

PROJECT

Presented by: Suresh Kumar Mishra



Greetings and welcome to Guvi.

Problem Statement: -

Breast cancer is a disease in which abnormal breast cells grow out of control and form tumours. If left unchecked, the tumours can spread throughout the body and become fatal.

Breast cancer cells begin inside the milk ducts and/or the milk-producing lobules of the breast. The earliest form (in situ) is not life-threatening. Cancer cells can spread into nearby breast tissue (invasion). This creates tumours that cause lumps or thickening.

According to World Health Organization, in 2020, there were 2.3 million women diagnosed with breast cancer and 685,000 deaths globally. As of the end of 2020, there were 7.8 million women alive who were diagnosed with breast cancer in the past 5 years, making it the world's most prevalent cancer. Breast cancer occurs in every country of the world in women at any age after puberty but with increasing rates in later life.

Type of Task: Classification (Binary)

Objective: To predict the target variable, that is, indicating whether a tumor is malignant (M) or benign (B). Malignant tumors are indicative of cancerous growth, while benign tumors are non-cancerous.

Dataset Explanation:

Features:

- 1. Radius Mean: Mean of distances from the center to points on the perimeter.
- 2. Texture Mean: Standard deviation of gray-scale values.
- 3. Perimeter Mean: Mean size of the core tumor.
- 4. Area Mean: Mean area of the core tumor.
- 5. Smoothness Mean: Mean of local variation in radius lengths.
- 6. Compactness Mean: Mean of perimeter 1/2 / area 1.0.
- 7. Concavity Mean: Mean severity of concave portions of the contour.
- 8. Concave Points Mean: Mean number of concave portions of the contour.
- 9. Symmetry Mean: Mean symmetry of tumor.
- 10. Fractal Dimension Mean: Mean "coastline approximation" 1.