- **Support Vector Machine Algorithm**

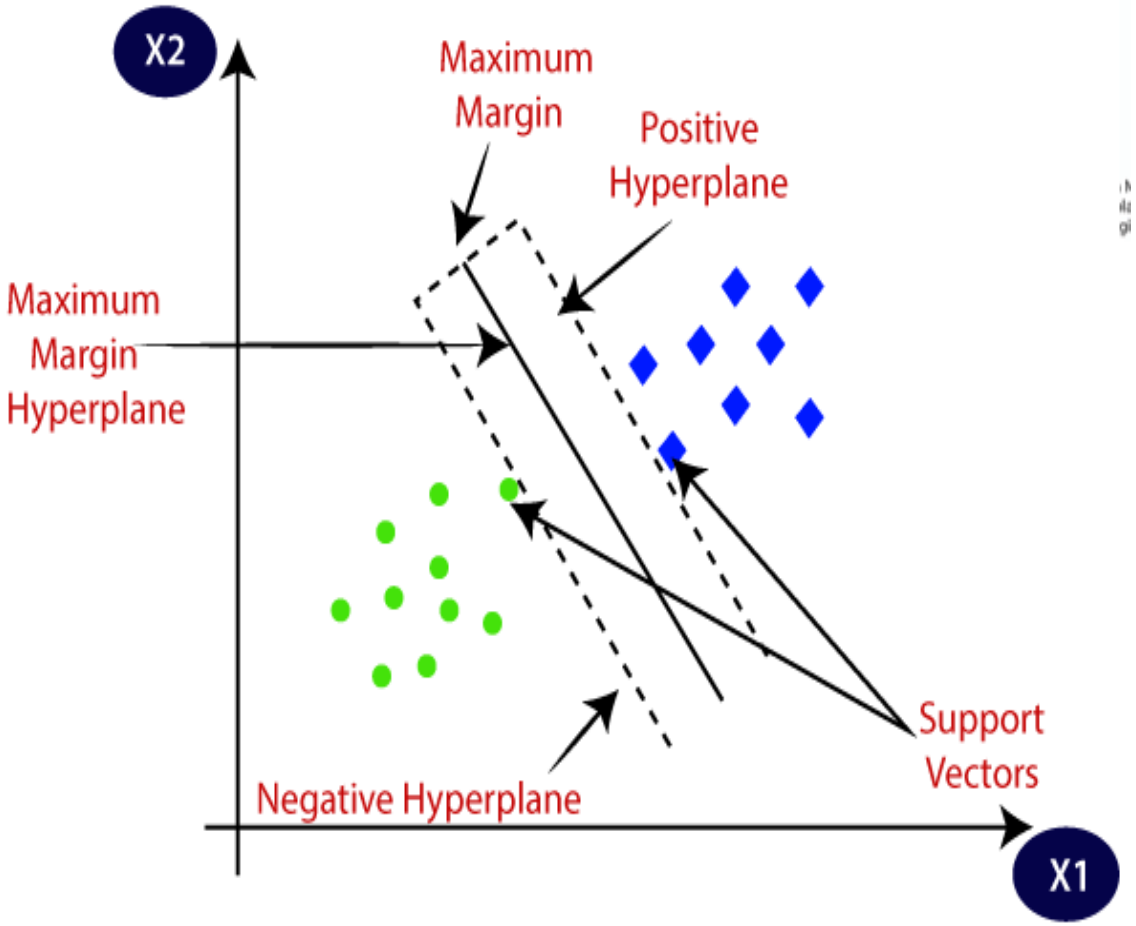
Machine Learning

Dr. Jagendra Singh

SUPPORT VECTOR MACHINE ALGORITHM

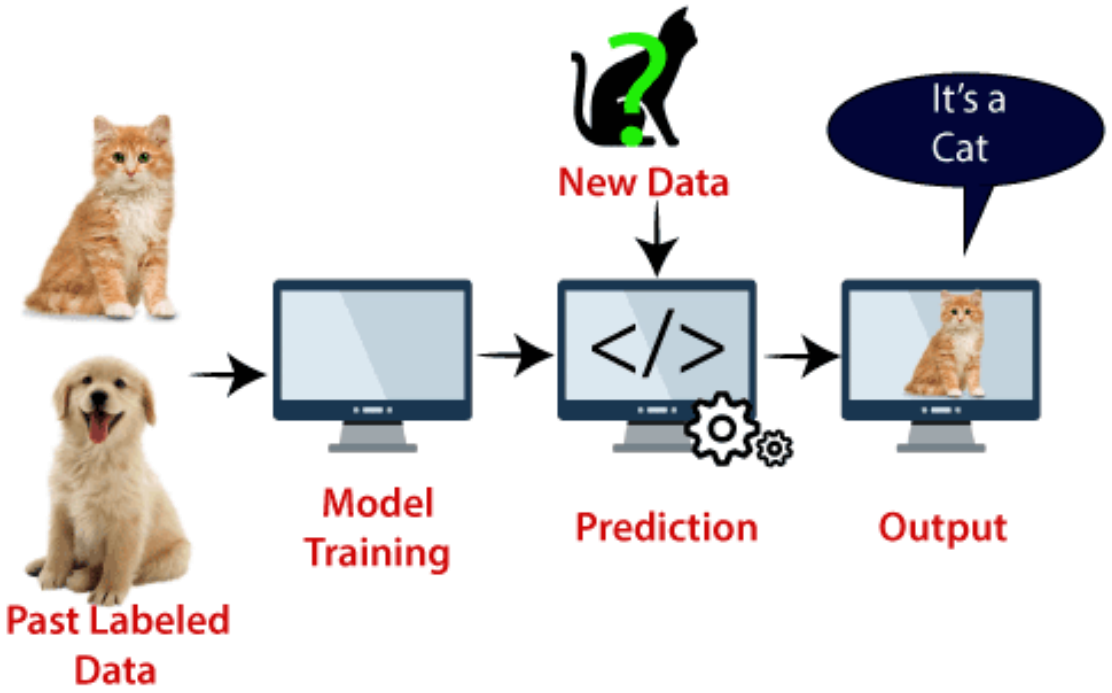
- Support Vector Machine or SVM is one of the most popular Supervised Learning algorithms, which is used for Classification as well as Regression problems.
- Primarily, it is used for Classification problems in Machine Learning.
- The goal of the SVM algorithm is to create the best line or decision boundary that can segregate n-dimensional space into classes so that we can easily put the new data point in the correct category in the future.

