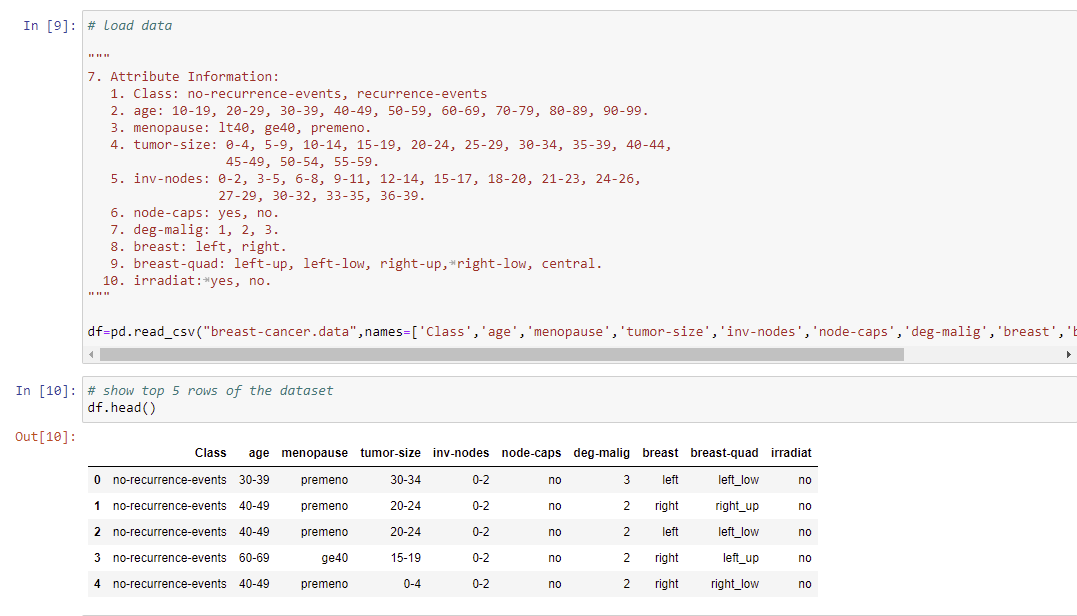
**Breast Cancer**

1. **Introduction**

The objective was to implement naïve Bayes model and predict the results by cross validation method. There are total 6 attributes in Breast Cancer dataset. The Dataset is shown below

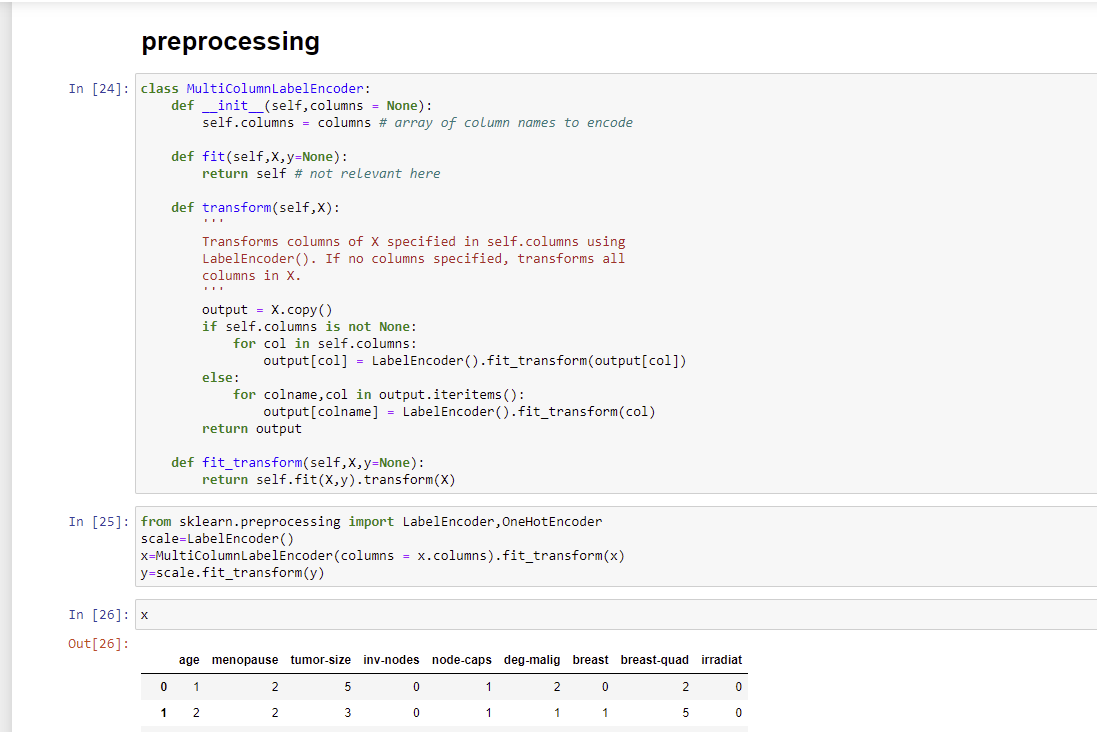


**Figure 1Dataset of** Breast Cancer

In Figure-1, the data is classified into different attributes and each attribute contain some essential data related to Breast Cancer dataset

