COURSE NAME: ARTIFICIAL INTELLIGENCE
COURSE CODE: CIS 412

SYED TANGIM PASHA

LECTURER,

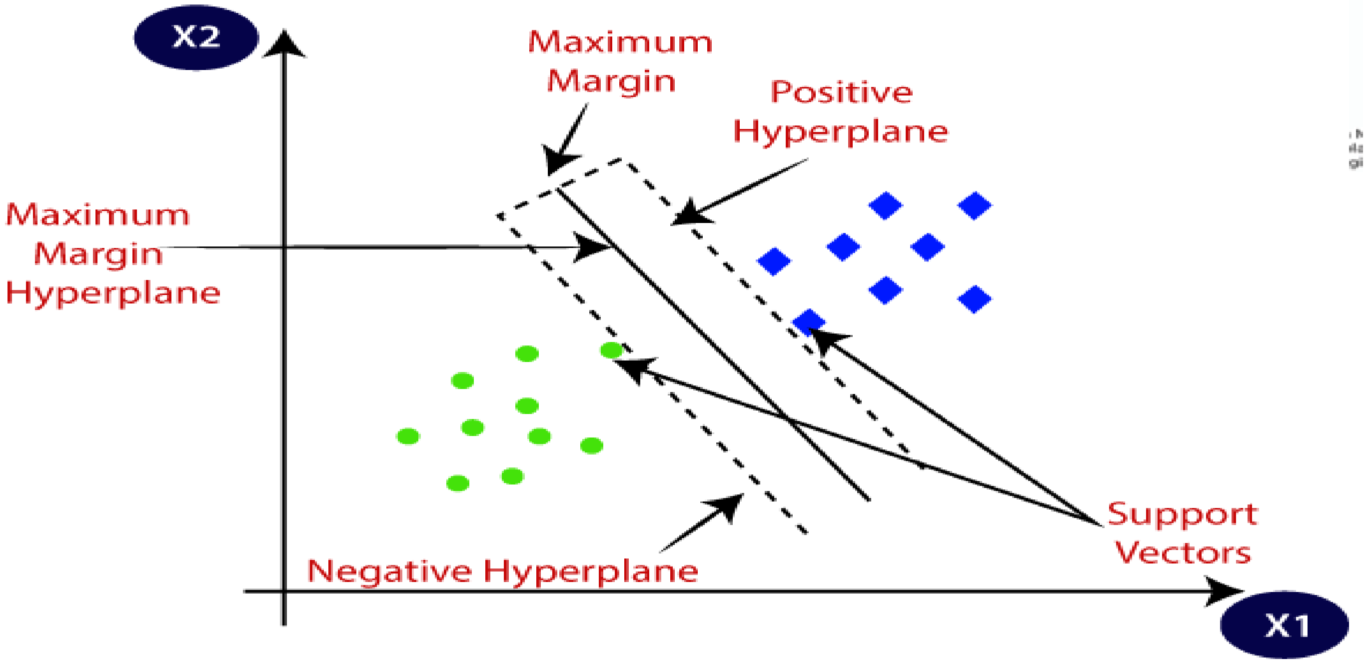
DEPARTMENT OF COMPUTING AND INFORMATION SYSTEM (CIS)

DAFFODIL INTERNATIONAL UNIVERSITY (DIU)

DHAKA,BANGLADESH

SUPPORT VECTOR MACHINE (SVM)