SUPPORT VECTOR MACHINE ALGORITHM



- This 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.
- Consider this diagram in which there are two different categories that are classified using a decision boundary or hyperplane

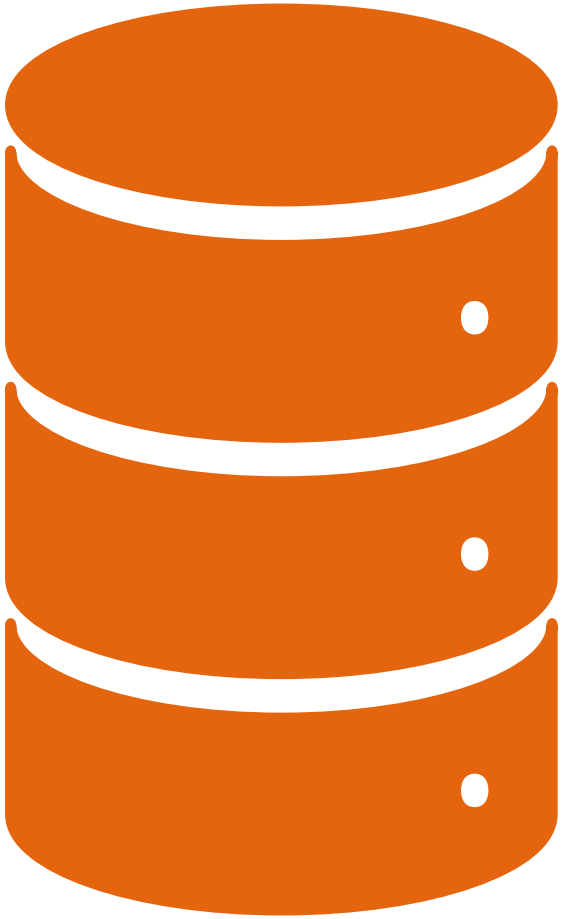
SVM: EXAMPLE



- SVM can be understood with the example:
- 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.

SVM: EXAMPLE

- 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. Considering diagram
- SVM algorithm can be used for **Face detection, image classification, text categorization,** etc.



TYPES OF SVM

SVM can 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 used called as Linear SVM classifier.
- **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 used is called as Non-linear SVM classifier.

HYPERPLANE AND SUPPORT VECTORS IN THE SVM ALGORITHM

Hyperplane:

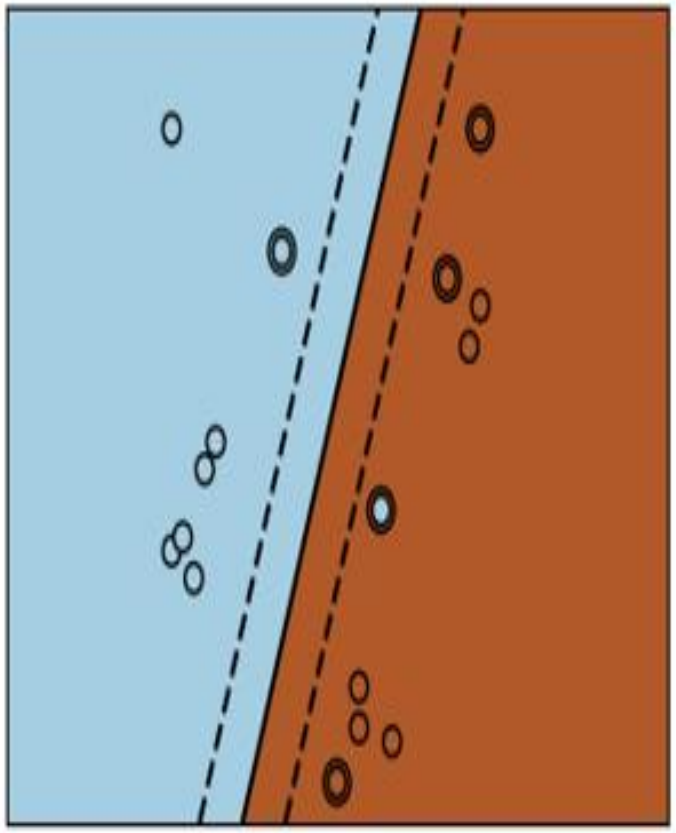
- There can be multiple lines/decision boundaries to segregate the classes in n-dimensional space, but we need to find out the best decision boundary that helps to classify the data points.
- This best boundary is known as the hyperplane of SVM.
- The dimensions of the hyperplane depend on the features present in the dataset, which means if there are 2 features (as shown in image), then hyperplane will be a straight line.
- And if there are 3 features, then hyperplane will be a 2-dimension plane.
- We always create a hyperplane that has a maximum margin, which means the maximum distance between the data points.

HYPERPLANE AND SUPPORT VECTORS IN THE SVM ALGORITHM

Support Vectors:

- The data points or vectors that are the closest to the hyperplane and which affect the position of the hyperplane are termed as Support Vector.
- These vectors support the hyperplane, hence called a Support vector.

