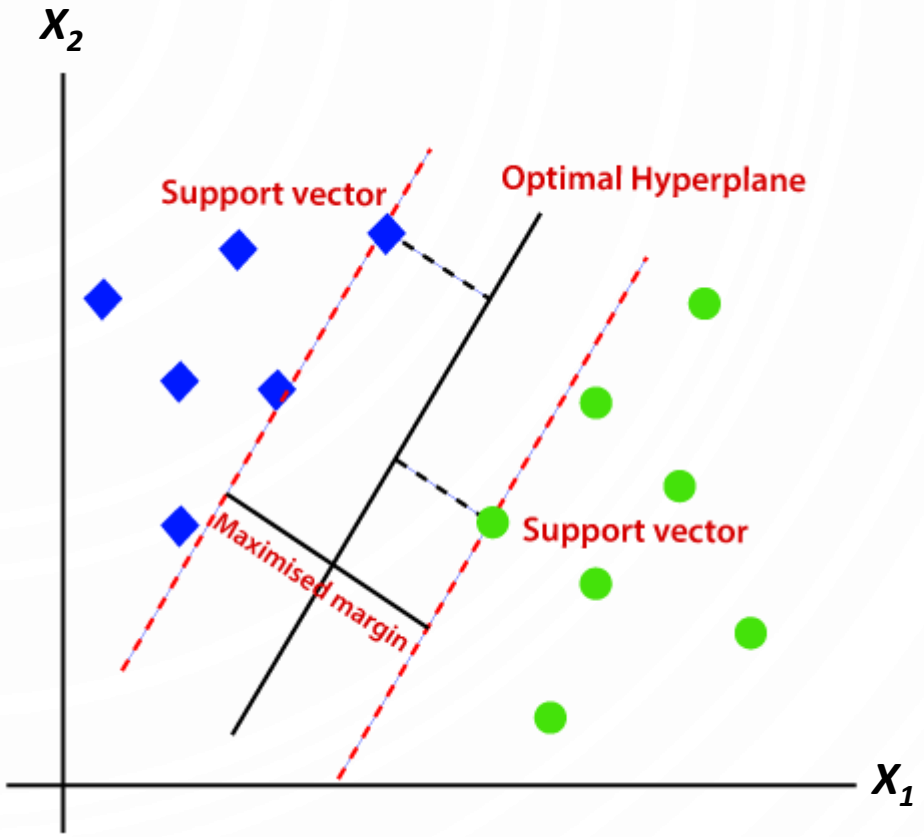
CSC 462 – Machine Learning

3.4 Support Vector Machine

Dr. Sultan Alfarhood

Support Vector Machine (SVM)

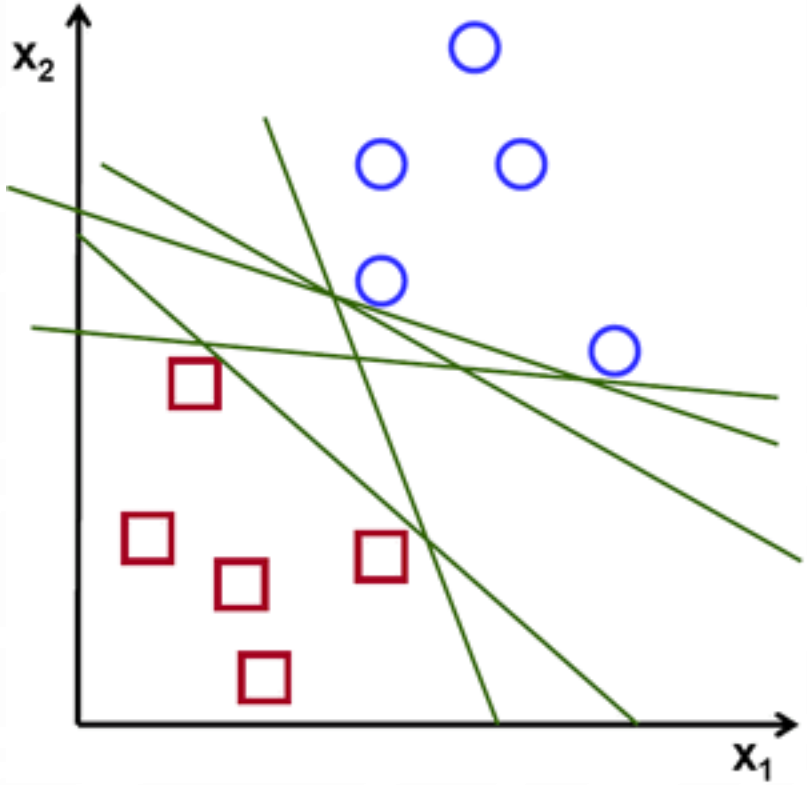
- Support Vector Machine (SVM) is one of the most popular Machine Learning Classifier.
- It uses the concept of Margin to classify between classes.