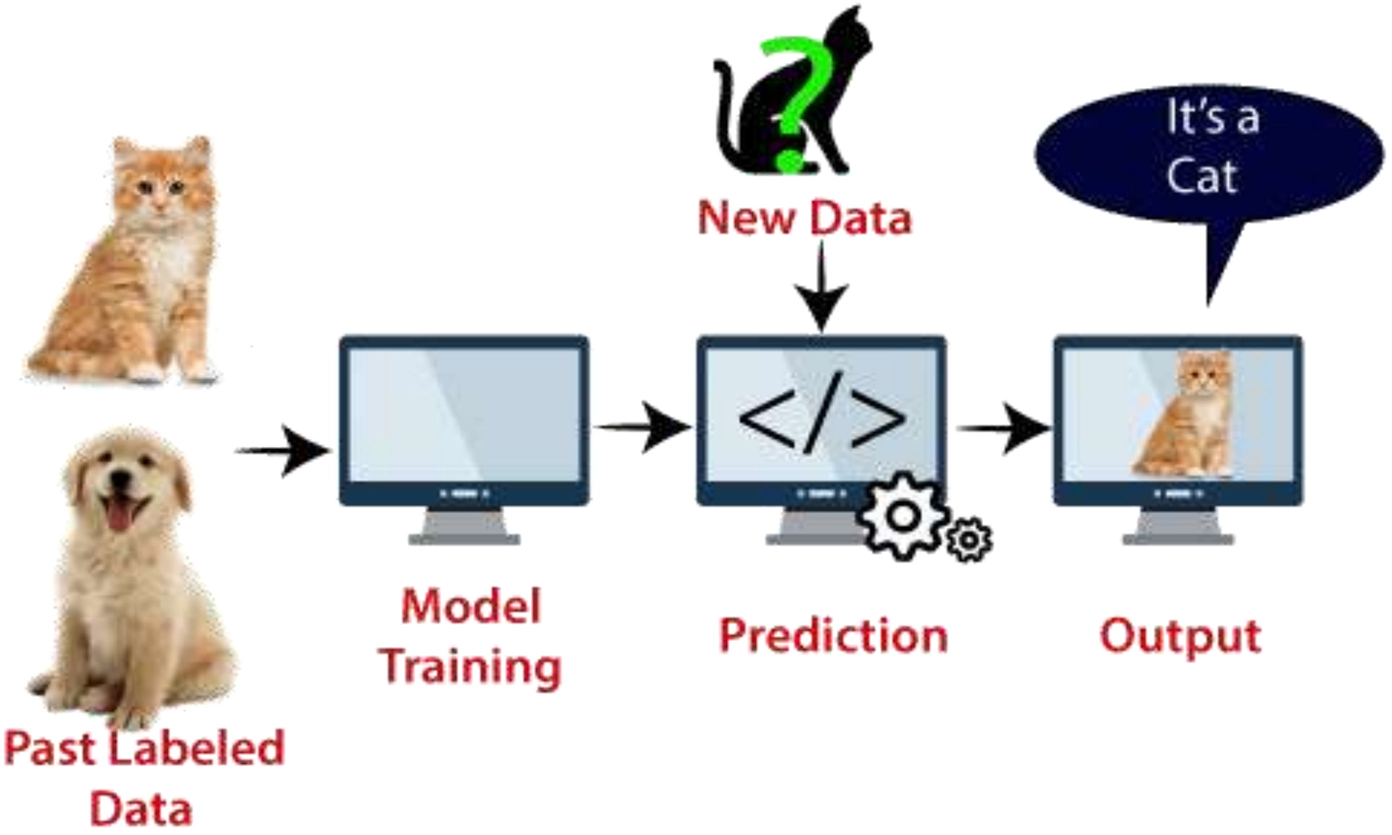
- **SVM:** Support Vector Machine or SVM is one of the most popular Supervised Learning algorithms, which is used for classification task.
- The goal of the SVM algorithm is to create the best line or decision boundary that can segregate different classes value and the best decision boundary is called a hyperplane.
- **SVM chooses the extreme points/vectors that help in creating the hyperplane. These extreme cases are called as support vectors, and hence algorithm is termed as Support Vector Machine.**



SUPPORT VECTOR MACHINE (SVM)

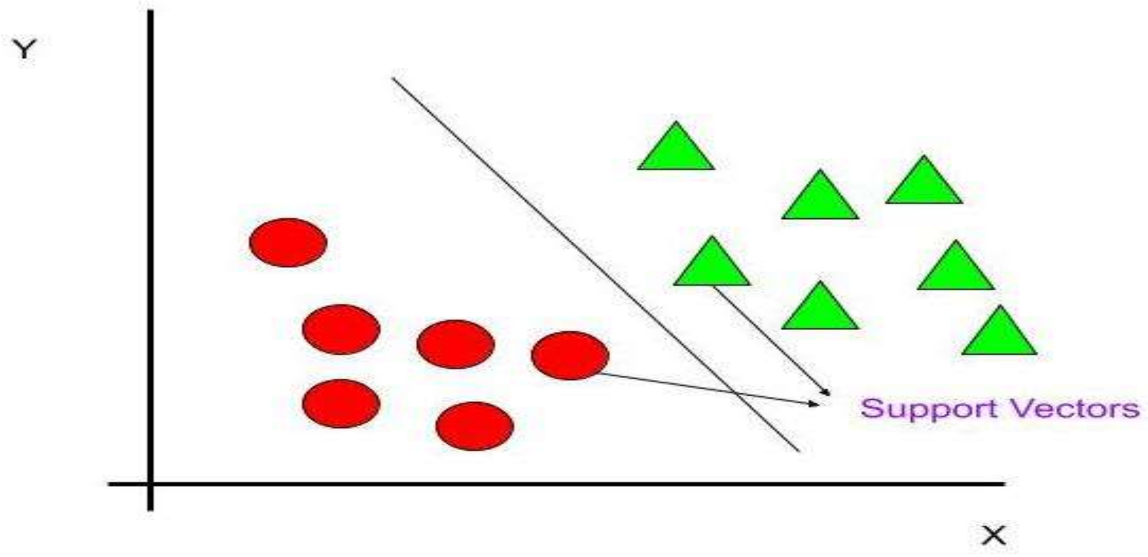
- **Example:** SVM can be understood with the example that we have used in the KNN classifier. Suppose we see a strange cat that also has some features of dogs, so if we want a model that can accurately identify whether it is a cat or dog, so such a model can be created by using the SVM algorithm. We will first train our model with lots of images of cats and dogs so that it can learn about different features of cats and dogs, and then we test it with this strange creature. So as support vector creates a decision boundary between these two data (cat and dog) and choose extreme cases (support vectors), it will see the extreme case of cat and dog. On the basis of the support vectors, it will classify it as a cat.

