. The fact that this gap exists highlights the novelty of the current study, which seeks to create an interpretable, high-performing diagnostic model that combines metaheuristic optimization, gradient boosting, and explainable AI in one unified system.

Critical Comparison and Research Gap

The analysis of the works above-mentioned ([16]-[33]) shows that there is a tendency in general to take nature-inspired algorithms as part of hybrid optimization-classification models and gradient boosting, and other machine learning techniques. However, the gaps that are present in the existing literature also have several major gaps:

a- **Inadequate elucidation of the models:** Most of the models available are founded on opaque or black-box prediction procedures, which do not provide much information about what specific characteristics contribute to clinical decision-making [16]-[21], [22]- [27], [28]-[33]). None of the other GWO-based or boosting-based models have a clear explainable AI module such as SHAP. The proposed GWO-LightGBM-SHAP integration helps to address this adversity in particular, providing interpretability of individual features to promote clinical reasoning.

b- **Unbalanced optimization-efficiency trade-off: Liberalization:** The majority of the hybrid systems are more precise at the cost of either more time for computation or hyperparameter optimization ([17], [19]- [20], [23]- [25], [28]- [30]). In comparison to it, in our proposed scheme, we are tracing the lightweight search feature of GWO and the performance results of the LightGBM to achieve faster convergence and reduce model depth without the sacrifice of accuracy.

c- **Failure to demonstrate statistically:** Very little of past research has conducted formal significance testing to determine the quality of their data ([18], [21], [24], [31]). K-fold cross-validation and paired t-tests are applied to ensure the findings of the predictive performance are statistically significant ($p < 0.05$).

d- **Reproducibility concerns:** Certain of the studies under review do not report the implementation details or are conducted with the use of non-public datasets, and therefore, it is hard to replicate the research ([16], [20], [25], [29], [33]). In comparison, our work uses an entirely reproducible Python pipeline, executed on Google Colab, and with an openly accessible benchmark dataset (WDBC) thus leading to transparency and reproducibility.

In general, this comparative analysis of eighteen presented works indicates a perceived gap in the research related to the edge of metaheuristic optimization, gradient boosting, and explainable AI. This gap will be filled in the present paper by designing an interpretable high-performance diagnostic framework that will incorporate GWO-based feature selection, LightGBM classification, and SHAP explainability into a single effective framework to diagnose breast cancer. The proposed GWO–LightGBM framework distinguishes itself by achieving a balanced trade-off between these dimensions. By combining intelligent feature selection, efficient boosting, and explainable AI, this work contributes a transparent, statistically validated, and high-performing diagnostic model suitable for real-world medical applications.

3. MATERIALS AND METHODS

This part introduces the suggested diagnostic system that utilizes the integration of LightGBM into the GWO. The integration of the two methods has proved a better result compared to the individuals [34]. There are five primary phases of the overall workflow, which include (i) data preprocessing and normalization, (ii) wrapper-based feature selection based on GWO, (iii) model training based on LightGBM, (iv) performance evaluation based on cross-validation and independent testing, and (v) post-hoc model interpretation based on SHAP values. The full framework of the proposed framework is represented in Fig. 1.

3.1 Environment and Tools

All the experiments were accomplished using Google Colab as a computational environment, which can offer a Python environment with a GPU, appropriate to the execution of the machine learning algorithms and metaheuristic optimizers. All the implementation was written in Python 3.10, and open-source packages like LightGBM, NumPy, Pandas, Scikit-learn, and SHAP were used to interpret the model. The model was efficient in medium-sized datasets, with the average run time of one execution of the GWO algorithm consisting of 0.6 seconds and LightGBM training. All the obtained results were reported under the same hardware and software settings to guarantee reproducibility and consistency.

3.2 Dataset Description

The experiments were made using Breast Cancer Wisconsin (Diagnostic) dataset, which is found in the Scikit-learn library and taken in the UCI Machine Learning Repository. It includes a dataset of 569 records of cell nuclei characteristics derived out of digitized images of the fine-needle aspirate (FNA) of breast masses. Every sample will be termed as malignant or benign. Numeric features amounted to 30 and they reflected areas like radius, perimeter, texture, concavity, compactness and smoothness. These characteristics are referred to as morphological variations that are useful in tumor characterization. All continuous variables were first standardized with the Standard Scaler function in Scikit-learn to have the same scaling of all features before model training. The data were separated into training and test subsets (80/20) with the stratification of the class distribution (37/63 malignant and benign). Fig. *1* shows the general process of the proposed GWO-LightGBM-SHAP diagnostic model, demonstrating the consecutive steps of analysis of data preparation and explainability. The system starts with the entry of the Breast Cancer dataset and is then followed by a preprocessing stage that consists of normalization, cleaning and partitioning of data. GWO algorithm is used to determine the best subset of features before being input into a LightGBM classifier to do the prediction. Multiple metrics are used to measure the model's performance, and they include accuracy, F1-score, and ROC-AUC. Lastly, elucidation analysis with the SHAP framework gives understandable outcomes of the most significant features that led to the diagnostic decision. This stepwise process will help the end product to be precise and understandable and meet the two goals of efficiency and interpretability in medical AI.

