International Journal of Computer Sciences and Engineering
Vol.**11**, Special Issue.**1**, pp.**316-322**, November **2023**
ISSN: 2347-2693 (Online)
Available online at: www.ijcseonline.org

IJCSE
ISSN: 2347-2693 (E)

Research Paper

Designing Interpretable and Transparent Machine Learning Models for Early Detection of Breast Cancer

Souvik Santra[1], Arindam Sarkar[2], Pradip Sahoo[3], Trishita Ghosh[4], Siddharth Gautam[5], Arunava Bhowmik[6], Soumya Majumdar[7], Abhishek Bhowmik[8], Sonali Bhowmik[9*], Paramita Sarkar[10]

[1,2,3,4,5,6,7,8,9,10]CSE, JIS University, Kolkata, India

**Corresponding Author: sonali.bhowmik@jisuniversity.ac.in*

Abstract: Annually, over 1.15 million individuals are diagnosed with breast cancer across the globe. Today, only a select few accurate prognostic and predictive indicators are utilized in the clinical management of these patients. Early recognition of this deadly disease is pivotal, as it can reduce mortality rates while enhancing the life expectancy of those affected by breast cancer. Women are severely impacted by this condition, which has a high incidence and lethality rate. The absence of sturdy prognosis models complicates the medical professional's ability to develop a treatment strategy that could extend the patient's lifespan. Consequently, the need of the hour is to devise techniques that minimize error to boost accuracy. This study contrasts four algorithms, that predict the outcome of breast cancer, utilizing diverse datasets. All trials are performed within a simulated environment and facilitated on the JUPYTER platform. The research objective is divided into three sectors. The first sector entails predicting cancer prior to diagnosis, the second involves forecasting diagnosis and therapy, and the third centres on outcome during treatment. The proposed initiative can be employed to predict the outcomes of various methods, with the appropriate techniques chosen based on necessity. This investigation is undertaken to predict accuracy. Future studies can focus on predicting other distinct parameters, and breast cancer research can be classified based on these other parameters.

Keywords: Machine Learning, Identification of Breast Cancer, Attribute Isolation, PCA, LDA, Decision Tree.

1. Introduction

Breast cancer stems from your breast tissue. It arises when cells in the breast transform abnormally, proliferating unrestrainedly, thus forming a clump of tissue (tumor). Similar to other types of cancer, breast cancer has the potential to penetrate and expand into the tissue around your breast. Typically, when cells in the breast turn malignant, they accumulate to form a noticeable tumor that can be seen through radiographic imaging or felt as a lump.. The propagation of breast cancer occurs when these malignant cells infiltrate the circulatory or lymphatic systems, leading to metastasis. DNA alterations and deviations are responsible for breast cancer's genesis. When this happens, it's called metastasis.

Breast cancer ranks among the prevalent cancers affecting women. Women aged over 50 are most susceptible to it. .
Breast cancer in its initial phase often exhibits no signs. Often, an abnormality might be present in a mammogram, even when a tumor is too tiny to detect by touch. The most

common initial symptom, if a tumor is palpable, is a novel lump in the breast that wasn't previously present. Nevertheless, not all breast lumps indicate cancer. There are numerous breast cancer types, each associated with a diverse range of symptoms. Many of these symptoms overlap, yet there are also distinct ones. The following are symptoms frequently associated with the most common types of breast cancer:

• The emergence of a breast lump or thickening that feels distinct from the tissue around it and is new
Discomfort in the breast
• The breast skin turning red, discoloured, or appearing pitted
• Inflammation in a part or all of the breast
• Any discharge from the nipple excluding breast milk
• A blood-streaked secretion from your nipple
• The nipple or breast skin peeling, scaling, or flaking
• The nipple inverting

• Alterations in the skin look on the breasts
• A swelling or lump in your underarm
Even though you may have these symptoms, breast cancer is not always the cause. For instance, breast pain or lumps may be brought on by benign cysts.