SUPPORT VECTOR MACHINE (SVM)



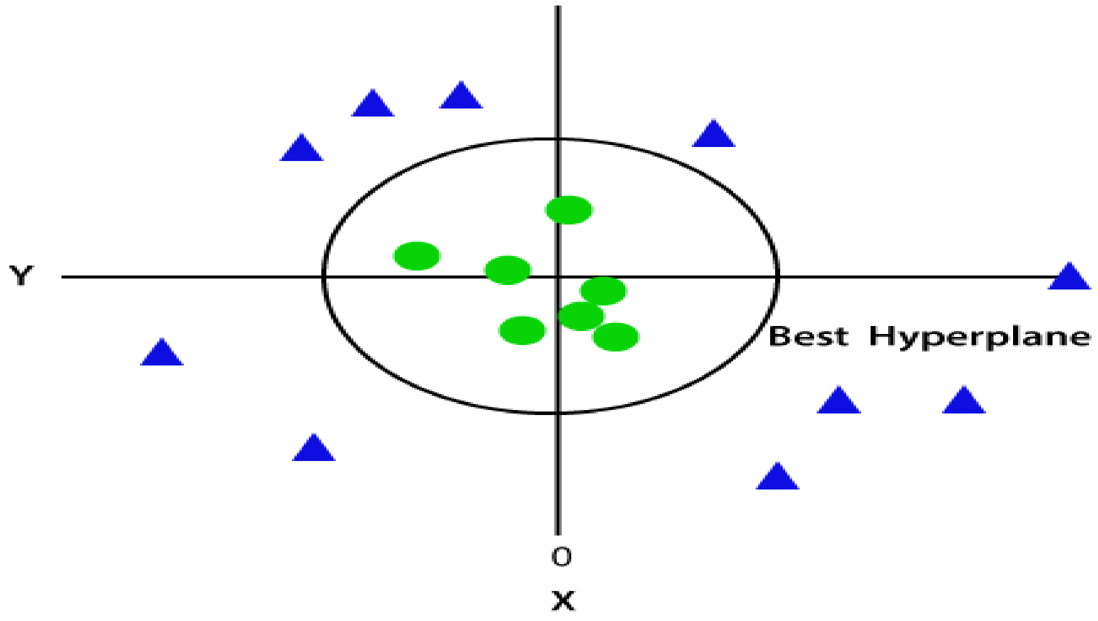
TYPES OF SVM

- SVM ca be of two types:
- **Linear SVM:** Linear SVM is used for linearly separable data, which means if a dataset can be classified into two classes by using a single straight line, then such data is termed as linearly separable data, and classifier is called as Linear SVM classifier.