1. **Data Pre-Processing**

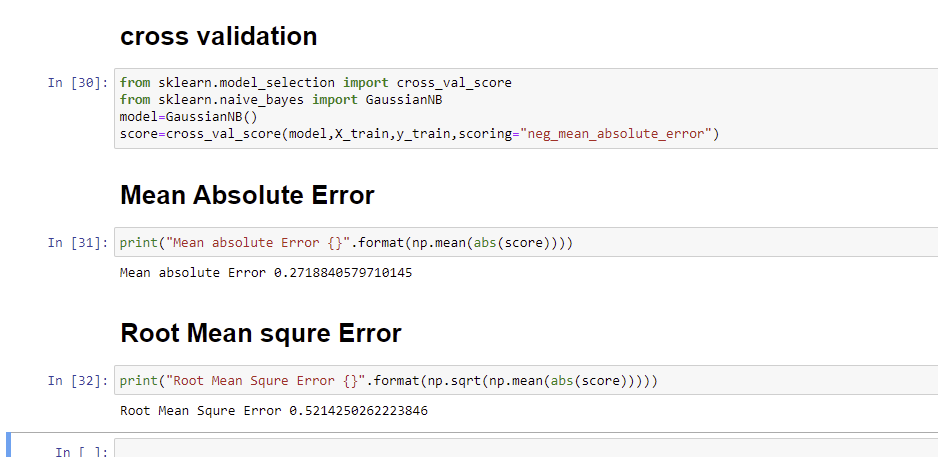
The data for Breast Cancer requires pre-processed data for implementation of Naïve Bayes model. The Python modules and libraries are required for Pre-processing. I will use feature scaling for scale the features and separate the label form the data.



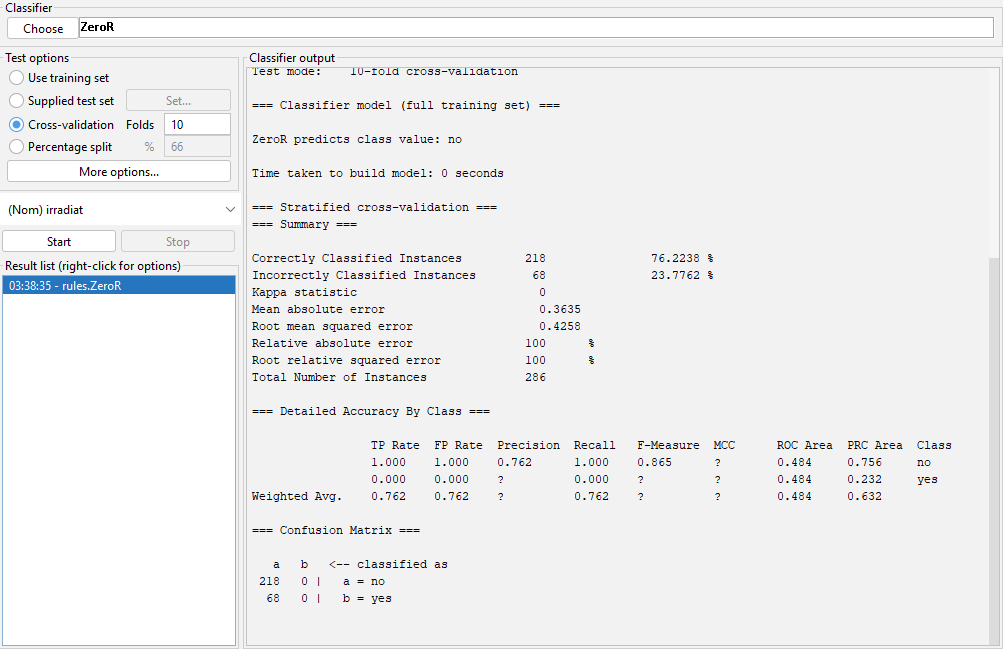
1. **K-Fold Cross Validation**

The K-folds requires cross validation for data splitting and validation. The comparison results are shown below





**WEKA Results for Naïve Bayes**



**Comparison of Results**

|  |  |  |
| --- | --- | --- |
| **Naïve Bayes** | **Python** | **WEKA** |
| **Mean absolute error** | 0.27 | 0.36 |
| **Root mean squared error** | 0.52 | 0.42 |

1. **Conclusion**

The results are compared between python and WEKA tool for cross validation to compute absolute mean error and root mean squared error. The root mean squared for python is 0.27 and WEKA is 0.36 and absolute mean error for Python is 0.52 and WEKA is 0.42