

A microscopic view of numerous red blood cells, which are spherical and have a reddish-pink hue. The cells are densely packed, with some in sharp focus in the foreground and others blurred in the background, creating a sense of depth. The lighting highlights the smooth, slightly reflective surface of the cells.

Breast Cancer Detection

SHOYAB SHAIKH

TANMAY KULKARNI

TEJASWINI UGALE

SHANTANU KUKADE

Aim

The goal is to classify whether the breast cancer is benign or malignant using machine learning.

Introduction

Breast cancer (BC) is one of the most common cancers among women worldwide, representing the majority of new cancer cases and cancer-related deaths according to global statistics, making it a significant public health problem in today's society.

Recommended Screening Guidelines:

Mammography. The most important screening test for breast cancer is the mammogram. A mammogram is an X-ray of the breast. It can detect breast cancer up to two years before the tumor can be felt by you or your doctor.

Dataset

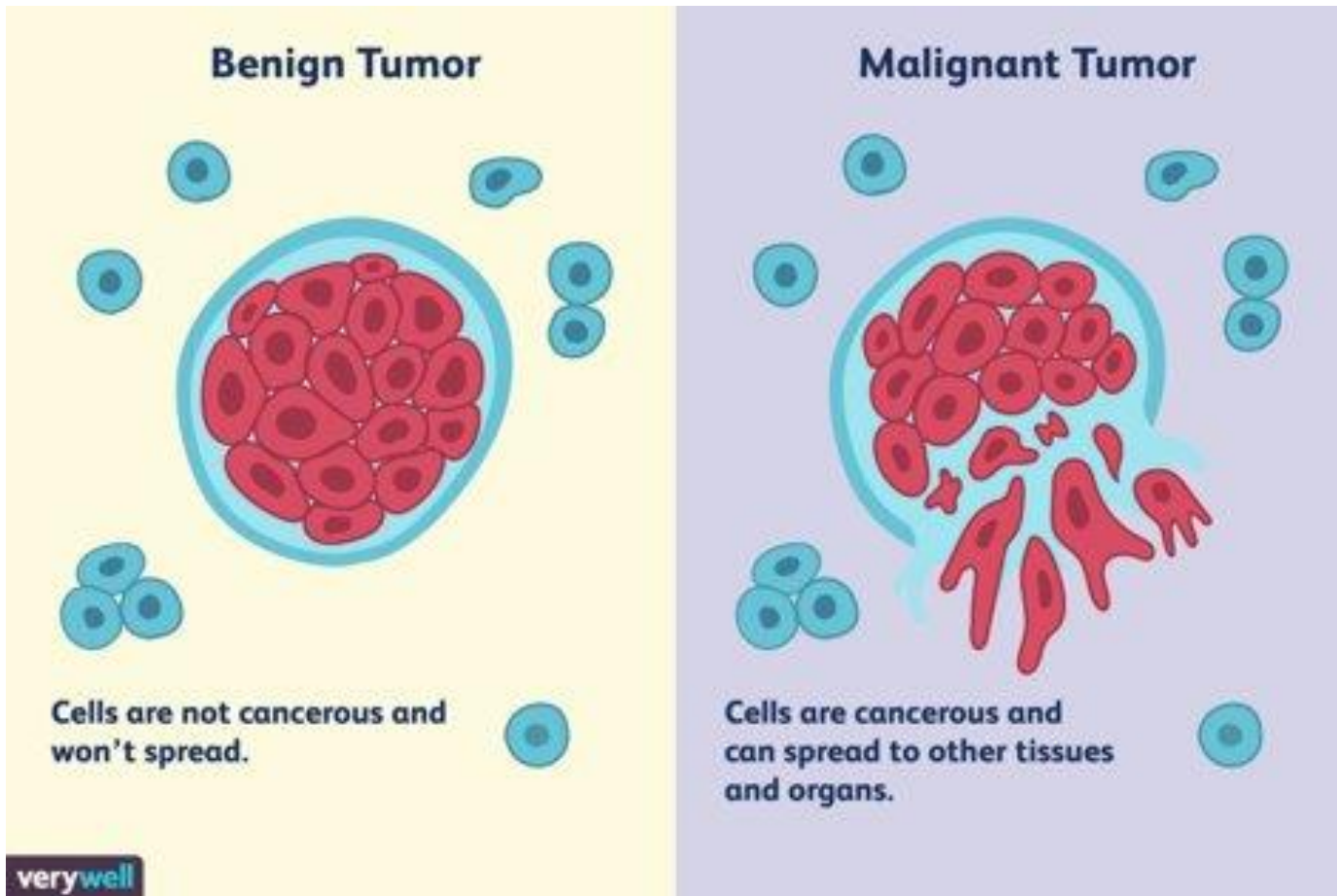
We used the UCI Machine Learning Repository for breast cancer dataset.