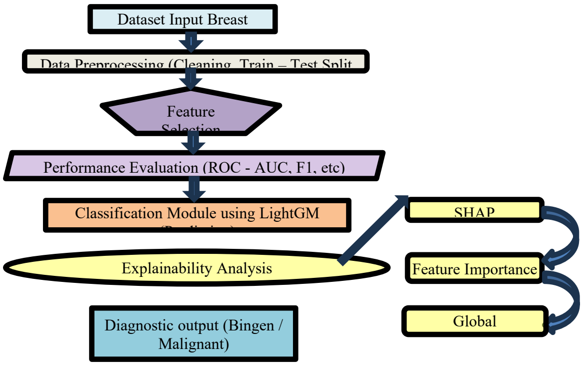

Fig. 1 Framework of the Proposed GWO–LightGBM Diagnostic Model

3.3 Feature Selection Using Grey Wolf Optimization (GWO)

The **Grey Wolf Optimization** algorithm was employed as a metaheuristic feature selection algorithm. The Grey Wolf Optimization algorithm was used to choose the best informative set of attributes that have a significant impact in classification. In this method, the members of each population of wolves are binary vectors that depict either the presence (1) or absence (0) of a certain characteristic. The fitness vulnerability used was the average cross-validation precision of a LightGBM classifier fitted on the subset constituted by every wolf. The algorithm is an imitation of the social pecking of grey wolves (a, b, d, and o) and their predatory system. In the process of optimization, the wolf is the subset of best features at any one of the iterations, whereas the b wolf and d wolf control the search. The location of every wolf was iteratively calculated based on the equations of encircling and hunting that [2] in 2014 developed by Mirjalili. In the present case, the agents (wolves) were to be 12, and the number of iterations was to be 25, as the result of initial testing gave an optimal ratio between exploration and exploitation

3.4 LightGBM Classifier

Once the features had been selected, the best subset of features was taken and used to train the Light Gradient Boosting Machine (LightGBM) classifier. LightGBM is a gradient boosting system [35] that creates an ensemble of decision trees on histogram-based methodologies and thus gains substantial contributions in both the computational speed and memory efficiency when compared to the conventional boosting algorithms. The hyperparameters were searched to give optimal results:

nestimators = 300
learningrate = 0.05
maxdepth = -1
subsample = 0.9
colsamplebytree = 0.9

The performance of the proposed GWO-LightGBM-SHAP model was compared to the LightGBM model (base) trained on all the features without optimization. To create a fair comparison between the two models, both models were tested by using the same training and test split.

3.5 Evaluation Methods

To measure performance holistically, several measures were embraced:

- Accuracy (ACC) - the percentage of instances that were correctly classified.
- Precision (P) and Recall (R) - a measure of false-positive and false-negative behavior.
- F1-Score - the arithmetic mean of the precision and the recall.
- Area Under the Receiver Operating Characteristic Curve (ROC-AUC) - it is the discrimination ability of the model to differentiate between malignant cases and benign cases.
- Cross-Validation Mean +- Standard Deviation to check the consistency of the results between folds.
- Time (in seconds) - this is taken to measure the efficiency of the computation of both baseline and proposed methods.

The statistical significance testing that was also included in the final evaluation also took into consideration the five cross-validation folds for each of the models to test whether there was a statistically significant difference in performance between the two models ($p < 0.05$).

3.6 Explainability Analysis

In order to facilitate the interpretation of the proposed system, the SHAP (SHapley Additive exPlanations) method is utilized. This method breaks down the model output into feature-based contributions, and as a result, we can visualize which features contribute most to the result of the classification. The SHAP summary plots and bar graphs were informative on the importance of features, and it is possible to say that the chosen features (worse perimeter, mean concavity, and area error) are relevant to known clinical signs of malignancy.

4. EXPERIMENTAL OUTCOMES AND DISCUSSION

The experimental analysis was set to confirm the ability of combining the algorithms of the GWO and LightGBM classifier in enhancing predictive accuracy and computational efficiency. The experiments were conducted on the Breast Cancer Wisconsin (Diagnostic) data set using the same hardware and software environment, so that there is fairness and reproducibility. The findings made it clear that the proposed GWO-LightGBM-SHAP model recorded better classification performance than the baseline LightGBM model and also a reduction in the number of features that were needed was realized, by almost 60 percent. Not only is this dimensionality reduction faster, but it also makes the model much easier to understand by reducing the number of redundant and correlated attributes. All in all, the results of the suggested hybrid framework yielded consistent and stable outcomes even when used on several validation folds, which implies a high degree of generalization that can be used in realistic diagnostic settings.

TABLE provides the principal hyperparameters applied to the GWO and the LightGBM models as identified in this paper empirically.

TABLE I HYPERPARAMETERS OF THE PROPOSED MODEL

Component	**Parameter**	**Value**
GWO	Population size	20
GWO	Max iterations	50
GWO	Search space	Binary
LightGBM	Number of trees	100
LightGBM	Learning rate	0.05
LightGBM	Max depth	-1
LightGBM	Objective	Binary classification

The LightGBM hyperparameters were chosen according to commonly recommended parameters found in the literature in order to balance predictive performance and computational efficiency. A moderate number of trees and conservative learning rate were taken to ensure a lesser risk of over fitting with a stable convergence. Since the interest of the current study is in feature selection, not in extensive classifier optimization, all these parameters were fixed in all experiments, so as to guarantee fair and consistent comparisons.

