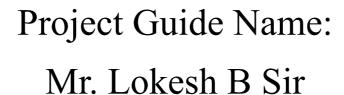# Project Title:

Analysis and Predication of "Breast Cancer Diagnosis Using Logistic Regression" to classify breast cancer tumors as either malignant or benign with high accuracy.

**Name:K.Pranavi**

**Roll No:20X01A0584**

**Year: IV**

Project Guide Name:

Mr. Lokesh B Sir

# Problem Statement

Breast cancer is a prevalent and potentially life-threatening disease. Early detection is crucial for successful treatment. This project leverages logistic regression and machine learning techniques to predict breast cancer diagnoses based on patient data. The Breast Cancer dataset is used, containing various features extracted from breast cancer biopsies.

The primary goal is to build a predictive model that can classify breast cancer tumors as either malignant or benign with high accuracy. We preprocess the data by splitting it into training and testing sets and standardizing the features to improve model performance. A logistic regression model is trained on the training data to learn the underlying patterns.

The project evaluates the model's performance using metrics such as accuracy, confusion matrix, and a classification report. Accurate breast cancer classification can assist medical professionals in making more informed decisions about patient care and treatment options, potentially saving lives through early diagnosis.

# ✨ **Logistic Regression Algorithm** ✨

## ▾ Project Title:

Analysis and Predication of "Breast Cancer Diagnosis Using Logistic Regression" to classify breast cancer tumors as either malignant or benign with high accuracy.

**Problem Statement:**

Breast cancer is a prevalent and potentially life-threatening disease. Early detection is crucial for successful treatment. This project leverages logistic regression and machine learning techniques to predict breast cancer diagnoses based on patient data. The Breast Cancer dataset is used, containing various features extracted from breast cancer biopsies.

The primary goal is to build a predictive model that can classify breast cancer tumors as either malignant or benign with high accuracy. We preprocess the data by splitting it into training and testing sets and standardizing the features to improve model performance. A logistic regression model is trained on the training data to learn the underlying patterns.

The project evaluates the model's performance using metrics such as accuracy, confusion matrix, and a classification report. Accurate breast cancer classification can assist medical professionals in making more informed decisions about patient care and treatment options, potentially saving lives through early diagnosis.

## ▾ Importing the Libreys and Dataset

```
import pandas as pd
import numpy as np

data=pd.read_csv("Breast_cancer_data.csv")
X=data.iloc[:,:-1].values
y=data.iloc[:,-1].values
```

## ▾ Split the data into training and testing sets

```
from sklearn.model_selection import train_test_split
X_train,X_test,y_train,y_test=train_test_split(X,y,test_size=0.2,random_state=17)
```

## Create and train the Logistic Regression model

```
from sklearn.linear_model import LogisticRegression
model=LogisticRegression()
model.fit(X_train,y_train)
```

```
  ▾ LogisticRegression
  LogisticRegression()
```

## Make predictions on the test set

```
y_pred=model.predict(X_test)
y_pred
```

```
array([1, 1, 1, 1, 1, 1, 1, 0, 1, 1, 1, 1, 1, 1, 0, 0, 1, 0, 1, 1, 1, 1,
       0, 1, 1, 1, 1, 1, 1, 0, 1, 0, 1, 0, 1, 1, 1, 1, 1, 1, 1, 1, 1, 0,
       1, 1, 1, 1, 1, 1, 1, 0, 1, 1, 0, 1, 0, 1, 1, 0, 0, 0, 0, 1, 1, 0,
       0, 1, 1, 1, 0, 1, 0, 0, 1, 1, 1, 0, 0, 1, 1, 0, 1, 1, 1, 0, 0, 1,
       0, 1, 0, 0, 1, 0, 1, 1, 0, 0, 0, 0, 1, 1, 0, 1, 0, 1, 1, 1, 1, 0,
       1, 1, 1, 1])
```

```
#import classes and method
from sklearn.metrics import accuracy_score,confusion_matrix,classification_report
# Calculate accuracy
accuracy=accuracy_score(y_test,y_pred)
print(f"Accuracy:{accuracy:.2f}")

# Confusion Matrix
conf_matrix=confusion_matrix(y_test,y_pred)
print("Confusion Matrix")
print(conf_matrix)
# Classification Report
class_report=classification_report(y_test,y_pred)
print("Classification Report:")
print(class_report)
```

```
Accuracy:0.96
Confusion Matrix
```

```
[[35  3]
 [ 2 74]]
Classification Report:
              precision    recall  f1-score   support

           0       0.95      0.92      0.93        38
           1       0.96      0.97      0.97        76

    accuracy                           0.96       114
   macro avg       0.95      0.95      0.95       114
weighted avg       0.96      0.96      0.96       114
```

# ✨Result / Conclusion

1. From my ML Model I concluded that , It's to predict whether a breast tumor is benign (non-cancerous) or malignant (cancerous) based on various features extracted from diagnostic..

2. Here's a brief description of the target variable:

   Target Variable (y):

   0: Malignant (cancerous)

   1: Benign (non-cancerous)

3. So, in this dataset, we are predicting the presence or absence of breast cancer (malignant or benign) based on features like tumor size, texture, perimeter, and other characteristics extracted from medical images. The logistic regression model is used to classify tumors into one of these two categories.

-----------------------Thank You! --------------------------

✓　0s　　　completed at 11:07 AM　　　　　　　　　　●　✕