IT341 ML Lab

Support Vector Machine

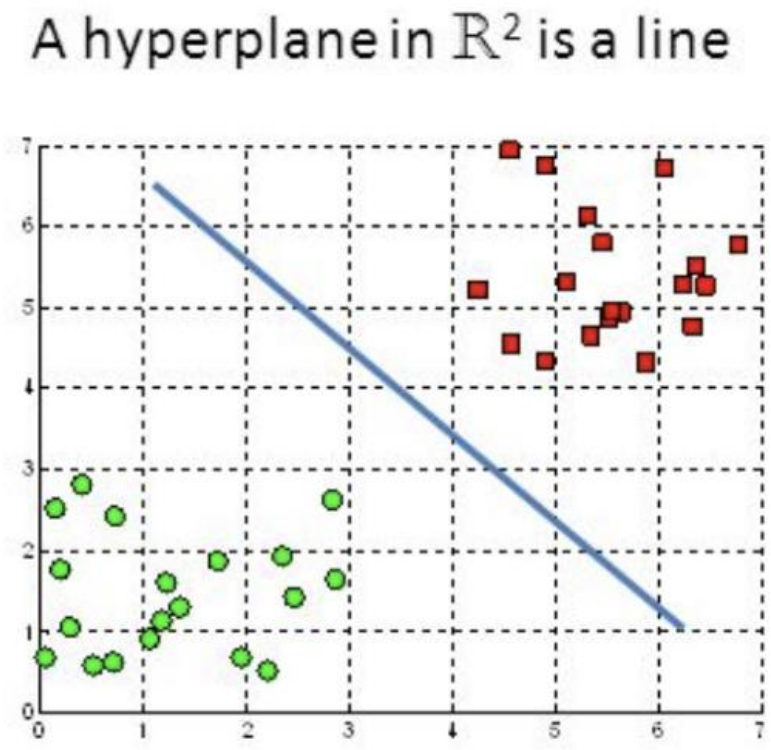
Classification with Support Vector Machine (SVM)

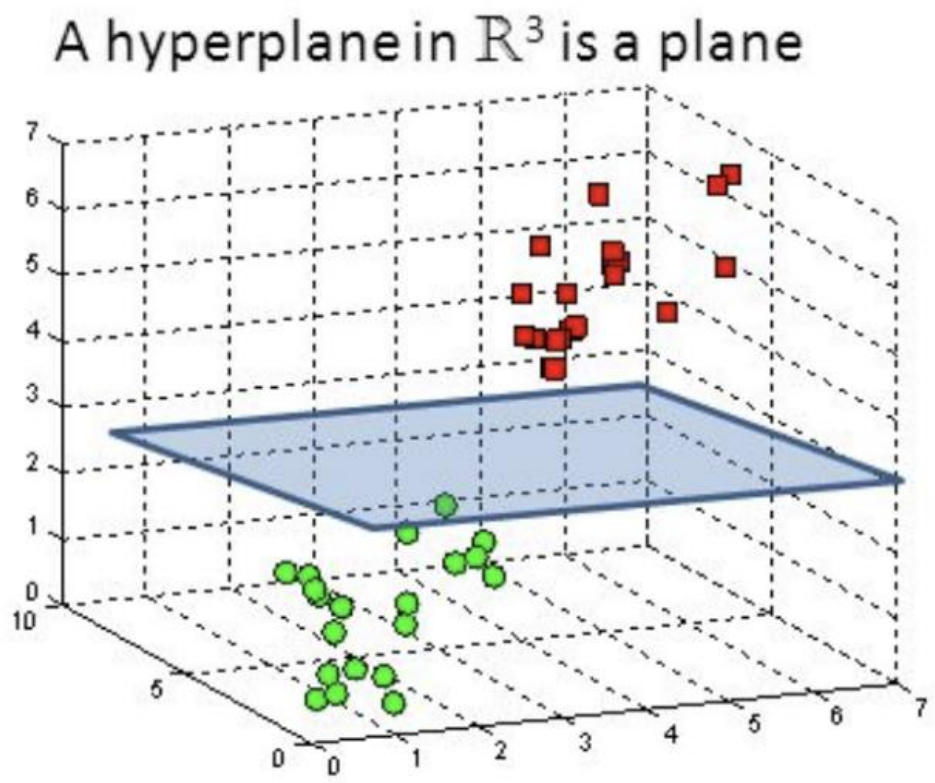
- SVM maps training samples to points in space so as to maximize the width of the gap between the two categories.
- The decision boundary created by SVMs is called the maximum margin classifier or the maximum margin hyper plane.
- Support vectors are data points that are closer to the hyperplane and influence the position and orientation of the hyperplane
- The distance between the vectors and the hyperplane is called as **margin**.
- Goal of SVM is to maximize this margin. The **hyperplane** with maximum margin is called the **optimal hyperplane.**

Contd.

- The data point is viewed as a p-dimensional vector and we want to know whether we can separate such points with a (p-1)-dimensional hyperplane.
- A good separation is achieved by the hyperplane that has the largest distance to the nearest training-data point of any class, since in general the larger the margin, the lower the generalization error of the classifier

Contd.





Source: towardsdatascience.com

Mathematical Representation

- We are given a training dataset of n points of the form $\{(X_1, y_1), (X_2,y_2), \ldots, (X_n, y_n)\}$ where y_i are either 1 or −1, each indicating the class to which the point belongs. Each X_i is a p-dimensional real vector.
- We want to find the "maximum-margin hyperplane" that divides the group of points for which y_i =1 from the group of points for which y_i =-1, which is defined so that the distance between the hyperplane and the nearest point from either group is maximized.
- Any hyperplane can be written as the set of x satisfying w.X – b = 0, where w is the normal vector to the hyperplane.
- The parameter b determines the offset of the hyperplane from the origin along the normal vector.

  w.X – b >= 1, for y_i = 1

  w.X – b <= -1, for y_i = -1

Model Hyperparameters

- **Kernel**: The main function of the kernel is to transform the given dataset input data into the required form. There are various types of functions such as linear, polynomial, and radial basis function (RBF). Polynomial and RBF kernels compute the separation line in the higher dimension.
- **Regularization**: For larger values of C, a smaller margin will be accepted if the decision function is better at classifying all training points correctly. A lower C will encourage a larger margin, therefore a simpler decision function, at the cost of training accuracy.
- **Gamma**: A high value of gamma considers only nearby points in calculating the separation boundary of the class regions, while a small value of gamma considers far away data points also.

Assignment 10

- Implement binary classification using SVM classifier
- Print accuracy, precision, f1-score and recall
- Examine effect of model hyperparameters

```
#import svm model
from sklearn import svm
#Create a svm Classifier

#import scikit-learn metrics module for performance evaluation
from sklearn import metrics
```