SVM

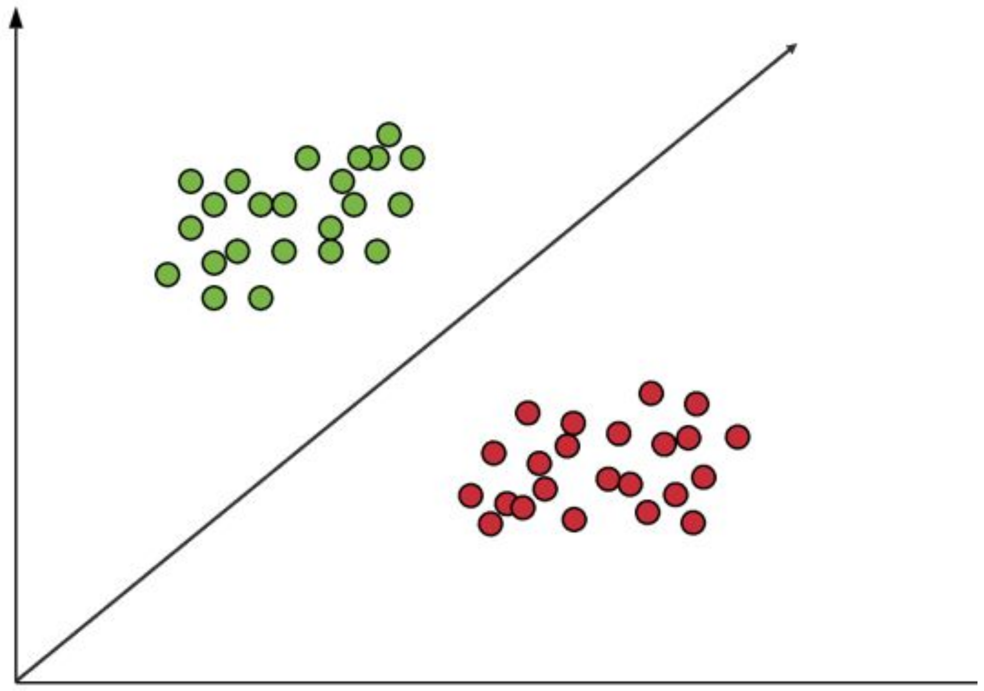
Support Vector Machine

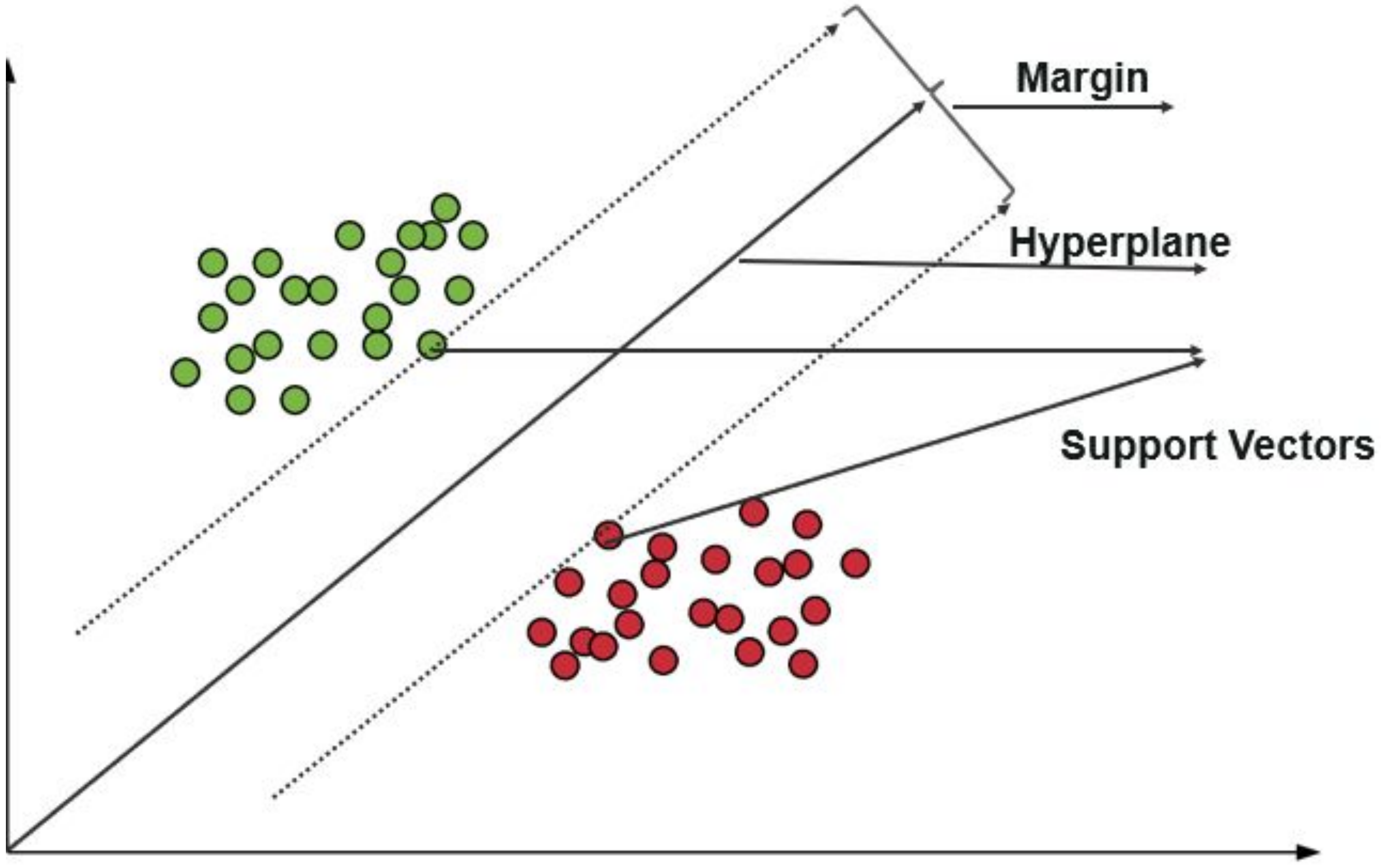
What is the Support Vector Machine?

