**SVM:**

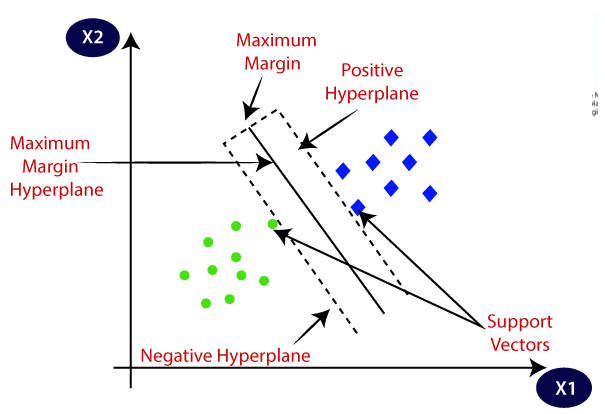
Support Vector Machine is a supervised learning algorithm which can be used for regression as well as classification problems.

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. 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 the below diagram in which there are two different categories that are classified using a decision boundary or hyperplane:

Befor going to SVM work first know some genral things:

1. Support Vector
2. Hyper Plane
3. Marginal Distance
4. Kernel
5. Linear and Non Linear Separable Data
6. Hyperplanes are decision boundaries that help classify the data points. Data points falling on either side of the hyperplane can be attributed to different classes. Also, the dimension of the hyperplane depends upon the number of features. If the number of input features is 2, then the hyperplane is just a line. If the number of input features is 3, then the hyperplane becomes a two-dimensional plane.
7. Support vectors are data points that are closer to the hyperplane and influence the position and orientation of the hyperplane. Using these support vectors, we maximize the margin of the classifier. Deleting the support vectors will change the position of the hyperplane. These are the points that help us build our SVM.
8. Marginal Distance is the two parallel line that create w.r.t most near point of +ve and -ve to plane distance. we always maximise the marginal distance.
9. See the below figure you better understood all these three term.

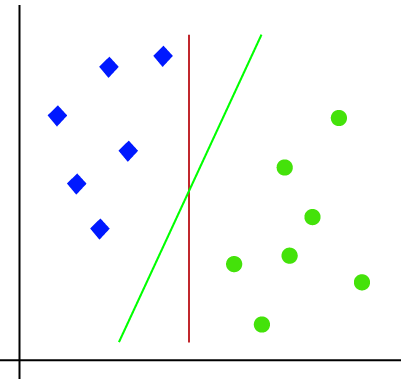
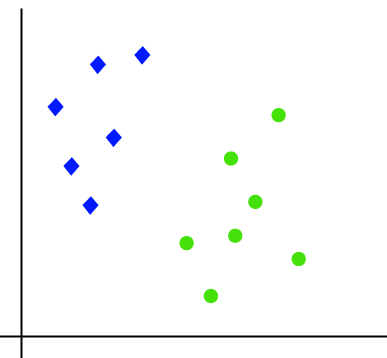


SVM algorithm can be used for Face detection, image classification, text categorization, etc.

Types of SVM

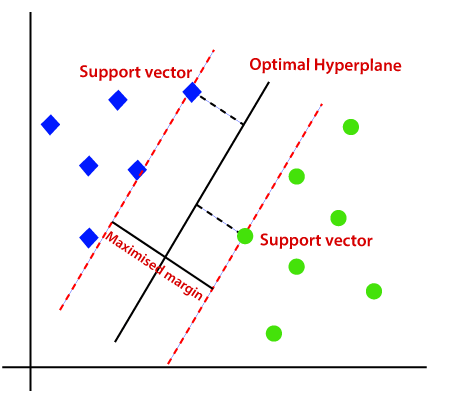
SVM can be of two types:

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.

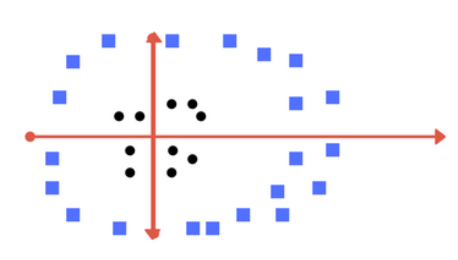


Suppose we have a dataset that has two tags (green and blue), and the dataset has two features x1 and x2. We want a classifier that can classify the pair(x1, x2) of coordinates in either green or blue.So as it is 2-d space so by just using a straight line, we can easily separate these two classes. But there can be multiple lines that can separate these classes

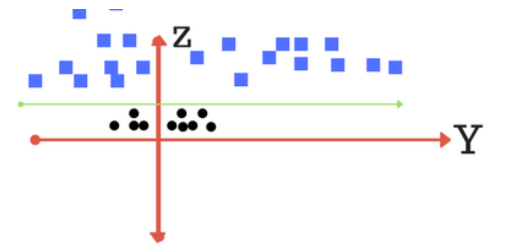
Hence, the SVM algorithm helps to find the best line or decision boundary; this best boundary or region is called as a **hyperplane**. SVM algorithm finds the closest point of the lines from both the classes. These points are called support vectors. The distance between the vectors and the hyperplane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



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.



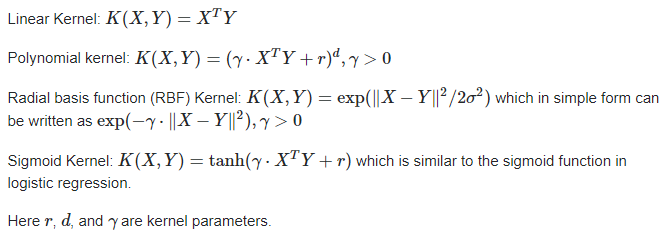
* Now consider what if we had data as shown in image below? This is a type of non linear data.
* Clearly See, there is no line that can separate the two classes in this x-y plane.
* So what do we do? We apply transformation and add one more dimension as we call it z-axis.
* Lets assume value of points on z plane, w = x² + y² - So, basically z co-ordinate is the square of distance of the point from origin. Let’s plot the data on z-axis.. After Transform



* These Transform Done By SVM kernels.

Types of Kernels in SVM:

1. Linear
2. Polynomial
3. Gaussian (RBF)
4. Sigmoid



**Pros of SVM classifiers**

* SVM classifiers offers great accuracy and work well with high dimensional space.
* SVM classifiers basically use a subset of training points hence in result uses very less memory.

**Cons of SVM classifiers**

* They have high training time hence in practice not suitable for large datasets.
* Another disadvantage is that SVM classifiers do not work well with overlapping classes.