ISSN:0976-0172
Journal of Bioscience And Technology
https://jbstonline.in/

A Smart Breast Cancer Prediction Model Using Ant Colony Optimization

K. Banerjee*, Anirud

Department of Information Technology, JIS College of Engineering, Kalyani, WB, India.

Article Info

Received: 30-06-2024 *Revised: 19-09-2024* *Accepted: 29-09-2024*

Abstract

It is crucial to identify breast cancer early and accurately in order to slow its development and improve outcomes, as it is still the most common cancer in women. When it comes to sifting through massive datasets in search of probable situations, machine learning (ML) approaches shine. Through the use of classifiers including K-Nearest Neighbor (KNN), Support Vector Machine (SVM), Random Forest, Gradient Boosting, and XGBoost, this work presents a machine learning (ML) model for the prediction of breast cancer. Hyperparameters were optimized using Ant Colony Optimization (ACO) to improve the models' accuracy. To improve model accuracy while decreasing computation, features were selected using the SelectKBest approach. The dataset, which was obtained from the UCI Machine Learning Repository, allowed for the production of reliable models. Notably, the highest prediction accuracy achieved by this approach is 99%, with the Random Forest classifier optimized through ACO. The dataset consisting of thirty neumeric features. This research highlights the potential of integrating ML and optimization techniques to enhance disease prediction capabilities by early prediction of disease in turn a better patient outcome.

Keywords: Ant Colony Optimization (ACO), Breast Cancer Prediction, Feature Selection (SelectKBest), Hyperparameter Optimization, Machine Learning Classifiers.

Introduction

The field of artificial intelligence (AI) known as machine learning focuses on creating models and algorithms that enable computers to learn from data and make educated predictions or judgments without requiring explicit programming for each job. Machine learning is based on the idea that computers may learn from their mistakes and get better with time. This allows them to process more data more efficiently and accurately. When it comes to healthcare, machine learning is one of the most significant applications. It's vital in illness diagnosis, treatment planning, and patient outcome prediction. In

For the purpose of breast cancer prediction, this research presents an improved machine learning model. This model intends to make a significant contribution to healthcare improvements by enhancing diagnostic accuracy and facilitating early detection via the use of data-driven learning approaches. The World Health Organization reports that 670,000 individuals lost their lives to breast cancer in 2022, and that 2.3 million women were diagnosed with the illness. Breast cancer is more common in older women, however it may affect women of any age after puberty worldwide. 1

The work's primary contributions are as what follows:

To predict the occurrence of breast cancer, researchers have integrated machine learning with Ant Colony Optimization (ACO).

• Proves that ACO works in various applications of machine learning.

• Enhances generalizability via enhanced prediction accuracy and hyperparameter optimization for classifiers.

ISSN:0976-0172
Journal of Bioscience And Technology
www.jbstonline.com

Provides an all-encompassing framework by combining several methodologies with advanced optimization.

A number of popular classifiers are included in the suggested model. These include K-Nearest Neighbor (KNN), Support Vector Machine (SVM), Random Forest, Gradient Boosting, and XGBoost. These algorithms are selected because they have a track record of success in classification tasks and can effectively identify intricate patterns in the data, which makes them ideal for predictive modeling. Following the first model design, the hyperparameters are fine-tuned using Ant Colony Optimization (ACO), which improves the model's performance by optimizing important parameters. The accuracy, reliability, and resilience of the model in detecting subtle data patterns are all enhanced by combining classifiers with ACO. This method is a metaheuristic optimization example from the family of ant colony algorithms used in swarm intelligence approaches. Originally published in 1992 as part of Marco Dorigo's PhD thesis. Two-Gradient Boosting and XGBoost are robust ensemble methods that can fix errors from earlier iterations and achieve high accuracy; Random Forest offers robust performance and is resistant to overfitting; Support Vector Machines excel at dealing with complicated, non-linear data and features with high dimensions; and K-Nearest Neighbors is simple and easy to grasp.

The goal is to enhance the accuracy of breast cancer predictions by merging the power of these machine learning algorithms with an optimization strategy inspired by ant foraging behavior. Utilizing K Nearest Neighbor in conjunction with Ant Colony Optimization yielded 94% accurate results, while Support Vector Machine yielded 96% accurate results, Gradient Boosting 96% accurate results, XGBoost 95.6% accurate results, and Random Forest with Ant Colony Optimization 96.0% accurate results.

A maximum accuracy of 99% and a false positive rate of 0.87% were achieved via optimization. In addition to a 98% AUC, 100% sensitivity, and 97% specificity, the model also has a perfect score for accuracy. For those working on medical diagnostics, the UCI Machine Learning Repository's breast cancer data collection is an invaluable resource. Digital pictures of breast masslesions acquired by fine needle aspirates (FNA) are the basis of this dataset, which include clinical and diagnostic characteristics. In all, 569 occurrences make up the dataset's 30 features, which include things like texture, perimeter, symmetry, and radius. For the purpose of training and assessing machine learning models, each event is classified as either benign or malignant.

