MACHINE LEARNING

Classification

Muhammad Afif Hendrawan, S.Kom., M.T.

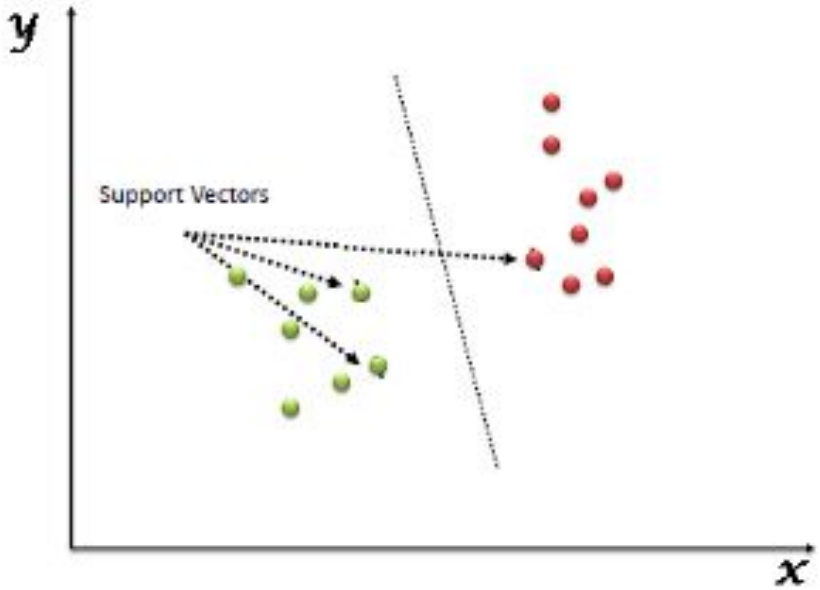
Outlines

- What is Support Vector Machines?
- The Kernel Tricks
- Maximum Margin Classification and Support Vectors

What is Support Vector Machine?

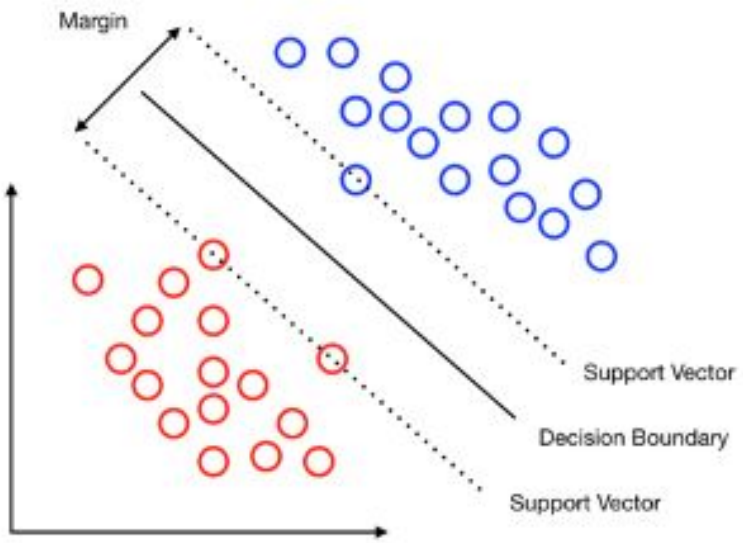
Intro (Again 😊) to SVM #1

- A supervised machine learning algorithm.
- Can be used for both classification (SVM) or regression (SVR), but it is mostly used for classification.
- Attempts to find the line that "best" separates two classes of points.
- "best“ means the line that results in the largest margin between the two classes.



Intro (Again 😊) to SVM #2

- In the SVM algorithm, plot each data item as a point in n-dimensional space (where n is number of features you have) with the value of each feature being the value of a particular coordinate.
- Perform classification by finding the hyper-plane that differentiates the two classes very well.