- A Support Vector Machine was first introduced in the 1960s and later improvised in the 1990s.
- It is a supervised learning machine learning classification algorithm that has become extremely popular nowadays owing to its extremely efficient results.
- An SVM is implemented in a slightly different way than other machine learning algorithms.
- It is capable of performing classification, regression and outlier detection.
- Support Vector Machine is a discriminative classifier that is formally designed by a separative hyperplane. It is a representation of examples as points in space that are mapped so that the points of different categories are separated by a gap as wide as possible. In addition to this, an SVM can also perform non-linear classification.

How Does SVM Work?

- The main objective of a support vector machine is to segregate the given data in the best possible way. When the segregation is done, the distance between the nearest points is known as the margin. The approach is to select a hyperplane with the maximum possible margin between the support vectors in the given data-sets.
-





- To select the maximum hyperplane in the given sets, the support vector machine follows the following sets:
- Generate hyperplanes which segregates the classes in the best possible way
- Select the right hyperplane with the maximum segregation from either nearest data points

SVM Kernels

- An SVM kernel basically adds more dimensions to a low dimensional space to make it easier to segregate the data. It converts the inseparable problem to separable problems by adding more dimensions using the kernel trick. A support vector machine is implemented in practice by a kernel. The kernel trick helps to make a more accurate classifier. The different kernels in the Support vector machine.

- **Linear Kernel** – A linear kernel can be used as a normal dot product between any two given observations. The product between the two vectors is the sum of the multiplication of each pair of input values. Following is the linear kernel equation.

$$f(x) = B(0) + sum(ai * (x,xi))$$

- **Polynomial Kernel** – It is a rather generalized form of the linear kernel. It can distinguish curved or nonlinear input space. Following is the polynomial kernel equation.

$$K(X_1, X_2) = (a + X_1^T X_2)^b$$

b = degree of kernel & a = constant term.

Radial Basis Function Kernel – The radial basis function kernel is commonly used in SVM classification, it can map the space in infinite dimensions. Following is the RBF kernel equation.

$$K(X_1, X_2) = exponent(-\gamma \|X_1 - X_2\|^2)$$

||X1 – X2 || = Euclidean distance between X1 & X2

Support Vector Machine Use Cases

- Face Detection
- Text And HyperText Categorization
- Classification Of Images
- Bioinformatics
- Protein Fold and Remote Homology Detection
- Handwriting Recognition
- Generalized Predictive Control

Advantages of SVM

- Effective in high dimensional spaces
- Still effective in cases where the number of dimensions is greater than the number of samples
- Uses a subset of training points in the decision function that makes it memory efficient
- Different kernel functions can be specified for the decision function that also makes it versatile

Disadvantages of SVM

If the number of features is much larger than the number of samples, avoid over-fitting in choosing kernel functions and regularization term is crucial.

SVMs do not directly provide probability estimates, these are calculated using five-fold cross-validation.