The main contribution of this study is the integration of several machine learning approaches with the Ant Colony Optimization methodology, which has been largely overlooked in the area of breast cancer prediction. Existing literature demonstrates that ACO has been beneficial in several machine learning-based applications. Applying ACO to Federated Searching yielded substantial improved outcomes, according to a study by Adamu Garba et al. 3 What makes this work unique is its use of ACO as a performance-enhancing approach. It enhances prediction accuracy and optimizes classifier hyperparameters for greater generalization. By combining several algorithms with a sophisticated optimization strategy, this methodology offers a robust framework for breast cancer prediction, in contrast to models that depend on individual classifiers or use simpler optimization methods. 3. Applying adaptive ACO has also been effective in a different study by Duygu Yilmaz Eroglu et al. 4. 4

Section 2 provides a detailed analysis of a small number of recent research in this area. The experimental results and explanation of the suggested procedure have been covered in sections 3 and 4. The study's conclusions and its potential future directions are laid forth in section 5.

Review of Relevant Literature

A large body of research has investigated the feasibility of using machine learning techniques for cancer prediction, with encouraging findings across a range of cancer types. Scientifically, a breast cancer prognosis model was created. model that accomplished a maximum accuracy of 97.14% by using Logistic Regression, Support Vector Machine (SVM), K-Nearest Neighbor (KNN), and Random Forest. 5 A different study used SVM, Decision Tree, Naïve Bayes, and KNN to forecast breast cancer, with SVM achieving the highest accuracy at 97.13%. An 86% success rate was achieved using SVM in a

ISSN:0976-0172
Journal of Bioscience And Technology
www.jbstonline.com

model that a scientist had devised that combined Random Forest and SVM. 6 researchers focused on lung cancer prediction using Random Forest, Decision Tree, Naïve Bayes, and Support Vector Machines; Random Forest attained a 96.8% accuracy rate. 7

The researchers used Naïve Bayes, SVM, Random Forest, and KNN in their paper on colorectal cancer, and Naïve Bayes had the highest accuracy rate of 92.83%. 8. For colorectal cancer, another researcher employed Logistic Regression, Random Forest, SVM, and KNN, reaching a maximum accuracy of 98.2%. 9. Using a transfer learning ResNet-50 model with global average pooling to identify brain tumors while avoiding overfitting and gradient vanishing. 10 With k=1 and k=3, the authors used KNN and Deep Neural Networks (DNN) to achieve a recall, accuracy, and F-measure of 0.97, and an AUC of 0.984 with DNN. They were able to get a recall of 0.955 and an accuracy of 0.95611 using KNN at k=1. An overall accuracy of 92.13% was produced by their model, which divided pictures into three groups: pituitary, meninoma, and no tumor. 12

They went on to create a model that used magnetic resonance imaging (MRI) to identify and categorize instructors, which greatly improved brain tumor diagnosis. 13 Researchers suggested a CNNleNetV2-based model for colon cancer detection that achieved a 99.67% success rate, surpassing competing designs. 14 Using neural networks trained on chest X-ray images, researchers were able to identify COVID-19 with an AUC score close to 0.999, indicating very good diagnosis accuracy. Using the LC250, researchers were able to classify pictures of lung and colon cancer as benign, adenocarcinoma, or squamous cell carcinoma with an accuracy of 97.20% in validation and 96.11% in training. Machine learning has shown promising results in disease prediction in a number of studies, with improved classifiers and feature selection approaches achieving robust performance across a range of malignancies. Novel illness prediction models based on machine learning16 have recently been presented in the literature, especially in cases of cardiovascular disease and cancer. The EFFI-CNN experiment, which showcases developments in CNN architecture17, was inspired by the ICDSSPLDCNN and EASPLD-CNN trials and is a new convolutional neural network. The author of the article on heart disease prediction used a concatenated ensemble classifier to build a model that achieved an accuracy of 86.69%. 18 Following this, the author and colleagues presented a model for predicting cardiac problems that combined SVM-XGBoost and Particle Swarm Optimization, used LASSO to pick features, and increased accuracy to 91.8%. 19

Using a wide variety of machine learning algorithms—including support vector machines (SVMs), logistic regression, decision tree classification, k-nearest neighbors (KNNs), Gaussian Naive Bayes, and artificial neural networks—a thorough model was given in a different research, and it achieved an impressive accuracy level of 99%20. Because the signs of early-stage lung cancer are often mild, researchers examined several classifiers to find the best method for early diagnosis. Their results showed that SVM was more successful than KNN, with an accuracy of 95.56% compared to 88.40% for KNN. feature selection approaches, such as Logistic Regression21 and Random Forest, were the primary emphasis of the researcher's study. The Radial Basis Function (RBF) classifier22 outperformed Naive Bayes, Support Vector Machines (SVM), and KNN in this study's lung cancer classification task, with an accuracy of 81.25 percent. The increasing significance of machine learning in healthcare is shown by these advancements. Tools for early diagnosis and effective intervention have been greatly improved in illnesses with complicated data patterns via the use of optimized classifiers, ensemble models, and enhanced feature selection, which have led to much higher prediction accuracy.

