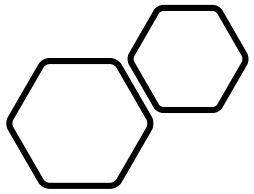
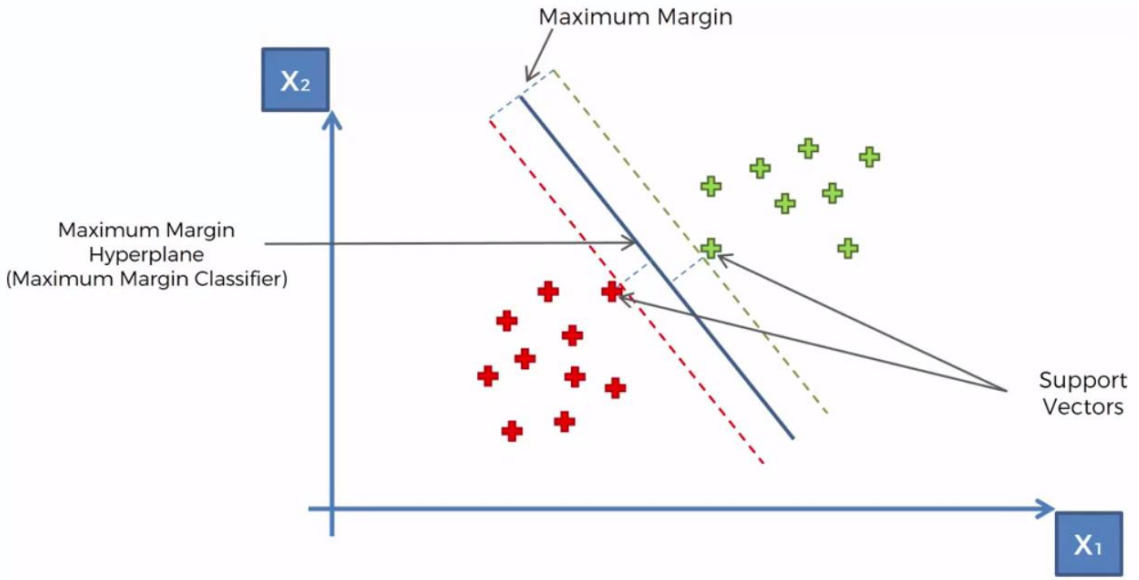


Basic Machine Learning: Support Vector Machine

Goal

Goal

Understanding the top five supervised algorithms which is Support Vector Machine algorithm.