Attribute Information

1. ID number
 2. Diagnosis (M = malignant, B = benign)
 3. Radius_mean
 4. Texture_mean
 5. .
 6. .
- 32 fractal_dimension_worst

Dependent Variable

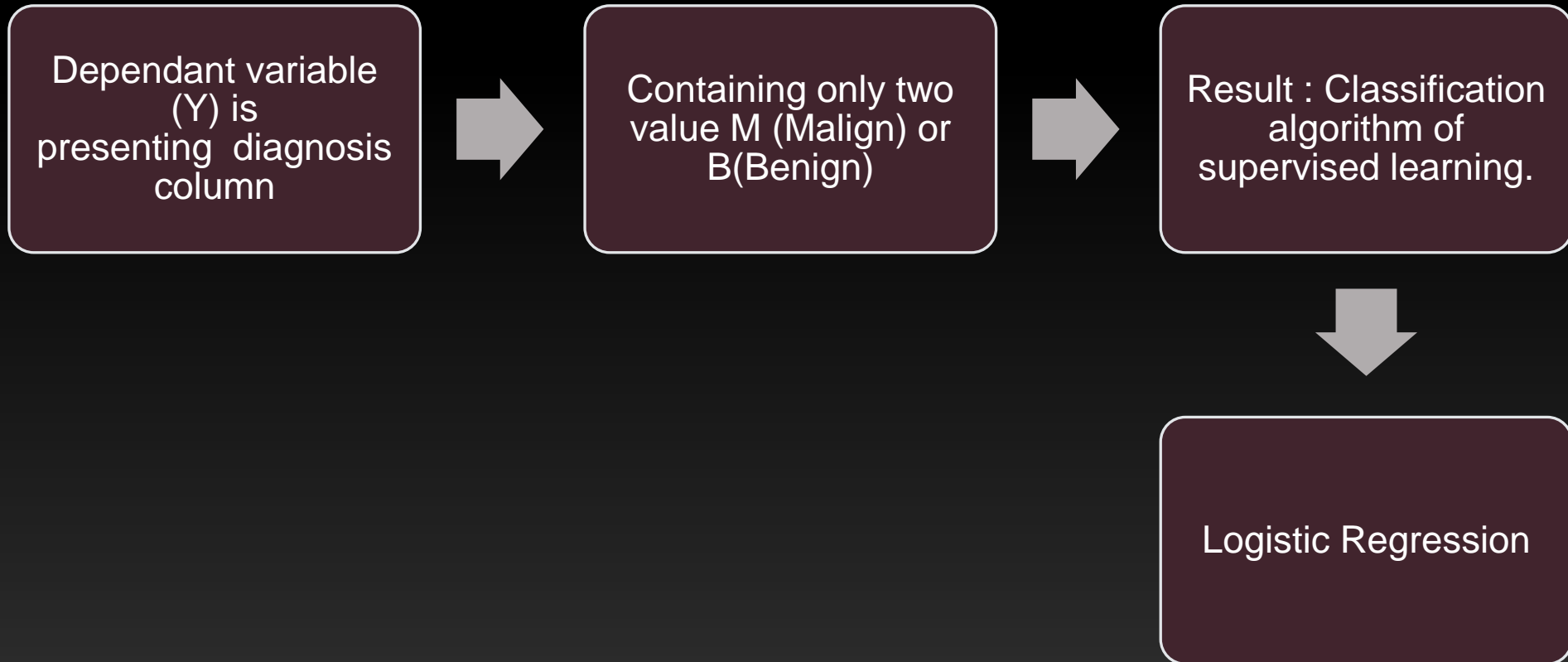
DIAGNOSIS COLUMN



Real-valued features

1. radius (mean of distances from center to points on the perimeter)
2. texture (standard deviation of gray-scale values)
3. perimeter
4. area
5. smoothness (local variation in radius lengths)
6. compactness ($\text{perimeter}^2 / \text{area} - 1.0$)
7. concavity (severity of concave portions of the contour)
8. concave points (number of concave portions of the contour)
9. symmetry
10. fractal dimension ("coastline approximation" — 1)

Model Selection



1. '*Diagnosis*' is the column which we are going to predict , which says if the cancer is M = malignant or B = benign.
2. We can identify that out of the 569 persons, 357 are labeled as B (benign) and 212 as M (malignant).
3. We converted labels into numeric data i.e. M = 1 and B=0

