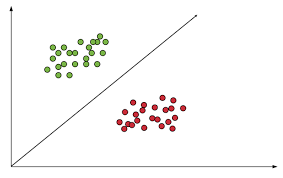
Support Vector Machine

Support vector machine is a supervised learning algorithm which can be used to solve both classification and regression, but popularly it is used for classification problems. In this algorithm we plot each data point in the n-dimensional space, here n= number of features given. After plotting the data-items, a hyper-plane is created which differentiates two classes to perform the classification. Svm is also used for outlier detection.



Important terminologies

Optimal Hyper plane- The best Hyper plane that will separate the classes.

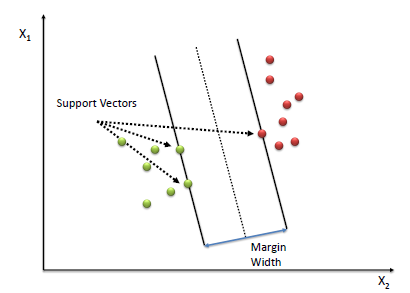
Marginal planes- Marginal planes are the pair of two parallel planes that are parallel to the hyper plane.

A marginal plane is selected in such a way that it must be very close to the support vectors.

Support vectors- support vectors are the data-points that are close to the marginal plane/decision boundary. Sometimes the data plane may pass through the support vectors

Marginal Distance- The distance between two marginal planes or support vectors

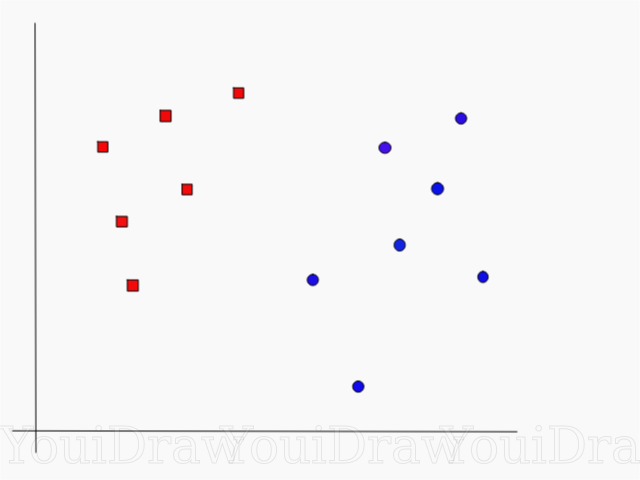
Kernel- Used for dimension transmission i.e. converting low dimension to high dimension.



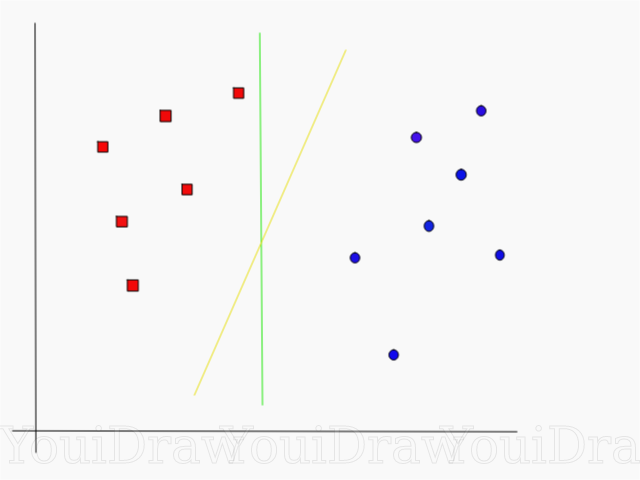
For better classification we need to maximize the marginal width. If marginal width is large then we can easily classify the new data item and designate where it belongs.

Let us understand SVM with an example

You have been getting many unwanted sales mails which are not useful to you. And you decided to separate the spam mails from ham mail. Plotting the data in dimensional space it looks somewhat like this.



You have decided to use svm classifier for classification.