Several models in the literature on machine learning techniques for cancer prediction achieve great accuracy using classifiers like SVM, Random Forest, and KNN. These findings demonstrate amazing outcomes across a number of cancer types. Even yet, there are a lot of holes and problems. Many studies utilize a single classifier or a mixture of basic classifiers, which may not provide the most accurate predictions or be applicable to other situations. Also, even though Ant Colony Optimization (ACO) and other sophisticated optimization methods have been developed, very few research have actually used them. The hyperparameters or characteristics can be optimized by some. Additionally, the problem of model robustness when dealing with complex and imbalanced data remains one of the most significant obstacles to early-stage cancer detection.

ISSN:0976-0172
Journal of Bioscience And Technology
www.jbstonline.com

This proposed research strengthens the existing breast cancer prediction model by integrating several strong classifiers as KNN, SVM, Random Forest, Gradient Boosting, and XGBoost, thereby addressing the shortcomings of the existing literature. Unlike previous work that depended on single classifiers or less advanced optimization approaches, this strategy brings together many classifiers to grasp a broader spectrum of complex patterns in the data. In addition, by using Ant Colony Optimization (ACO) to tune the model's hyperparameters, we increase its performance and eliminate the optimization issue that plagues many traditional methods. By using advanced optimization and classifier integration, this model aims to improve diagnostic accuracy and generalizability, providing a stronger solution for breast cancer prediction compared to existing techniques.

MATERIALS AND METHODS

This study proposes a machine learning model for breast cancer prediction, utilizing data sourced from the UCI Machine Learning Repository. Various classifiers such as K-Nearest Neighbor (KNN), Support Vector Machine (SVM), Random Forest, Gradient Boosting, and XGBoost have been implemented to construct the predictive model. Each classifier was selected for its demonstrated effectiveness in classification tasks and ability to identify complex patterns in medical data. Following model construction, Ant Colony Optimization (ACO) was applied to fine-tune parameters, further enhancing model performance. The structural flow and processes of the proposed model are illustrated in the design diagram (Fig.1.), providing a clear overview of the integration of classifiers and optimization techniques.The data has been collected, then data preprocessing has been done, followed by applying different classifiers. After that ACO has been applied as shown in the Figure-1.This approach aims to improve diagnostic accuracy, contributing valuable insights to the field of machine learning- based disease prediction.

Data Collection

The dataset utilized to train the model has been collected from the UCI data repository.[24] This dataset is made up of clinical and diagnostic characteristics that were taken from digital images of breast mass lesions using fine needle aspirations (FNA). Thirty parameters, including texture, radius, perimeter, and symmetry, are included in the dataset, which consists of 569 examples. Machine learning (ML) models can be trained and evaluated with ground truth information provided by the labeling of each instance as benign or malignant. To guarantee the precision and applicability of the features taken from the breast mass pictures, technicians and

Fig. 1. Design diagram of the proposed model

ISSN:0976-0172
Journal of Bioscience And Technology
www.jbstonline.com

medical specialists worked together during the data collecting process. The primary fields in the dataset include various attributes relevant to breast cancer diagnosis. The *Diagnosis* field categorizes cases as either malignant (M) or benign (B). Key features include *radius*, representing the mean distance from the center to points on the tumor perimeter, and *texture*, which captures the standard deviation of gray-scale values. Additional attributes consist of *perimeter* and *area*, both fundamental to tumor size and shape assessment. The dataset also includes *smoothness*, indicating the local variation in radius lengths, and *compactness*, a measure of shape regularity. Furthermore, *concavity* and *concave points* describe the extent and number of concave regions along the tumor boundary, respectively. *Symmetry* and *fractal dimension*, the latter reflecting "coastline approximation" minus one, provide further structural information. Together, these features offer a comprehensive profile for distinguishing between malignant and benign cases.

Data Pre-processing

The data pre-processing module is an acute component in the development of ML models.

Converting Categorical filed into Numerical filed

In the above-mentioned dataset, the dependent variable diagnosis is a categorical filed representing either by 'M' or 'B'. This filed has been converted into numerical value applying Label Encoding. In machine learning, label encoding is a technique that transforms category data into numerical form so that models can process it. This approach is a straightforward and effective technique to manage categorical data since it gives each separate category a unique integer value. When categorical data contains an intrinsic order, label encoding is especially helpful since it maintains this order, which is desirable for some algorithms, such as tree-based models. Moreover, it consumes less memory and processing power than more sophisticated techniques like one-hot encoding.