Outline

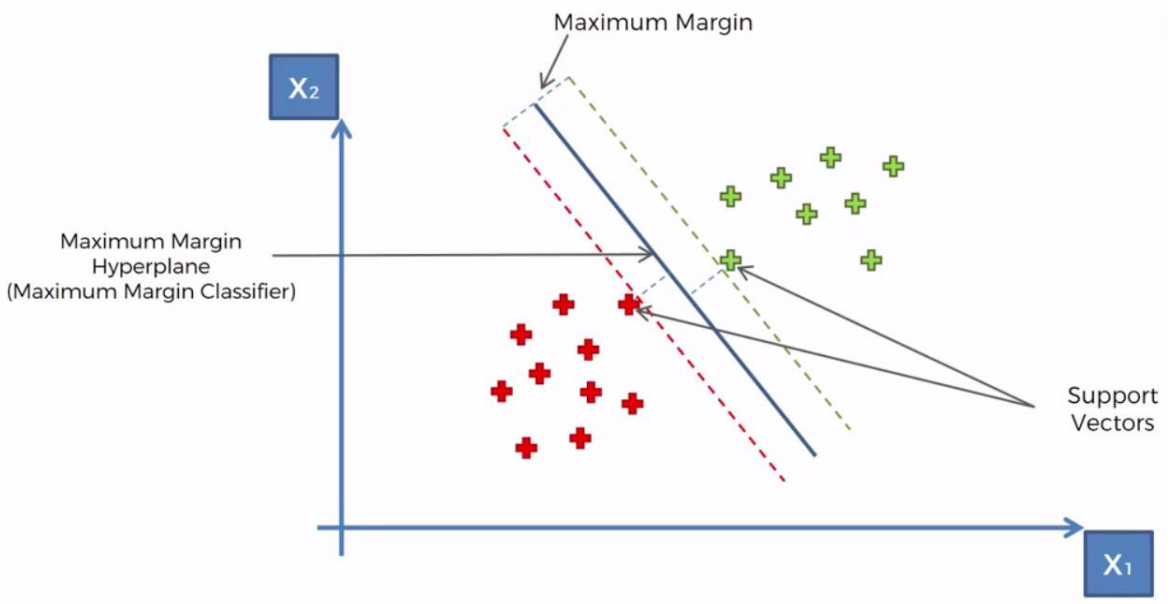
Outline

- Support Vector Machine Algorithm
  - Concept
  - Application

Content

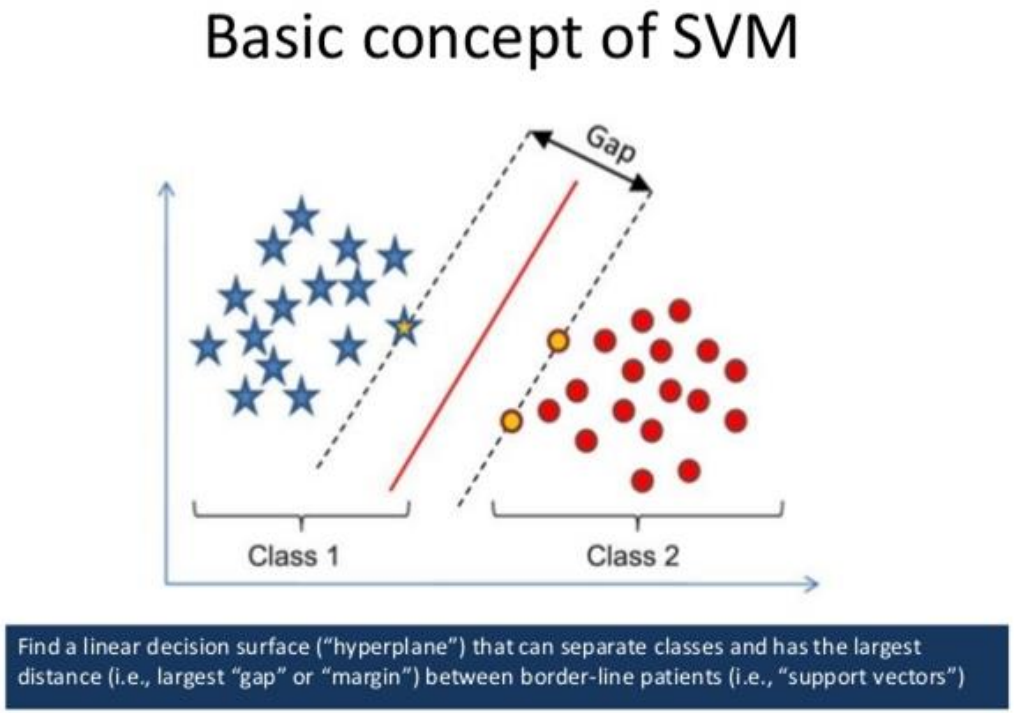
Support Vector Machine (Concept)

- Top 5 supervised algorithms
- Can be used for Classification & Regression problems
- Can generalize linear & nonlinear model
- Names in scikit-learn:
  - SVC (SVM Classifier)
  - SVR (SVM Regressor)



Support Vector Machine (Concept)

