DATASET DESCRIPTION

- In This Dataset we will talk about Breast Cancer.

  Features are computed from a digitized image of a fine needle aspirate (FNA) of a breast mass. They describe characteristics of the cell nuclei present in the image.

  n the 3-dimensional space. This Dataset consists of 33 Atrributes and 1000 Records. The Atrributes of The Dataset Are (ID Number, diagnosis, Radius_Mean, Texture_Mean, Perimeter_Radius, Area_Mean, Smoothness_Mean, Compactness_Mean, Concavity _Mean, Concave_Points, Symmetry _Mean, Fractal _Dimension, Radius_Se, Texture_Se, Perimetere_Se,

Area_Se, Smoothness_Se, Compactness_Se,

Concavity_Se, Concave_Se, Symmetry_Se, Fractal

_Dimension, Radius_Worst, Texture_Worst,

Perimeter_Worst, Area_Worst, Smoothness_Worst,

Compactness_Worst, Concavity_Worst,

Concave_Worst, Symmetry_Worst,

Fractal_Dimension_Worst).

- **Ten real-valued features are computed for each cell nucleus:**

a) radius (mean of distances from center to points on the perimeter)
b) texture (standard deviation of gray-scale values)

c) perimeter

d) area

e) smoothness (local variation in radius lengths)

f) compactness (perimeter^2 / area - 1.0)

g) concavity (severity of concave portions of the contour)

h) concave points (number of concave portions of the contour)

i) symmetry

j) fractal dimension ("coastline approximation" - 1).

- The mean, standard error and "worst" or largest (mean of the

  three largest values) of these features were computed for each image, resulting in 30 features. For instance,

field 3 is Mean Radius, field 13 is Radius SE, field 23 is Worst Radius.

- All feature values are recoded with four significant digits.

- Missing attribute values: none

- Class distribution of diagnosis: 357 benign, 212 malignant

- Who create This Dataset: **Dr. William H. Wolberg.**
- WE Get This Dataset From Kaggle Website.