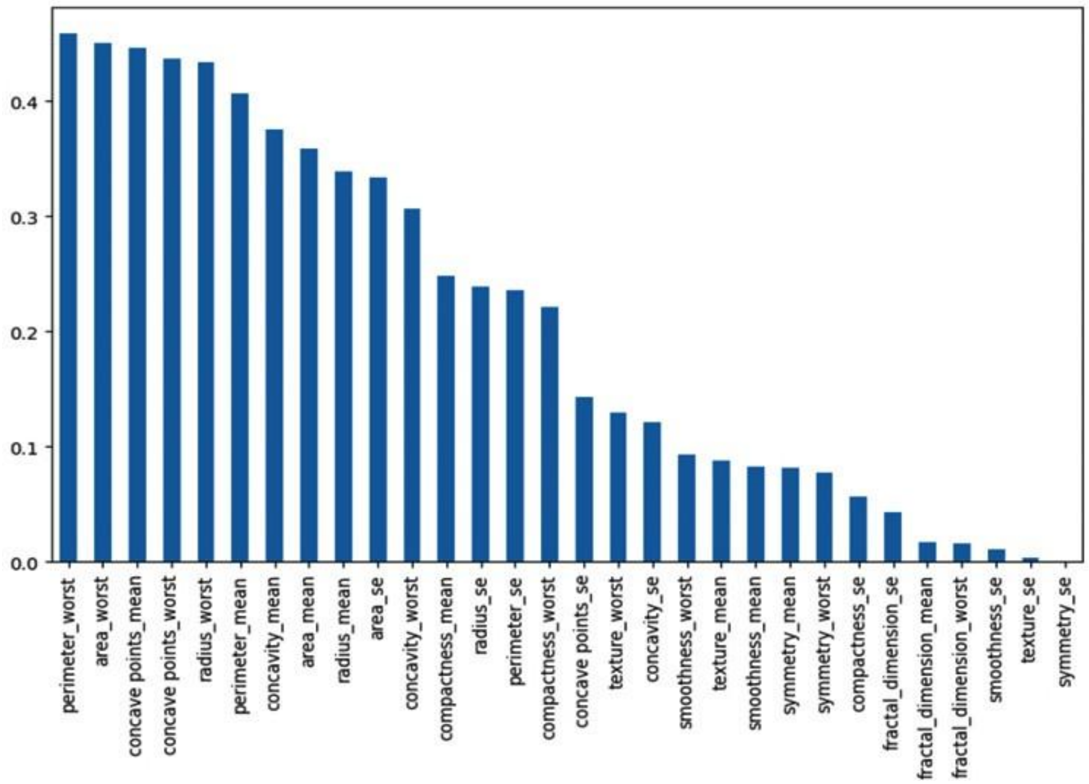

Fig. 2. Feature importance applying SelectKBest

ISSN:0976-0172
Journal of Bioscience And Technology
www.jbstonline.com

Feature Selection

The process of feature selection is an integral aspect in developing a machine learning model. There is a lack of parity in the contribution of current features to decision making, and the introduction of superfluous features might occasionally lower the model's performance. The feature selection in this work was done using SelectKBest.

One well-known method for selecting features in machine learning is SelectKBest. It takes a model's performance into account when deciding which "K" features to use. This method prioritizes characteristics according to their relevance to the target variable, which is achieved by assigning a score to each feature and then selecting the ones with the highest scores. In order to reduce the data's dimensionality, SelectKBest highlights the most important characteristics. More accurate and generally applicable models are the outcome of this process, which lessens the likelihood of overfitting and noise. When dealing with massive datasets, SelectKBest is invaluable as it can reduce the amount of characteristics to significantly quicken the training process. It also makes administration and understanding easier by simplifying the produced models. With its ability to run several statistical tests depending on the data type, SelectKBest becomes quite versatile when it comes to machine learning problems. The fact that it is easy to use adds to its appeal as it allows for seamless integration into machine learning processes for efficient and effective feature selection.

The model's training set includes the top 18 characteristics chosen using SelectKBest. As illustrated in Figure 2, the feature relevance was determined using SelectKBest. Based on the priorities assigned, perimeter_worst is at the top and symmetry_se is at the bottom.

Applying Classifiers

Multiple classifiers have been applied for prediction of breast cancer as follows:

Table 1. Psesudo code of KNN Classifier

```
Knn_clf=KNN_Classifier(No of neighbours, p,weights)
Knn_clf.fit(x_train_selected,y_train)
y_prediction=Knn_clf.predict(x_test,y_prediction)
accuracy=accuracy_score(y_test,y_prediction)
print(accuracy)
```

Table 2. Psesudo code of SVC Classifier

```
Svc_clf=SVC(C=100,gamma='scale',kernel='liner',random_state=42)
Svc_clf.fit(x_train_selected,y_train)
y_prediction=Svc_clf.predict(x_test_selected)
accuracy=accuracy_score(y_test,y_prediction)
print(accuracy)
```

Table 3. Psesudo code of SVC Classifier

```
Rf_clf=RandomForestClassifier(max_depth,min_samples_leaf,min_samples_split, n_estimators,random_state)
Rf_clf.fit(x_train_selected,y_train)
y_prediction=Rf_clf.predict(x_test_selected)
accuracy=accuracy_score(y_test,y_pred)
print(accuracy)
```

Table 4. Psesudo code of XGBoost Classifier

```
Xgb_clf=xgb.XGBClassifier(learning_rate,n_estimators,random_state)
Xgb_clf.fit(x_train_selected,y_train)
y_prediction=Xgb_clf.predict(x_test_selected)
accuracy=accuracy_score(y_test,y_prediction)
```

ISSN:0976-0172
Journal of Bioscience And Technology
www.jbstonline.com

KNN Classifier

The KNN is a simple learning algorithm used for classification and regression tasks especially for instance based. It classifies data points based on the majority vote of their KNN in the feature-space.The designing of KNN model has been shown in the Table-1 below.

