**Project Idea**

**Binary Classification Prediction for type of Breast Cancer**

### Description:

Breast cancer is the most common cancer amongst women in the world. It accounts for 25% of all cancer cases, and affected over 2.1 Million people in 2015 alone. It starts when cells in the breast begin to grow out of control. These cells usually form tumors that can be seen via X-ray or felt as lumps in the breast area.

The key challenges against it’s detection is how to classify tumors into malignant (cancerous) or benign (non cancerous).

**Dataset:** a dataset contains data from patients with breast tumors, either benign or malignant. The data contains many features and each patient has been labelled as ‘B’ (Benign) or ‘M’ (Malignant).

**The project implementation steps will go as follows:**

* Understand the dataset features (attributes)
* Apply the required data preprocessing methods
* Detect outliers in data
* Build classification models to predict whether the cancer type is Malignant or Benign and compare the evaluation metrics of various classification algorithms.
* Mention your observations and determine which factors are more prominent in deciding the type of Breast Cancer.
* Visualize the result of data in plots (discover the potential plots to describe result )