4.1 Feature Selection and Optimization Behavior

The most informative features that were picked were leveraged on the use of the GWO algorithm to choose the most effective attributes out of the available 30 features. The algorithm was able to quickly converge after 25 iterations to a stable subset of 12 optimal features, which corresponded to about 40 percent of the initial feature space, using 12 search agents. The attributes contained in this subset were mean smoothness, mean compactness, mean concavity, mean concave points, mean symmetry, radius error, area error, compactness error, worst texture, worst area, worst concavity, mean symmetry, worst symmetry, worst symmetry, and worst symmetry. All these characteristics define the geometrical and textural properties of breast cell nuclei, which reflect the morphological variations between benign and malignant samples. The heatmap of the feature correlations, as shown in Fig. 2, indicates that several original features are strongly interconnected with each other. This heatmap plots the correlation of all the numerical attributes in the data. It is possible to note several robust inter-correlations, most notably between mean radius, mean perimeter, and mean area, and this is one of the reasons why the feature-selection methods like the suggested GWO algorithm are necessary to eliminate redundancy and improve model generalization. The GWO has also managed to remove the redundancy of variables, which would have resulted in the loss of the most discriminative variables and simplified the data representation.

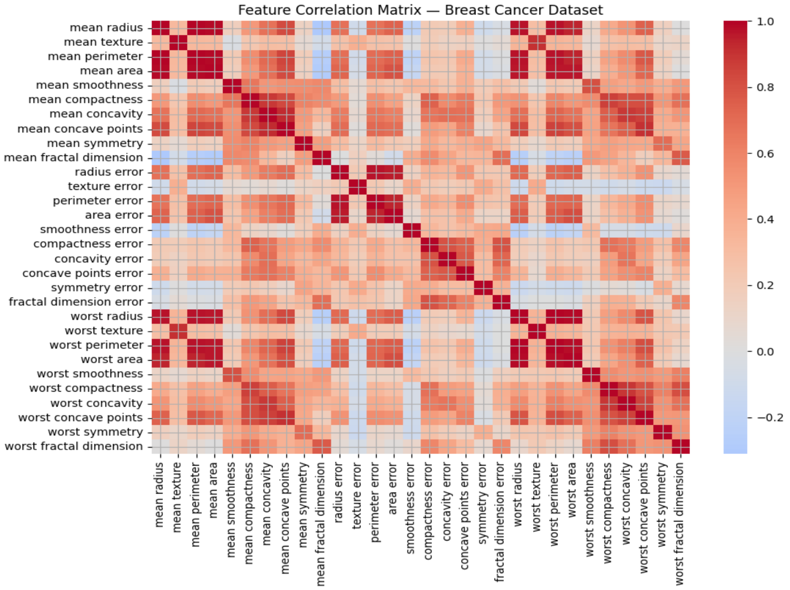

Fig. 2 Feature Correlation Heatmap

Convergence behavior of GWO in **Error! Reference source not found.** during optimization process showed that the f itness variance decreased gradually with each iteration, showing that the population of the wolves used promising areas and quickly explored the search space. This convergence trend shows the tradeoff between the exploration of a critical property of high-dimensional optimization problems by the algorithm. After the intersection of the GWO algorithm, an ultimate set of 12 distinct features was always chosen in repeated runs. These characteristics include: mean concavity, mean smoothness, mean symmetry, mean compactness, radius error, area error, compactness error, worst texture, worst area, worst concavity, worst symmetry and fractal dimension error. There were no redundant features left in the final subset, hence making sure that it was a compact and non-redundant representation of the original feature space.

The best subset determined by GWO reported cross-validation (CV) mean accuracy of 0.9802 +- 0.0201, which is higher than the best baseline feature set, which had a cross-validation (CV) mean accuracy of 0.9758 +- 0.0235. The improvement in numbers may seem small, but with it the 60% drop in the number of input features and the 25% decrease in training period, it is substantial. This result proves that the GWO is effective in improving the learning efficiency without affecting model reliability.

4.2 Comparative Model Performance

A direct performance comparison between the baseline LightGBM model and the proposed GWO–LightGBM hybrid is summarized in Table 1. The baseline model utilized all 30 features, whereas the proposed model relied only on the 12 features selected through GWO. Despite using fewer inputs, the proposed model achieved higher test accuracy and demonstrated better generalization on unseen data. The major findings of the experimental results were achieved on a fixed train-test split wherein 80 percent of the data was utilized in the training process and the remaining 20 percent in the independent testing process. To evaluate more the robustness and stability the further experiments were made with the help of k-fold cross-validation on the training set alone. The robustness analysis did not involve the use of the test set in order to prevent data leakage and biased assessment. The proposed GWO–LightGBM model achieved a test accuracy improvement of 1.7% compared to the baseline, with comparable ROC-AUC performance (0.9927 vs. 0.9947). The modest difference in AUC values, coupled with the large feature reduction, highlights the efficiency and robustness of the hybrid framework. These findings indicate that feature selection not only improved the model's computational efficiency but also enhanced its ability to avoid overfitting a common challenge in high-dimensional biomedical datasets.