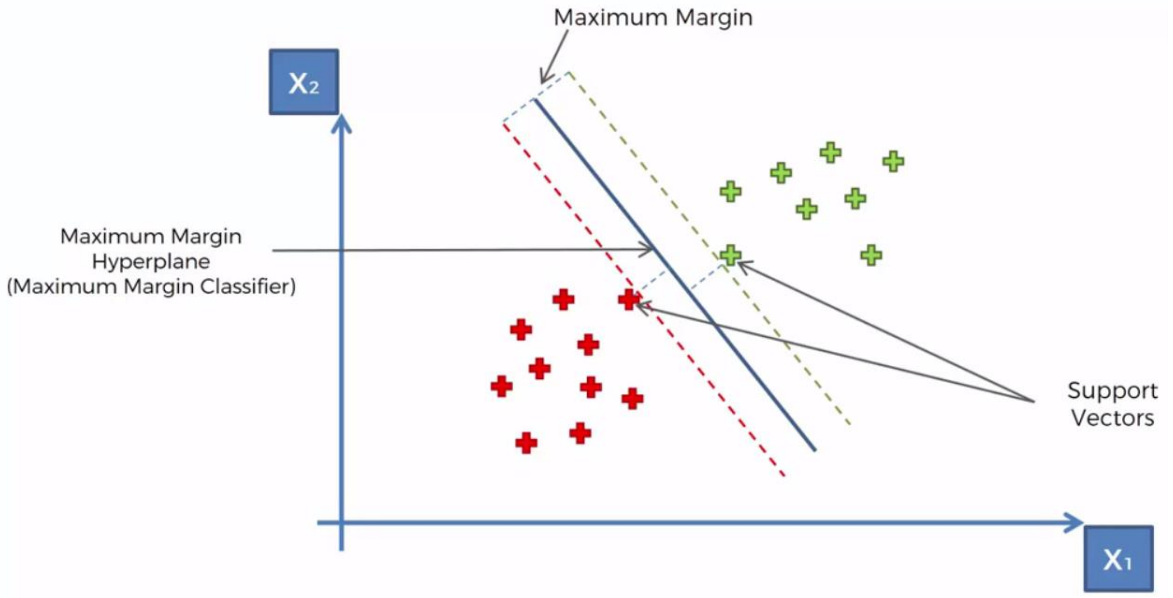
- Find solution (hyperplane) with fewest errors
- Maximize margin separator to improve generalization



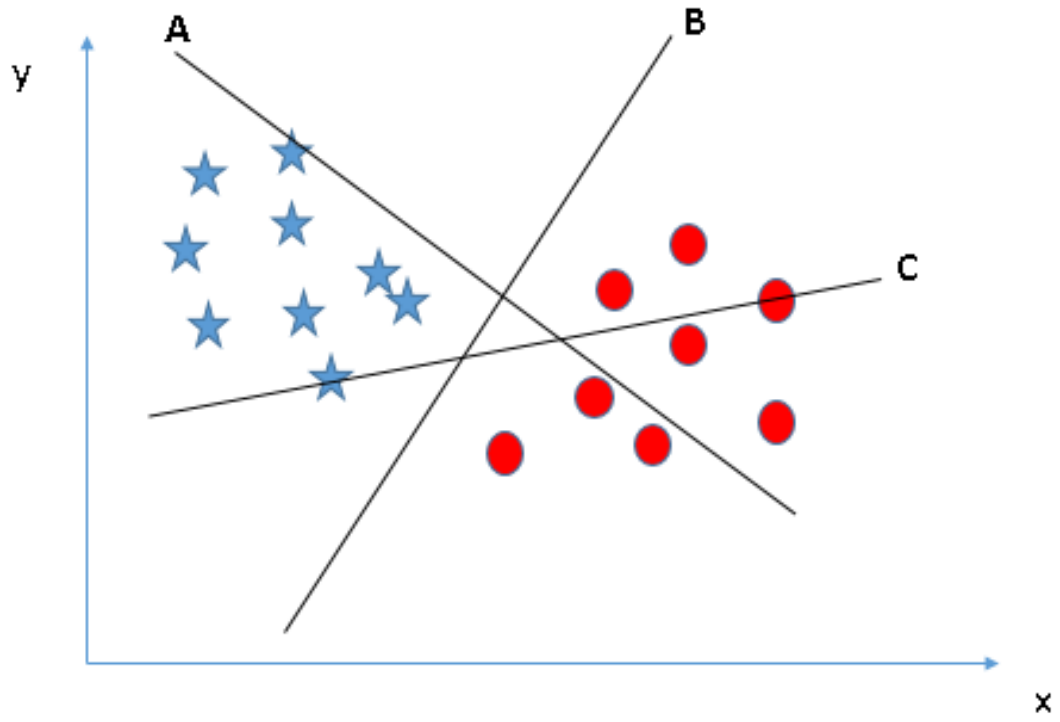
Find a linear decision surface ("hyperplane") that can separate classes and has the largest distance (i.e., largest "gap" or "margin") between border-line patients (i.e., "support vectors")