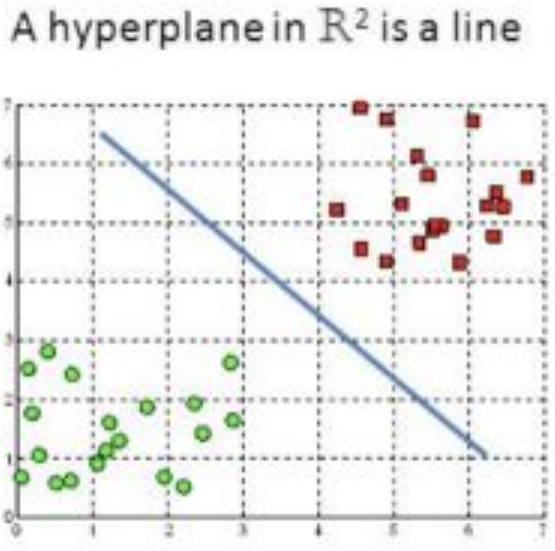
Why SVM?

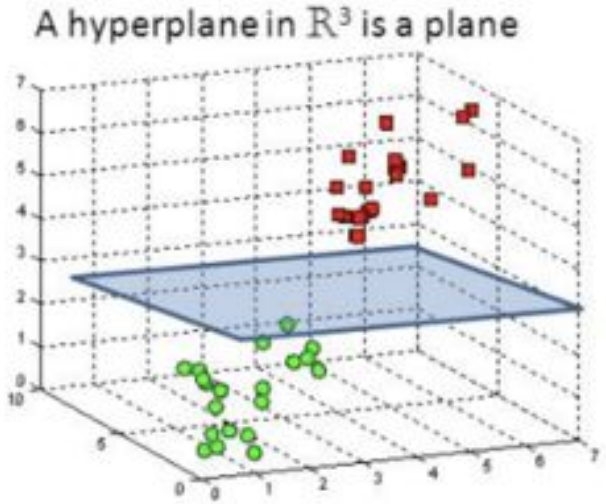
- Modeling a linear relationship between the response variable and the features in the higher dimensional feature space has two problems:
  - Computational problem, i.e. computing the mapped features and working with larger vectors requires more computing power.
  - Generalization, i.e. increasing the dimensions of the feature representation exacerbates the curse of dimensionality, because learning from high-dimensional feature representations requires exponentially more training data to avoid overfitting.

Hyperplanes

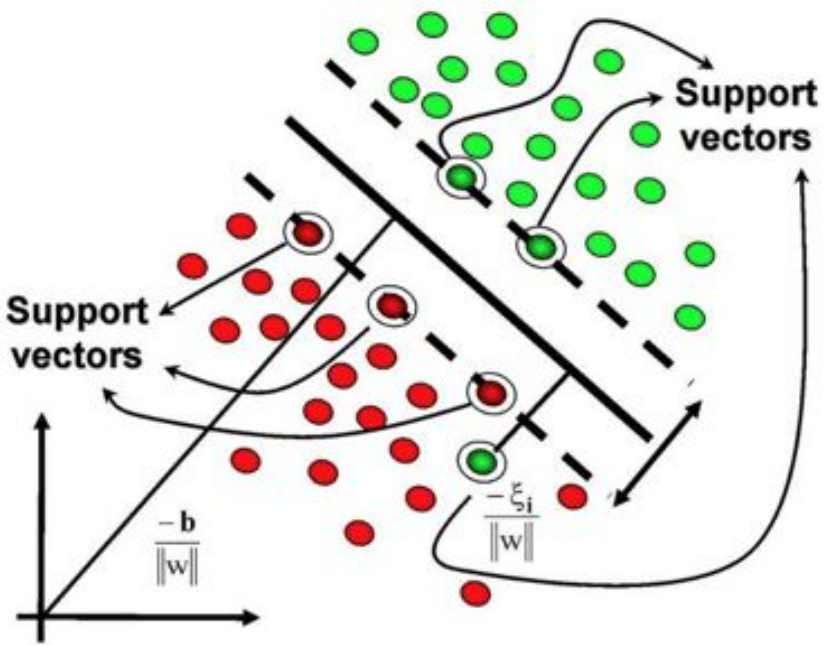
- Can be considered decision boundaries that classify data points into their respective classes in a multi-dimensional space.
- Data points falling on either side of the hyperplane can be attributed to different classes.
- A generalization of a plane:
  - in two dimensions, it's a line.
  - in three dimensions, it's a plane.
  - in more dimensions, you can call it a hyperplane.

