**Lab 5**

**To implement SVM**

**Support Vector Machine**aka **Support Vector Network**is a supervised machine learning algorithm used for classification and regression problems.

**Hyperplane:** A hyperplane or decision boundary/surface is an n-dimensional Euclidean space that distinctly separates the data points. The data points on either side of the hyperplane belong to different classes. If we have a p-dimensional feature set, the dimension of the hyperplane will be p-1.

**Support Vectors:** The individual data points that are close to the hyperplane are called the support vectors.

**Margin:**The width that the boundary could be increased before hitting a data point. In simple terms, the distance between the hyperplane and the support vectors is referred to as the Margin.

**Linear separability:**A dataset is linearly separable if there is at least one line that clearly distinguishes the classes.

**Non-linear separability:** A dataset is said to be non-linearly separable if there isn’t a single line that clearly distinguishes the classes.

## How does SVM work?

## Maximum margin hyperplane

In SVM, the data of finite-dimensional space is mapped to much a higher dimension (p-dimension) and aims at finding the p-1 dimension hyperplane called a linear classifier. If the data is linearly separable, unlike logistic regression, in addition to finding the p-1 dimension hyperplane, SVM creates two parrel hyperplanes on either side that passes through the nearest data points(Support Vectors). The region bounded by these two hyperplanes is called the margin.

# If the data is non-linearly separable, we need to apply transformations that map the original data to a much higher dimensional space. After transformation, the data would be linearly separable and we can easily find a hyperplane/decision boundary to classify.

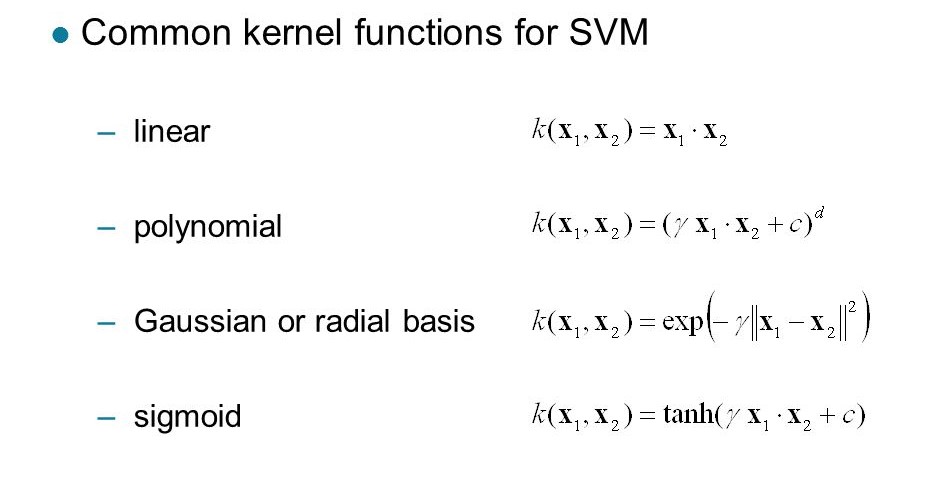
Popular SVM Kernel functions:

1. Linear Kernel: It is just the dot product of all the features. It doesn’t transform the data.

2. Polynomial Kernel: It is a simple non-linear transformation of data with a polynomial degree added.

3. Gaussian Kernel: It is the most used SVM Kernel for usually used for non-linear data.

4. Sigmoid Kernel: It is similar to the Neural Network with sigmoid activation function.



In simple terms, a kernel is nothing but a transformation that we apply to the existing features so that we can draw a classifier easily for non-linearly separable datapoints.

# Kernel trick

Steps:

1. Read the database and perform pre-processing
2. Perform feature scaling (transforming a dataset to fit within a specific range)
3. Split the database into training and testing set (80:20)
4. Instantiate Linear SVC object and train the linear SVC classifier using the training data (Choose C=1)
5. Test the model and compute test accuracy
6. Evaluate test accuracy using C = 1, 5, 10, 4, 70, 100
7. Perform similar experiment using RBF kernel (Choose different C and gamma ) and compute accuracy
8. Also, compute confusion matrix and ROC curve