TYPE OF KERNAL: LINEAR KERNEL

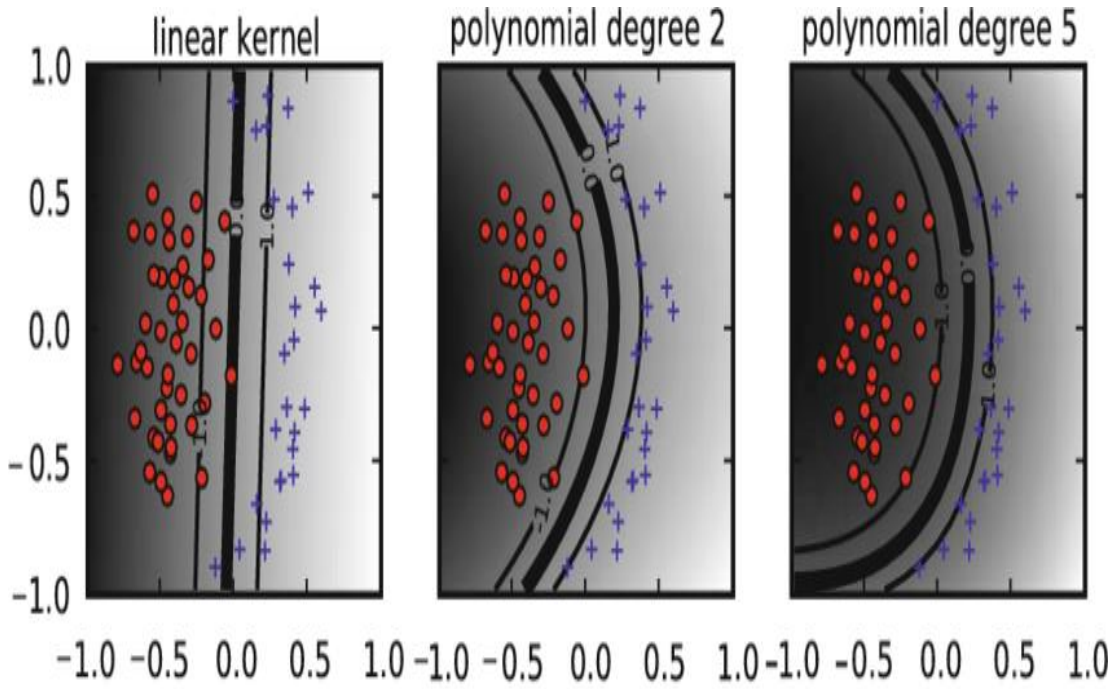


In linear kernel, the kernel function takes the form of a linear function as follows-

- **linear kernel : K(xi , xj) = xiT xj**
- Linear kernel is used when the data is linearly separable.
- It means that data can be separated using a single line. It is one of the most common kernels to be used.
- It is mostly used when there are large number of features in a dataset.
- Linear kernel is often used for text classification purposes.
- Training with a linear kernel is usually faster, because we only need to optimize the C regularization parameter.

TYPE OF KERNAL: POLYNOMIAL KERNEL

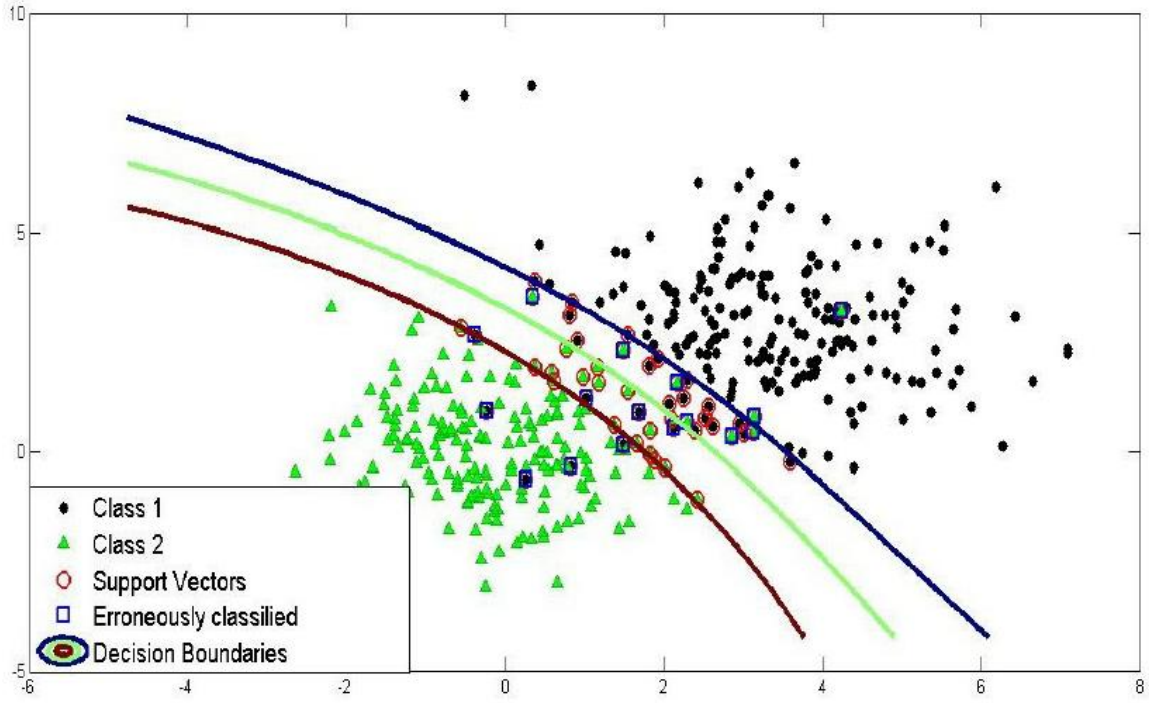
- The polynomial kernel looks not only at the given features of input samples to determine their similarity, but also combinations of the input samples.
- For degree-d polynomials, the polynomial kernel is defined as follows –
- Polynomial kernel : $K(xi , xj) = (\gamma.xiT\ xj + r)d , \gamma > 0$
- Polynomial kernel is very popular in Natural Language Processing.
- The most common degree is d = 2 (quadratic), since larger degrees tend to overfit on NLP problems. It can be visualized with this diagram.