Breast cancer can present in a variety of ways, which are often categorised into two groups: invasive and noninvasive. In situ breast cancer is another name for noninvasive breast cancer.

Noninvasive cancer remains within the original tissue, whereas invasive cancer spreads to neighbouring breast tissues from the milk ducts or glands.
.
A cancer that starts in the glands that make milk in your breast is called lobular carcinoma in situ (LCIS). Cancer cells have not spread to neighbouring tissues, unlike DCIS. IDC, or invasive ductal carcinoma, however some can develop into cancer. Angiosarcoma: This type of breast cancer appears on the blood arteries or lymphatic vessels.
Breast cancer is divided into phases by medical professionals according to the size of the tumour and the degree of its dissemination. Clinicians must be cognizant of the following while staging breast cancer:

• If the carcinoma is aggressive or contained • The dimensions of the malignant growth? • Involvement of lymphatic nodes? • Has the carcinoma spread to adjacent tissues or organs? There are five primary stages of breast cancer: from phase 0 to Phase 4.
Stage 0: DCIS is the term used for this stage. The malignant cells are still contained in DCIS.
Breast cancer is classified as invasive at this stage, meaning it has escaped its confines and is now attacking healthy tissue.

Stage 1A: At this point, the cancer has spread to the breast's fatty tissue. The primary tumour is no more than 2 centimetres (cm) in diameter. No lymph nodes are impacted.

Stage 1B: At this point, nearby lymph nodes have been shown to have cancer. The tumour in the breast is either undetectable.
Cancer in stage 2 has either grown, spread, or both.

Stage 2A refers to tumours that are less than 2 cm in size and have not yet spread to any adjacent lymph nodes, or that are larger than 2 cm but have not yet done so.
Stage 2B: The cancer has grown to a size higher than 5 cm but hasn't yet harmed any lymph nodes, or it has grown to a size between 2 and 5 cm but has already reached one to three axillary (underarm) lymph nodes.
Stage 3: The cancer is progressed and more challenging to treat, even though it hasn't spread to the bones or internal organs.
Stage 3A: Any size primary tumour is possible. The size of tumours exceeds 5 cm. Any breastbone node or one to three axillary lymph nodes have been affected by the malignancy.

Stage 3B: The skin or chest wall has been pierced by the tumour, and it may or may not have infiltrated up to nine lymph nodes.

Stage 4: By this stage, lymph nodes surrounding and the breast have not been spared by the spread of breast cancer cells. Most frequently, the brain, liver, lungs, and bones are affected. This stage is known as "metastatic," indicating that it has spread outside of the area of the body where it was first discovered.

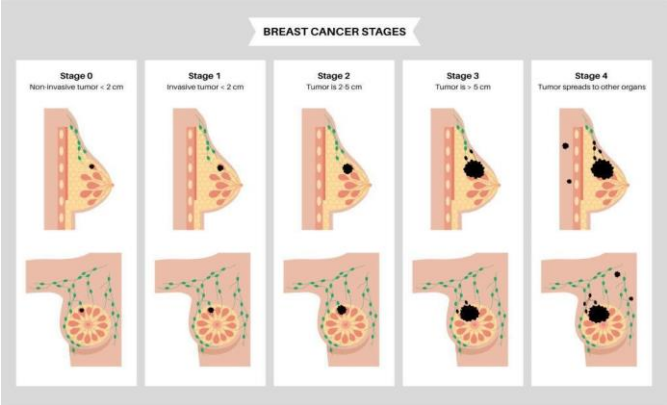

Figure 1. Stages of Breast Cancer [1]

2. Literature Review

In the realm of Supervised Learning classification, various methodologies like Neural networks, We compare the performance of Naive Bayes, CART Classifier, Decision Tree, and Support Vector Machine with Radial Basis Function (SVM-RBF) for datasets including breast cancer. Quinlan's study [2] utilized the C4.5 Decision Tree model, achieving an impressive 94.74% classification precision.

