**DATA SCIENCE AND MACHINE LEARNING**

**Internship – 1**

**YBI FOUNDATION**

**Project -2**

**Cancer Prediction**

Dataset Information:

Target Variable (y):

* Diagnosis (M = malignant, B = benign)

Ten features (X) are computed for each cell nucleus:

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 (perimeter^2 / 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)

For each characteristic three measures are given:

a. Mean

b. Standard error

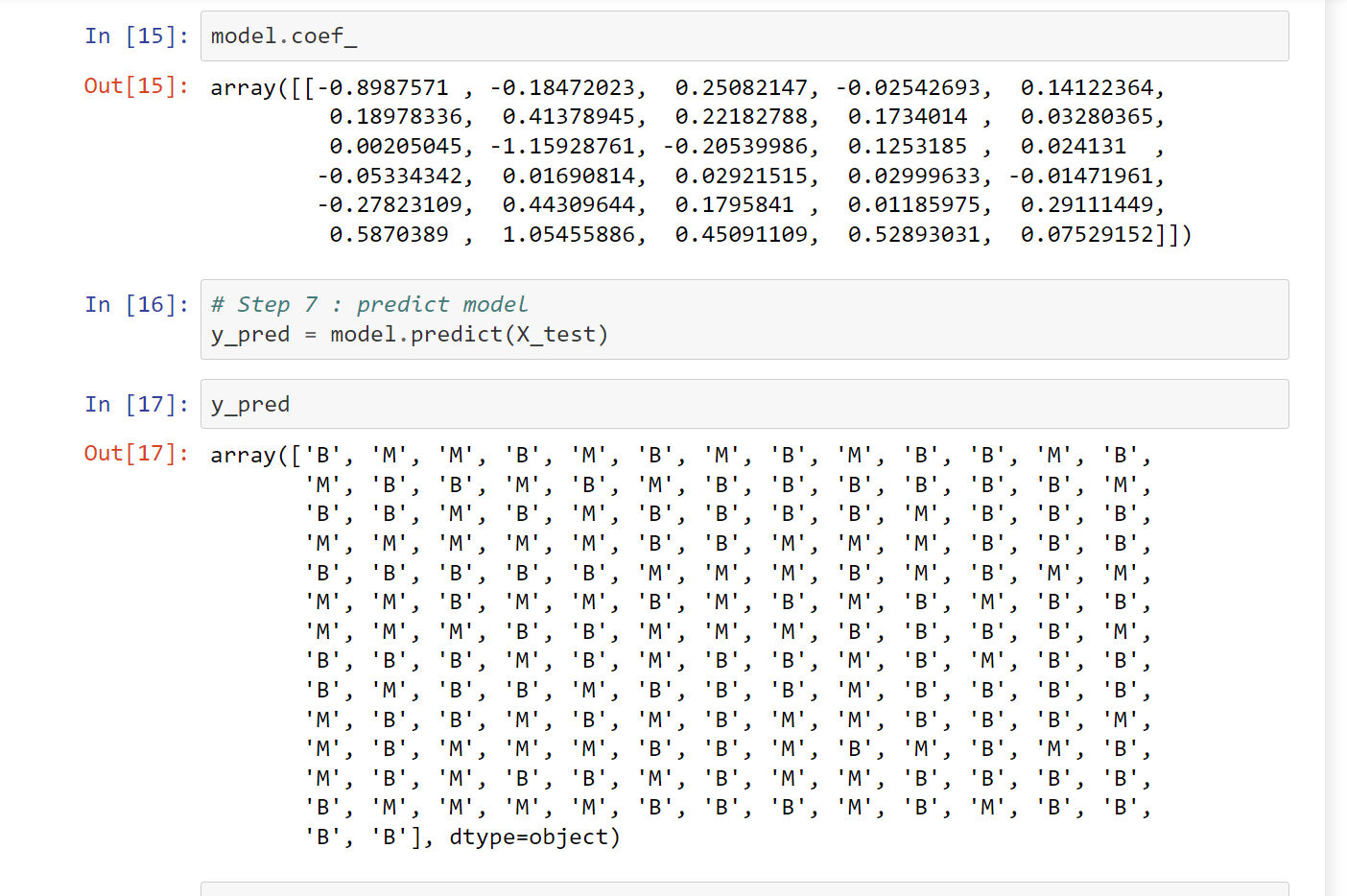
c. Largest/ Worst





A screenshot of a computer program

Description automatically generated



A screenshot of a computer program

Description automatically generated