Hyperplanes Illustration



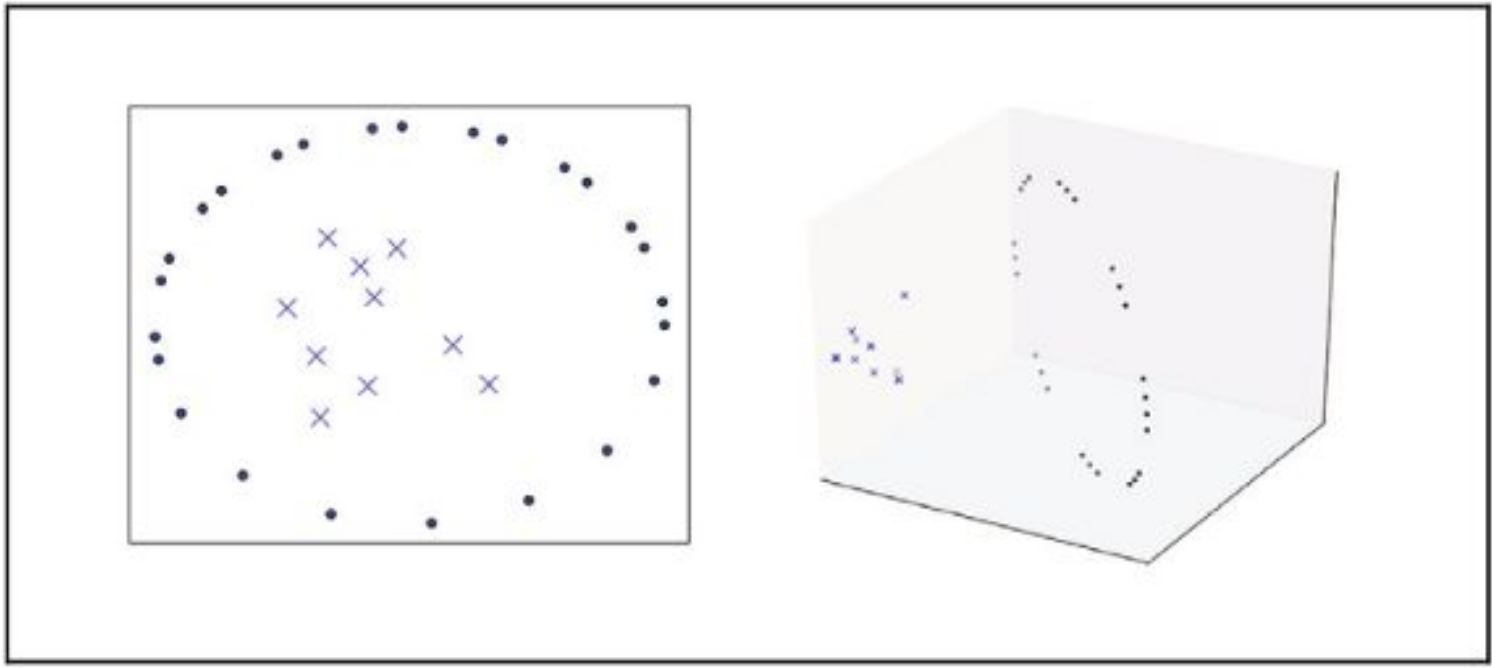


Support Vector Concept



- The points that lie on or fall within the margin.
- Used to formulate to find the "margin hyperplane“, the line that best separates the two classes.
- So-called as dual form.
- Can be used to apply the kernel trick to effectively transform the SVM into a non-linear classifier.