TYPE OF KERNAL: RADIAL BASIC KERNEL

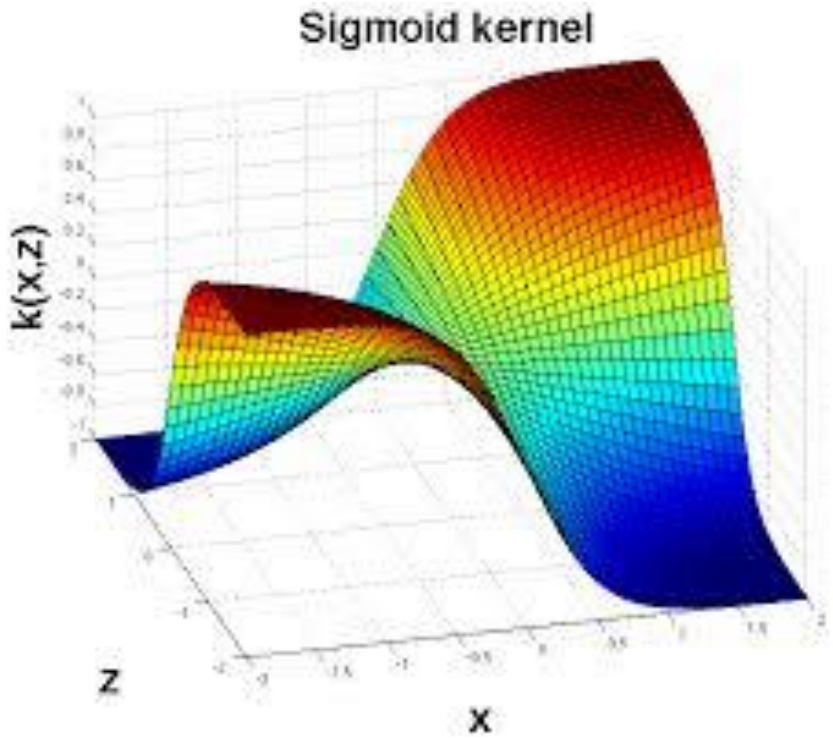
- Radial basis function kernel is a general purpose kernel.
- It is used when we have no prior knowledge about the data.
- The RBF kernel on two samples x and y is defined by the following equation -

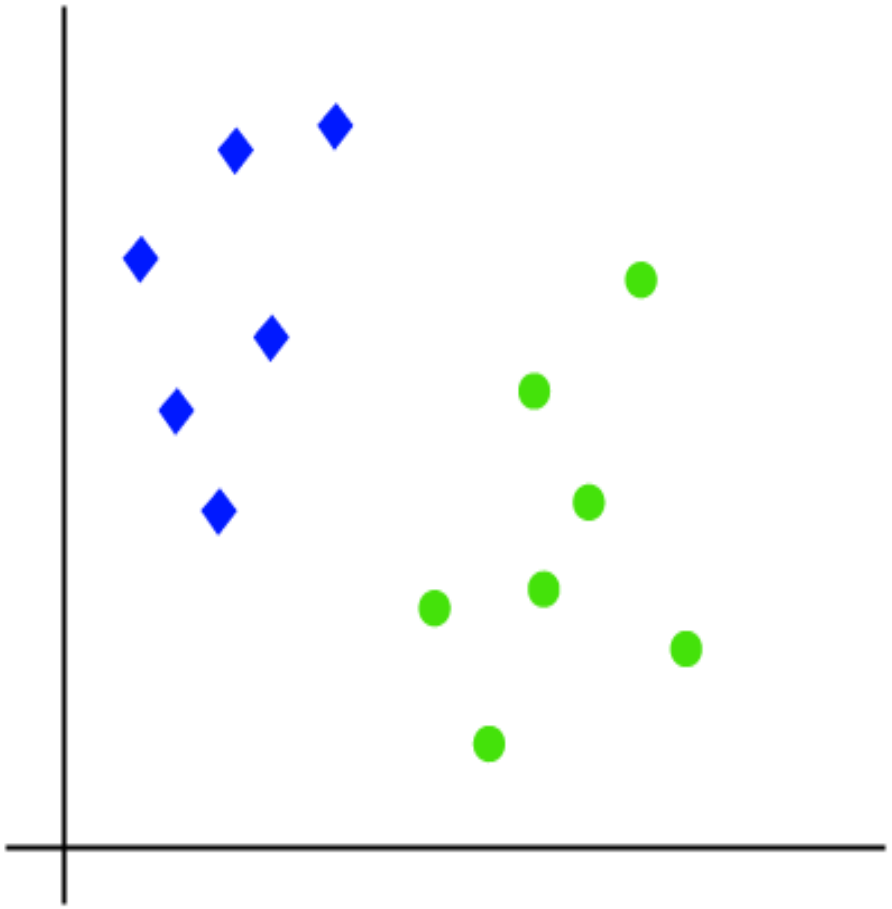
$$k(x, y) = \exp\left(-\frac{\|x - y\|^2}{2\sigma^2}\right)$$



TYPE OF KERNAL: SIGMOID KERNEL

- Sigmoid kernel has its origin in neural networks.
- We can use it as the proxy for neural networks. Sigmoid kernel is given by the following equation –
- **sigmoid kernel : k (x, y) = tanh(α.xTy + c)**
- Sigmoid kernel can be visualized with the following diagram-

