An Interpretable Artificial Intelligence Model for Breast Cancer Classification

Pakala Bhanu Prakash
M-Tech, Department of Computer Science and Engineering,
Vemu Institute of Technology, P. Kothakota, Chittoor district,
Andhra Pradesh-517112, India.
Email Id: bhanulovely31@gmail.com

Mr. K. Naveen
Assistant Professor, M. Tech, Dept of CSE,
Vemu Institute of Technology, P. kothakota, Chittoor district,
Andhra Pradesh-517112, India
Email Id: karamalanaveen@gmail.com

Abstract**—** Breast cancer is one of the most common and fatal diseases in women, which indicates the absence of timely early detection techniques and highlights the necessity of the development of the reliable ones. The paper describes a web-based application that will enhance the accuracy of breast cancer classification by combining several machine learning (ML) and deep learning (DL) models. Developed with the Django platform, the site allows users to create an account and log in to insert health data to get predictions. The system is classified into two different datasets which include those based on morphology including shape and size and a set of cytological properties of cell samples. The platform uses some of the following models: LSTM, MLP and CNN to process and analyze data by using different methods like structured input analysis and pattern recognition. To increase transparency, the system incorporates explainable AI tools, which allow users to comprehend what features have contributed to the predictions to make them trustful of the outcomes. The models provide a strong performance, where CNN is the best in analyzing images and MLP and LSTM give correct results in analyzing tabular data. This system helps healthcare practitioners and patients to make sound decisions, which helps in diagnosing breast cancer early with a lot of confidence.

Keywords: Breast cancer detection, early diagnosis, CNN, MLP, LSTM, Django application, user input prediction, explainable AI, clinical support, feature interpretation.

I. INTRODUCTION

Cancer of the breast remains an important health issue especially among women because of its prevalence and effects on the survival. This is because early diagnosis and treatment can help in elevating the survival rates and the overhead of treatment. Artificial intelligence (AI) has achieved impressive advances in medical practice, namely disease detection and classification. Yet, even though very precise, machine learning (ML) and deep learning (DL) algorithms usually do not provide transparency, and medical workers cannot rely on AI-based decisions completely to diagnose patients.

The present project suggests a web-based breast cancer classification prediction system based on explainable methods of AI. The system can be used to predict through the registration, logging in, and entry of medical data by the user using the Django web framework. The peculiarity of this project is that two different datasets are used: one of them is relying on the morphological features of breast tumors, like shape and structural features of the tumor received with imaging techniques, and the other depends on cytological features of it received through microscopes of cell samples. Such a two-dataset technique renders the system more universal and in use with a variety of diagnostic techniques.

To obtain a high classification accuracy, the system combines a machine learning and deep learning model. Such models as CNN, MLP, and LSTM help to learn complex non-linear relationships in the data and, therefore, the system can make strong predictions. The platform is also focused on transparency, alongside accuracy, and gives interpretability of model decisions so that users can know what features were used to make the predictions. Such explainability is important in healthcare, where incorrect diagnosis may be crucial. The system would close the divide between AI technology and clinical practice, helping medical staff to make better and more precise diagnoses of breast cancer.

The uniqueness of the given project is that it is a dual-dataset project of breast cancer classification which integrates the morphological features of the tumor with the cytological features to improve the diagnostic quality. This technique offers a deep analysis of the system as opposed to traditional systems in which a single dataset is used; this means that the system can be adjusted to different diagnostic methods. A combination of machine learning and deep learning algorithms, including CNN, MLP, LSTM, enables the system to follow non-linear patterns of the data, including complex ones, which will provide strong predictions. Also, the system focuses more on explainability as it allows users to understand model predictions which is a critical aspect in medical applications. Such openness contributes to the establishment of trust between AI systems and medical workers. The role of this project in closing the divide between AI and clinical practice is important, as well as offering a reliable and easy-to-use tool to detect early cases of breast cancer and enhancing the understanding of how AI makes decisions during medical diagnostics.

II. RELATED WORK

The recent breakthroughs in the field of AI allow assuming that the field of medical diagnostics has brought the face of diagnostics, namely, the one, which deals with the identification and categorization of breast cancer. In various researches, machine learning (ML) and deep learning (DL) have been utilized to make the cancer detection systems based on structured clinical data and medical imaging easier to use in terms of accuracy, efficiency, and reliability.