Note: In this graph, both X_1 and X_2 axes represent **features**, and the **target** is represented by the sample color (blue or green).

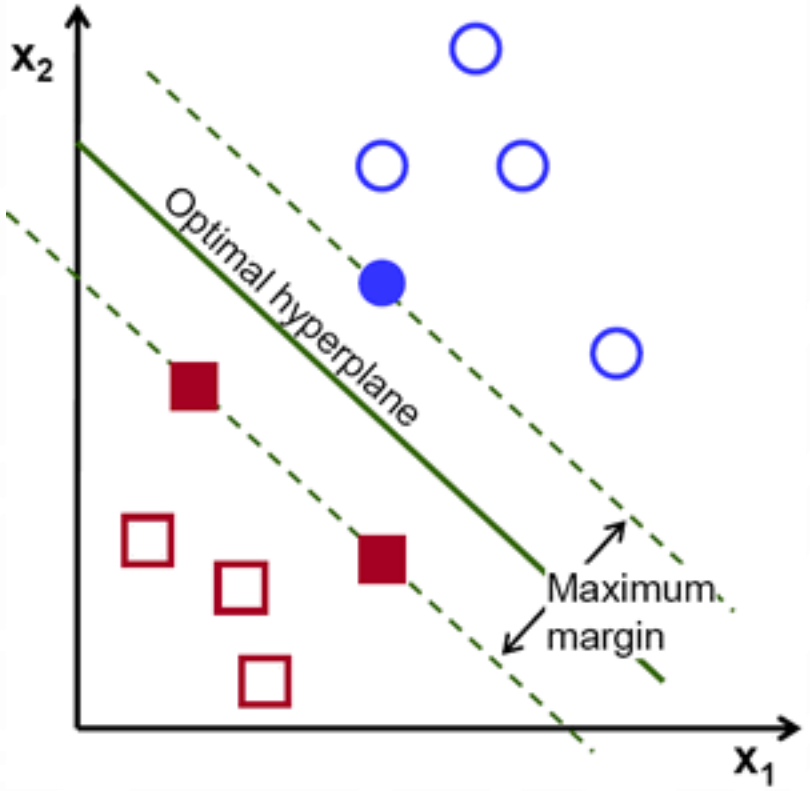
Support Vector Machine (SVM)

It tries to find the decision boundary which separates the classes the best (i.e., with minimum error).



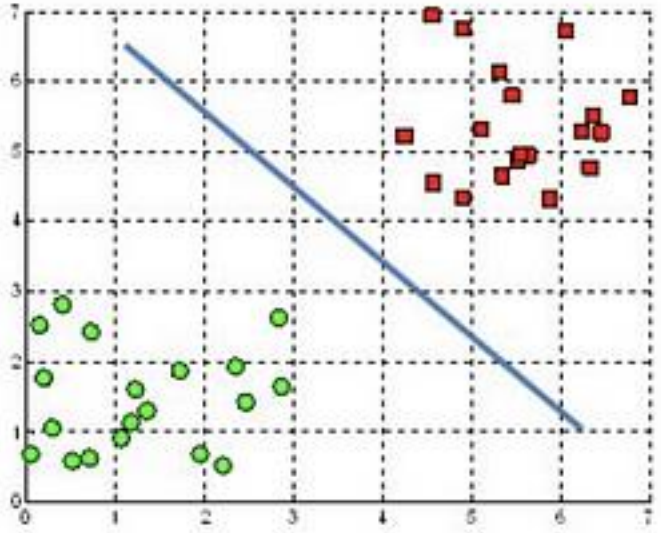
Support Vector Machine (SVM)

- A **hyperplane** is a decision plane which separates between a set of objects having different class memberships.
- The nearest points from the line are called **support vectors**.
- The most optimal line is the one with **maximum margin**.