Support Vector Machine (Concept)

- What is Support Vector?
- The borderline data points that is “supporting” the construction of the hyperplane
- More support vectors doesn’t mean better
- What is Margin?
- (Perpendicular) gap between classes
- Higher margin = better separating hyperplane

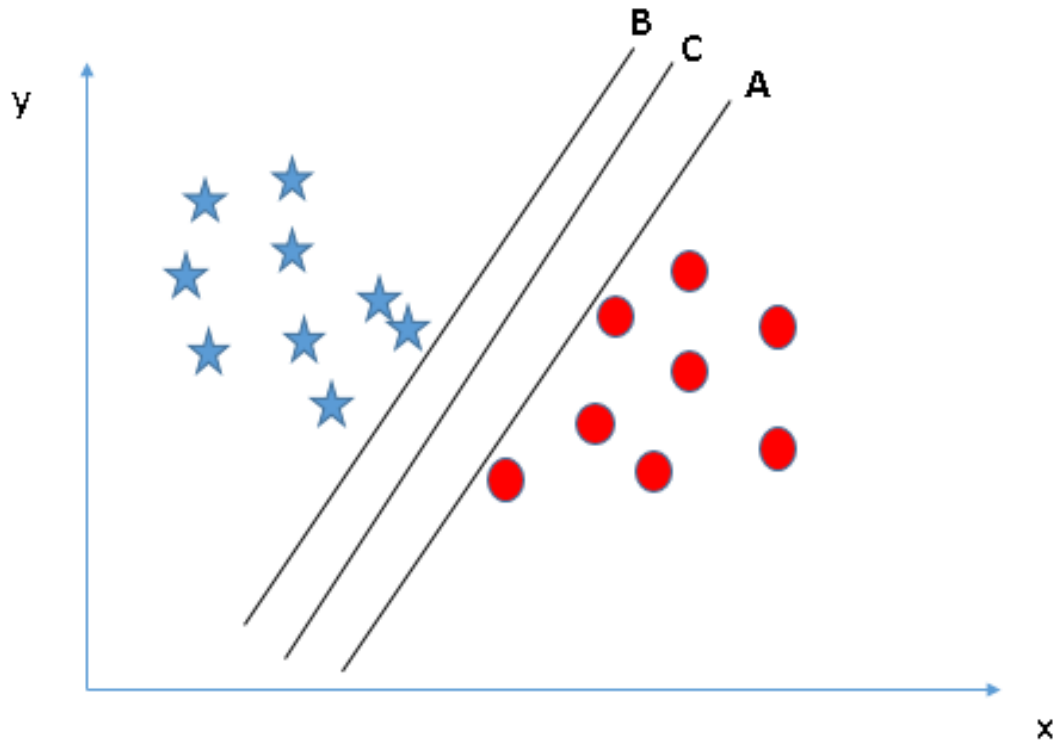


Support Vector Machine (Concept)