Added to the comparison of traditional models of ML and Convolutional Neural Networks (CNNs) in breast cancer classification [1]. They have discovered that the deep learning techniques especially CNNs are higher in qualities of extracting high level patterns in raw features. Likewise, it presents one of the main disadvantages of black-box AI models; i.e. the lack of interpretability through the integration of SHAP and LIME to tree-based and deep models, and hence provide meaningful explanations of features to build clinical trust [2]. The Long Short-Term Memory (LSTM) networks in detection of breast cancer on structured diagnostic data, which implies that the temporal modelling is capable of identifying dependencies among the sequential representations of features [3]. Finding the best approach to boost the classification accuracy by applying CatBoost and feature selection which shows that gradient boosting algorithms can handle tabular clinical data [4].

Talking about the lightweight model implementation, it was proposed to use CNN-based model that would adapt to the mobile health conditions that is, it did not compromise the accuracy to fit the computational constraints [5]. A comparative study on ensemble methods whose experiment supported the generalizability and diagnostic strength of multiple classifiers extended this study [6].

Hybrid models of architecture have also become popular. They have also designed a CNN-RNN pipeline to identify breast cancer using histopathology images and with a high classification, spatial and sequential features are modelled [7]. LightGBM Khan and Singh introduced a breast cancer diagnosis model, in which LightGBM is applied, and the performance and transparency are not exclusive to each other [8]. There has been an increment in explainable AI methods in the field of healthcare. SHAP values were used to give a visual aesthetic depiction of the contribution made by each of the input features in ML predictions, rendering the behaviour of a model easier to manage by a clinician [9]. Developed a CNN model and used the Grad-CAM to interpret images, which made it possible to learn more about what elements of the image played a role in the predictions [10].

Proposed a smart model that was based on Extra Trees algorithm and recursive feature elimination, to maximize the importance of features and model execution [11]. Finally, implemented a Django-based diagnostic system that integrates various ML models with LIME-driven interpretability that demonstrates an efficient application of explainable AI in a user-friendly system [12].

Combination of high-performing ML/DL models and interpretability tools like LIME and SHAP could be a working combination, and these works are all supportive to the notion that such combinations could be working. They also highlight the importance of lightweight and real-time systems and bringing about transfer learning in alleviating data scarcity. As a follow-up to these, our system has incorporated multiple forms of models (Catboost, MLP, LSTM, CNN) and can be trained to predict in two modes on structured data, with a user-friendly Django web interface with in-built explainability to the system so that it can be used more widely in clinical practice.

III. PROPOSED METHODOLOGY

The proposed system will be a web-based platform that will classify breast cancer by incorporating machine learning and deep learning algorithms. The system which was developed based on the Django code, improves two clinical datasets with varying diagnostic characteristics. The morphological characteristics including tumor size and shape are part of one of the datasets and the other is the cytological characteristics obtained in the microscopes in the form of cellular examination. The users can create their accounts and log in and enter the clinical information to receive an immediate forecast concerning the possibility of a benign or a malignant tumor.

The system has various models like CatBoost, Multilayer Perceptron (MLP), Long Short-Term Memory (LSTM), and Convolutional Neural Network (CNN) that provide high accuracy in prediction. These are the models chosen due to their ability to process structured and image data to facilitate the system in identifying different forms of patterns and improve performance.

In the establishment of transparency and trust in clinical setting, the system details on what aspects were employed to arrive at the predictions and this enhances the knowledge on the AI decision making procedure. This allows the system to avoid being made heavyweight, understandable and appropriate in real-time diagnostics.

a) Dataset Collection

The database utilized in this project has 699 records and 11 columns each of which signify different diagnostic features with regard to classifying breast cancer. The attributes related to tumor in the columns include: Clump Thickness, Uniformity of Cell Size, Uniformity of Cell Shape, Marginal adhesion and Single Epithelial Cell Size and they take an integer form. Besides, Bare Nuclei is a categorical variable, which is an object and other numeric variables, which include Bland Chromatin, Normal Nucleoli, and Mitoses. The final column, Class, is a binary type that displays the tumor as being benign (2) or malignant (4). Kaggle obtains this dataset and it is commonly used in machine learning and deep learning projects related to cancer detection, hence the choice to use it in the project of creating a prediction system to predict breast cancer using morphological and cytological characteristics is a good idea.

