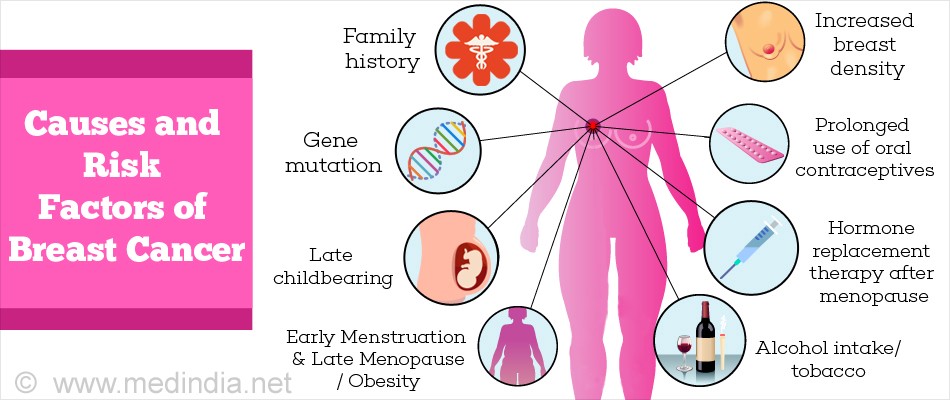
**Breast Cancer Detection Using Machine Learning**

*Breast cancer the most common cancer among women worldwide accounting for 25 percent of all cancer cases and affected 2.1 million people in 2015 early diagnosis significantly increases the chances of survival.  
The key challenge in cancer detection is how to classify tumors into malignant or benign machine learning techniques can dramatically improves the accuracy of diagnosis.*

Research indicates that most experienced physicians can diagnose cancer with 79 percent accuracy while 91 percent correct diagnosis is achieved using machine learning techniques.

*In this case study, our task is to classify tumors into malignant or benign tumors using features of pain from several cell images.*

Let’s take a look at the cancer diagnosis and classification process.  
So the first step in the cancer diagnosis process is to do what we call it final needle aspirate or if any process which is simply extracting some of the cells out of the tumor. And at that stage, we don’t know if that human is malignant or benign. When you say malignant or benign as you guys can see these are kind of the images of the this would be benign tumor and this is the malignant tumor. And when we say benign that means that the tumor is kind of not spreading across the bodies of the patient is safe somehow.  
It’s if it’s malignant That means it’s a cancerous.



That means we need to intervene and actually stop the cancer growth  
And what we do here in the machine learning aspect so now as we extracted all these images and we wanted to specify if that cancer out of these images is malignant or benign that’s the whole idea.  
So what we do with that we extract out of these images some features when we see features that mean some characteristics out of the image such as radius, for example the cells such as texture perimeter area smoothness and so on. And then we feed all these features into kind of our machine learning model in a way which is kind of a brain in a way.

The idea is to teach the machine how to basically classify images or classify data and tell us OK if it’s malignant or benign for example in this case without any human intervention which is going to change the model once the model is trained we’re good to go we can use it in practice to classify new images as we move forward. And that’s kind of the overall procedure or the cancer diagnosis procedure.