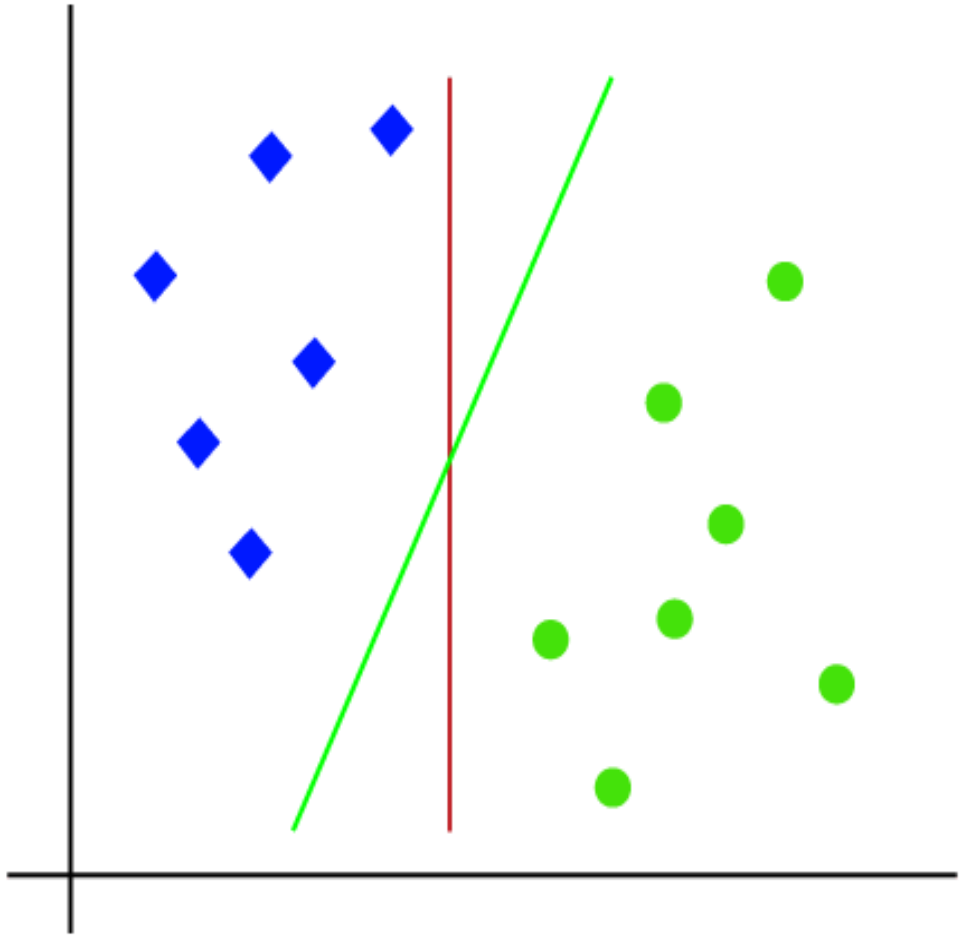




HOW DOES SVM WORKS?

Linear SVM:

- The working of the SVM algorithm can be understood by using an example.
- Suppose we have a dataset that has two tags (green and blue), and the dataset has two features x1 and x2.
- We want a classifier that can classify the pair(x1, x2) of coordinates in either green or blue. Consider this image:

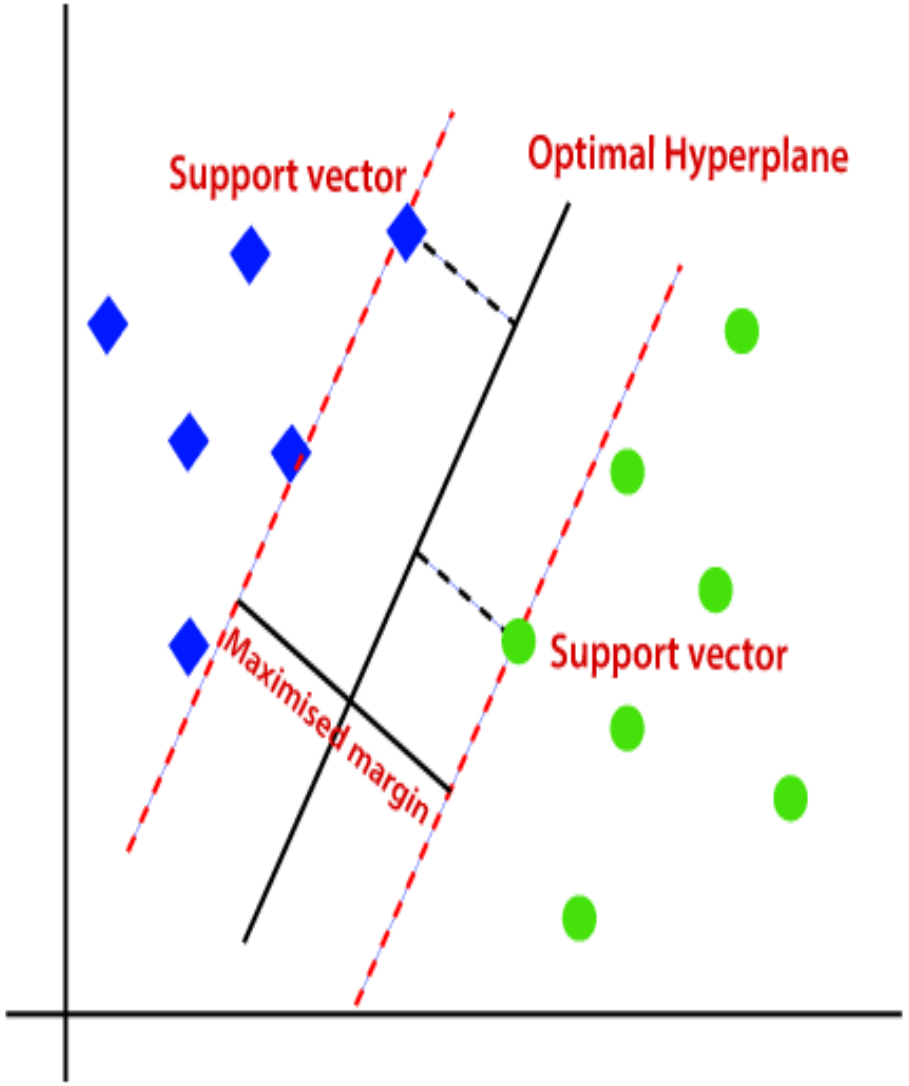


HOW DOES SVM WORKS?

Linear SVM:

- So as it is 2-d space so by just using a straight line, we can easily separate these two classes.