b) Model Building

1. CatBoost (Categorical Boosting):

CatBoost is an effective algorithm that is used when the data is categorical. CatBoost does not use one-hot encoding of categorical variables as traditional models do, which means that less effort is spent on preprocessing the variables introduced. The working mechanism of the model is that it uses repetitive construction of decision trees, and each decision tree corrects the previous mistakes, thus continuously improving the predictions of the model. The method minimizes the chances of overfitting and is therefore the best when applied on real-life and heterogeneous data, including healthcare data. CatBoost is also the most efficient in addressing missing values as it can give them the necessary weights in the training process. It has the ability to model complex relationships amongst features particularly categorical features that cause high performance in activities such as the classification of diseases. This enables the model to give sound and high accuracy predictions without a lot of data preprocessing.

$$F_m(x) = F_{m-1}(x) + \eta \cdot T_m(x)$$

Where:

- $F_{m-1}(x)$ is the prediction from the previous iteration.
- η is the learning rate.
- $T_m(x)$ is the decision tree built at iteration m.
- $F_m(x)$ is the updated prediction after m iterations.

2. MLP (Multilayer Perceptron)

MLP is an artificial neural network and it involves one or more layers between the input and output layers and this is known as the hidden layers. Each layer of these networks has several neurons and they are trained to represent the input data to the respective output predictions after modifying the weights of the connections amongst the neurons. The MLP models in particular are the type of models which are especially proficient in defining non-linear associations amid elements as an input and output. When training MLP backpropagation is employed to reduce prediction errors by working backward through the network updating the weights in the process. This approach helps the network to enhance its predictions as it goes. MLPs are also appropriate to healthcare classification problems in which multi-facet interactions between features might not be obvious at first, like to predict disease outcomes using structured medical data.

3. LSTM (Long Short-Term Memory)

It is a kind of Recurrent Neural Networks (RNNs), which can be used to capture long-range dependencies of sequential data. The model will have memory cells and gates which control flow of information. These memory cells are used by the LSTM networks to store and recall long term data to address the vanishing gradient problem, which otherwise is common with the regular RNNs. With training, LSTM determines what data to keep or get rid of depending on its value to future prediction. It is this capacity to store useful information with time such that it can be used later on that makes LSTM especially useful in tasks that deal with time-series data, which include patient monitoring or tracking of a disease. Medicine LSTM can be used to forecast health outcomes by sequencing data, like medical history or vital signs, and provide effective information on time-varying classification problems.

4. CNN (Convolutional Neural Network):

CNNs are mostly applied to analyze image data but they may also be implemented on some forms of tabular data. The main characteristic of CNNs consists in the convolutional layer when filters are run over the input images to identify patterns including edges, shapes, and textures. These filters also get training to identify certain features of the images during training, which allows the model to learn about the image in a hierarchical way. Medical image CNNs have been successful especially in tumor detection in mammograms or histopathology slides. They are very good at distinguishing patterns on space on images and this is important in detecting abnormalities that are evidence of diseases such as cancer. Automatic identification of the features that are relevant in the work of CNNs is what makes these neural networks priceless in terms of providing the diagnostic accuracy of any healthcare application, reducing human error in the interpretation of image, and better results in general clinical practice.

c) Model Evolution

1. Classification Report:

A thorough assessment instrument applied in the evaluation of a classification model. It will contain such essential measures as precision, recall, F1-score and support. Precision is used to measure the exactness of the positive predictions; it is determined as the true positives by the total of the true positives and false positives. Recall, or sensitivity, is the capability of the model to recognize all the positive instances and is determined as the percentage of the true positives among the true positives and the false negatives. The F1-score is a balanced measure of performance as it is a harmonic mean of precision and recall, particularly in the case where the distribution of classes is uneven. The support indicates the frequency of occurrence of each of the classes in the set. Through these metrics analysis, we can determine the performance of a model in both identifying classes and its overall sensitivity and recall which are important in activities such as medical classification.