Support Vector Classifier (SVC)

The Support Vector Classifier (SVC) is a supervised machine learning algorithm widely applied in classification tasks. Its primary goal is to identify the optimal hyperplane that separates different classes within the feature space, ensuring maximum margin between the classes. By maximizing this margin, SVC improves classification accuracy and generalizes effectively to new data. The design process and functional architecture of SVC is illustrated in Table-2, demonstrating how the algorithm delineates the decision boundary between classes. This approach is particularly valuable in cases where a clear separation between classes is crucial for accurate predictions.

Random Forest

Random Forest is an ensemble learning technique that builds multiple decision trees during the training process and makes predictions based on the majority vote of these trees. This method employs bagging (Bootstrap Aggregating) and random feature selection, which help mitigate overfitting and enhance model accuracy. By averaging the results from diverse trees, Random Forest achieves robust performance and generalizes well to new data. Table 3 illustrates the design of the K-Nearest Neighbor (KNN) model, showcasing its approach to classification based on the proximity of data points in the feature space. Both Random Forest and KNN offer distinct, effective strategies for handling classification tasks.

XGBoost

XGBoost, or Extreme Gradient Boosting, is a highly optimized library built to enhance the

Table 5. Psesudo code of Gradientboosting Classifier

```
Gb_clf=GradientBoostingClassifier(learning_rate,max_depth,n_estimator, random_state)
Gb_clg.fit(x_train_seelcted,y_train)
y_prediction=Gb_clf.predict(x_test_selected)
accuracy=accuracy_score(y_test,y_predcition)
```

Table 6. Psesudo code of Gradientboosting Classifier

```
Aco=aco(fitness_function,dimensions=X.shape[1],colony_size,max_iter)
best_solution=Aco.Optimize()
```

Table 7. Overall experimental Outcomes of the proposed work

Algorithm	Best Parameter	Accuracy	AUC	Sensitivity	Specificity
KNNWith ACO	{'n_neighbors':7,'p':1, 'weights':'distance'}	94%	0.9231	0.9859	0.8604
SVC with ACO	{'C':100, 'gamma': 'scale', 'kernel':'linear'}	96%	0.9464	0.9859	0.9069
Rando mForestwith ACO	{'max_depth':20, 'min_samples_leaf': 1, 'min_samples_split': 2, 'n_estimators':50}	99%	0.9883	1	0.9767
Gradient Boosting Classifier with ACO	{'learning_rate': 0.1,'max_depth':3, 'n_estimators':200}	96%	0.951	0.9302	0.9718
XG Boost with ACO	{'learning_rate': 0.1, 'n_estimators':100}	95.60%	0.951	0.9718	0.9302

ISSN:0976-0172
Journal of Bioscience And Technology
www.jbstonline.com

gradient boosting framework. It is specifically engineered for efficient, scalable implementation of gradient boosting algorithms, renowned for delivering exceptional performance and accuracy. The design and architecture of the XGBoost model are illustrated in Table-4.

Gradient Boosting

Gradient Boosting is a machine learning approach that constructs an ensemble of weak learners, often decision trees, in a sequential manner. In this process, each new learner is added to the ensemble to reduce the loss function,