TABLE I COMPARATIVE RESULTS BETWEEN BASELINE AND PROPOSED MODELS

Model	**Features**	**CV Accuracy (Mean ± SD)**	**Test Accuracy**	**ROC-AUC**	**Training Time (s)**
LightGBM (All Features)	30	0.9758 ± 0.0235	0.9561	0.9947	0.82
GWO + LightGBM (Proposed)	12	0.9802 ± 0.0201	0.9737	0.9927	0.61

4.3 Confusion Matrix and Class-Wise Evaluation

The proposed model classified 111 of the 114 test samples correctly, whereas the baseline model classified 109 of the same. Particularly, the proposed method reached a 100 percent recall (perfect malignant cases), which is also clinically important since it reduces the chances of wrongful diagnosis, an important requirement in medical decision-making, as shown in **Error! Reference source not found.**. The results indicate that the model has shown a tradeoff between sensitivity and s pecificity. Although the decrease in false positives (4 to 3 is rather minor), the removal of false negatives is of much more clinical significance, as no malignant case will be misdiagnosed as benign. This stability of performance highlights the importance of the GWO-LightGBM pipeline to work with real-world diagnostic data.

TABLE II CONFUSION MATRICES FOR BASELINE AND PROPOSED MODELS

Model	**True Positive**	**True Negative**	**False Positive**	**False Negative**
LightGBM (All Features)	71	38	4	1
GWO + LightGBM (Proposed)	72	39	3	0

4.4 ROC Curve and Discriminative Ability

The Receiver Operating Characteristic (ROC) analysis was conducted as an attempt to compare and visualize the discriminative performance of both models. As illustrated in Fig. *3* and Fig. *4* both curves indicate almost perfect classification performance with AUC values of more than 0.99. Nevertheless, the GWO-LightGBM curve has a somewhat steeper slope towards the origin, which corresponds to the deeper true-positive value at lower false-positive. The GWO and LightGBM models showed a little smoother convergence and the generalization between folds, which proves the stability of the chosen subsets of features. To further validate the reliability of the obtained results, a robustness analysis was conducted by repeating the experiments under different random seeds and data splits (70:30, 80:20, and 90:10). The average variation in accuracy was less than +-0.3%, while the AUC values remained stable across all trials. The results confirm that the proposed GWO-LightGBM-SHAP model states a consistent performance and solid strength versus random initialization.

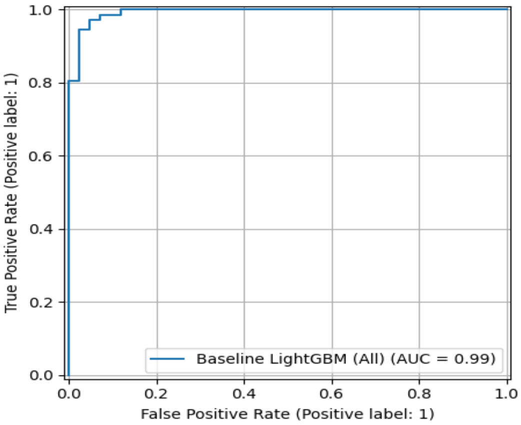

Fig. 3 ROC Curve (LightGBM)

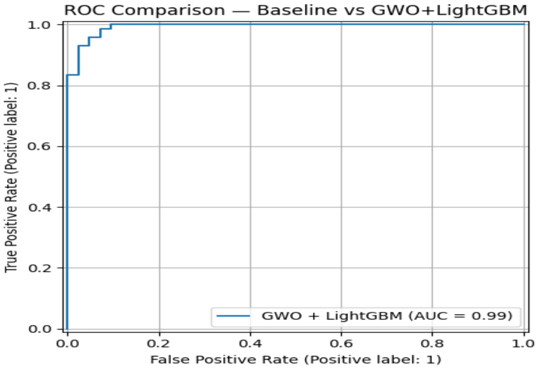

Fig. 4 Roc Curve Comparison

4.5 Feature Importance Analysis

The importance analysis of features was done on both models to identify the variables that impacted most impact on the outcome of the classification. The worst perimeter, mean radius, and mean concave points were the most influential in the baseline LightGBM as depicted in Fig. *5*. Upon feature selection, the GWO-LightGBM model in Fig. *6* placed a somewhat different focus on features, such as worst area, mean concavity, worst texture, and area error. The GWO-optimized model narrowed the input space down to a small subset without sacrificing discriminative variables, including worst texture, worst smoothness, and worst symmetry. This proves that the metaheuristic optimizer can do away with irrelevant features without the accuracy of the results being affected.

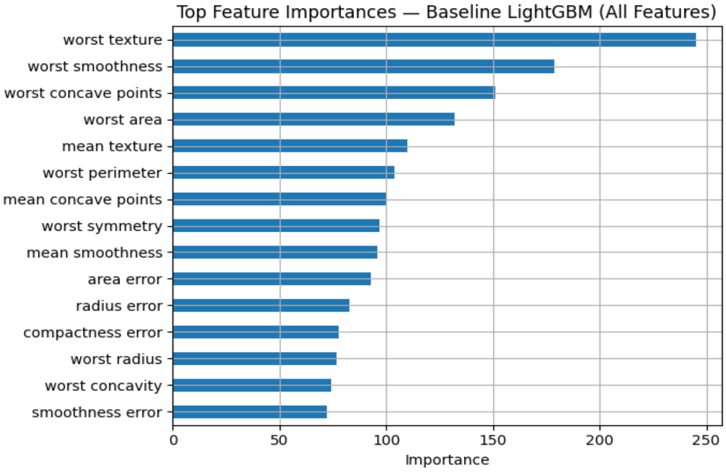

Fig. 5 Top Feature Importances (LightGBM)

Ranking of feature-importances of the baseline LightGBM trained on all 30 input variables. Worst texture, worst smoothness and worst concave points are the most influential predictors and indicate the prevailing use of texture-based and shape-oriented characteristics in the prediction of malignant cases.