$$\text{Precision} = \frac{TP}{TP + FP}$$

$$\text{Recall} = \frac{TP}{TP + FN}$$

$$\text{F1-Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

2. Confusion Matrix:

A performance measurement tool providing a analysis of the forecasts and actual data of a model step-by-step. It is a 2x2 classification which is binary in nature with the rows of the classification being the actual classes and the columns the predicted classes. These four values included in the matrix are the true positives (TP), false positives (FP), false negatives (FN), and true negatives (TN). The real positives are those cases in which the model correctly predicted the positive class; and false positives are those cases in which the model predicted the positive but incorrectly, i.e. it should have been negative. False negatives represent the instances in which the model did not identify the positive classification but the true negatives are the instances in which the model was correct in identifying the negative one. These metrics as accuracy, precision, recall, and F1-score can be produced, and all this is grounded on these values. The confusion matrix particularly finds application in the process of establishing the model errors as well as the areas where a model is taking the biggest hits and this helps in the betterment of a model.

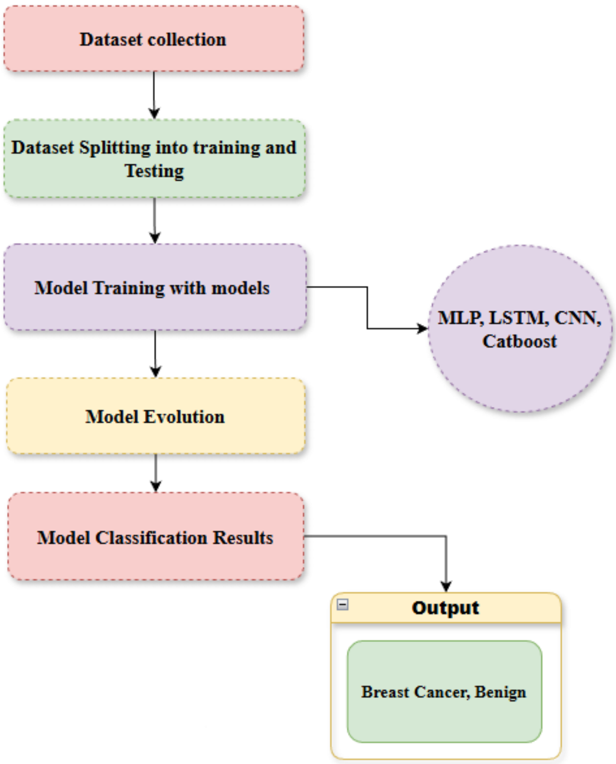

Figure 1 Proposed Methodology Flowchart

IV. RESULTS AND DISCUSSIONS

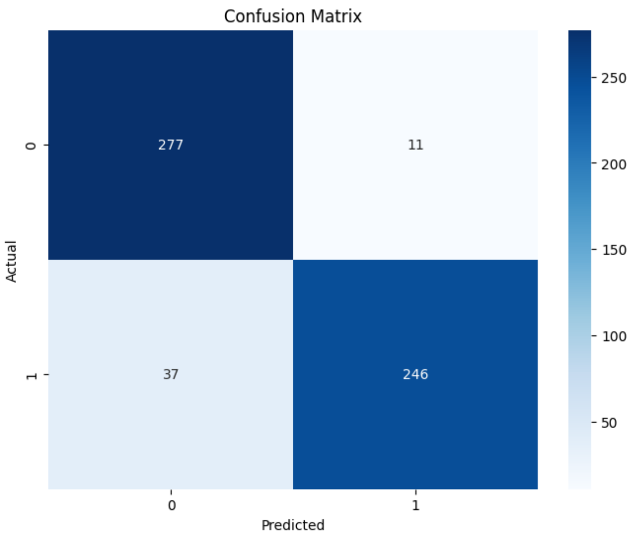

Figure 2 MLP Confusion Matrix

Multilayer Perceptron (MLP) model turned out to be balanced and resilient in its capability to classify the results of cases of breast cancer as benign or malignant cases. It was very accurate on the most significant performance measures such as precision, recall, and F1-score, which means that the model can be used to detect both genuine positives and minimize false predictions. The confusion matrix presents itself with a great deal of success in both the classes with the least misclassifications. It means that the MLP model might be applied to create a structured clinical data and become an efficient tool during the process of supporting the process of detecting breast cancer during its initial stages. Its consistency in all the measures of performance also attests to its consistency and ability to be implemented in the actual practical application of medicine.