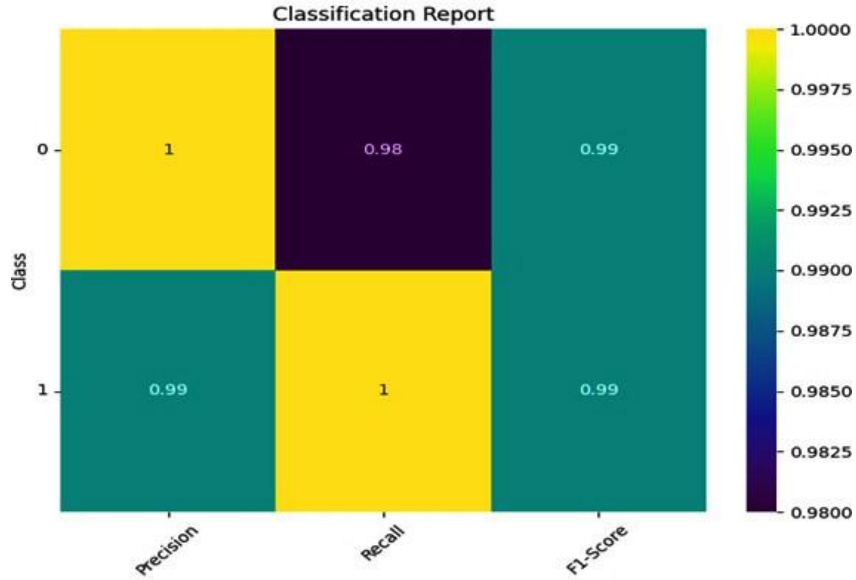

Fig. 3(a). Confusion Matrix of the Model utilizing Random Forest and ACO

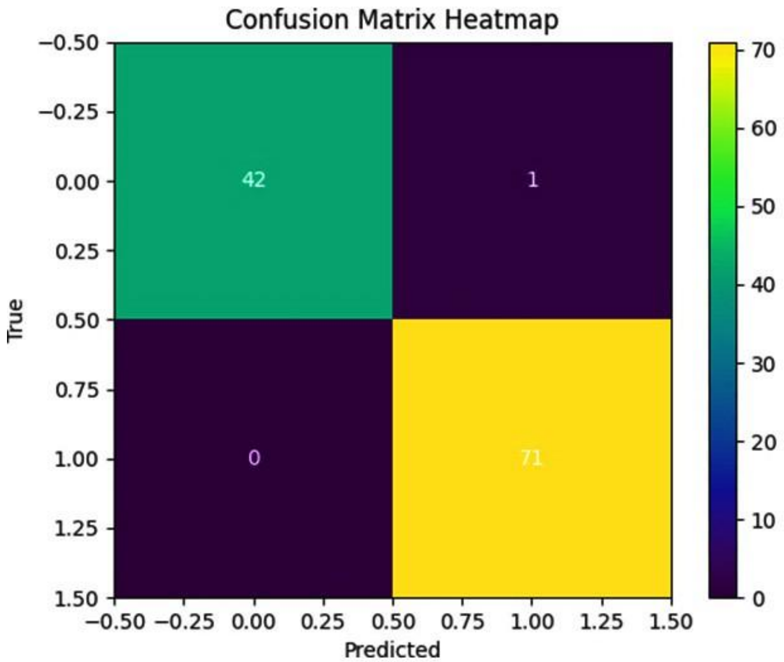

Fig. 3(b). Classification report of the Model utilizing Random Forest and ACO

ISSN:0976-0172
Journal of Bioscience And Technology
www.jbstonline.com

iteratively enhancing the model's prediction accuracy. The structure of the Gradient Boosting model is illustrated in Table-5.

Mode Optimization

For performance optimization Ant Colony Optimization has been applied. Some ant species' foraging strategies serve as the foundation for ant colony optimization, or ACO. When these ants locate a good path, they mark it for other ants in the colony to follow by leaving pheromone markers on the ground. An analogous approach is used in ant colony optimization to address optimization issues. Within the context of swarm intelligence techniques, this algorithm belongs to the family of ant colony algorithms and represents a few metaheuristic optimizations. The first algorithm was first presented by Marco Dorigo in 1992 in his PhD thesis.[25,26] Its goal was to find the best path through a graph by mimicking the actions of ants as they navigated a path among their colony and the

Table 8. Comparative Analysis

Works	Observations
Breast Cancer Prediction: A Comparative Study Using Machine Learning Techniques[5]	In this study, lung cancer prediction was explored using a combination of machine learning models, including Logistic Regression, Support Vector Machine (SVM), K-Nearest Neighbor (KNN), and Random Forest. Among these methods, the highest prediction accuracy achieved was 97.14%.
Prediction of Survival Rate from Non-Small Cell Lung Cancer using Improved Random Forest[7]	Another approach for lung cancer prediction employed Support Vector Machine (SVM), Naïve Bayes, Decision Tree, and Random Forest models. The highest accuracy recorded in this combination was 96.8%, showcasing the effectiveness of ensemble methods for disease prediction.
A novel transfer learning approach for the classification of histological images of colorectal cancer[8]	For colorectal cancer prediction, a blend of Naïve Bayes, Support Vector Machine (SVM), Random Forest, and K-Nearest Neighbor (KNN) algorithms was applied. This approach yielded a top accuracy of 92.83%, highlighting the models' effectiveness in handling colorectal cancer datasets.
Classification using deep learning neural networks for brain tumors[10]	In this analysis, Deep Neural Networks (DNN), K-Nearest Neighbor (KNN) with parameters k=1k=1k=1 and k=3k=3k=3, Linear Discriminant Analysis (LDA), and Sequential Minimal Optimization (SMO) were applied to predict cancer. The model achieved an Area Under the Curve (AUC) of 0.97, indicating high discriminatory power for classification.
Diagnosing COVID-19 Infection in chest X-Ray Images using Neural Network[14]	A prediction model for COVID-19 diagnosis was developed using Support Vector Machine (SVM) and a Neural Network. This model achieved an impressive Area Under the Curve (AUC) of 0.99, reflecting its strong capability for accurate COVID-19 prediction.
Heart Disease Prediction Using Concatenated Hybrid Ensemble Classifiers[18]	For heart disease prediction, a concatenated ensemble classifier was implemented, which combines multiple classifiers to improve accuracy. This method resulted in an overall accuracy of 86.89%, demonstrating the utility of ensemble approaches for complex health condition prediction.
Lung cancer Prediction and Classification based on Correlation Selection method Using Machine Learning Techniques[21]	Support Vector Machine (SVM) and K-Nearest Neighbors (KNN) algorithms were applied for disease prediction, achieving a peak accuracy of 95.56%. This indicates the models' suitability for robust prediction in medical datasets.
Breast Cancer Prediction Using Machine Learning Techniques[27]	To enhance the prediction of lung cancer, this research applied Support Vector Machine (SVM), Decision Tree, Naïve Bayes, and K-Nearest Neighbor (KNN) algorithms. This combination of models reached a peak accuracy of 97.13%, demonstrating strong predictive potential in lung cancer diagnosis.
Our Proposed Model	A comprehensive model for breast cancer prediction incorporated K-Nearest Neighbor (KNN), Support Vector Machine (SVM), Random Forest, Gradient Boosting, XGBoost, and Ant Colony Optimization algorithms. The highest accuracy achieved was 99%, underscoring the effectiveness of this diverse ensemble for breast cancer diagnosis.