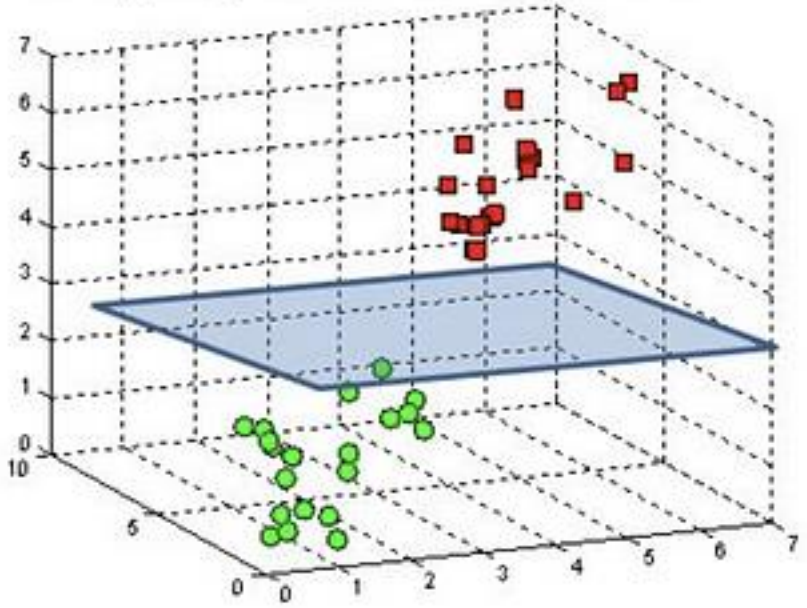


Hyperplanes in 2D and 3D feature space

A hyperplane in $\mathbb{R}^2$ is a line

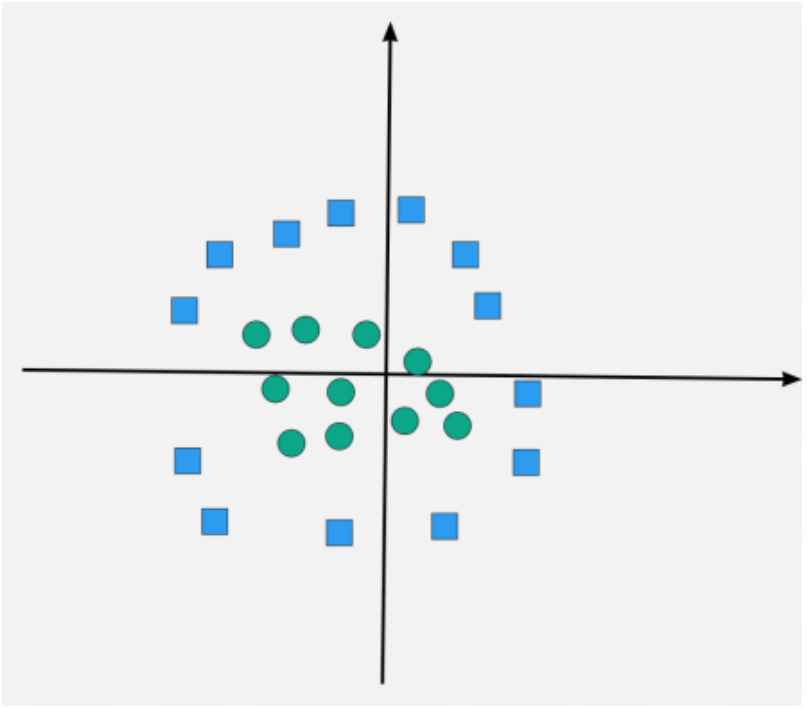


A hyperplane in $\mathbb{R}^3$ is a plane



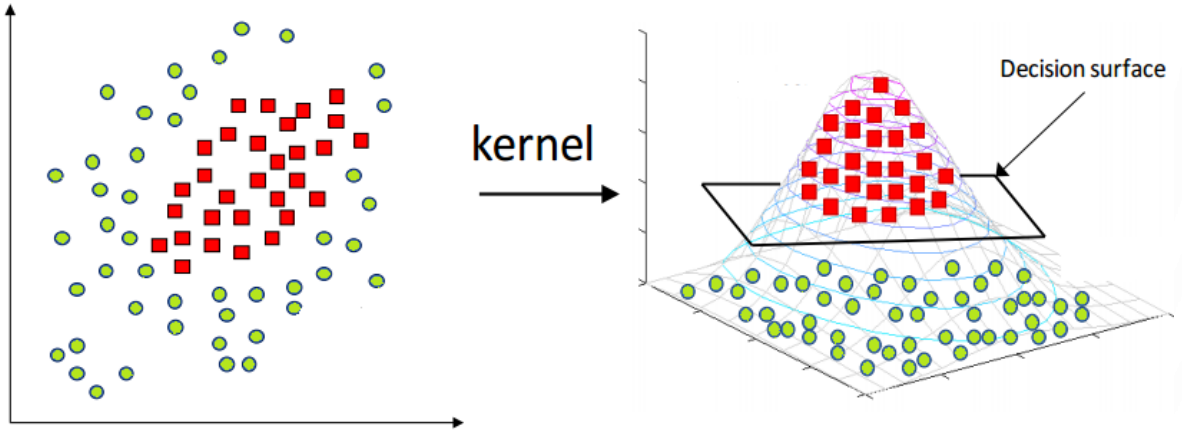
Support Vector Machine (SVM)

- What if the data is inseparable?
  - Things can become difficult when the data is non-linear.
- SVM can use the **kernel function** to make non-linear data linear.