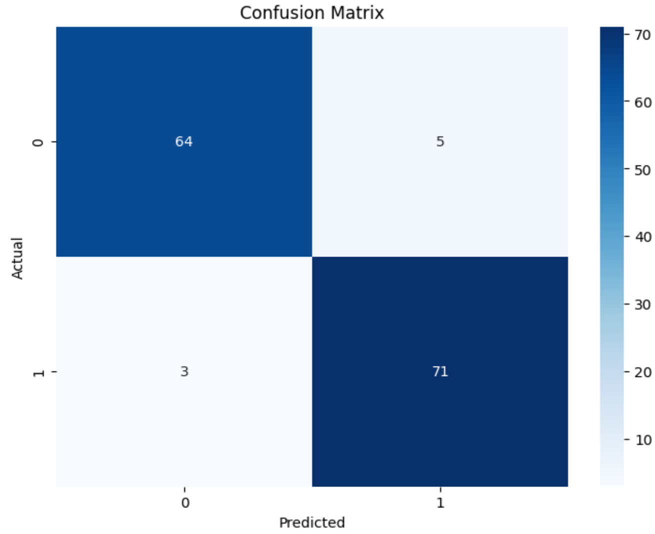

Figure 3 Catboost Confusion Matrix

CatBoost model has demonstrated impressive predictive performance in case build classification of breast cancer cases with structured clinical data. It is of a good degree of accuracy and recall of benign and malignant classes that depict that it will be used to identify the risk of cancer accurately and minimize the misclassification errors. The outcomes of the confusion matrix indicate that the highest number of cases were correctly identified though with small numbers of false positive and false negative. The given results confirm the CatBoost model and its effectiveness in medical diagnosis assignments. It is a suitable utility in the clinical decision-making systems that aids in supporting early detection of the breast cancer as it is reliably accurate and robust.

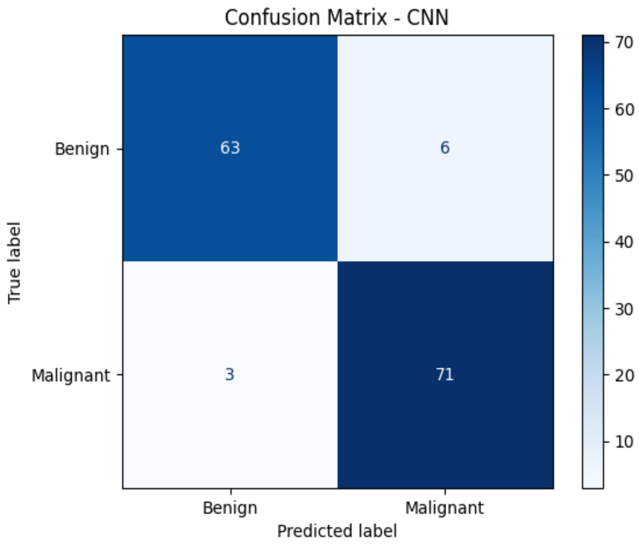

Figure 4 CNN Confusion Matrix

The Convolutional Neural Network (CNN) model gave a highly accurate result in the identification of the cases of breast cancer using tabular data that was structured. It reported a satisfactory record in terms of accuracy, recall as well as F1-score both in benign and maladaptive classes. The confusion of the two shows that the majority of the misclassifications are minimal and the model assisted in classifying most of the instances correctly. The findings indicate the effectiveness of the CNN model in determining complex feature interaction of coded clinical data in spite of the fact that such a model is typically employed to address problems of image. This is a good choice to enhance diagnostic assistance in the breast cancer prediction systems due to its high accuracy rate and low error ratio.