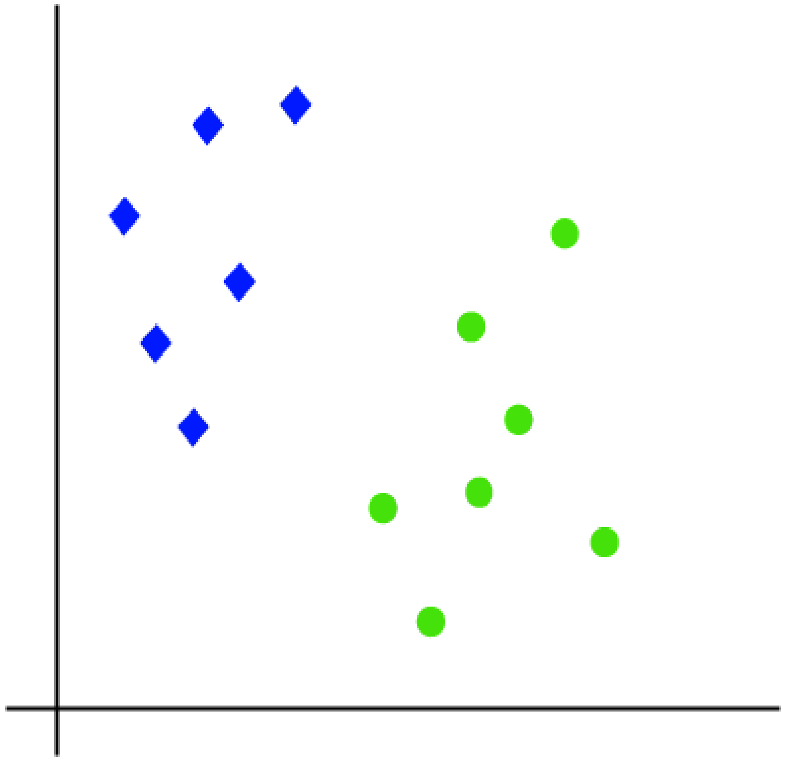


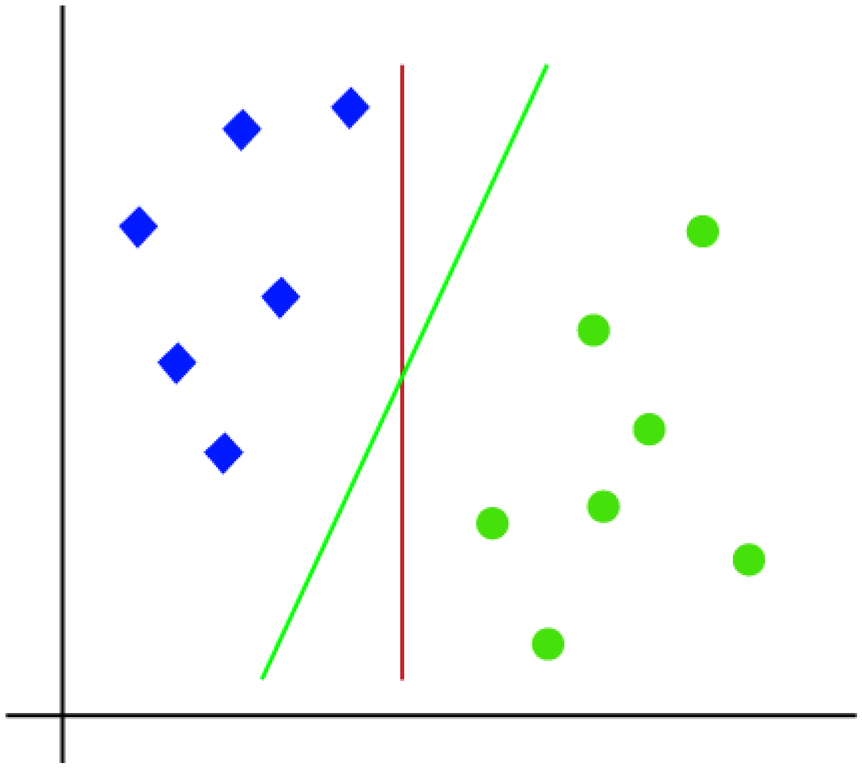
TYPES OF SVM

- SVM ca be of two types:
- **Non-Linear SVM:** Non-Linear SVM is used for non-linearly separated data, which means if a dataset cannot be classified by using a straight line, then such data is termed as non-linear data and classifier is called as Non-Linear SVM classifier.