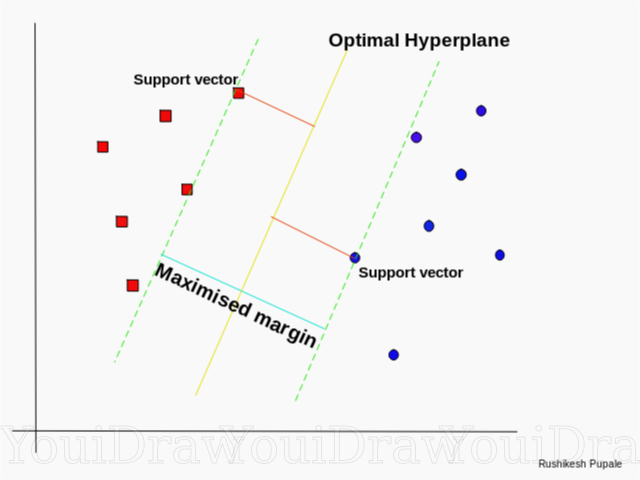


As we can see there are two data plane which can separate both the classes.

Which one to choose?

Here comes support vectors and marginal planes. You need to calculate marginal distance for all the data planes which are present in n dimensional space. The one which gives the high marginal distance can be called as Optimal Hyper plane.

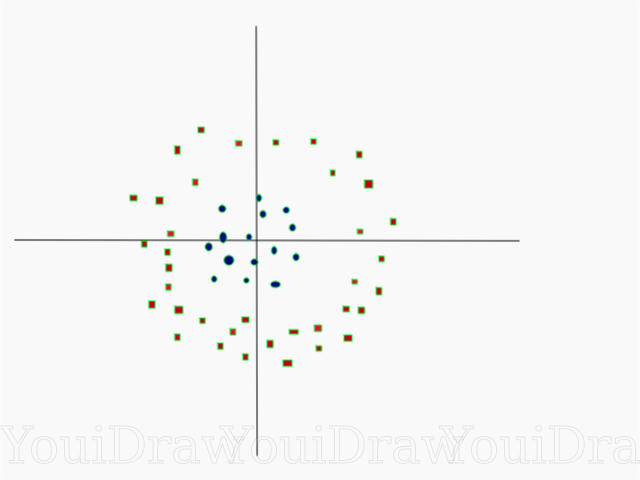
With help of the OPH we can easily classify the upcoming data points.



As we can see that the data can be classified by linear decision boundary, we can say that these kinds of problems can be solved by Linear Support vector machine.

The real life datasets are more complex than the above example

!!! look carefully in the below image!!!



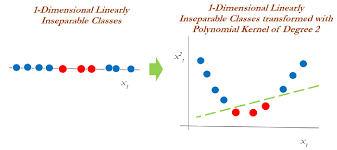
Try to solve this with LVSM? We Can’t.

Clearly, we cannot classify the dataset by a [**LVSM**](https://dataaspirant.com/how-logistic-regression-model-works/), but this data can be converted into a linear one using **higher dimensions**.

With help of kernels, we can classify these kinds of data.

Kernels are used in dimension transmission. 1d->2d,2d->3d

Example-1

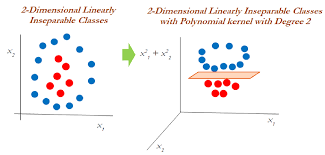


After plotting data points in dimensional space. If all the data points lie in a single line. Then we cannot separate the classes. With help of Polynomial kernel of degree, we can convert 1d to 2d and classify the problem.

Example-2

Here also we cannot classify by creating a decision boundary/ hyper plane. Even though it is in 2d form, It is not linearly separable.

By using polynomial kernel of degree 3 we can convert the 2d to 3d and classify.



Types of kernels

Polynomial Kernel

Linear kernel

RBF kernel

Sigmoid Kernel

Advantages of SVM

More effective in high-dimensional spaces.

Used for the data which is not regularly distributed.

Input data can be converted into high dimensional data.

Memory efficient.

When compared to other algorithms it has more computational complexity.

Better accuracy, Fast prediction

Applications of SVM

-face detection

-text and hypertext categorization

-classification of images

-bioinfamatics