**Assignment 1**

By Harsh Sharma

**Part III. Dimensionality Reduction**

Use PCA to reduce the dataset’s dimensionality from 30 to 2, and then generate the two-dimensional scatter plot of the breast cancer dataset using the first two principal components, where each data point is shaped and colored according to its class label, i.e., blue dot for M and red triangle for B.

Solution:

I’m sharing the scatter plot below. Also, I’ll be attaching the code as jupyter notebook named as ‘part 3 – jupyter notebook’. The dataset after reduced dimensionality is also saved as **wdbc\_2D.data** and will be attached along in the same folder.

A diagram of a graph

Description automatically generated with medium confidence