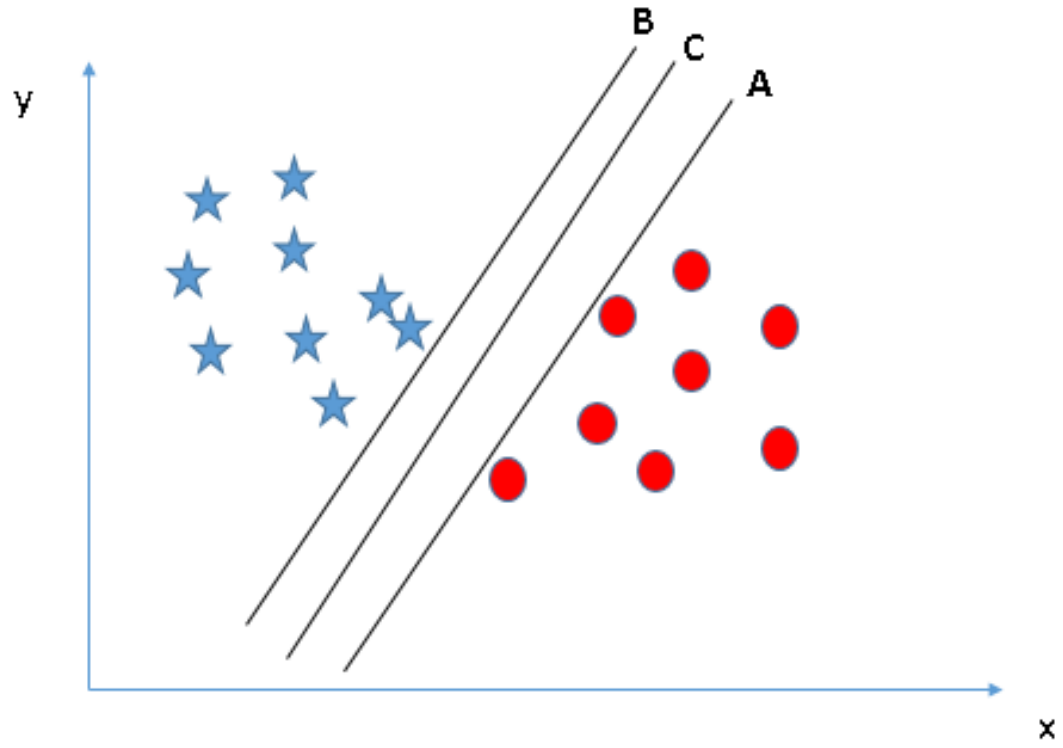
Scenario 1

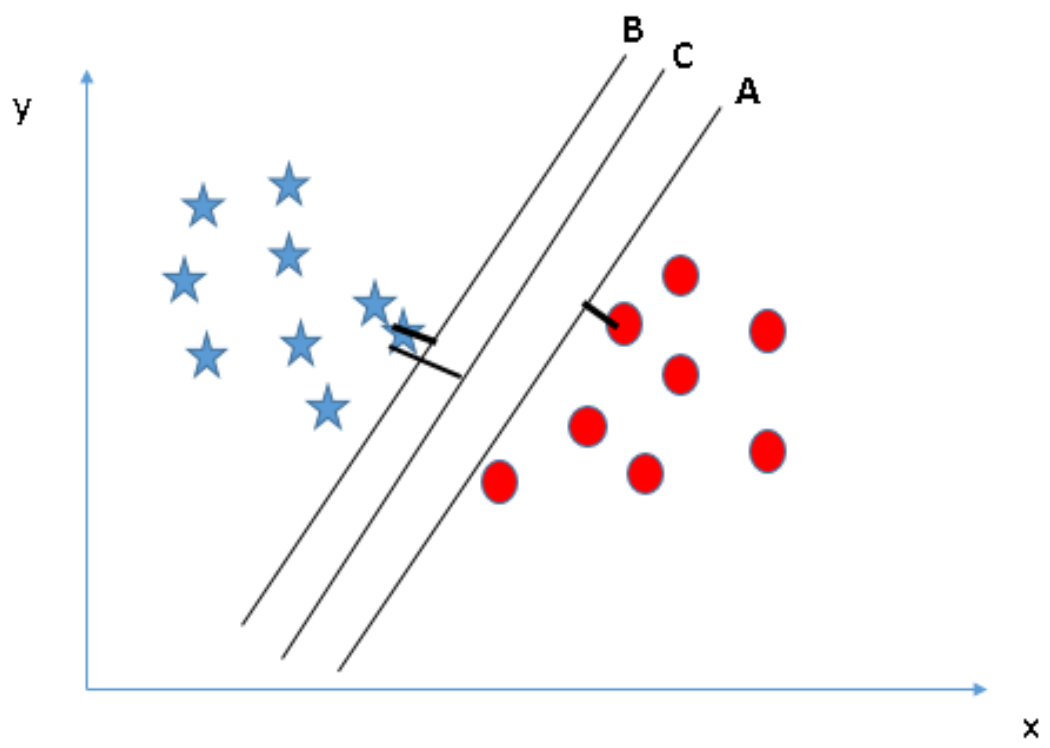
Support Vector Machine (Concept)



