Non - Linear SVM

Introduction to SVM: Why SVM?

- Working with neural networks for supervised and unsupervised learning showed good results while used for such learning applications.

- MLP's uses feed forward and recurrent networks.

- Multilayer perceptron (MLP) properties include universal approximation of continuous nonlinear functions and include learning with input-output patterns and also involve advanced network architectures with multiple inputs and outputs.

Introduction to SVM: Why SVM?

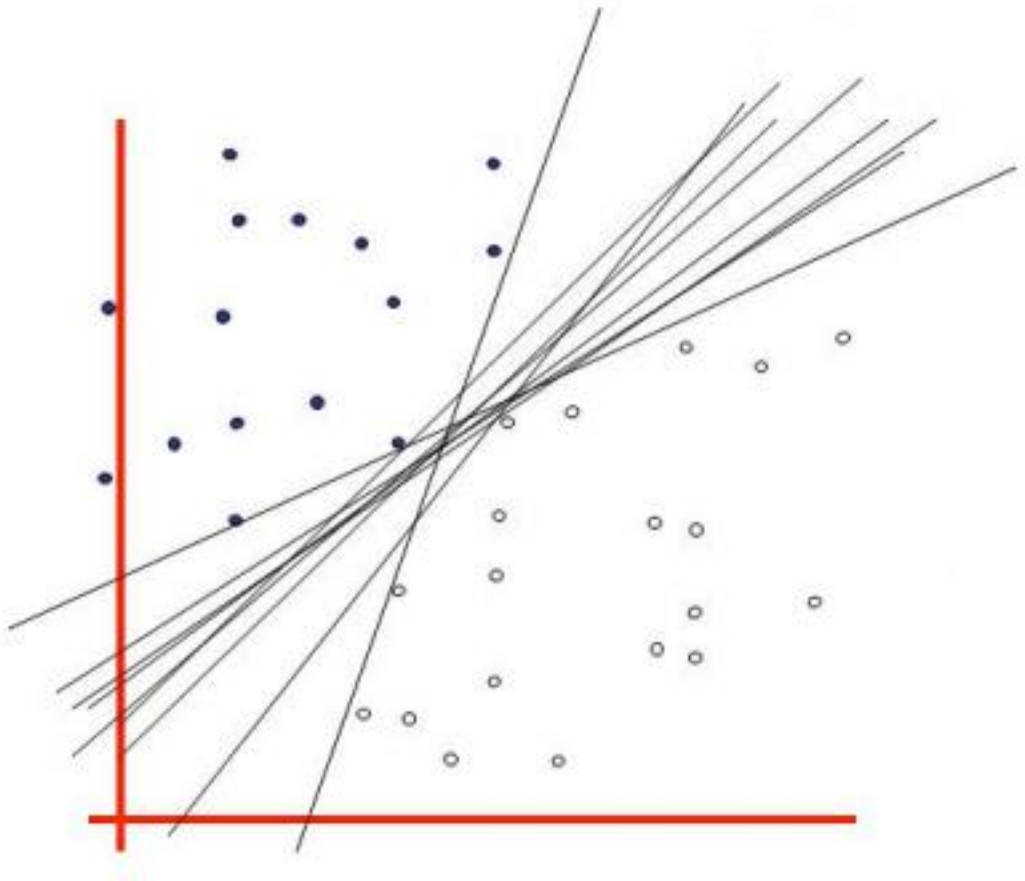
- There can be some issues noticed. Some of them are having many local minima and also finding how many neurons might be needed for a task is another issue which determines whether optimality of that NN is reached.

- Another thing to note is that even if the neural network solutions used tends to converge, this may not result in a unique solution.

Introduction to SVM: Why SVM?

- Now let us look at another example where we plot the data and try to classify it and we see that there are many hyper planes which can classify it.

- But which one is better?