Mapping to Higher Dimensional Space

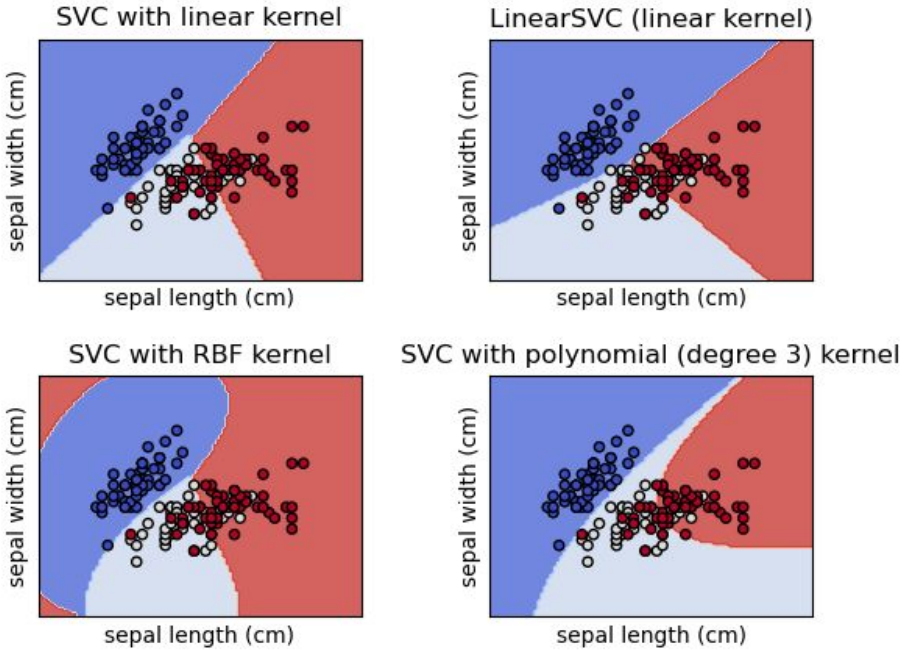


The Kernel Tricks

Kernel Tricks

- A kernel is a function that, given the original feature vectors, returns the same value as the dot product of its corresponding mapped feature vectors.
- Kernels do not explicitly map the feature vectors to a higher dimensional space or calculate the dot product of the mapped vectors.
- Kernels produce the same value through a different series of operations that can often be computed more efficiently.

Type of Kernels #1



Types of Kernels #2

- Polynomial

$$K(x, x') = (\gamma\langle x - x'\rangle + r)^k$$

- Sigmoid

$$K(x, x') = tanh(\gamma\langle x - x'\rangle + r)$$

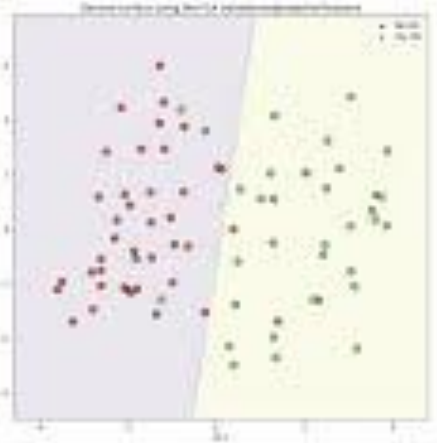
- Gaussian or Radial Basis Function (RBF)

$$K(x, x') = exp(-\gamma|x - x'|^2)$$

  - Where γ and r are hyperparameters that can be tuned from cross-validation.

Wrap Up! - SVM by StatQuest

Support Vector Machines in Python....



...From Start To Finish!!!

Let’s get your hands dirty!

Classification Practice!