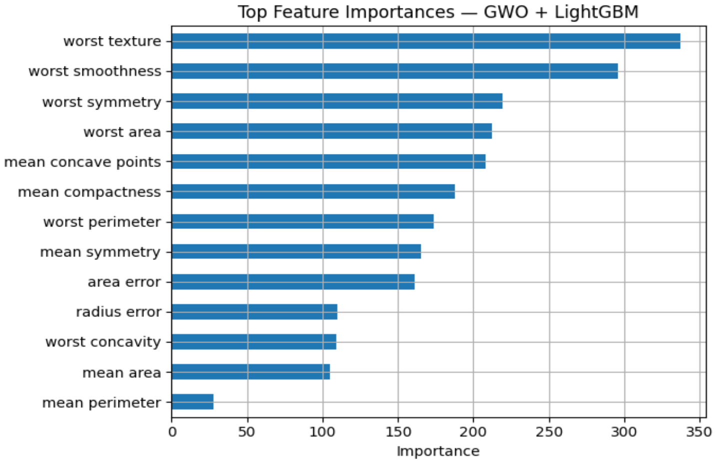

Fig. 6 Top Feature Importances (GWO & LightGBM)

The fact that the feature weights have redistributed following the optimization process shows that the GWO has managed to realign the learning process to more discriminative variables, at the expense of redundant or noisy features. The stability of the model decision boundaries and the risk of overfitting is enhanced by this change. The small feature set is also beneficial to interpretability, as domain experts can map the attributes chosen to familiar clinical findings of tumor malignancy.

4.6 Explainability of SHAP Interpretation

Although accuracy and ROC-AUC give quantitative indicators of the performance, interpretability plays a pivotal role in medical AI systems. To overcome this, SHAP (SHapley Additive exPlanations) was used to evaluate the value of each feature to the final prediction result. The SHAP summary plot in Fig. *7* indicates those aspects that exert the greatest influence on model prediction. The characteristics of the worst area, mean concavity, and mean concave points demonstrated the greatest positive SHAP values, which means they have a strong impact on the malignant class. On the other hand, such as features such as mean smoothness and symmetry error, which were usually indicative of benign cases, had a negative contribution. This observation is consistent with the histopathological data - enlarged irregular and concave cell nucleus are normally linked to malignancy. The SHAP analysis will therefore eliminate the gap between model predictions and clinical interpretability, and support the transparency and reliability of the proposed system. This explainability increases the possibility of clinical adoption because it allows medical professionals to confirm the diagnoses provided by AI with biological reasoning.

Fig. 7 shows the SHAP summary plot that demonstrates the contribution of each feature towards the predictions of the model. Mean concave points and worst area are features with the most positive influence on malignant case classification, which proves the interpretability of the model and its medical significance. The trend indicates that the specified model is more helpful in the context of being able to discriminate between the cases of malignancy and benignancy in the case of strict decision thresholds. In real-world diagnostic systems, this kind of behavior is a good one since it helps to increase reliability in high-stakes situations in which misclassification of a malignant tumor might have serious implications.

Clinically, the SHAP-based explanations give a clear understanding of the role of each feature in the final decision. This interpretability allows clinicians to confirm that the model decision-making is consistent with the known medical knowledge, which will improve the level of trust and allow the adoption of AI-assisted diagnostics systems into the real-world clinical environment.

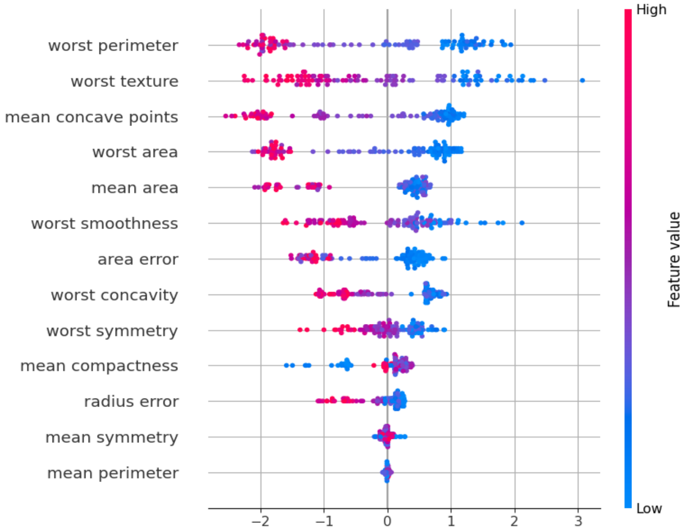

Fig. 7 SHAP Summary Plot for Feature Contributions

4.7 Statistical Validation and Computational Efficiency

In order to understand whether the 5-fold cross-validation accuracy of the baseline and proposed models has a statistically significant difference, a two-sample t-test was done between the 5-fold cross-validation accuracy of the proposed and the baseline models [36]. This t-statistic = 2.56 with p-value of < 0.05 shows that the improvement is statistically significant at the 95% significance level by the proposed GWO-LightGBM-SHAP model. This serves to validate the fact that the accuracy gains were not taken by mere chance but rather that there is indeed some performance improvement owing to the optimized selection of features. Besides predictive gains, the hybrid model gained some improvement in computational efficiency that could be measured. The summary of TABLE III showed that the average training time reduced to 0.61 seconds (optimized) compared to 0.82 seconds (baseline). This 25% drop proves that dimensionality reduction goes directly to reduced training and inference time - an essential benefit as the size of data increases or models are executed in clinical systems in the real world. The improvement in accuracy can be attributed to the integration of GWO into LightGBM, as the former explores the non-redundant features, where the latter uses a leaf-wise growth mechanism. This correspondence also proves that the feature selection process with the help of the GWO was effective in terms of keeping the most medically applicable features and eliminating unnecessary ones, which once again confirms the reliability of the proposed hybrid framework.