**Features Of This Project**

**Id**: ID number

**Diagnosis**: The diagnosis of breast tissues (M = malignant, B = benign)

**Radius mean**: mean of distances from center to points on the perimeter

**Texture mean**: standard deviation of gray-scale values

**Perimeter mean**: mean size of the core tumor

**Area mean**: mean area of the breast

**Smoothness mean**: mean of local variation in radius lengths

**Compactness mean**: mean of perimeter^2 / area - 1.0

**Concavity mean**: mean of severity of concave portions of the contour

**Concave points mean**: mean for number of concave portions of the contour

**Symmetry mean:** mean of symmetry.

**Fractal dimension mean**: mean for "coastline approximation" – 1

**Radius se**: standard error for the mean of distances from center to points on the perimeter

**Texture se**: standard error for standard deviation of gray-scale values

**Perimeter se**: standard error for the size of core tumor.

**Area se**: standard error for the area of the breast.

**Smoothness se**: standard error for local variation in radius lengths

**Compactness se**: standard error for perimeter^2 / area - 1.0

**Concavity se**: standard error for severity of concave portions of the contour

**Concave points se**: standard error for number of concave portions of the contour

**Symmetry se**: standard error for symmetry.

**Fractal dimension se**: standard error for "coastline approximation" – 1

**Radius worst**: "worst" or largest mean value for mean of distances from center to points on the perimeter

**Texture worst**: "worst" or largest mean value for standard deviation of gray-scale values

**Perimeter worst**: “worst” or largest mean value for the size of core tumor.

**Area worst**: “worst” or largest mean value for the area of the breast.

**Smoothness worst**: "worst" or largest mean value for local variation in radius lengths

**Compactness worst**: "worst" or largest mean value for perimeter^2 / area - 1.0

**Concavity worst**: "worst" or largest mean value for severity of concave portions of the contour

**Concave points worst**: "worst" or largest mean value for number of concave portions of the contour

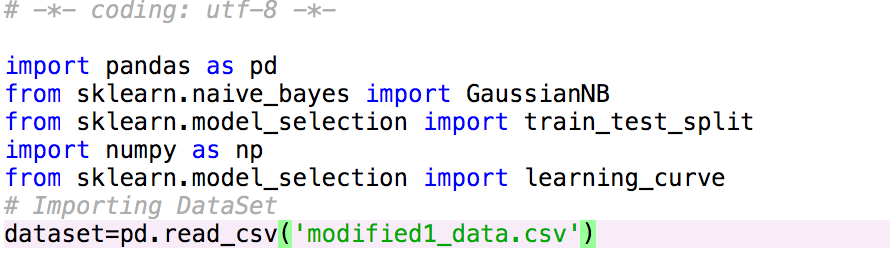
**Symmetry worst**: "worst" or largest mean value for symmetry mean.

**Fractal dimension worst:** "worst" or largest mean value for "coastline approximation" – 1

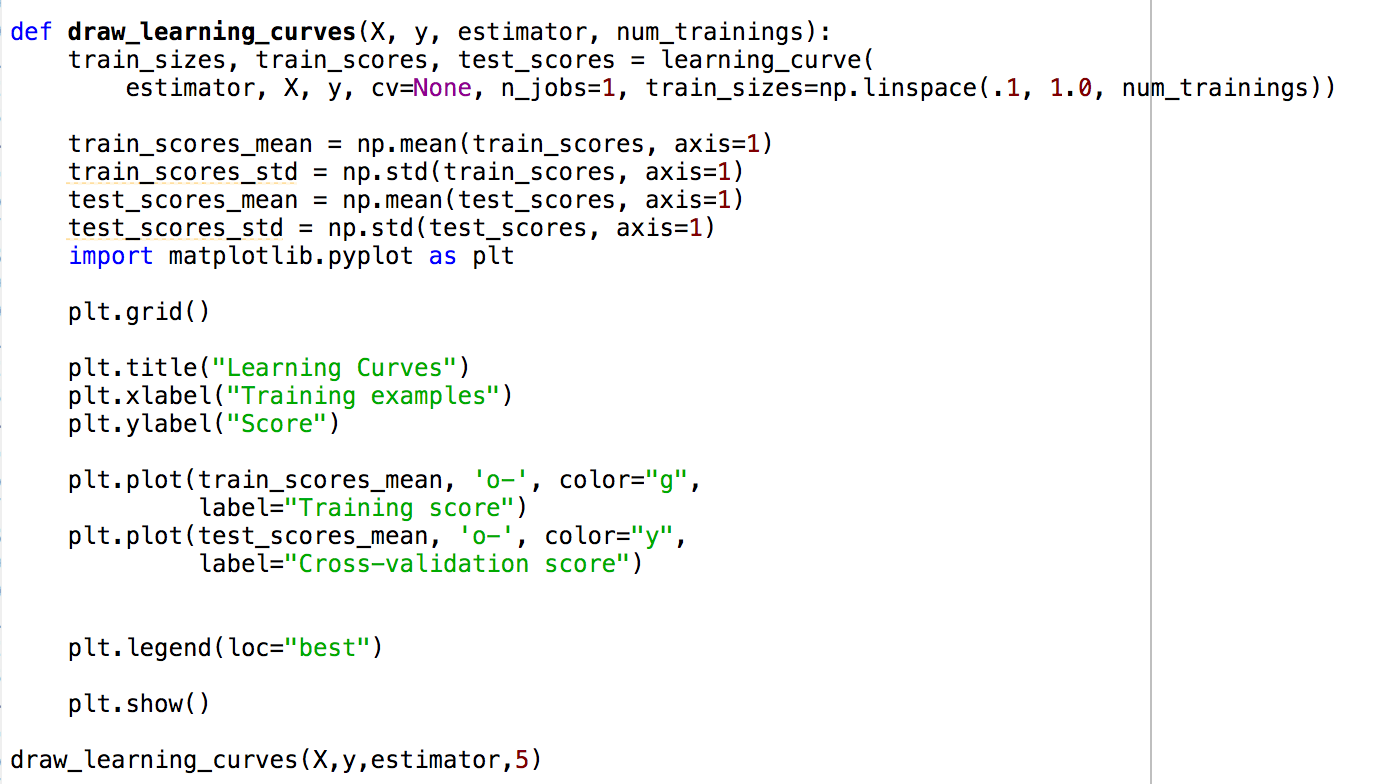
**STEP #1: PROBLEM STATEMENT**

* Predicting if the cancer diagnosis is benign or malignant based on several observations/features
* 30 features are used, examples:
* - radius (mean of distances from center to points on the perimeter)
* - texture (standard deviation of gray-scale values) - perimeter
* - area - smoothness (local variation in radius lengths)
* - compactness (perimeter^2 / area - 1.0)
* - concavity (severity of concave portions of the contour)
* - concave points (number of concave portions of the contour)
* - symmetry
* - fractal dimension ("coastline approximation" - 1)
* Datasets are linearly separable using all 30 input features
* Number of Instances: 569
* Class Distribution: 212 Malignant, 357 Benign
* Target class:
* - Malignant – Benign