In a related vein, work from Hamlin et al. [3] applied the RIAC technique, achieving an impressive accuracy of 96%. Sudarshan Nayak's study [4] elucidates the role of distinct Forest methodologies, tracing back to when Breiman [5] introduced this approach to machine learning. The process begins by selecting m variables at random from a set of independent guided machine learning methodologies. Nayak proficiently categorized breast cancer from 3D imagery, identifying SVM as the superior performer. Breiman's goal was to refine the estimation method that had been previously employed by the Decision Tree and CART.

Machine learning (ML) explainability has emerged as subject of considerable fascination in recent times. The emergence of Shapley Additive explanations (SHAP) and Local Interpretable Model-Agnostic Explanations (LIME) [6] [6] are recent efforts to increase the interpretability of ML models. However, the appraisal of these tools is still in nascent stages. While some efforts have been made to assess the understandability or intuitiveness of model predictions by allowing users to experiment with different explanations [7], more formal, standardized assessments are required. In our study.

3. Proposed Method

The dataset for breast cancer was subjected to machine learning classifiers, and the results were examined in order to develop a model with the highest level of accuracy. Finding a reliable and accurate algorithm to diagnose breast cancer is the major objective of this work.

In this investigation, we have capitalized and encompasses 699 records, each comprising of nine distinct characteristics. These characteristics in WDBC include:

- Clump Density
- Consistency of Cell Dimensions
- Consistency of Cell Formation
- Peripheral Adherence
- Singular Epithelial Cell Dimensions
- Bare Cellular Nuclei
- Bland Chromosomal Structure
- Normal Nuclear Bodies
- Cell Division Rate.

Out of the 699 records, 458 or 65.5% of the records are Benign, while 241 or 34.5% of the records are Malignant.

The suggested framework is as follows:

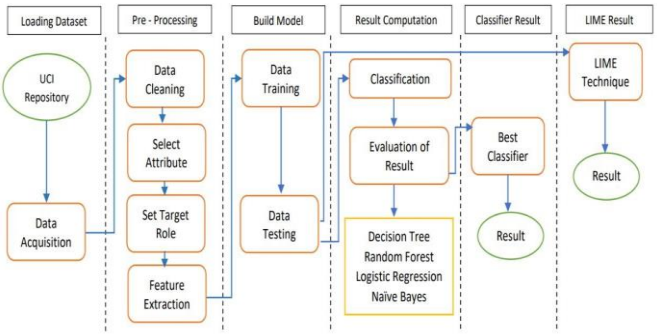

Figure 2. Proposed Model

3.1 Feature Extraction:

In the modern era, dealing with datasets boasting hundreds or thousands of attributes is the new norm. Encountering a number of attributes comparable to, or exceeding, the dataset's recorded observations could potentially result in a machine learning model prone to overfitting. Countermeasures against such issues often involve the application of regularization or dimensionality diminution tactics, such as Feature Extraction.

The newly condensed feature subsets should encapsulate a significant portion of the information found in the initial set. Consequently, a condensed version of the initial attributes can be synthesized from an amalgamation of the initial group. The most prevalent feature extraction techniques encompass:

Principal Component Analysis (PCA)
PCA, a widely utilized linear dimensionality reduction technique, processes the original data to find an amalgamation of input attributes that compactly represents the initial data distribution, effectively diminishing its original dimensions. It achieves this by magnifying variances and reducing the reconstruction error by evaluating pairwise distances. Being an unsupervised learning algorithm, PCA focuses exclusively on variation, regardless of data labels, which can occasionally lead to data misclassification.

Linear Discriminant Analysis:
LDA, a machine learning classifier and supervised learning dimensionality reduction algorithm, tries to maximise the mean distance between each class while minimising the dispersion within the class. Consequently, LDA makes use of both intra- and inter-class measures. The advantage of this approach is that it projects data into a lower-dimensional space while increasing the mean distance between each class, which can enhance classification outcomes due to the less overlap between the various classes.