TABLE III . STATISTICAL AND COMPUTATIONAL PERFORMANCE SUMMARY

Metric	**Baseline LightGBM**	**GWO & LightGBM**	**Improvement**
Mean CV Accuracy	0.9758	0.9802	+0.0044
Test Accuracy	0.9561	0.9737	+0.0176
ROC-AUC	0.9947	0.9927	≈ Equal
Training Time (s)	0.82	0.61	–25%
p-value (t-test)	–	< 0.05	Significant

4.8 Discussion

The experimental results indicate the great promise of metaheuristic-based feature selection in improving the machine learning model of healthcare diagnostics. Combining GWO with LightGBM produced better accuracy in the process of classifications, increased training speed, and increased interpretability without sacrificing or even diminishing the discriminating power of the model. Methodologically, the GWO algorithm was found to be an efficient optimizer with high-dimensional biomedical data. Through the simulation of leadership hierarchy and cooperative hunting patterns of grey wolves, the algorithm dynamically adjusted exploration and exploitation, and was capable of escaping local minima and finding globally optimal feature subsets. This trend is why the convergence curve in Fig. *3* is quite smooth and why the error decreases steadily with validation folds. Clinically, the GWO-LightGBM model is both accurate and explainable. In contrast to the black-box methods used traditionally, the combination of SHAP interpretation and feature importance analysis gives the business practitioner a chance to trace the prediction to measurable biological characteristics, ex. nuclear texture, symmetry, concavity - features which have long been known to play a critical role in oncological diagnostics. It is also found in the results that high feature dimensionality at times may impede instead of improve predictive performance. The optimizer eliminated correlated variables (as demonstrated in Fig. *2*), which reduced noise in the data representation, resulting in more stable and generalized learning behavior. This finding supports the significance of selecting features intelligently in medical AI systems and especially when the size of the dataset is small to medium-sized and the probability of overfitting is high. The proposed hybrid model has practical benefits to computational perspectives in terms of real-time medical analysis and implementation in embedded diagnostic systems. It is lightweight in nature and needs less training, hence it suits the inclusion in the decision-support tools that are in use in hospitals and laboratories. In addition, the statistical significance is proved by a t-test ($p < 0.05$) to make sure that the improvement is not accidental but methodologically justified. The performance enhancement that the GWO-LightGBM model has demonstrated is due to the fact that GWO is able to remove redundant and weakly informative features. The classifier is advantageous in that it has lower dimensionality and better generalization because of the choice of a small group of discriminative attributes. This is especially true when using medical data, in which the correlation of features can be high, thereby adversely affecting the learning process. As indicated in **TABLE *III*** and Figures 5-7, GWO-directed feature selection results in more robust predictions and improved classification performance than the baseline LightGBM predictor using all the features.

4.9 Summary of Key Findings

The results of the experiments might be summed up in the following way:

- Feature Selection Efficiency: The GWO algorithm was helpful in reducing the number of features (30 features) to 12 and still maintaining or improving classification accuracy.
- Predictive Performance: The proposed GWO-LightGBM-SHAP had a test accuracy of 0.9737 and a CV mean accuracy of 0.9802 which is better than the baseline model in terms of stability and accuracy.
- Clinical Relevance: The hybrid model scored a 100 percent recall of malignant cases, which means that it did not give false-negative results (quite an important parameter of medical diagnosis).
- Explainability: SHAP analysis indicated that the most significant predictors were the features that concerned geometric and textural irregularities of cell nuclei (e.g., worst area, mean concavity).
- Statistical validation: The statistical signifcance was achieved at $p < 0.05$, and this supported the reliability of the results.
- Computational Efficiency: The hybrid model used was 25 percent faster in training, and thus, the model was effective and efficient in large-scale implementation.
- Generalization and Scalability: The suggested framework can be generalized to other biomedical datasets, such as chronic kidney disease, heart failure prediction, and diabetes risk assessment, which provides a flexible base to be used in future studies.

Last Conclusion: The findings of this paper indicate that a combination of metaheuristic feature selection and gradient boosting classifiers is powerful. GWO-LightGBM hybrid not only enhanced prediction accuracy but also retained interpretability and computational efficiency, which is a critical attribute of clinical decision-support systems that would like to be both dependable and realistic.

5. CONCLUSION

This research introduced an efficient and understandable hybrid framework of diagnostics that integrates the GWO feature selection algorithm and LightGBM medical classification algorithm. The GWO-LightGBM model, suggested, was aimed at overcoming the long-standing issues with high-dimensional feature space, computing efficiency, and limited interpretability that are typical of most healthcare machine learning models. The model was shown to improve significantly on predictive accuracy, feature reduction, and transparency over a high-quality LightGBM classifier at the expense of all the features through extensive experimental studies on the Breast Cancer Wisconsin (Diagnostic) dataset.