Scenario 2

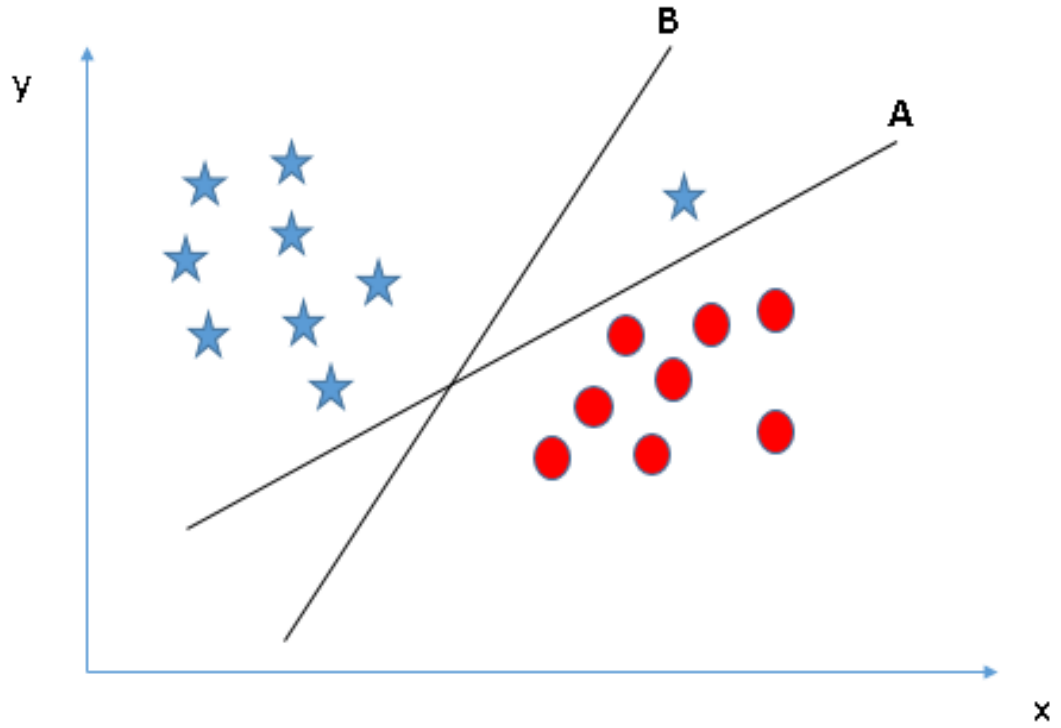
Support Vector Machine (Concept)