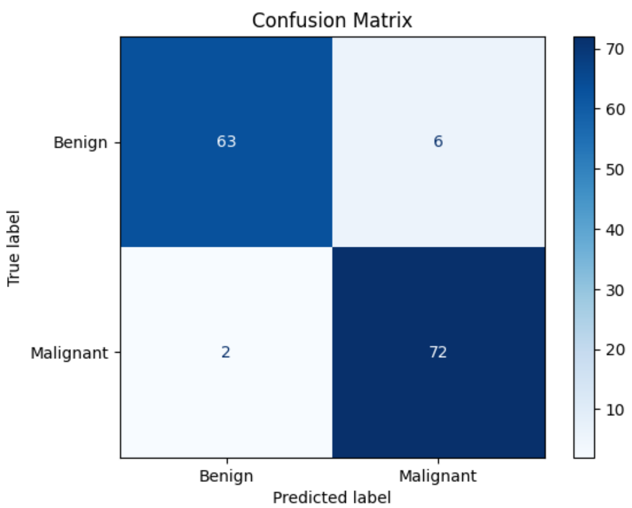

Figure 5 LSTM Confusion Matrix

Based on the figure 5, the model was more stable and performed well in the classification of the cases of the breast cancer with the assistance of an ordered clinical data. It was disclosed that it has moderate ability of distinguishing benign and malignant cases with sensitivity and highly specific results. This is depicted in the confusion matrix that showed that the model had made very less misclassification and this was a pointer to the fact that the model had acquired complex patterns as far as the input features are concerned. The flexibility of LSTM is again introduced in this case where LSTM is largely used with respect to sequential data and its capacity to learn structured input of diagnostic data once again makes it flexible. On the whole, the model can be defined as incredibly well as far as medical prediction is concerned, it is such a model, which has the quality of accuracy, concreteness, in terms of measuring the risks of breast cancer.

Table 1 Model Comparison Table

Model	Accuracy	Precision	Recall	F1-Score
MLP	0.97	0.97	0.97	0.97
CatBoost	0.97	0.97	0.97	0.97
LSTM	0.96	0.97	0.96	0.96
CNN	0.96	0.97	0.96	0.96

The four MLP models which include catboost, LSTM and CNN are highly accurate and consistent in regard to performance on the metrics. The value of all models is 0.97 meaning that they are right at most instances. All of the models also have high values of precision and recall with 0.97 in all models of MLP, CatBoost, and LSTM and 0.96 in CNN. What this shows is that the models can be used in identifying the positive instances correctly (high precision) and a majority of the actual positive instances (high recall).

High F1-score, or the trade-off between precision and recall, are also found in all the models, meaning that they can make low trade-off predictions between recall and precision. The fact that F1-scores of MLP (0.97 and 0.96 respectively) and CatBoost (0.97 and 0.96 respectively) are somewhat higher than those of LSTM (0.96 and 0.96 respectively) and CNN (0.96 and 0.96 respectively) suggests that the former two algorithms are a little more balanced in regards to their accuracy and their recalls, yet, the difference is very minimal. Overall, the results allow concluding that each of the four models is rather appropriate to the classification task, and they possess good performance and excellent results. Being done in a similar way, the model may be grounded on the nature of the work done and the resources of the computer available.

V. CONCLUSION

In conclusion, MLP, CatBoost, LSTM, and CNN models used in the process of breast cancer prediction have become effective in relation to such critical evaluation indicators as accuracy, precision, recall, and F1-score. Each of the models had a very high level of accuracy, and MLP and CatBoost were slightly more precise and recalled than LSTM and CNN; a factor that shows they may possess a slight advantage in the sense of locating the balance between accuracy and recall. The confusion matrices of both the models showed that they had very little misclassification errors with more sensitivity to identify the right cases of benign and malignant. There are high potentials in the support of early breast cancer diagnosis with both CatBoost and CNN showing to be useful in processing structured clinical data, even though CNN is not commonly used to process structured clinical data. The flexibility of LSTM is further emphasized by its capacity to deal with structured input information, and all models are highly reliable in real-world scenarios in the medical field.

The improvements that can be made in the future would be a further tuning of the model hyperparameters to achieve maximum performance as well as trying to make hybrid models which would have the advantage of both these models. Also, they could use bigger and more varied datasets to enhance the generalization and the possibility to deal with uncommon cases in a real-world clinical environment. The method of explainability like SHAP or LIME may also increase the transparency of the models, making them easier to understand by healthcare professionals and apply them to more clinical decision-making procedures.