Code Explanation

Result and Analysis

Cancer Predictor

Enter Name Of Patient

Enter Your Name



Alexa

SUBMIT ➤

0
0

1
1

2
2

3
3

4
4

5
5

6

7

8

9

10

11

Enter Name Of Patient

Enter Your Name



Alexa

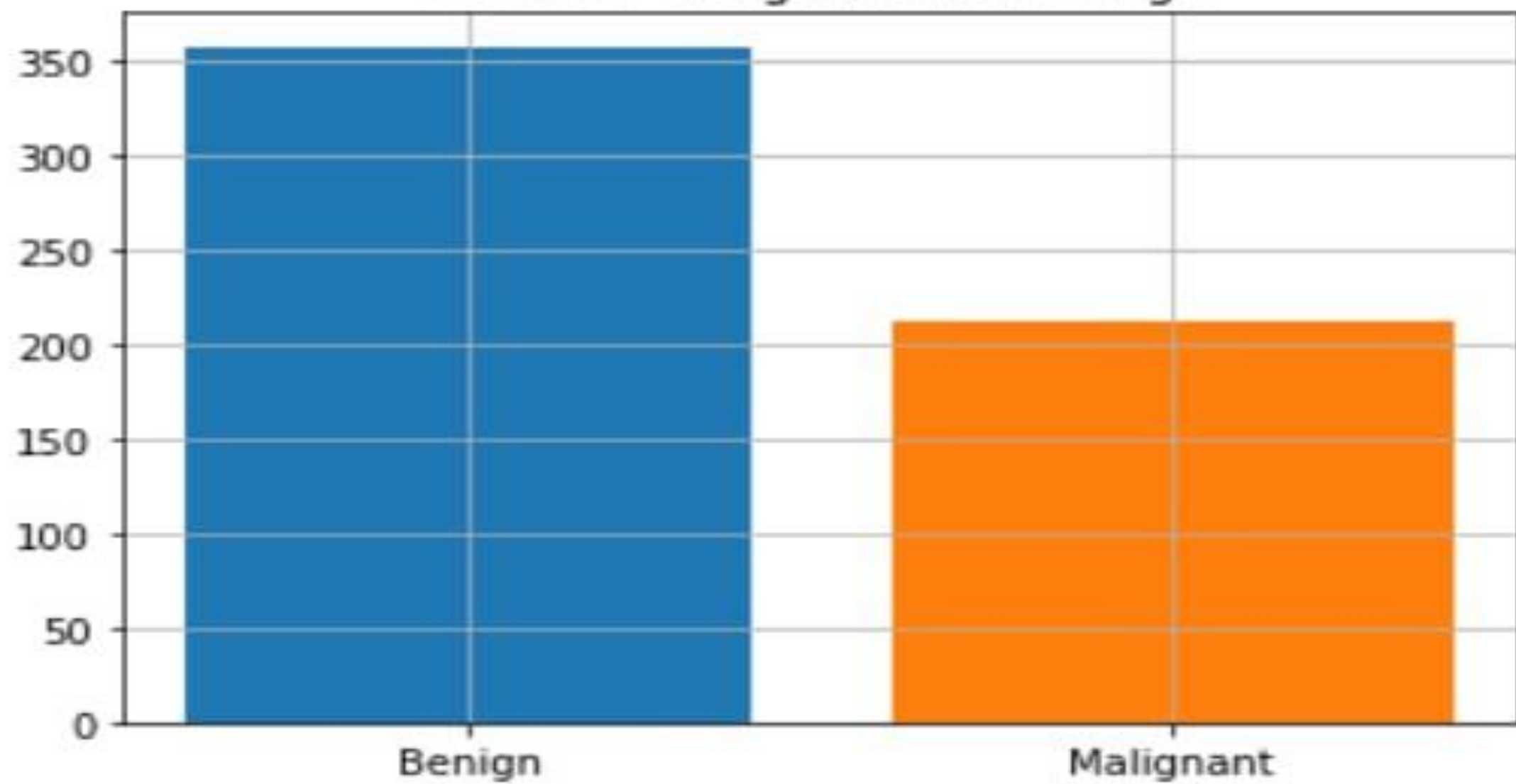
SUBMIT ➤

Result is Positive

The Patient : Alexa

Have 95% chance of Breast cancer
The Treatment should start Urgently

Count of Malignant and Benign



References

<https://www.verywellhealth.com/>

<https://archive.ics.uci.edu>

UCI Machine Learning Repository,
<http://archive.ics.uci.edu/ml/> (5-15) several
classification algo.

McCarthy et al. Applications of Machine
Learning and High -Dimensional
Visualization in Cancer Detection,
Diagnosis, and Management.

Conclusion

Results conclude that Simple Logistic regression method obtains the Best Model to predict breast cancer by means of different data mining techniques. Results indicate that Simple Logistic regression obtained best performance in general compared to the other classifiers in terms of classification accuracy.

THANK YOU