Linear Classifier

- Decision boundary: Hyperplane

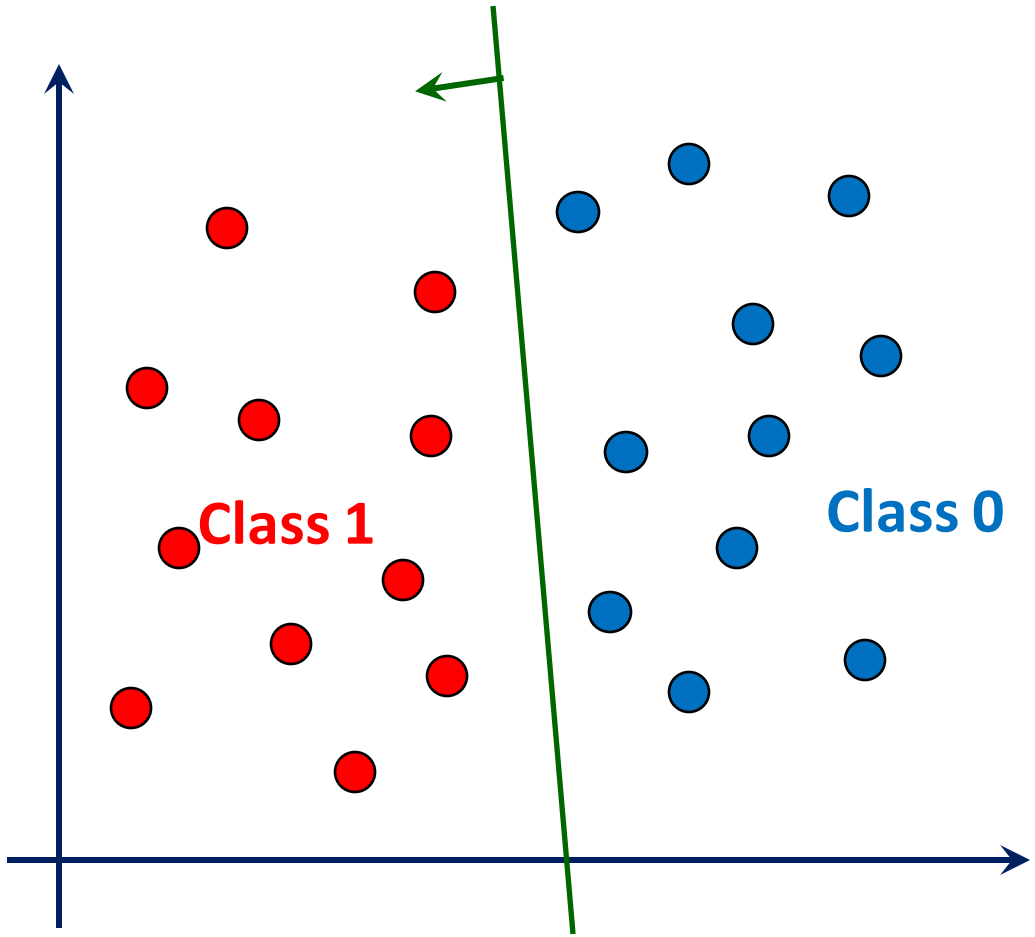
$$w^T x = 0$$

- Class 1 lies on the positive side

$$w^T x > 0$$

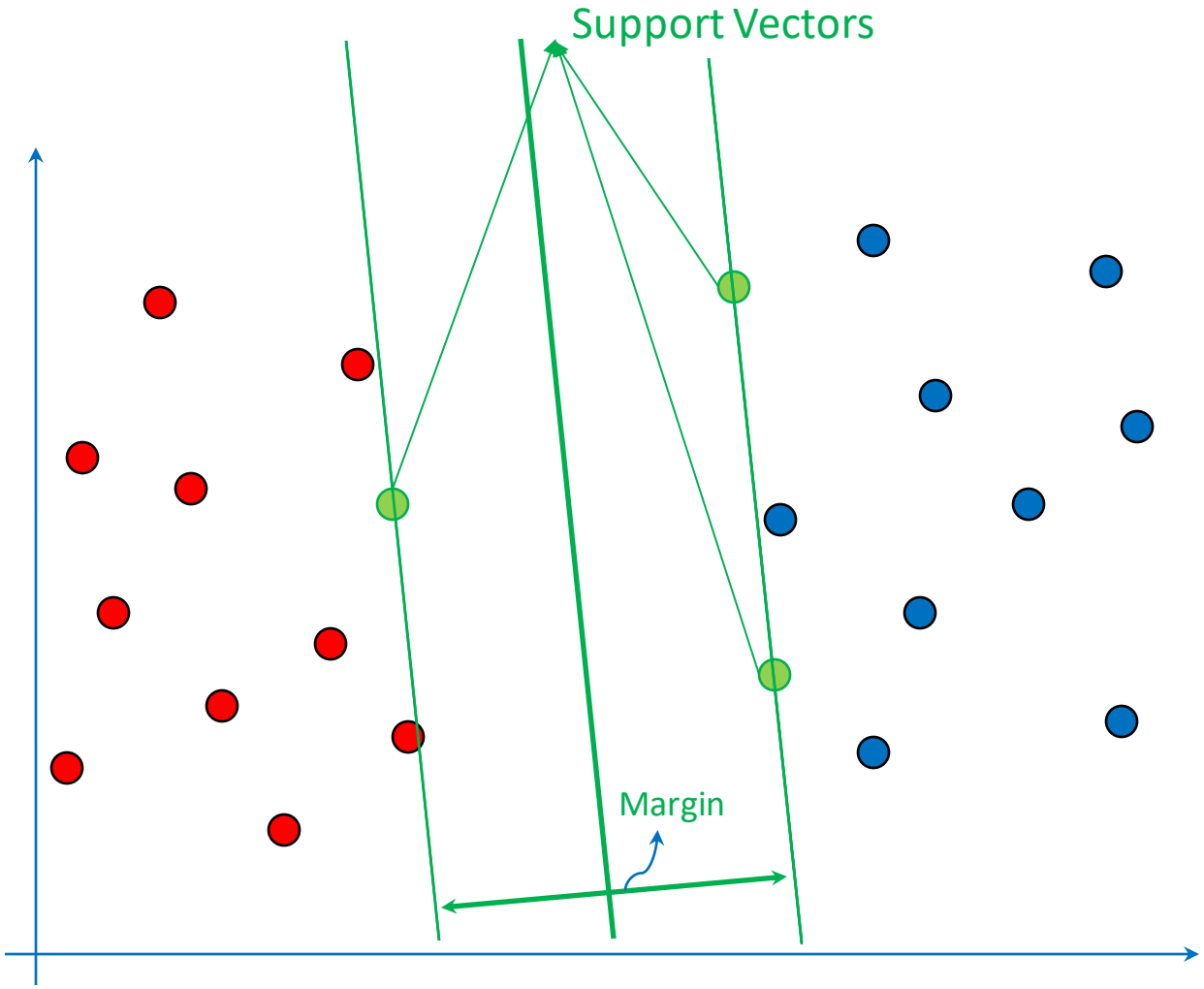