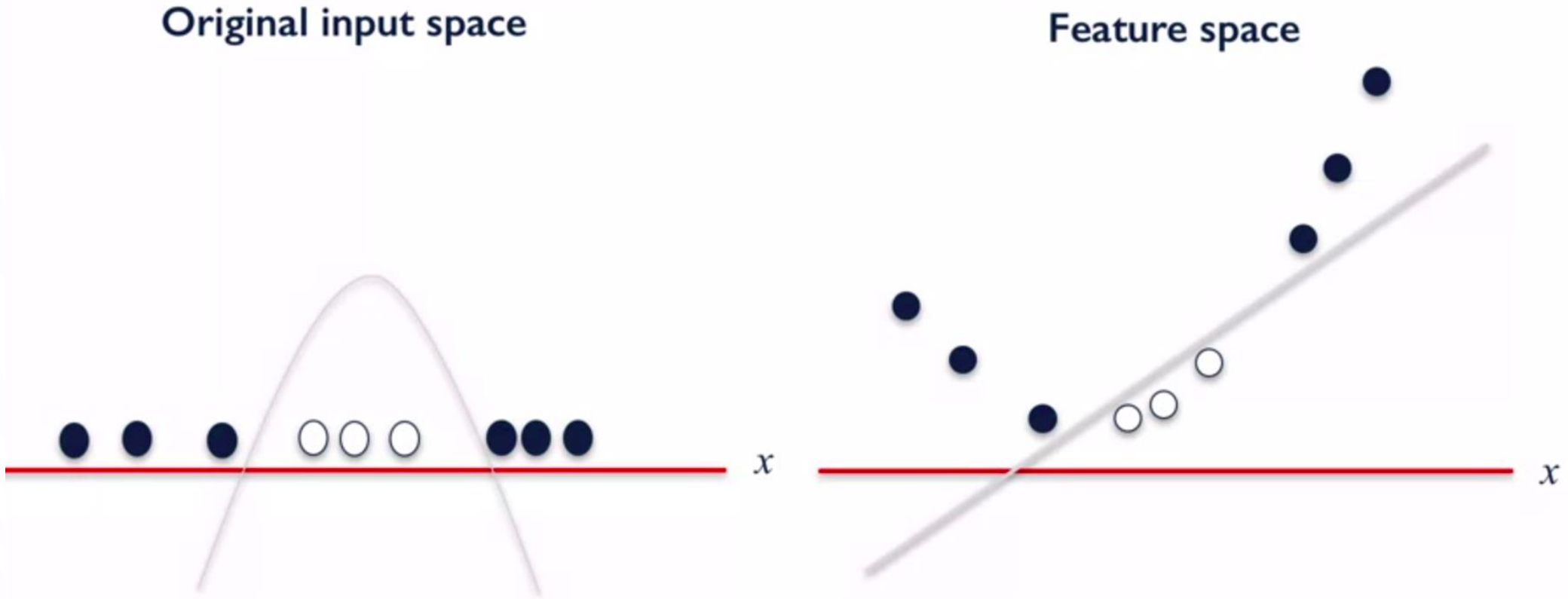


Kernel Function

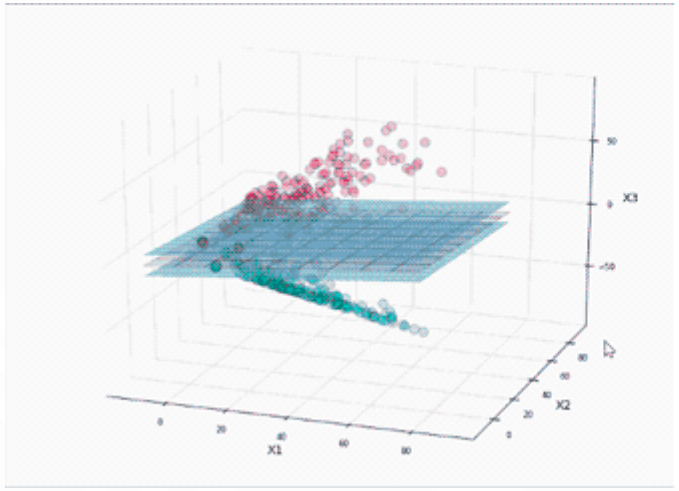
- The **kernel function** separates the data by **adding dimensions to the problem**.
- Kernel Function generally transforms the training set of data so that a non-linear decision surface can transform to a linear equation in a higher number of dimension spaces.