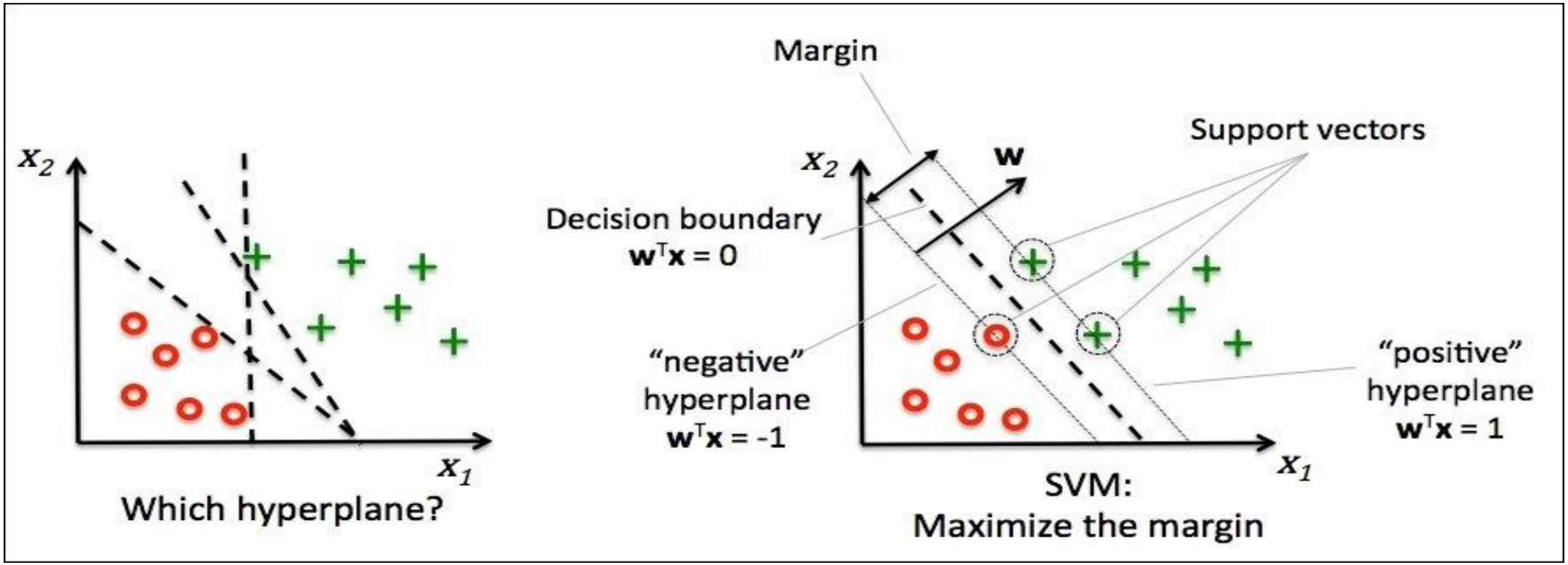
- But there can be multiple lines that can separate these classes. Consider this image.

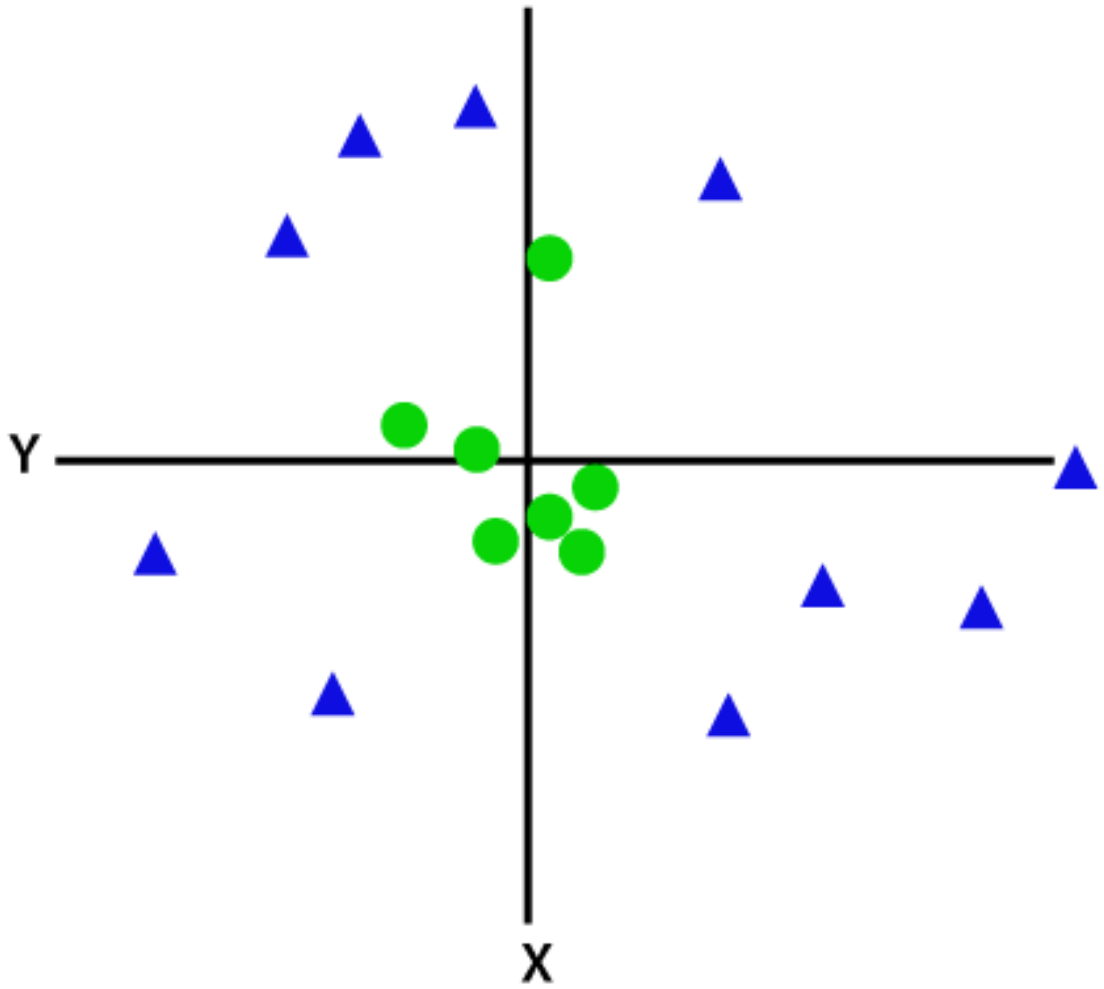


HOW DOES SVM WORKS?

- The SVM algorithm helps to find the best line or decision boundary; this best boundary or region is called as a **hyperplane**.
- SVM algorithm finds the closest point of the lines from both the classes. These points are called support vectors.
- The distance between the vectors and the hyperplane is called as **margin**.
- And the goal of SVM is to maximize this margin.
- The **hyperplane** with maximum margin is called the **optimal hyperplane**.