- Class 0 lies on the negative side

$$w^T x < 0$$



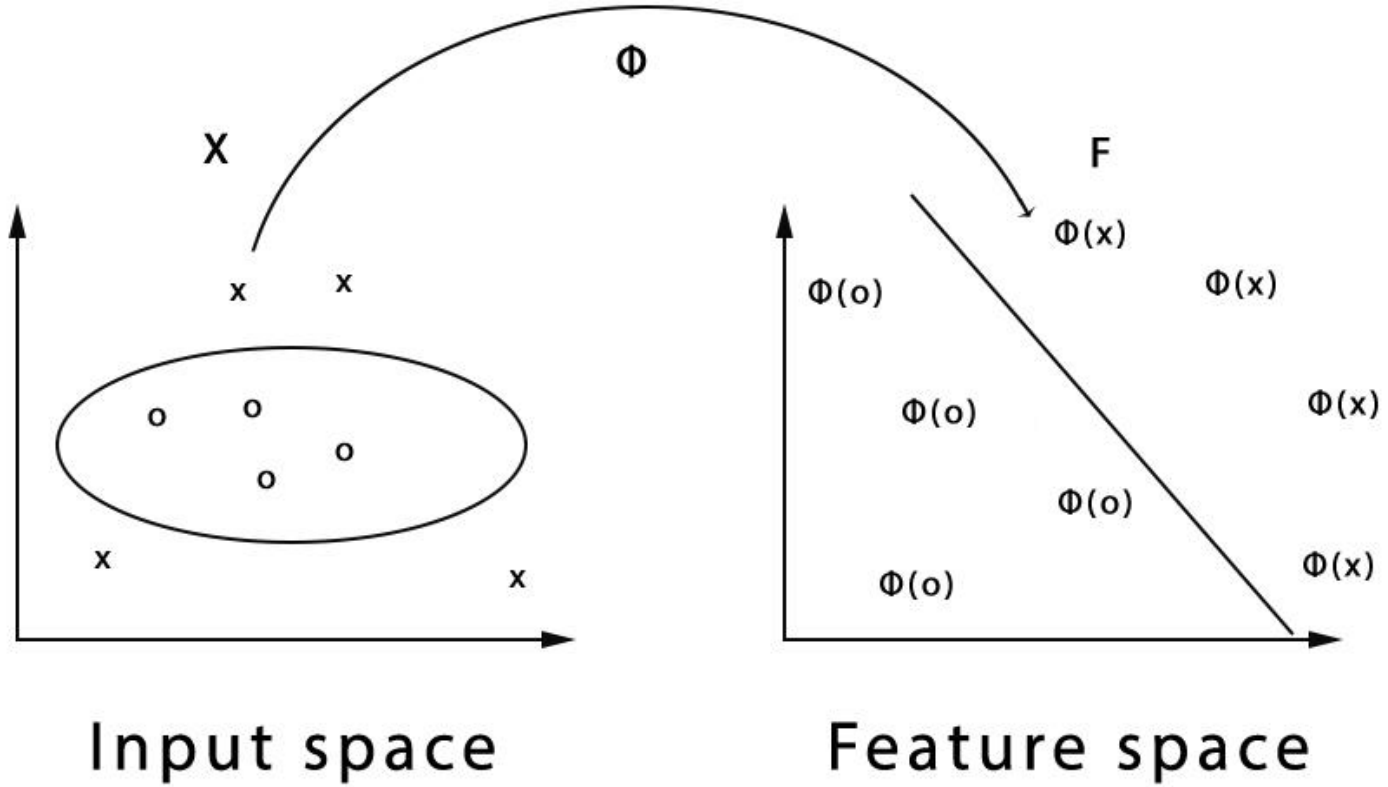
Summary: Max-Margin Classification

- A Large Margin will reduce the chance of misclassifying future test samples
- In other words, large-margin classifiers will generalize better.
- Samples at the boundary support the margin: called Support Vectors