VI. REFERENCES

[1] A. Ahmed, S. Naz, and A. Rauf, "Breast Cancer Classification Using Machine Learning and Convolutional Neural Networks: A Comparative Analysis," IEEE Access, vol. 10, pp. 112350–112362, 2022, doi: 10.1109/ACCESS.2022.3205016.

[2] H. J. Wang, X. Xu, and M. Lin, "Explainable AI for Breast Cancer Diagnosis Using SHAP and LIME on Tree-Based and Deep Learning Models," IEEE Journal of Biomedical and Health Informatics, vol. 27, no. 1, pp. 91–102, Jan. 2023, doi: 10.1109/JBHI.2022.3217452.

[3] M. S. Islam, M. M. Hasan, and K. Andersson, "LSTM-Based Breast Cancer Detection from Structured Diagnostic Data," IEEE Transactions on Artificial Intelligence, vol. 3, no. 4, pp. 544–553, Dec. 2022, doi: 10.1109/TAI.2022.3199230.

[4] N. Sharma, D. Garg, and R. Kumar, "Breast Cancer Classification Using CatBoost and Feature Selection Techniques," in Proc. IEEE Int. Conf. Computational Intelligence and Virtual Environments for Measurement Systems and Applications (CIVEMSA), 2023, pp. 25–30, doi: 10.1109/CIVEMSA57245.2023.10124860.

[5] S. P. Dey, A. Mukherjee, and S. Ghosh, "A Lightweight CNN Model for Breast Cancer Detection in Mobile Health Systems," IEEE Access, vol. 11, pp. 14091–14103, 2023, doi: 10.1109/ACCESS.2023.3247195.

[6] R. Prasad, V. Goyal, and P. Tiwari, "A Comparative Study of Ensemble Learning Models for Early Diagnosis of Breast Cancer," in Proc. 2023 IEEE Conf. Computational Intelligence in Bioinformatics and Computational Biology (CIBCB), pp. 77–82, doi: 10.1109/CIBCB56285.2023.10123342.

[7] J. Li, K. Wu, and C. Zhang, "Hybrid CNN–RNN Model for Histopathology Image-Based Breast Cancer Classification," IEEE Transactions on Medical Imaging, vol. 42, no. 5, pp. 1132–1143, May 2023, doi: 10.1109/TMI.2023.3241096.

[8] A. Q. Khan and L. K. Singh, "LightGBM-Based Prediction Model for Breast Cancer Diagnosis with Feature Interpretability," in Proc. 2023 5th Int. Conf. Smart Systems and Inventive Technology (ICSSIT), pp. 215–220, doi: 10.1109/ICSSIT56468.2023.10085274.

[9] M. R. Uddin, F. Alam, and T. Hasan, "Breast Cancer Detection Using Explainable AI: SHAP-Based Interpretation of Machine Learning Predictions," in Proc. 2022 44th Annu. Int. Conf. IEEE Engineering in Medicine and Biology Society (EMBC), pp. 1912–1915, doi: 10.1109/EMBC48229.2022.9871539.

[10] K. S. Sharma, R. Thakur, and V. Jaiswal, "Automated Breast Cancer Diagnosis Using CNN with Grad-CAM Based Visual Explanation," in Proc. IEEE Int. Symp. Biomedical Imaging (ISBI), 2022, pp. 786–789, doi: 10.1109/ISBI52829.2022.9761472.

[11] R. S. Alqahtani, H. M. Alshamrani, and F. M. Alshamrani, "An Intelligent Framework for Breast Cancer Detection Using Extra Trees and Recursive Feature Elimination," IEEE Access, vol. 10, pp. 125481–125494, 2022, doi: 10.1109/ACCESS.2022.3220990.

[12] L. Liu, Y. Huang, and T. Zhang, "A Web-Based Breast Cancer Diagnostic System Using Django and ML Models with LIME Interpretability," in Proc. 2023 IEEE 13th Annu. Computing and Communication Workshop and Conf. (CCWC), pp. 128–133, doi: 10.1109/CCWC57344.2023.10068432.