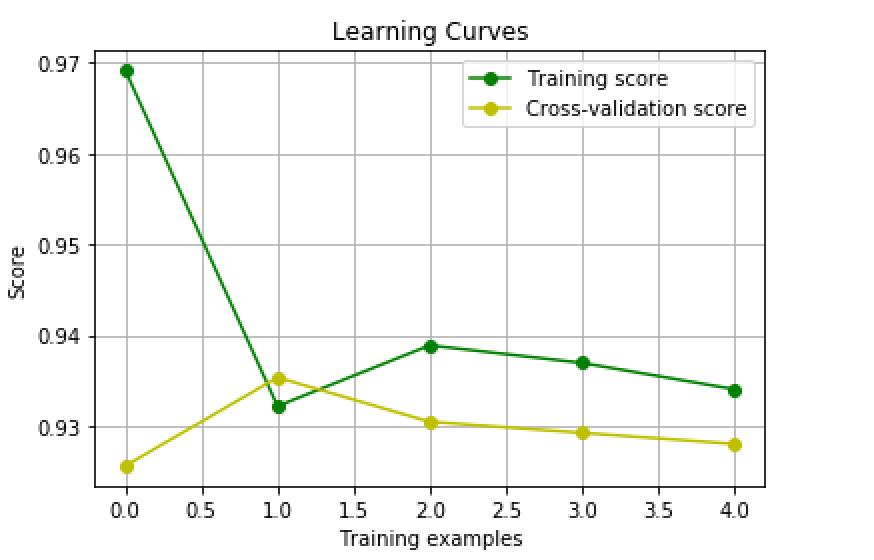
### STEP #2: IMPORTING DATA



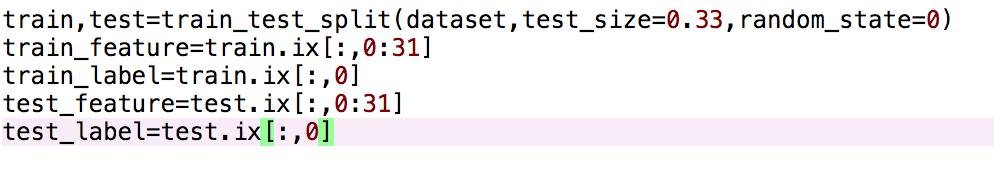
**STEP #3: Draw Curve for Cross-Validation**