- The idea of the **kernel trick** and the reason why we use kernel functions in SVM is the following:
  - Although we transformed the data to a higher dimension, the kernel only calculates the relationship between the data as if its in a higher dimension but **doesn't change it**.

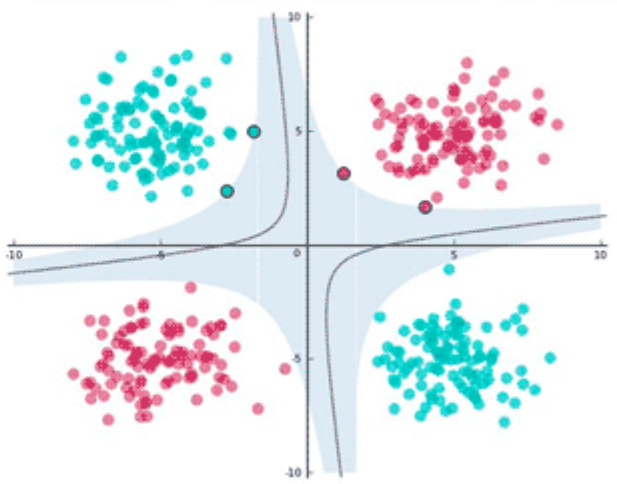


Kernel Transformation Example



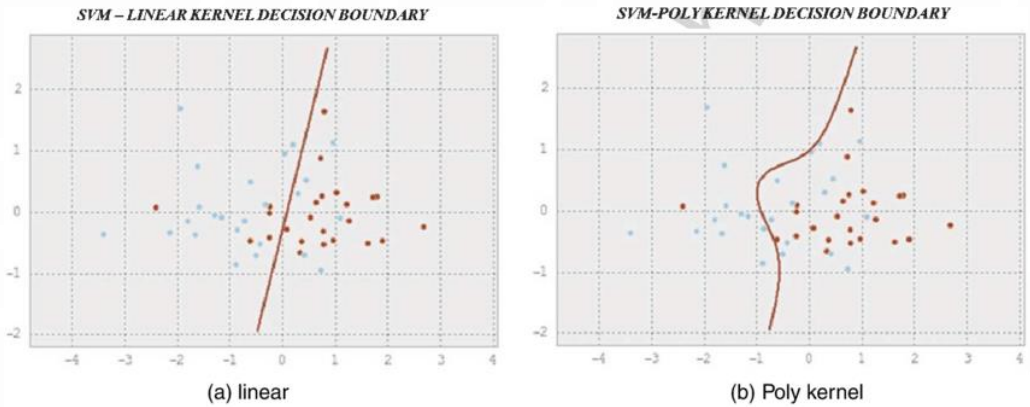
Kernel Transformation Example 2 (Animated)