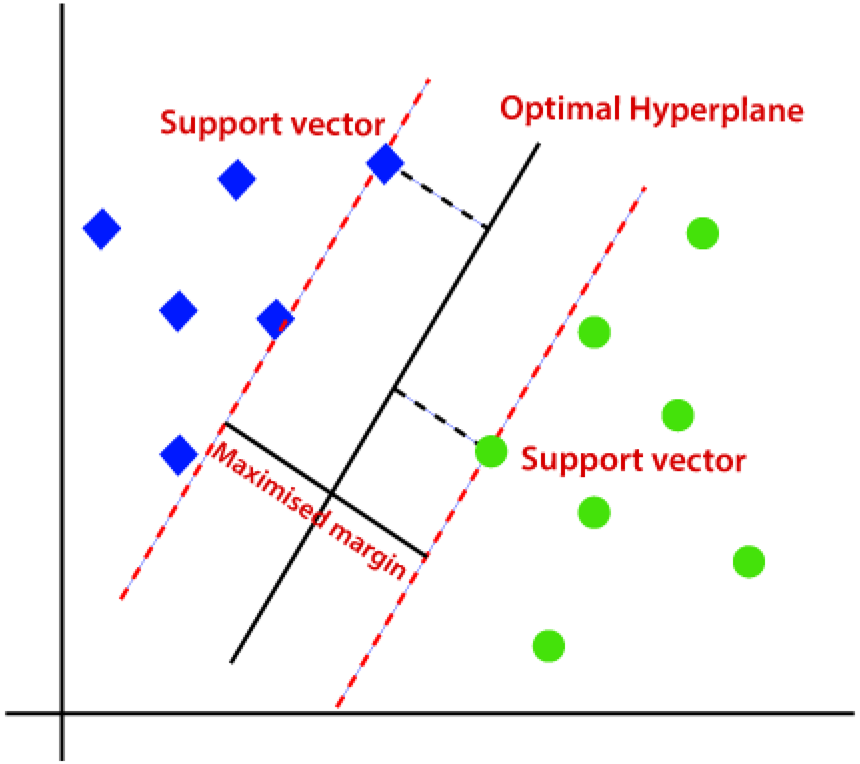
LINEAR SVM





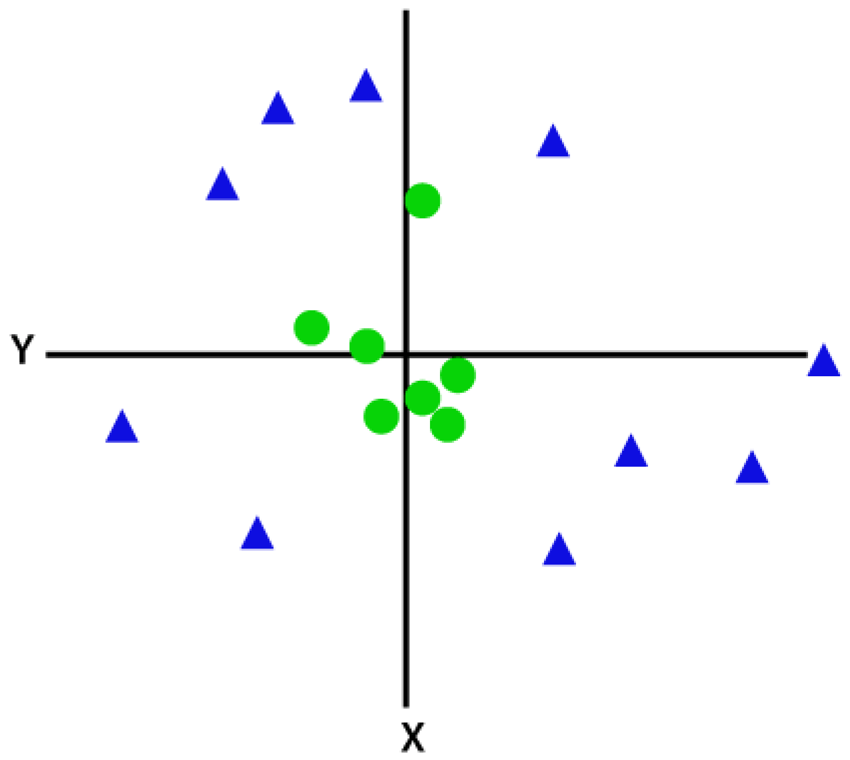
LINEAR SVM

- The distance between the vectors and the hyperplane is called as margin. The goal of SVM is to maximize the margin. The hyperplane with maximum margin is called the optimal hyperplane.



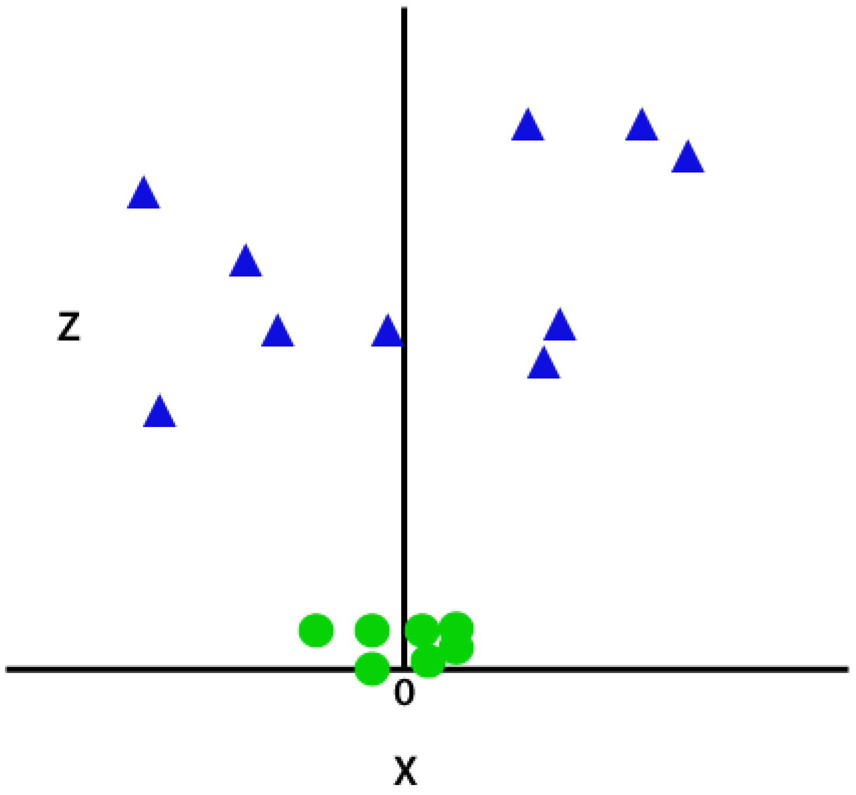
NON-LINEAR SVM

- If data is linearly arranged, then we can separate it by using a straight line, but for non-linear data, we cannot draw a straight line.