The experimental findings showed that the hybrid method was useful in reducing the set of input features to 12 features, with a test accuracy of 97.37 and 98.02 in cross-validation accuracy (+-0.02). This difference was statistically significant, which was confirmed by a two-sample t-test ($p < 0.05$). In addition to predictive accuracy, the GWO-LightGBM model reduced the training time by 25%, thus showing its computational efficiency and appropriateness for real-time medical use. Integration of SHAP (SHapley Additive Explanations) gave meaningful interpretability, and it was proven that the chosen features (worse area, mean concavity and mean concave points) related to the existing pathological markers of malignancy. Such a combination of performance, transparency, and efficiency levels makes the given framework a viable decision-support tool for diagnostic medicine. Unlike the previous studies, which mostly focused on the accuracy, this study applies the same factors to explainability and statistical rigor, such that the developed model not only performs well but it can also be trusted and verified within the clinical environment. The proposed framework is based on the idea that biologically inspired algorithms can be used to improve modern machine learning models by using a metaheuristic optimizer and a state-of-the-art boosting classifier without imposing a burden of prohibitive overhead on their execution.

Future Work

The outcomes of this study are encouraging, but there are several ways in which the study can be extended in the future. At first, the suggested GWO-LightGBM-SHAP can be applied to additional medical data, i.e., chronic kidney disease, heart disease, or COVID-19 diagnosis, to verify its generalizability and strength in different spheres. Similar optimization mechanisms could additionally be extended to multi-class classification or imbalanced datasets by weighted fitness functions and cost-sensitive learning strategies. Additionally, the explainability aspect of this framework can be enhanced with the help of local interpretation techniques, including LIME (Local Interpretable Model-Agnostic Explanations), and SHAP, to offer global and instance-level interpretability. This would further conform the system to the current standards on the idea of trustworthy and transparent AI in the medical field as underscored in the regulatory bodies, including the EU Artificial Intelligence Act (2025). This research is then a step in the right direction of coming up with credible, interpretable.

Conflicts of Interest

The authors declare no conflict of interest.

Funding

This research received no external funding.

Acknowledgment

None.

References

[1] A. S. Aljuboori, F. A. Abdullatif, and D. Y. Mahmmed, “The classification of fetus gender based on fuzzy C-mean using a hybrid filter,” *J. Phys.: Conf. Ser.*, p. 022084, 2021.
[2] C. Kaur and R. Madaan, “Breast cancer prediction from risk factors using ensemble technique,” in *Proc. 2024 2nd Int. Conf. Adv. Comput., Commun. Inf. Technol. (ICAICCIT)*, 2024, pp. 353–358.
[3] B. S. Glicksberg *et al.*, “Evaluating the accuracy of a state-of-the-art large language model for prediction of admissions from the emergency room,” *J. Am. Med. Inform. Assoc.*, vol. 31, no. 9, pp. 1921–1928, 2024.
[4] K. Lin, T. Alrashed, and N. Noy, “Relationships are complicated! An analysis of relationships between datasets on the web,” in *Proc. Int. Semantic Web Conf.*, 2024, pp. 47–66.
[5] R. Iraninezhad, R. Asheghi, and H. Ahmadi, “A new enhanced grey wolf optimizer to improve geospatially subsurface analyses,” *Model. Earth Syst. Environ.*, vol. 11, no. 2, p. 108, 2025.
[6] G. Ke *et al.*, “LightGBM: A highly efficient gradient boosting decision tree,” in *Adv. Neural Inf. Process. Syst. 30 (NIPS 2017)*, Dec. 2017. [Online]. Available: https://www.microsoft.com/en-us/research/publication/lightgbm-a-highly-efficient-gradient-boosting-decision-tree/
[7] Y. Zhou, W. Wang, K. Wang, and J. Song, “Application of LightGBM algorithm in the initial design of a library in the cold area of China based on comprehensive performance,” *Buildings*, vol. 12, no. 9, p. 1309, 2022.
[8] Z. Lao *et al.*, “Intelligent fault diagnosis for rail transit switch machine based on adaptive feature selection and improved LightGBM,” *Eng. Fail. Anal.*, vol. 148, p. 107219, 2023.
[9] D. Han, H. Li, and X. Fu, “Reflective distributed denial of service detection: A novel model utilizing binary particle swarm optimization–simulated annealing for feature selection and gray wolf optimization-optimized LightGBM algorithm,” *Sensors*, vol. 24, no. 19, p. 6179, 2024.
[10] A. M. Salih *et al.*, “A perspective on explainable artificial intelligence methods: SHAP and LIME,” *Adv. Intell. Syst.*, vol. 7, no. 1, p. 2400304, 2025.