Non Linear SVM

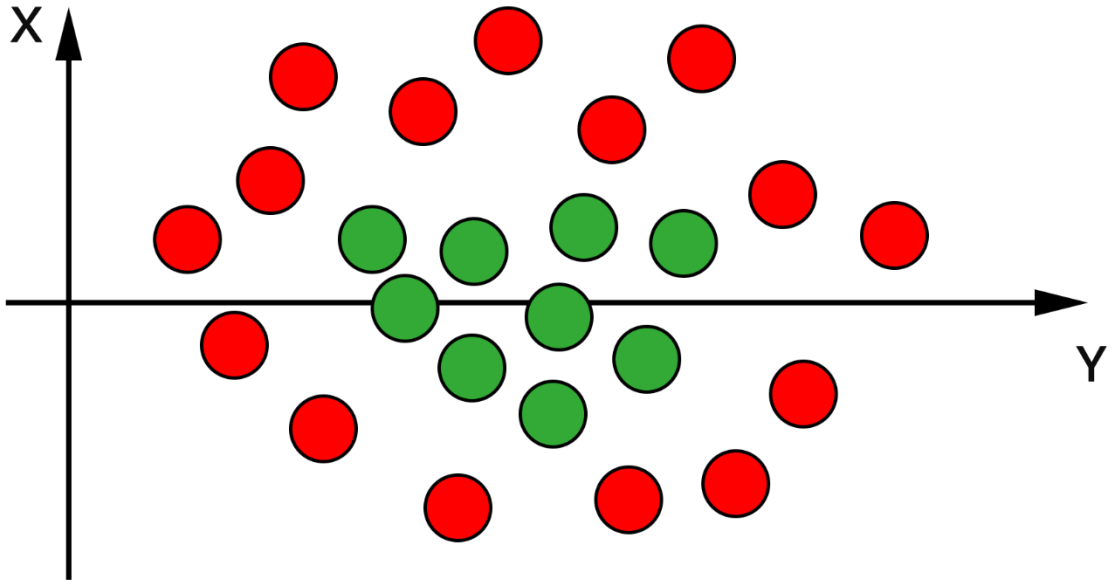
- Some data points are not linear separable
- **Intuition:** to transform the data to a high dimension space



Support Vector Machines

- SVM uses **kernel** to transform the data and then based on these transformations it finds an optimal hyperplane that distinctly classifies the data points.

- Example:
  - No Straight line that can separate the two groups

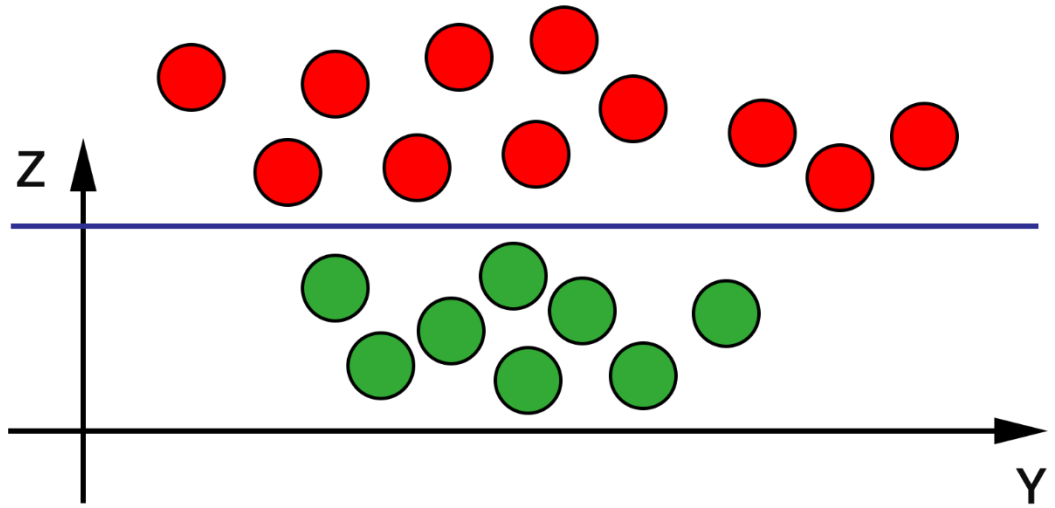


Transform

- Transform the data by adding one more dimension as z,
  - Where $z = x^2 + y^2$

- If we plot in z-y axis, a clear separation is visible and a line can be drawn.

- The line is the hyperplane we want and the transformation is a kind of kernel.



Code Available

Demo Experiments

Demo_Non_Linear_SVM_multi-class

The aim of this experiment is to perform binary and multi-class classification on Iris dataset and get the support vectors that are the basis for the max margin.

Demo_Non-Linear_SVM_RBF

The aim of this experiment is to understand how non-linear separable data can be visualized linearly in a higher dimensional space

Thanks!!

Questions?