With LDA, we presuppose that the data adheres to a Gaussian Distribution (as in this instance). Hence, implementing LDA on non-Gaussian data may yield less than optimal classification results.

3.2 Model Assembly
The prepared data is used to create machine learning algorithms that can predict breast cancer from a new set of measurements after feature extraction. By subjecting the model to fresh data with established labels, the algorithm's efficacy is assessed. This is commonly done by using the Train_test_split function to divide the collected labelled data into two portions. The training data, also known as the training set, makes up 80% of the data used to create our machine learning model. On the basis of the residual 20%, the model's effectiveness is assessed.

3.3 Algorithms for Machine Learning
Predictive analysis is carried out in this project using machine learning methods. In our project, machine learning methods like these were employed:

Decision Tree:
The Decision Tree algorithm is a member of the family of supervised learning algorithms. In contrast to other supervised learning methods, the decision tree methodology may be utilised for both classification and regression tasks.

The initial set S acts as the root node in the algorithm.
During each iteration, the method calculates the entropy (H) and information gain (IG) of each unused property in set S.
Then, the property with the greatest information gain or the lowest entropy is selected.

The chosen property produces a subset of the data in set S.
The method repeatedly iterates over each subset, only taking previously unselected features into account.

Random Forest:
"Random Forest, as the name suggests, is a classifier that comprises multiple decision trees on different subsets of the given dataset, and averages to improve the dataset's predictive accuracyRather of relying on a single decision tree, the random forest uses the forecasts from each tree and calculates the final output based on the majority of predictions. Following is the Random Forest classifier's working algorithm:

Algorithm:

Step 1: involves selecting K data points at random from the training set.

step 2: build the decision trees that correspond to the subsets of selected data points.

Step 3: Pick the N-th node in the decision tree you wish to build.

step 4: repeat steps 1 and 2.

Step 5: After receiving the forecasts from each decision tree for the new data points, classify the new data points by assigning them to the category that has received the most votes. Logistic Regression:

The statistician DR Cox introduced Logistic regression in 1958, predating the field of machine learning. Logistic Regression utilizes an equation akin to Linear Regression. In a categorical dependent variable, the output is predicted via logistic regression. As a result, the outcome must be a discrete or categorical value. It can be True or False, 0 or 1, Yes or No, etc. However, it delivers probabilistic values rather than precise 0 and 1 values.
values that range from 0 to 1.

(Sigmoid Function) Logistic Function:

A mathematical formula called the sigmoid function is used to convert expected values into probabilities.

It converts any real value between 0 and 1 to another value.

An "S"-shaped curve is produced by logistic regression if the value is greater than the range of 0 and 1. The Sigmoid function or logistic function is the name for this S-shaped curve.

The threshold value notion, which establishes the likelihood of either 0 or 1, is used in logistic regression. Values approaching the threshold value, for instance, tend towards 1, whereas values approaching the threshold value, tend towards 0.

Then, the property with the highest information gain or lowest entropy is picked.
To create a subset of the data, the set S is then divided by the chosen attribute.

The method keeps repeating for each subset, considering only attributes never selected before.

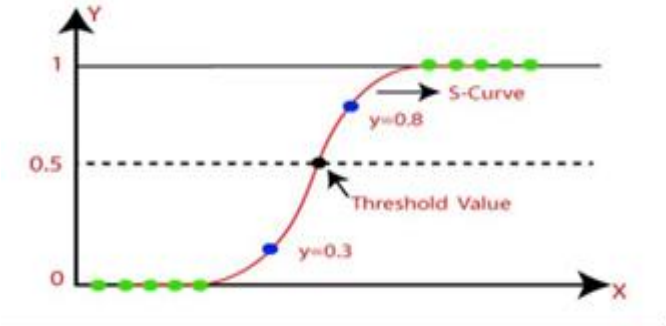

Figure 3. The Sigmoid Curve of Logistic Regression

Formula for Logistic Regression:
The logistic regression equation is built upon the linear regression equation. The mathematical procedures to create the Logistic Regression equations are as follows:

We are all aware with the straight line equation, which is written as y = b0 + b1x1 + b2x2 + b3x3 +.. + bnxn.