ISSN:0976-0172
Journal of Bioscience And Technology
www.jbstonline.com

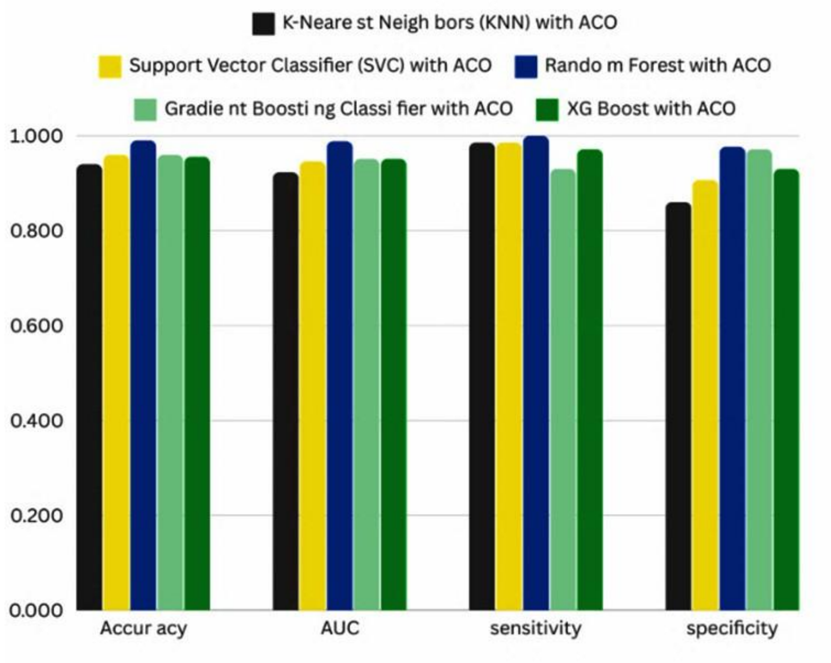

Fig. 4. Oveall outcome metrics

food source. Sample code snapshot of designing ACO has been represented in Table-6 below

RESULTS

In the proposed model Ant Colony Optimization has been applied over KNN, SVM Classifier, Gradient Boosting, Random Forest, and XGBoost. It has been identified that Random Forest Classifier with Ant Colony Optimization has achieved highest accuracy of 99%. The details observations of investigational results have been shown in the Table 7 below.

DISCUSSION

As seen in Figure 4, a significant achievement in machine learning was the combination of Random Forest (RF) and Ant Colony Optimization (ACO) to attain 99% accuracy in predictive modeling. This combination has several significant ramifications. To begin, the model's robustness is shown by its constant performance across several data instances. The model's reliability for real-world applications is enhanced by its ability to generalize effectively to fresh data. To add to it, the model's

The model's outputs may be trusted for key judgments in several sectors, such as banking and healthcare, because to its high prediction accuracy, which instills significant confidence in decision-makers. Because of its efficiency and scalability, the model can handle enormous datasets and provide accurate predictions in realistic durations. This makes it valuable for real-time and high-throughput situations. And lastly, decision-making methods are significantly affected by the high accuracy.

Random Forest with ACO models with 99% accuracy might revolutionize healthcare by delivering more reliable early diagnosis, personalized treatment, and enhanced predictive analytics. Better results and faster treatments may be possible if the model can accurately predict illnesses like cancer at an early stage. Also, doctors would be able to maximize their resources, personalize their treatment plans for each patient, and reduce the likelihood of human mistake with its assistance. The integration of Random Forest with ACO has the potential to enhance healthcare efficiency and decision-making via improved patient outcome prediction.

Ultimately, it may lead to better-informed, meaningful, and consequential choices by optimizing outcomes in domains like fraud

ISSN:0976-0172
Journal of Bioscience And Technology
www.jbstonline.com

detection, resource allocation, and personalized recommendations.

As you can see from Table 8, we compared the planned study to a small number of previous studies.

CONCLUSION

Because it can process massive volumes of data, machine learning is an ideal identifying tool. This paper presents an AnML-based model for the expectation of breast cancer. Classifiers such as KNN, SVM, Gradient Boosting, Random Forest, and XGBoost have all been used. We have used ACO to these models to improve their performance. The 99% accuracy achieved using Ant Colony Optimization with Random Forest concludes that the model has exceptional learning and generalization capabilities. Opens up new opportunities and promotes advances across a number of sectors by providing decision-makers with an effective tool to make extremely confident and accurate data-driven judgments.