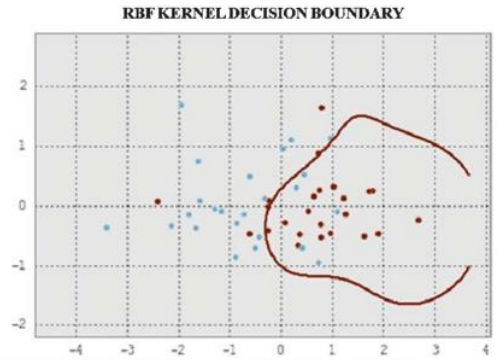




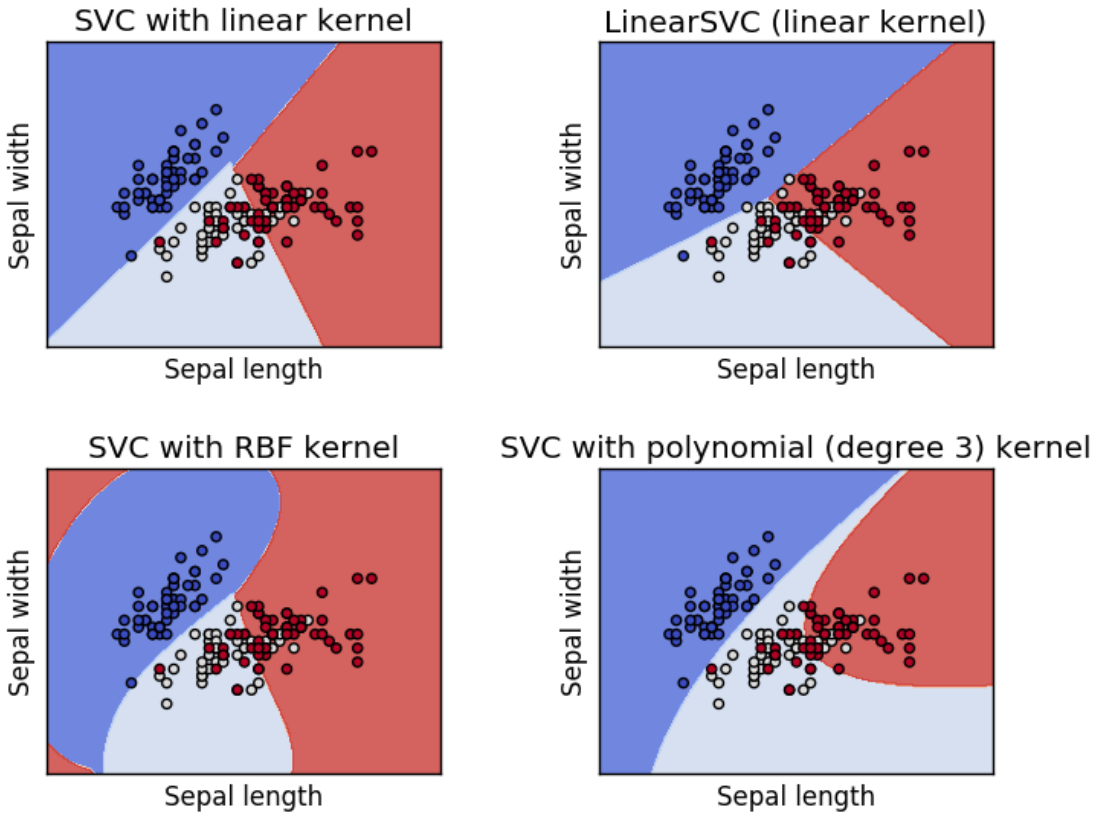
Kernels

- There are many kernel functions; most common:
  - Linear Kernel
  - Polynomial Kernel
  - Radial Basis Function (RBF) Kernel





Kernels



Plot different SVM classifiers in the iris dataset

https://scikit-learn.org/0.18/auto_examples/svm/plot_iris.html

SVM Advantages & Disadvantages

- Advantages
  - High accuracy
  - Works well with non-linear data.
- Disadvantages
  - Creating the model takes a long time (especially in high degree polynomial)
  - The kernel used can greatly change the accuracy of the model, so the right kernel must be chosen.

Thank You