The only valid values for y in Logistic Regression are 0 and 1. Let's modify the equation above by dividing it by (1-y) to account for this:

When y = 0, 1-yy=0, and when y = 1, infinity are true.

But since we need a range from -[infinity] to +[infinity], let's take the equation's logarithm, which gives us:

Log[1-yy]=B0 + B1x1, B2x2, B3x3, Bnxn,

Naïve Bayes:
A supervised learning algorithm is the Naive Bayes one. It is based on the Bayes theorem and is mostly applied to classification issues. The Nave Bayes Classifier is one of the most straightforward but effective Classification algorithms that makes it easy to build quick machine learning models that can make quick predictions. As a probabilistic classifier, it makes predictions based on the likelihood that an object will occur.

Spam filtration, sentiment analysis, and article categorization are a few popular uses of the Naive Bayes algorithm.

The terms Nave and Bayes, which make up the phrase "Nave Bayes algorithm," can be described as follows:

Because it is based on the Bayes' Theorem, the word "Bayes" is employed. The Bayes theorem's formula is as follows: Where,

P(A|B) is the Posterior probability: The probability of hypothesis A given the observed event B.

P(B|A) is Likelihood probability: Probability of the evidence given that the probability of a hypothesis is true.

P(A) denotes Prior Probability: Likelihood of hypothesis prior to the observation of evidence.

P(B) signifies Marginal Probability: Likelihood of evidence.

Algorithm:

Initially, it transmutes the provided dataset into frequency tables.

Create a Likelihood table by determining the probabilities of the given attributes.

Subsequently, employ Bayes theorem to compute the posterior probability.

3.4 Model Clarification via LIME
Clarifying the operations of a machine learning model entails understanding the process from input to output. This eradicates the obscurity of the model, addressing the so-called black box issue. Successful implementation of explainable AI (XAI) results in several key advantages:

Builds trust in ML models:
When stakeholders comprehend the mechanisms behind an ML model's final output, they exhibit increased trust in AI-driven systems. Explainable AI tools serve to elucidate the rationale leading to the model's result. For example, if a deep learning model is used for medical image analysis, such as X-rays, explainable AI could generate saliency maps, spotlighting the pixels crucial for diagnosis.

Elevates troubleshooting capabilities:

Explainable AI enables improved debugging and assessment of a model's functionality. Consider a model designed to recognise animals in images. You observe a recurring error where dogs in snow are misidentified as foxes. Explainable AI tools simplify diagnosing this persistent issue. As you investigate, it's found that the model uses image backgrounds for animal identification, mistakenly equating snow backgrounds with fox presence.

Improved control:
Comprehending your models' decision-making uncovers unseen vulnerabilities and mistakes. Enhanced control is facilitated with this knowledge, enabling quick error correction, which is particularly beneficial when implemented across all production models.

Explainable Model Techniques in ML:
We will focus on LIME (Local Interpretable Model-Agnostic Explanations), a tool from University of Washington researchers for dissecting algorithm internals. The nomenclature 'Local Interpretable Model-

Agnostic Explanations' highlights the desirable characteristics in explanations. 'Local' signifies local fidelity - an explanation accurately mirroring the classifier's behaviour in the vicinity of a predicted instance. These explanations are ineffective if not interpretable - humans must grasp their meaning. LIME can clarify any model without the need for internal access, hence it is model-agnostic.

