

# CDSS Project 1

# Team 6

**Breast Cancer Classification**

**Using Different ML Models**

|  |  |
| --- | --- |
| Name | ID |
| Aya Amr | 9202342 |
| Sara Ayman | 9202615 |
| Shuaib Abdulsalam | 9204062 |
| Abdelrahman Saeed | 9202839 |
| Mahmoud Rabea | 9203396 |

# Table of Contents

|  |  |
| --- | --- |
| Subject | Page |
| Introduction About Breast Cancer | 3 |
| Dataset Description | 3 |
| Used ML Models | 4 |
| Results Comparison | 8 |
| Links | 9 |

# Introduction About Breast Cancer

Breast cancer is a cancer that develops in the breast cells and progresses in stages. Few early symptoms may include new lump in the underarm or in breast, itching or discharge from the nipples, and skin texture change of the nipple or breast.

Cancer begins when healthy cells in the breast change and grow out of control, forming a mass or sheet of cells called a tumor. A tumor can be cancerous or benign. A cancerous tumor is malignant, meaning it can grow and spread to other parts of the body.

The exact cause of breast cancer is not known but risk factors include family history, hormonal changes, age (at more risk after 40 years of age), personal history of breast cancer, lifestyle including excess of alcohol consumption, environmental factors including exposure to radiations, obesity and overweight, menarche having periods at younger age and menopause at an older age, pregnancy becoming pregnant at an older age or never being pregnant, hormone use including long-term contraceptive use or postmenopausal hormone therapy.

Preventive measures involve healthy habits such as eating healthy and nutritious food, avoiding alcohol, practicing gentle exercises upon doctor’s advice, visiting doctor for regular examination, preventive surgery may be recommended in women with high risk. To reduce the risk of developing cancer, get the pre-screening done.

# Dataset Description

The dataset we used is used to predict the type of cancer either it is malignant or benign based on the input parameters like radius mean, texture mean, perimeter mean, area mean, smoothness mean and other parameters shown below.

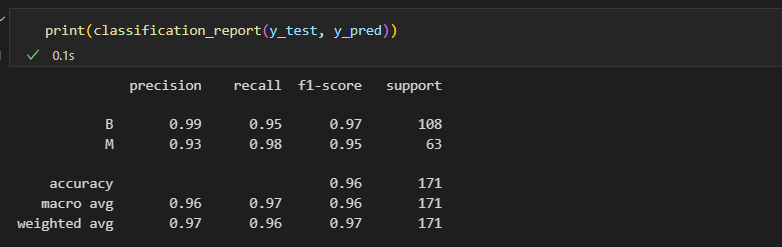
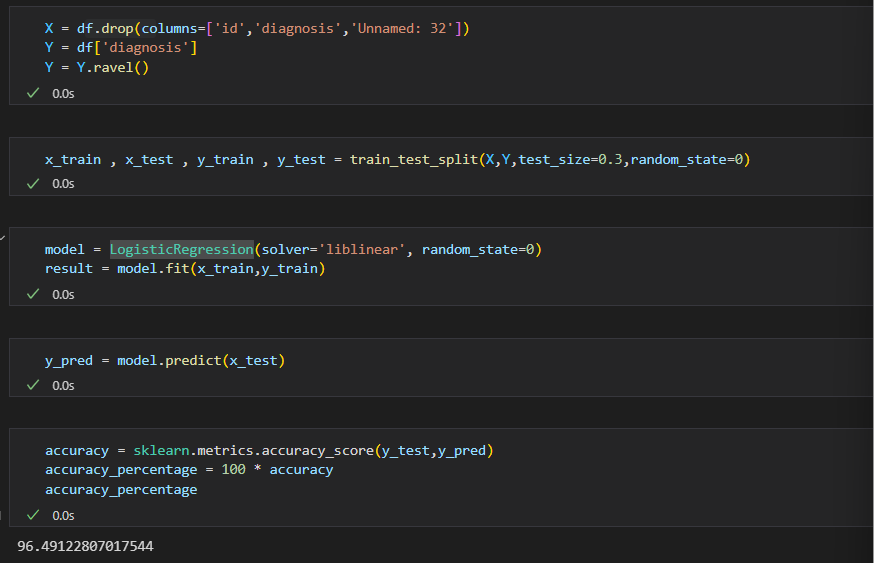
Text

Description automatically generated

The dataset has 357 records with benign tumor and 212 records with malignant tumor, where 0 or M presents that the patient has malignant tumor while 1 or B represents that the patient has benign tumor.

# Used ML Models

* Logistic Regression



* Decision Tree

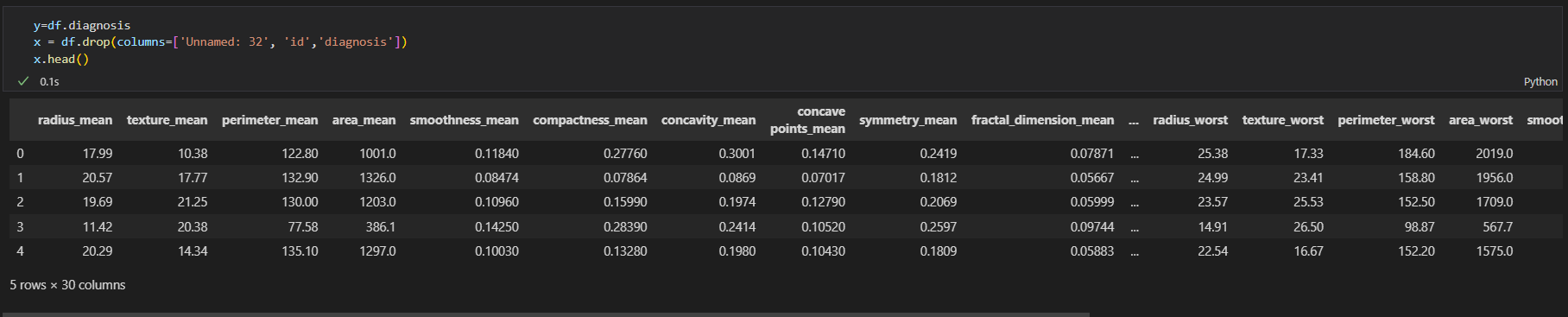
Graphical user interface

Description automatically generated

Graphical user interface

Description automatically generated with medium confidence

* Random Forest



A picture containing text

Description automatically generated

A screenshot of a computer screen

Description automatically generated with medium confidence

Graphical user interface, application

Description automatically generated

A picture containing text

Description automatically generated

* SVM (Support Vector Machine)

Text

Description automatically generated

Graphical user interface, text

Description automatically generated

Graphical user interface

Description automatically generated with low confidence

Text

Description automatically generated

Graphical user interface

Description automatically generated with medium confidence

# Results Comparison

|  |  |  |
| --- | --- | --- |
|  | Our results | Paper results |
| LR |  |  |
| DT |  |  |
| RF |  |  |
| SVM | Without data normalization |  |
| With data normalization |  |
|  | | |

# Links

* Paper link

<https://www.researchgate.net/publication/346617710_Breast_cancer_classification_using_machine_learning_techniques_a_comparative_study>

* Dataset Link

<https://www.kaggle.com/datasets/uciml/breast-cancer-wisconsin-data>