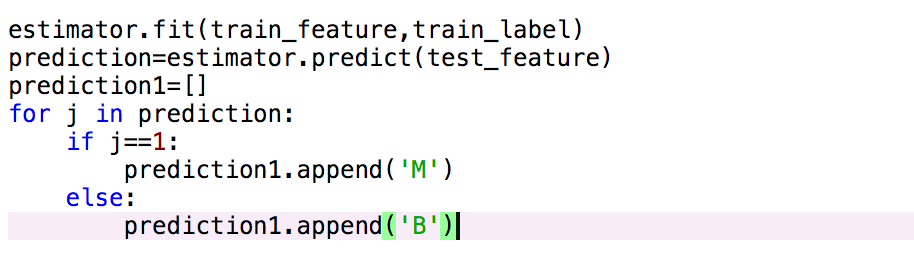
Step #4: Curve Is Showing Below



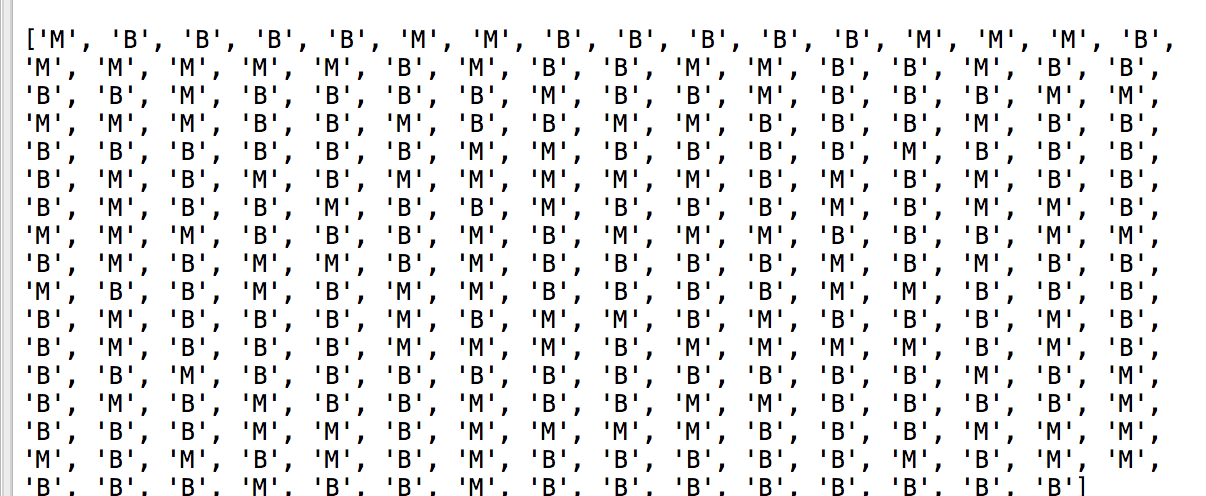
Step #5: Train and Test Set



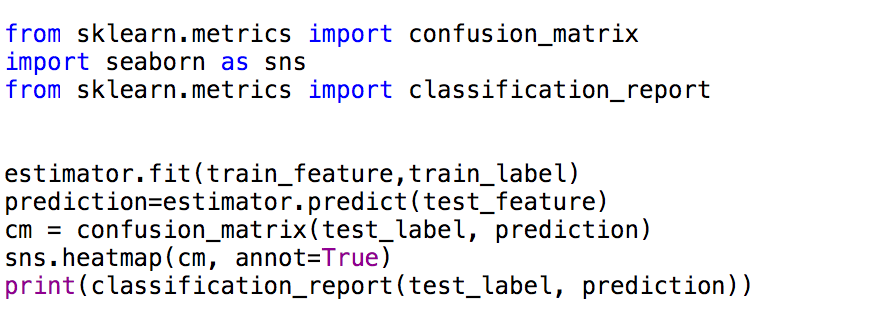
Step #6: Get Prediction From The Test Set

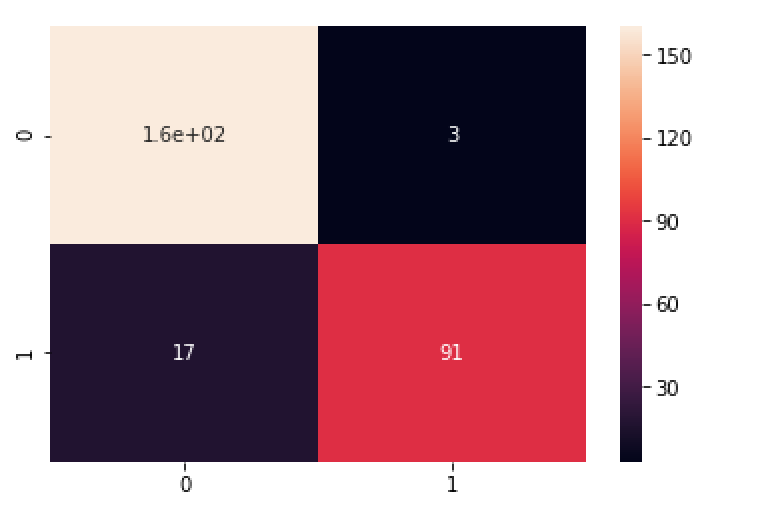


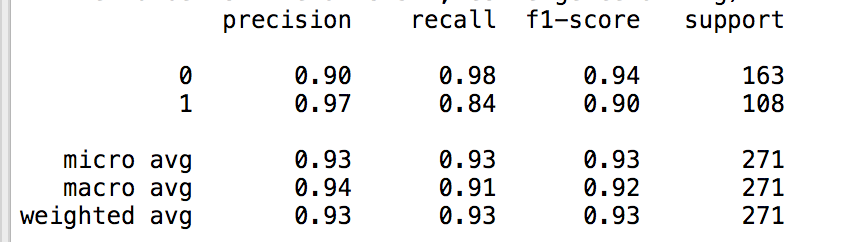
Step #7: The Predicted Value



Step #9: Confusion Matrix







Step #8: Conclusion

This project helps us to detect breast cancer on the basis of following given features which have particular maximum and minimum values stored in dataset.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.It needs to be linked to an early detection programme so that cases are detected at an early stage, when treatment is more effective and there is a greater chance of cure. It also needs to be integrated with a palliative care programme, so that patients with advanced cancers, who can no longer benefit from treatment .Identifying and studying these subtypes has potential in planning more effective treatment optand developing new therapies.