LIME offers a localized interpretation by adjusting single data sample feature values and monitoring the output effect. It constructs a surrogate model from the input (sample generation) and model predictions. An understandable model can serve as a surrogate model. As a model-agnostic tool, LIME can be applied to any model. Steps involved in LIME:

It generates a permutation (imitation) of the existing data.
It computes the distance between permutations and original observations. Distance measures can be specified.
Predictions are made on the new data using a black-box model.
It selects "m" features that best depict the complex model, using the permuted data via the maximum likelihood approach. The number of features, i.e., the value of "m" is flexible.

The “m” features are chosen, and a simple model is fitted to the permuted data with similarity scores as weights.
The weights from the simple model elucidate the complex model's local actions.

4. Results and Discussion

After applying Machine Learning Algorithms on Breast Cancer dataset. We used Accuracy, as performance metrics to evaluate and compare the models and identify the best algorithm for the breast cancer Prediction. Accuracy is the most common performance metric for classification algorithms. It is defined as the number of correct predictions made as a ratio of all predictions made.

Table 1. Breast cancer dataset accuracy %

Algorithms		Set of Precision Training (%)
Decision Tree		99.8 %
Logistic Regression		96.2 %
Random Forest		99.6 %
Naive Bayes		96.4 %

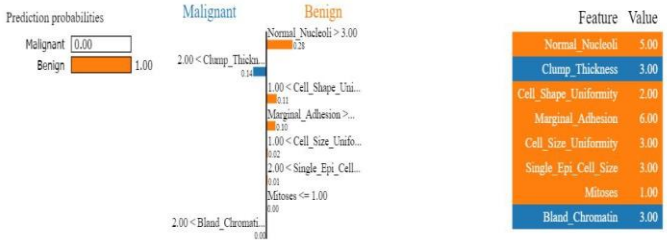

Figure 4. A comparison of several classifiers

Figure 5. Benign Prediction Using LIME

Here, we have a data point that is both anticipated to be benign and is really benign. We can see on the left side that the supervised model predicts with near certainty (~100% probability) that this data point is benign. Interestingly, LIME is capable of elucidating this prediction, using each attribute of the relevant data point in order of significance.

According to the insights derived from LIME,

The fact that the sample exhibits a value exceeding 3.00 for Normal Nucleoli increases its probability of being benign.

Given that the Cell_Shape_Uniformity falls within the 1 to 5 range, it is more inclined towards benignity than malignancy.

Conversely, the sample's elevated Marginal_Adhesion value leans it more towards benignity.

Because the sample's value for mitoses is less than 1, its likelihood of being benign is enhanced.

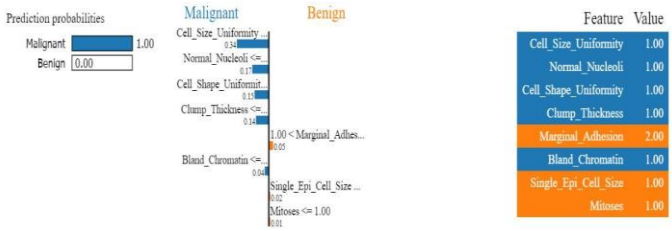

Figure 6. Malignant Prediction Using LIME

In this case, the data point is both malignant in actuality and malignant in prediction. The supervised model predicted that this point would be likely to be malignant with a probability of over 100%, as shown in the left panel. In the centre, we see that LIME is able to explain this prediction utilising each relevant aspect of the data point, in order of relevance.

According to LIME

Take a look at the first three features they all are having value <= 1.00 they are all increasing the chance of being classified as Malignant.

On the other hand, the fourth sample Clump_Thickness has a value <= 2.00 that makes it more likely to be Malignant.

5. Conclusion

In our analysis of the Breast Cancer dataset, we utilized four core techniques, namely: Decision Tree, Random Forests, Logistic Regression, and Naive Bayes. We performed, assessed, and compared varied outcomes based on precision to determine the superior machine learning algorithm that offers accuracy, dependability, and yields the topmost precision.