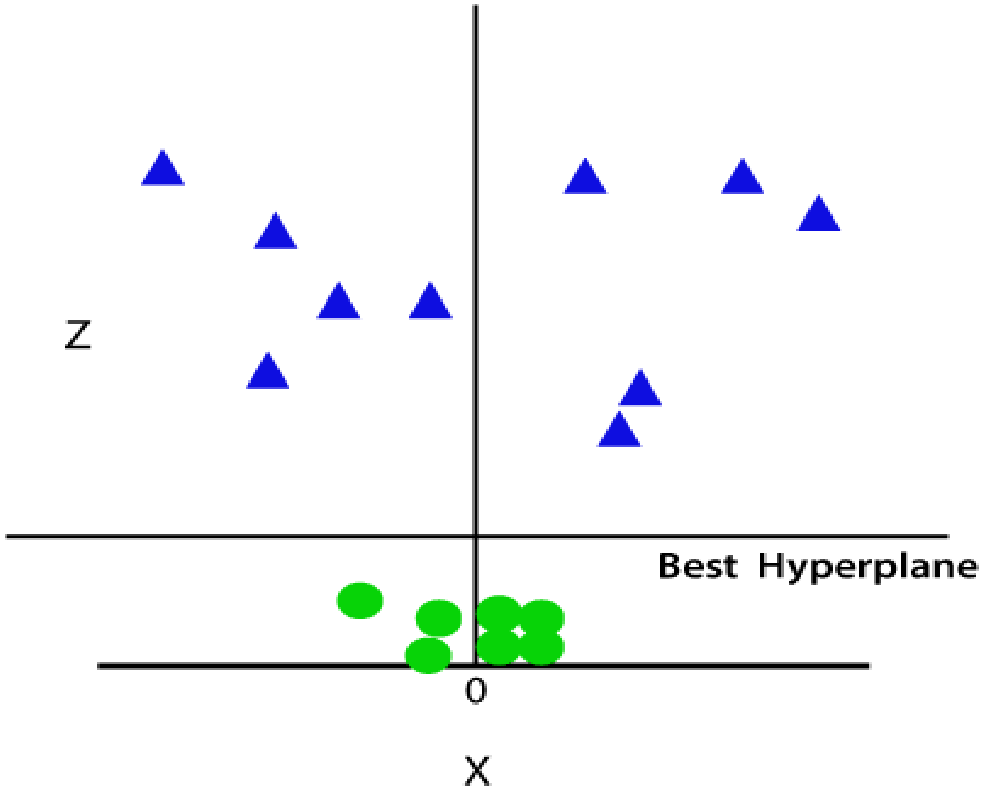
NON-LINEAR SVM

- So to separate these data points, we need to add one more dimension. For linear data, we have used 2 dimensions x and y, so for non-linear data, we will add a 3rd dimension Z. it can be calculated as, $Z = x^2 + y^2$.



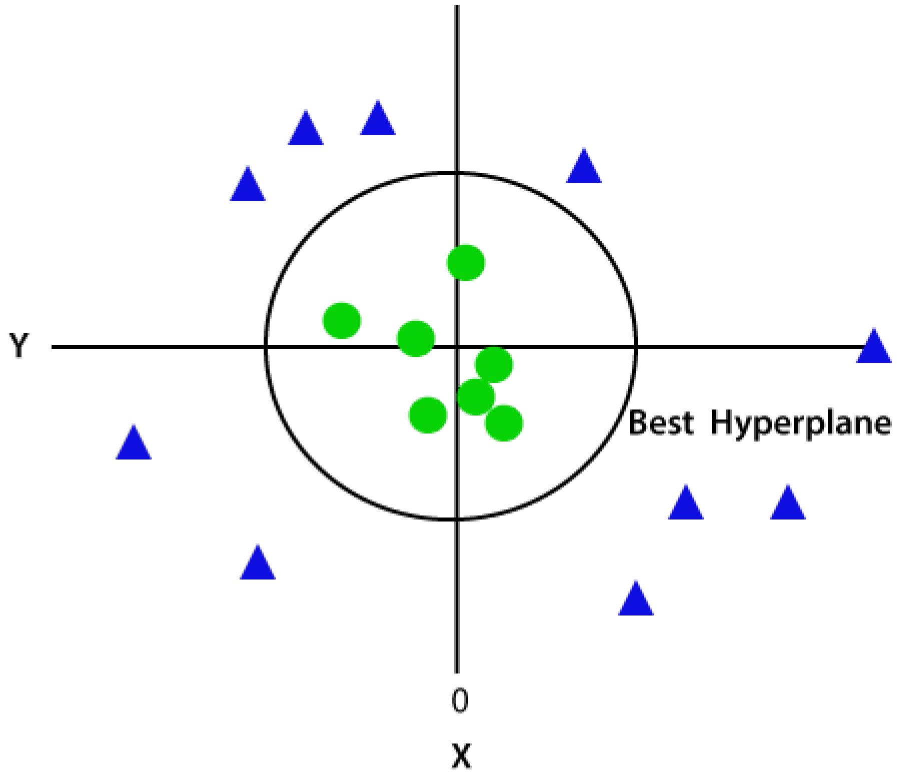
NON-LINEAR SVM

- So now, SVM will divide the datasets into classes in the following way. Consider the below image:

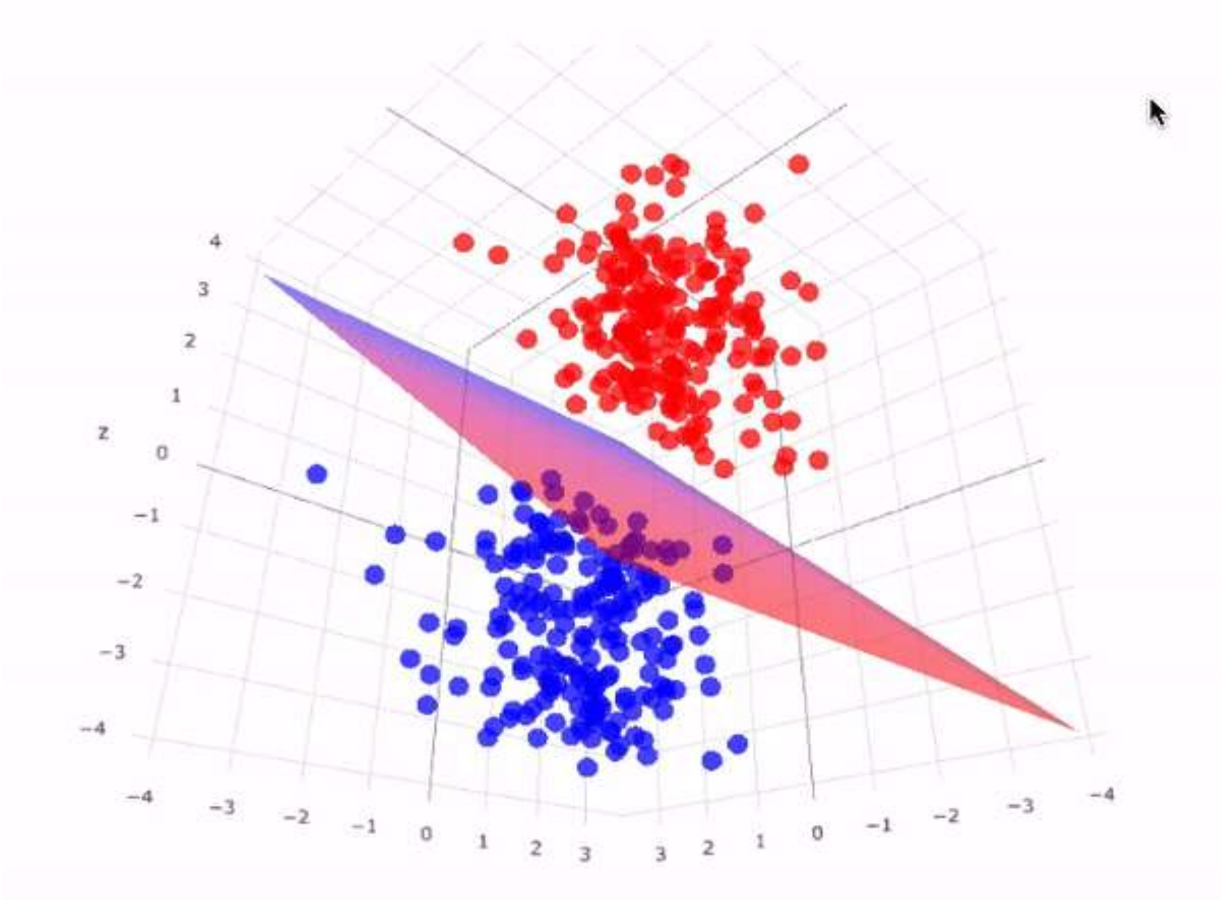


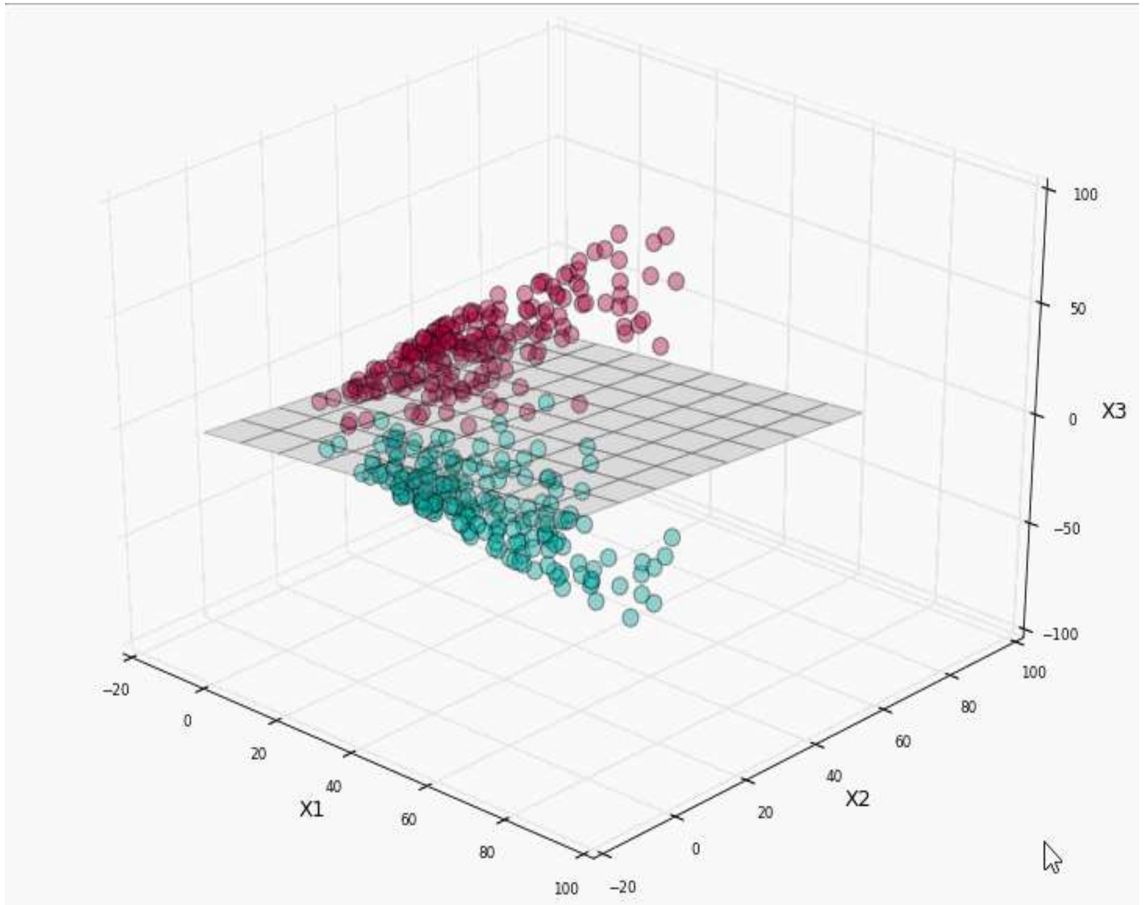
NON-LINEAR SVM

- Since we are in 3-d space, hence it is looking like a plane parallel to the x-axis. If we convert it in 2d space with z=1, then it will become as:



3D SPACE FOR NON-LINEAR DATA





COST FUNCTION

$$\underbrace{min}_{w} \frac{1}{m}\left[\sum_{i=1}^{m} \{y^{(i)}(Cost_1(z)) + \left(1 - y^{(i)}\right)(Cost_0(z))\}\right] + \frac{\lambda}{2m}\sum_{j=1}^{n} w_j^{\,2}$$

LOGISTIC REGRESSION COST FUNCTION

$$\underbrace{min}_{w}\left[\sum_{i=1}^{m} \{y^{(i)}\left(Cost_1(z)\right) + \left(1 - y^{(i)}\right)\left(Cost_0(z)\right)\}\right] + \frac{\lambda}{2}\sum_{j=1}^{n} w_j^{\,2}$$

REMOVE (m) from both sides

COST FUNCTION

$$\underbrace{min}_{w}\left[\sum_{i=1}^{m}\{y^{(i)}(Cost_1(z)) + (1-y^{(i)})(Cost_0(z))\}\right] + \frac{\lambda}{2}\sum_{j=1}^{n} {w_j}^2$$

$$A + \lambda B$$

$$A = \underbrace{min}_{w}\left[\sum_{i=1}^{m}\{y^{(i)}(Cost_1(z)) + (1-y^{(i)})(Cost_0(z))\}\right]$$

$$B = \frac{1}{2}\sum_{j=1}^{n} {w_j}^2$$

COST FUNCTION

$$CA + B \text{ যেখানে, } C = \frac{1}{\lambda}$$

$$\underbrace{min}_{w}\ C\left[\sum_{i=1}^{m}\{y^{(i)}(Cost_1(z)) + (1 - y^{(i)})(Cost_0(z))\}\right] + \frac{1}{2}\sum_{j=1}^{n} {w_j}^2$$

SVM COST FUNCTION