About Breast Cancer Wisconsin (Diagnostic) Data Set:

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 is that described in: [K. P. Bennett and O. L. Mangasarian: "Robust Linear Programming Discrimination of Two Linearly Inseparable Sets", Optimization Methods and Software 1, 1992, 23-34].

This database is also available through the UW CS ftp server:
 ftp ftp.cs.wisc.edu
 cd math-prog/cpo-dataset/machine-learn/WDBC/

Also can be found on UCI Machine Learning Repository:
https://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+%28Diagnostic%29

Attribute Information:

1) ID number
 2) Diagnosis (M = malignant, B = benign)
 3-32)

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: 357 benign, 212 malignant