REFERENCES

1. http://www.who.int/news-room/fact-sheets/ detail/breast-cancer is the World Health Organization's website on breast cancer. (Retrieved on January 23, 2025) In 1991, A. Colorni, M. Dorigo, and V. Maniezzo published Distributed Optimization by Ant Colonies in the proceedings of the first Européan conference on artificial life (134–142) in Paris, France, published by Elsevier Publishing.

3. Garba et al. (2024) were co-authored by Khalid et al. and Mumin et al. Merging Results in Federated Search with the Use of Ant Colony Optimization. Published in Engineering, Technology & Applied Science Research, Volume 14, Issue 4, August 20, 2024, pages 14832–14839. this article's DOI is 10.48084/etasr.7302.

4. Aydınya Eroglu and Uğur Akcan (2024). Selecting Features using an Adapted Ant Colony Optimization Mechanism. Artificial Intelligence in Practice, 38(1). This article is available online at https://doi.org/10.1080/08839514.2024.2335098..

5. Muhammad Islam, Malik Haque, Iqbal Hussain, Hasan Muhammad, Hasan Muhammad, and Kabir Hasan Muhammad. A Comparative Study on the Use of Machine Learning Techniques in Breast Cancer Prediction. publication date: 2020-01-29, DOI: 10.1007/s42979-020-00305-w. SN Computer Sci.

6. Banerjee N, Das S. Learning Machines for Lung Cancer Prediction. With the following DOI: 10.1109/ICCSEA49143.2020.9132913, this article is part of the 2020 International Conference on Computer Science, Engineering, and Applications (ICCSEA).

7. Use of Improved Random Forest for Survival Rate Prediction in Non-Small Cell Lung Cancer by Nanda and Duraipandian. Published in the 2020 International Conference on Inventive Computation Technologies (ICICT), pages 93–97, with the DOI 10.109/ICICT48043.2020.9112558.

Bezerra GM, Ohata EF, Chagas JVS das, de Albuquerque VHC, Hassan MM, and Filho PPR are the authors of the eighth article. Classification of colorectal cancer histology pictures using an innovative transfer learning method cancer(2021). Journal of Supercomputing, 77(9), 9494–9519. with the DOI 10.1007/s11227-020-03575-6.

9. Xiao Y, Li H, Lin J, Machine Learning Models for Colorectal Cancer Detection Using Traditional Laboratory Test Evidence. Modern Cancer Research and Treatment. 2021;20:15330338211058352. 15330338211058352 doi:10.1177/15330338211058352

RS Nayak, Amin J, Sharif M, Haldorai A, Yasmin M. Machine learning for brain tumor identification and classification: a thorough review | Complex & Intelligent Systems | This is EBSCOhost. article DOI: 10.1007/s40747-021-00563-y

El-Dahshan ESA, Mohsen H, El-Horbaty ESM, and Salem ABM were on the 11th. Automated brain tumor classification using DLNNs. Upcoming Computer Citation: Inform J. 2018;3(1):68-71. Published online: December 1, 2017, doi:10.1016/j. fcij.2017.12.001

Intelligent Model for Brain Tumor Identification Using Deep Learning, by Khan AH, Abbas S, and Khan MA, published in Appl Comput Intell Soft Comput in 2022 with the DOI 8104054. 2022/8104054 (doi:10.1155/2022.810454)

Avºar and Salçin (2013). Teh Glas. 2019;13(4):337-342. Using quicker R-CNN for tumor detection and

ISSN:0976-0172
Journal of Bioscience And Technology
www.jbstonline.com

classification in MRI images of the brain. 301707 (20190712095507)
Authors: Zaki SM, Jaber MM, and Kashmoola MA. Neural Network-Based COVID-19 Infection Diagnosis in Chest X-Ray Pictures. "Baghdad Science Journal" published in 2022, volume 19, issue 6, pages 1356–1356. DOI: 10.21123/ bsj.2022.5965
Himal Chand Thapa BKH 15. The Use of Convolutional Neural Networks on Histopathological Images for the Detection of Lung Cancer. The research team at Seventh Sense! Retrieved on September 30, 2024. Viewed at: https://dev.ijcttjournal.org//archives/ijctt-v68i10p104
Das, Saha, Nath, and Dipansu (2016). Investigating Machine Learning Approaches to Create a Parkinson's Disease Prediction Model. Science and Technology Research in Asia. 2024;21:569-582. Article DOI: 10.13005/bbra/3248
17. Srinivasu DrSVN, Professor of Computer Science and Engineering at Narasaraopeta Engineering College in Narasaraopet, Andhra Pradesh, India; Ponnada VT, Research Scholar at Acharya Nagarjuna University in Nagarjuna Nagar, Guntur, 522510, India. Intelligent Convolutional Neural Networks for Detecting Lung Cancer. Int