MAXIMUM MARGIN HYPERPLANE

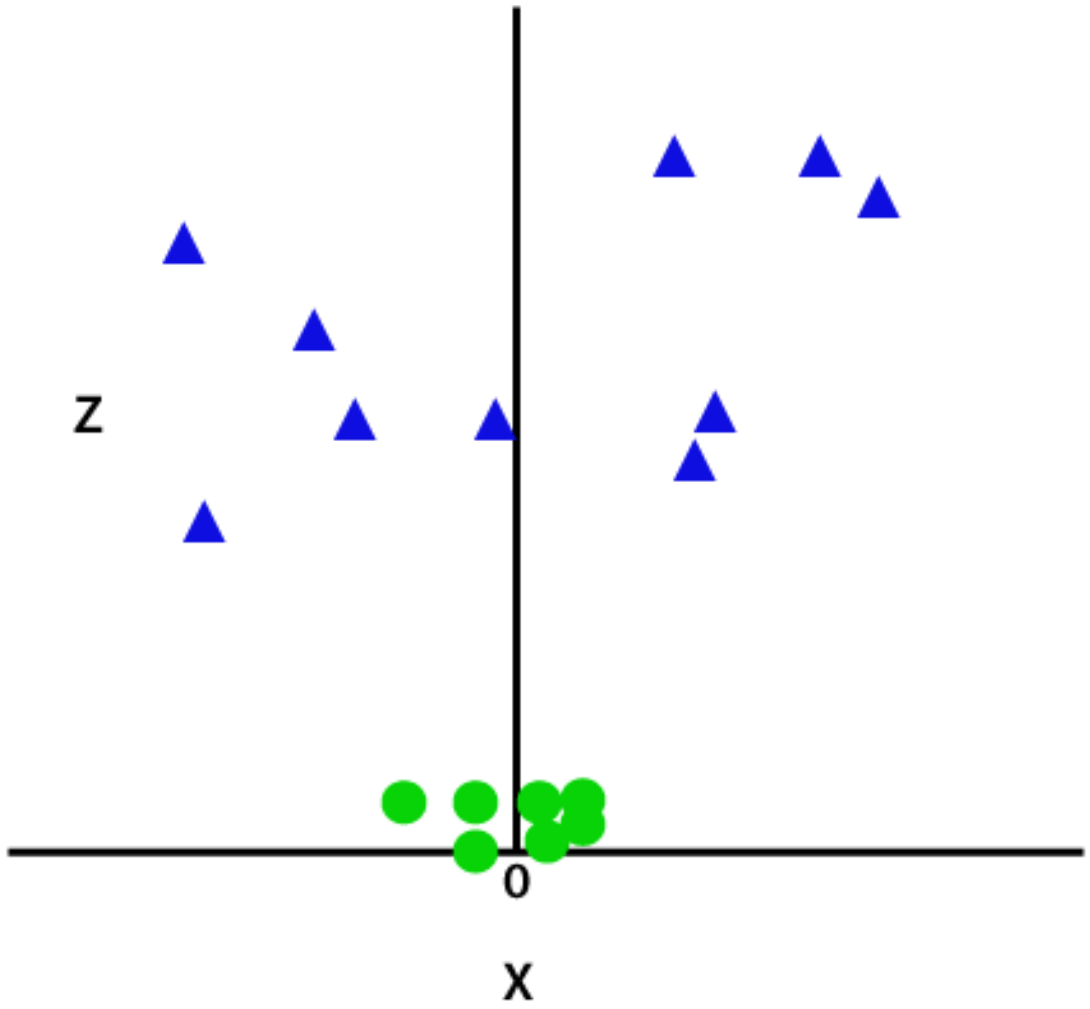




HOW DOES SVM WORKS?

Non-Linear SVM:

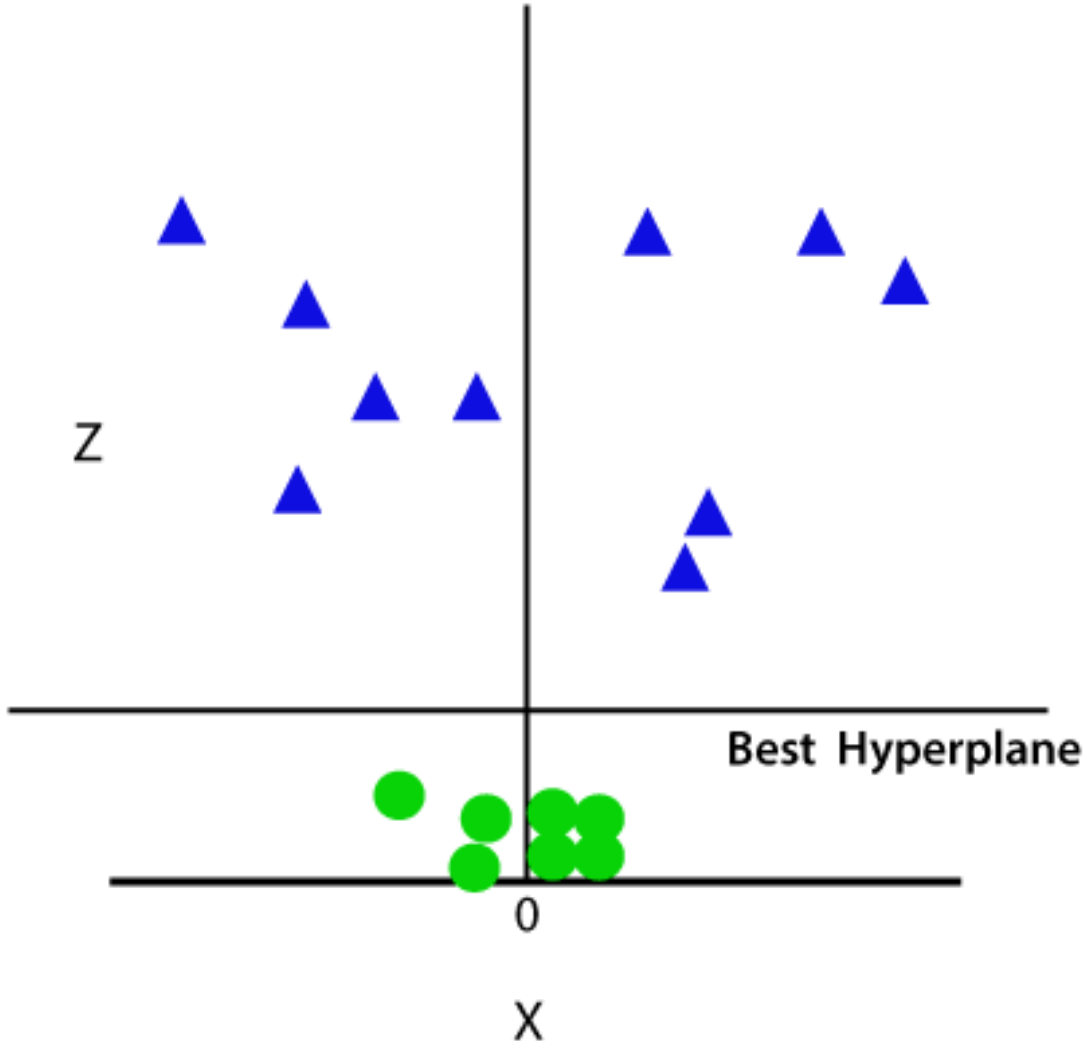
- If data is linearly arranged, then we can separate it by using a straight line, but for non-linear data, we cannot draw a single straight line.
- Consider this image
- So to separate these data points, we need to add one more dimension.
- For linear data, we have used two dimensions x and y, so for non-linear data, we will add a third dimension z. It can be calculated as:
- $z=x^2 +y^2$