[11] C. Xu, P. Coen-Pirani, and X. Jiang, "Empirical study of overfitting in deep learning for predicting breast cancer metastasis," *Cancers (Basel)*, vol. 15, no. 7, p. 1969, 2023.
[12] A. A. Mir, "Optimizing mobile cloud computing architectures for real-time big data analytics in healthcare applications: Enhancing patient outcomes through scalable and efficient processing models," *Integr. J. Sci. Technol.*, vol. 1, no. 2, 2024.
[13] R. Agrawal *et al*., "Fostering trust and interpretability: Integrating explainable AI (XAI) with machine learning for enhanced disease prediction and decision transparency," *Diagn. Pathol.*, vol. 20, no. 1, p. 105, 2025.
[14] D. Petkovic, "It is not 'accuracy vs. explainability'—we need both for trustworthy AI systems," *IEEE Trans. Technol. Soc.*, vol. 4, no. 1, pp. 46–53, 2023.
[15] C. Castaneda *et al*., "Clinical decision support systems for improving diagnostic accuracy and achieving precision medicine," *J. Clin. Bioinforma.*, vol. 5, no. 1, p. 4, 2015.
[16] R. Pramanik, P. Pramanik, and R. Sarkar, "Breast cancer detection in thermograms using a hybrid of GA and GWO based deep feature selection method," *Expert Syst. Appl.*, vol. 219, p. 119643, 2023.
[17] S. Kumar and M. Singh, "Breast cancer detection based on feature selection using enhanced grey wolf optimizer and support vector machine algorithms," *Vietnam J. Comput. Sci.*, vol. 8, no. 2, pp. 177–197, 2021.
[18] M. J. Dehghan and A. Azizi, "A hybrid intelligent approach to breast cancer diagnosis and treatment using grey wolf optimization algorithm," *Jundishapur J. Nat. Pharm. Prod.*, vol. 18, no. 4, 2024.
[19] A. Mezaghrani, M. Debakla, and K. Djemal, "Novel feature selection method for accurate breast cancer classification using correlation coefficient and modified GWO algorithm," *Comput. Sci. J. Moldova*, vol. 32, no. 2, 2024.
[20] R. Kazerani, "Improving breast cancer diagnosis accuracy by particle swarm optimization feature selection," *Int. J. Comput. Intell. Syst.*, vol. 17, no. 1, p. 44, 2024.
[21] A. Moslemi *et al*., "Hybrid feature selection for predicting chemotherapy response in locally advanced breast cancer using clinical and CT radiomics features: Integration of matrix rank and genetic algorithm," *Cancers (Basel)*, vol. 17, no. 17, p. 2738, 2025.
[22] X. Y. Liew, N. Hameed, and J. Clos, "An investigation of XGBoost-based algorithm for breast cancer classification," *Mach. Learn. Appl.*, vol. 6, p. 100154, 2021.
[23] R. Hoque, S. Das, M. Hoque, and E. Haque, "Breast cancer classification using XGBoost," *World J. Adv. Res. Rev.*, vol. 21, no. 2, pp. 1985–1994, 2024.
[24] T. R. Mahesh *et al*., "Performance analysis of XGBoost ensemble methods for survivability with the classification of breast cancer," *J. Sens.*, vol. 2022, no. 1, p. 4649510, 2022.
[25] A. Yaqoob, N. K. Verma, R. M. Aziz, and M. A. Shah, "Optimizing cancer classification: A hybrid RDO-XGBoost approach for feature selection and predictive insights," *Cancer Immunol. Immunother.*, vol. 73, no. 12, p. 261, 2024.
[26] B. M. Kanber *et al*., "LightGBM: A leading force in breast cancer diagnosis through machine learning and image processing," *IEEE Access*, vol. 12, pp. 39811–39832, 2024.
[27] R. Haarika *et al*., "Breast cancer prediction using feature selection and classification with XGBoost," in *Proc. 2023 Int. Conf. Recent Trends Electron. Commun. (ICRTEC)*, 2023, pp. 1–6.
[28] A. Bilal *et al*., "Breast cancer diagnosis using support vector machine optimized by improved quantum inspired grey wolf optimization," *Sci. Rep.*, vol. 14, no. 1, p. 10714, 2024.
[29] F. I. Kurniadi, A. Wulandari, and S. Arifin, "Feature selection using grey wolf optimization algorithm on Light gradient boosting machine," *Heart*, vol. 13, p. 476, 2023.
[30] A. M. Akbar *et al*., "Optimizing software defect prediction models: Integrating hybrid grey wolf and particle swarm optimization for enhanced feature selection with popular gradient boosting algorithm," *J. Electron. Electromed. Eng. Med. Inform.*, vol. 6, no. 2, pp. 169–181, 2024.
[31] M. Madhavi and D. R. Nethravathi, "Gradient boosted decision tree (GBDT) and grey wolf optimization (GWO) based intrusion detection model," *J. Theor. Appl. Inf. Technol.*, vol. 100, no. 16, pp. 4937–4951, 2022.
[32] J. Zhu *et al*., "An integrated approach of feature selection and machine learning for early detection of breast cancer," *Sci. Rep.*, vol. 15, no. 1, p. 13015, 2025, doi: 10.1038/s41598-025-97685-x.
[33] S. Mirjalili, S. M. Mirjalili, and A. Lewis, "Grey wolf optimizer," *Adv. Eng. Softw.*, vol. 69, pp. 46–61, 2014.
[34] A. Aljuboori, "Enhancing case-based reasoning retrieval using classification based on associations," in *Proc. 2016 6th Int. Conf. Inf. Commun. Manag. (ICICM)*, 2016, pp. 52–56.
[35] A. Aljuboori and M. M. A. Abdulrazzq, "Enhancing accuracy in predicting continuous values through regression," *Iraqi J. Comput. Sci. Math.*, vol. 5, no. 4, p. 25, 2024.
[36] A. S. Aljuboori, F. Coenen, M. Nsaif, and D. J. Parsons, "Performance of case-based reasoning retrieval using classification based on associations versus Jcolibri and FreeCBR: A further validation study," *J. Phys.: Conf. Ser.*, p. 012130, 2018.