# “BREAST CANCER DETECTION “

##### 

##### NATASHA SHARMA

2106200

***Abstract*** Using transformer models like DT, Random Forest, LR, Naïve Bayes, KNN and SVM, as well as tools like Seaborn, Pandas, matplot, numpy and sklearn . This research looks at mean radius, mean texture and mean perimeter and detects the cancer.

### Individual Contribution:

***Model Implementation – DT, LR, Naïve Bayes, KNN and SVM, Connection UI***

***to web app, ROC Curve***

My primary contribution to this project was the implementation of models – DT, LR,

Naïve Bayes, KNN and SVM that is, create all the models using Python language, to

be applied on the Breast Cancer dataset in order to predict the cancer. I had also

created a graph depicting importance of feature on result. And then for comparision of

different model and to find which model worked better, I created ROC Curve and also

took out the accuracy score for all the models and found that LR Models worked best for

Breast Cancer dataset. I had also worked on connection of UI interface to the selected

model i.e LR model .

### Findings:

In implementing these models, a variety of tools were utilized, predominantly Python libraries such as scikit-learn for model implementation, Pandas for data manipulation, and Matplotlib and Seaborn for visualization purposes. The importance of model diversity became evident,

as each model offered unique perspectives on the data and contributed to a more robust analysis. Furthermore, through model evaluation using various metrics provided insights into performance strengths and weaknesses, guiding the selection of the most appropriate model for the task at hand.

## Individual contribution to project report preparation

## Report review and formatting

## Individual contribution for project presentation and demonstration:

Model Selection , ROC Curve, Model Selected , Result and conclusion

# BREAST CANCER DETECTION“

##### AMMAR YASIR

2106184

**Abstract** Using transformer models like DT, Random Forest, LR, Naïve Bayes, KNN and SVM, as well as tools like Seaborn, Pandas, matplot, numpy and sklearn . This research looks at mean radius, mean texture and mean perimeter and detects the cancer.

### Individual Contribution:

***Dataset Finding, Model Implementation – Random Forest , UI Design, Connection***

***of UI to web app***

My primary contribution were, finding dataset which locate a reliable and well-structured dataset containing breast cancer information. Ensure sufficient data for training and testing

model effectively. I implemented a Random Forest classifier, an ensemble learning

method that combines decision trees for robust and accurate predictions. I had create a

user interface (UI) that allows users to interact with the model. I also worked on integrating

the UI with a web application framework for broader accessibility.

### Findings:

From this project I gain practical experience in data acquisition, preprocessing, model

selection, training, evaluation, and interpretation. Understand how Random Forest works, its strengths (handles imbalanced data, robust to outliers), and potential limitations. Learn to connect machine learning models to user interfaces for real-world applications.  Consider the importance of clear communication in presenting model results and potential limitations, emphasizing that the model should be used as a decision support tool alongside medical expertise.

## Individual contribution to project report preparation:

Project Planning, SRS, UML, Implementation

## Individual contribution for project presentation and demonstration:

Tools and Techniques, UML,Dataset explanation



**SAMPLE INDIVIDUAL CONTRIBUTION REPORT:**

# BEAST CANCER DETECTION “

##### SAGAR SINGH

2106245

**Abstract** Using transformer models like DT, Random Forest, LR, Naïve Bayes, KNN and SVM, as well as tools like Seaborn, Pandas, matplot, numpy and sklearn . This research looks at mean radius, mean texture and mean perimeter and detects the cancer.

### Individual Contribution:

### Dataset Finding, Data Cleaning and Preprocessing

My contributions of dataset finding and data pre-processing in a machine learning project for breast cancer detection. The contribution lies in sourcing a dataset that is comprehensive, accurately labeled, and representative of the target population.I had also worked on data preprocessing i.e preparing the dataset for analysis and model training. This involves several key steps, including cleaning noisy data, handling missing values, removing outliers, and normalizing or scaling features.

### Findings:

The quality of the dataset directly impacts the performance of machine learning models. Finding a high-quality dataset is essential for accurate predictions. Cleaning and

preprocessing data require attention to detail and knowledge of various techniques such

as handling missing values and outliers. Working with healthcare data provides valuable insights into the domain, including understanding medical terminology and the intricacies of disease diagnosis.

## Individual contribution to project report preparation:

Introduction, Basic Concepts/literature View , Machine learning Algorithms, Conclusion and

Future Scope

## Individual contribution for project presentation and demonstration:

## Introduction to team and domain , problem Statement