HOW DOES SVM WORKS?

Non-Linear SVM:

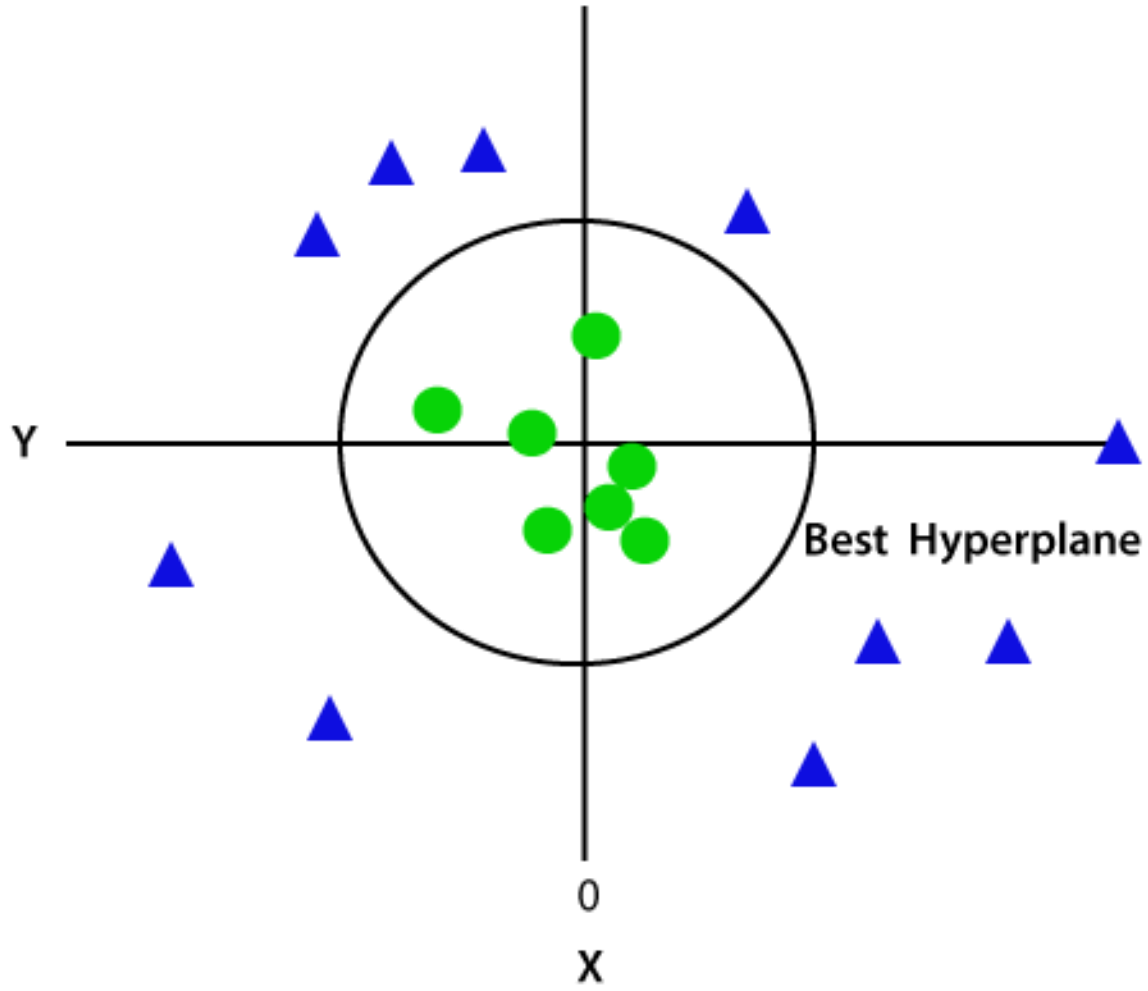
- By adding the third dimension, the sample space will become as this image
- So now, SVM will divide the datasets into classes in the following way.



HOW DOES SVM WORKS?

Non-Linear SVM:

- Consider this image
- Since we are in 3-d Space, hence it is looking like a plane parallel to the x-axis.



HOW DOES SVM WORKS?

Non-Linear SVM:

- If we convert it in 2d space with z=1, then it will become as:

- Hence we get a circumference of radius 1 in case of non-linear data.

PYTHON IMPLEMENTATION OF SUPPORT VECTOR MACHINE

- Now we will implement the SVM algorithm using Python.
- Here we will use the same dataset **user_data**, which we have used in Logistic regression.

THANK YOU