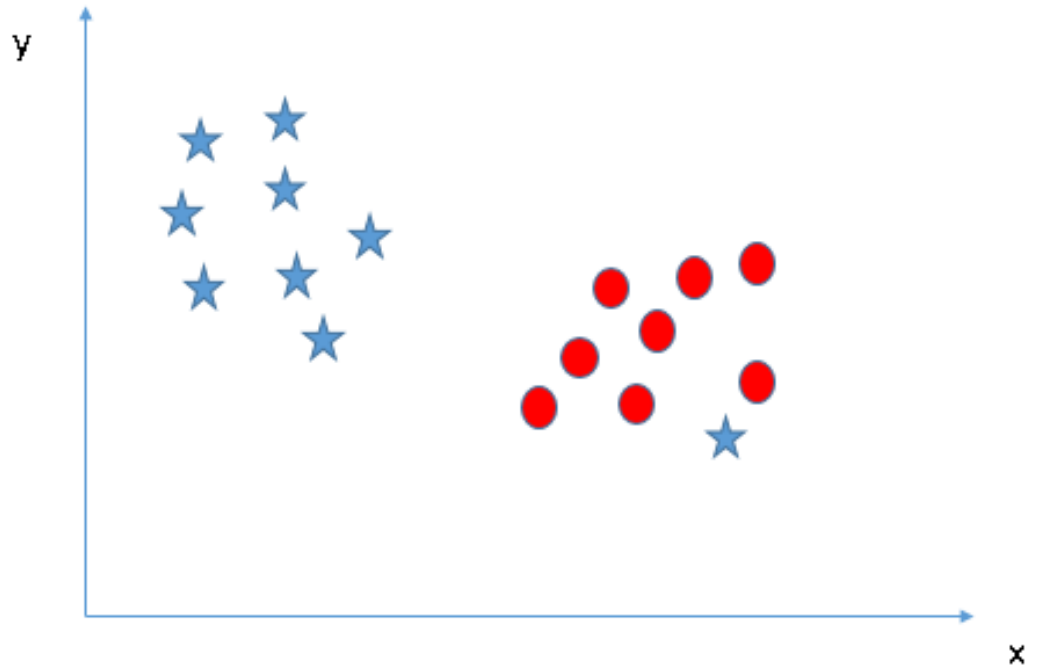
Scenario 2

Support Vector Machine (Concept)



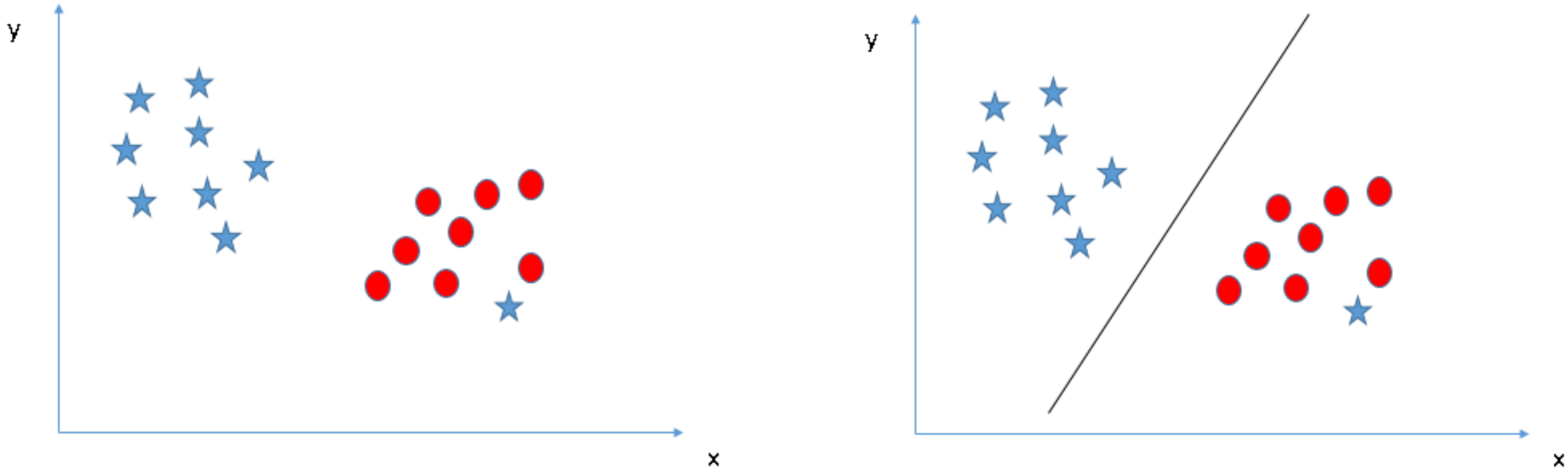
Scenario 3

Support Vector Machine (Concept)



