SIKSHA 'O' ANUSANDHAN

(Deemed to be University)

Faculty of Engineering & Technology (ITER)

Department of Computer Science and Engineering

Project Synopsis Form

SENIOR DESIGN PROJECT-2023

SECTION: K **GROUP NO: K-14**

PROJECT TITLE : Machine Learning Model to Diagnose Breast Cancer Using Mammogram Images

---

PROJECT ABSTRACT:

Cancer is one of the common diseases affecting people's life which has no cure yet. Aim of the study is to find model that will be helpful to predict breast cancer as the early detection of this disease could help slow down the disease. So, the patients will be saved from the effort of visiting a medical center, consulting doctor, and from various complications that occur if breast cancer remains untreated. The project's development and training of machine learning models will be done utilizing cutting-edge algorithms like Naive Bias, Decision Trees, Support Vector Machines. The study is done using the PIMA Indians for breast Cancer which are available publicly at the UCI machine learning repository. The likelihood of a successful therapy and the necessity for more intrusive operations can both be considerably increased with early detection of breast cancer. While working on this project we will take the dataset of the breast cancer patient and put the dataset in to the algorithms in order see which algorithm is giving highest accuracy, in order to examine huge patient data sets,including mammography pictures, demographic information, family history, and lifestyle factors, the research will use machine learning techniques. Among all the machine learning algorithms SVM is usedto perform classification. SVM obtains an accuracy of 94.28% which is then compared with Naive Biasand Decision Tree classifier algorithms. Accuracy of the obtained results can be increased in the future by using a larger dataset to train the model and obtain the result. After finishing the algorithm comparison. we will get a highest accuracy showed algorithm. So, that algorithm can be used for detection of breast cancer. The system will offer a trustworthy anduseful tool for classifying patients into high-risk groups, prioritizing patients for additional screening, and making tailored treatment recommendations.

(1) SOFTWARE, HARDWARE OR METHODS/ALGORITHMS SPECIFICATIONS:

a. Python as the mainframe Language

b. Datasets of breast cancer patients

c. Jupyter Notebook for code tweaking and dataset analysis

d. We use KNN, SVM, Naïve Bayes, Decision tree, Random Forest, Logistic regression algorithm to analyze the dataset and find the accuracy.

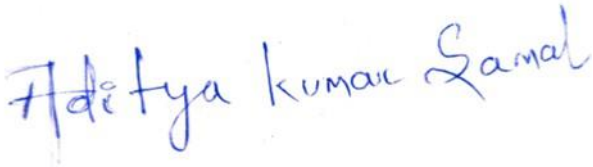
(2) NAME, REG. NO AND SIGNATURE OF GROUP MEMBERS:

1) Priyambada Sahu (2051012015)

2) Madhusmita Jena (1941012932)

3) Swayam Sarthak Rout (1941012604)

4) Aditya Kumar Samal (1941012931)



(3) APPROVAL STATUS (To be filled in by the Section Coordinator of SDP):

.....	.....	.....
Project Supervisor	**Section Coordinator, SDP**	**SDP Coordinator**