The methods were implemented in Python leveraging the scikit-learn library within the Visual Studio Code setting using Jupyter Notebook.

Upon meticulous comparison among our models, it emerged that the Random Forest Algorithm obtained an elevated testing accuracy of 95.7%, surpassing all other techniques Hence, it's crucial for future investigations to deploy these same techniques and procedures on different datasets to verify the outcomes produced by this dataset. Furthermore, in our upcoming studies, we intend to utilize other machine learning techniques with novel parameters on extensive datasets with more disease categories to achieve increased precision.

References

[1] The Massivebio, accessed 06 June 2022 <https://massivebio.com/breast-cancer-stage-4/>.
[2] J. R. Quinlan, “Improved Use of Continuous Attributes in C4.5,” J. Artif. Intell. Res., vol. 4, pp. 77–90, Mar. 1996.
[3] H. J. Hamilton, N. Shan, and N. Cercone, “RIAC: A Rule Induction Algorithm Based on Approximate Classiication,” 1996.
[4] S. Nayak and D. Gope, "Comparison of supervised learning algorithms for RF-based breast cancer detection," 2017 Computing and Electromagnetics International Workshop (CEM), Barcelona, 2017, pp.
[5] L Breiman, Machine Learning 45, 5-32 (2001).
[6] M. T. Ribeiro, S. Singh, and C. Guestrin, “Why should I trust you?: Explaining the predictions of any classifier,” in Proceedings of the 22nd ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining, ACM, 2016, pp. 1135–1144.
[7] A. A. Freitas, “Comprehensible classification models: A position paper,” ACM SIGKDD explorations newsletter, vol. 15, no. 1, pp. 1–10, 2014.
[8] UCI Repository, Wisconsin Breast Cancer Database (WBCD),2018.

AUTHOR’S PROFILE

Pradip Sahoo is currently pursuing Bachelor of Technology in Computer Science and Engineering from JIS University, batch of 2019-2023.

Souvik Santra is currently pursuing Bachelor of Technology in Computer Science and Engineering from JIS University, batch of 2019-2023.

Trishita Ghosh is currently pursuing Bachelor of Technology in Computer Science and Engineering from JIS University, batch of 2019-2023.

Siddharth Gautam is currently pursuing Bachelor of Technology in Computer Science and Engineering from JIS University, batch of 2019-2023.

Arunava Bhowmik is currently pursuing Bachelor of Technology in Computer Science and Engineering from JIS University, batch of 2019-2023.

Arindam Sarkar is currently pursuing Bachelor of Technology in Computer Science and Engineering from JIS University, batch of 2021-2025.

Abhishek Bhowmik earned his B. Tech and M.Tech from West Bengal University of Technology. He is currently working as an Asst. Professor at JIS University, Kolkata. His total teaching experience is 7 years

Sonali Bhowmik earned her B. Tech and M.Tech from West Bengal University of Technology. She is currently working as an Asst. Professor at JIS University, Kolkata. Her total teaching experience is 8 years.

Paramita Sarkar earned her B.Tech in C.S.E. and M.E. in C.S.E. She has been working as an Assistant Professor in CSE for last 17 years in Engineering colleges and University.
She has submitted her Thesis of Ph.D.(Tech) in the University of Calcutta in 2022. She has published 11 research papers in many national and International conferences, workshops. She has published papers in WoS, SCIE, SCOPUS indexed journals. Her area of specializations are Artificial Intelligence, Computer Networking, Resource scheduling optimization, Cloud computing.

Soumya Majumdar earned his B. Tech., M. Tech. in Computer Science and Engineering and Multimedia Information Processing from IIIEST Shibpur and IIT Kharagpur in 2014 and 2017 respectively. He is currently working as Assistant Professor in Department of CSE from JIS University, Kolkata since 2022. Besides, he is pursuing PhD in the Department of CSE in IIT Kharagpur.