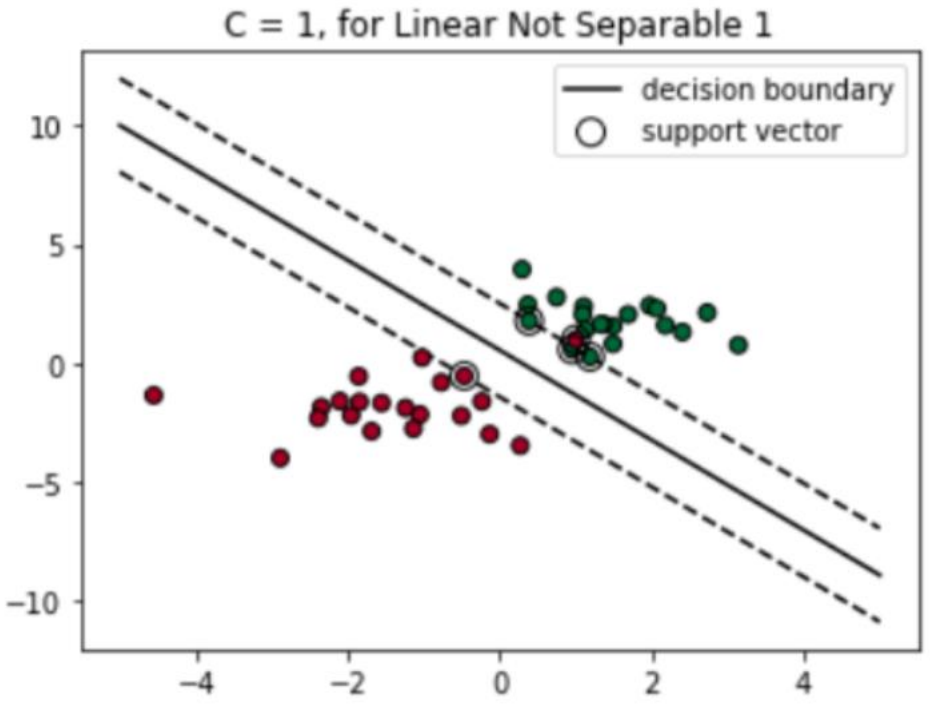
Scenario 4

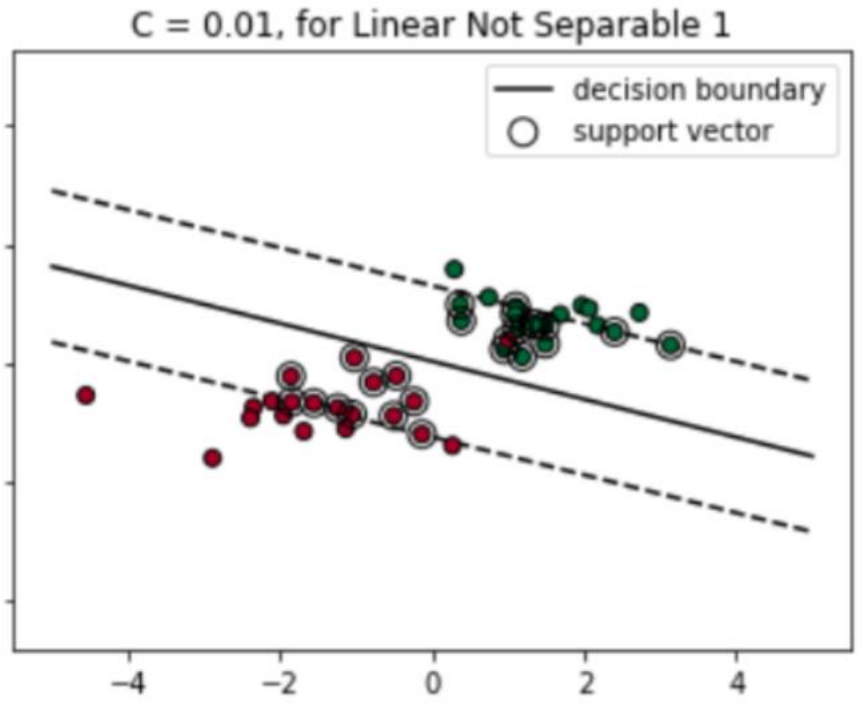
Support Vector Machine (Concept)



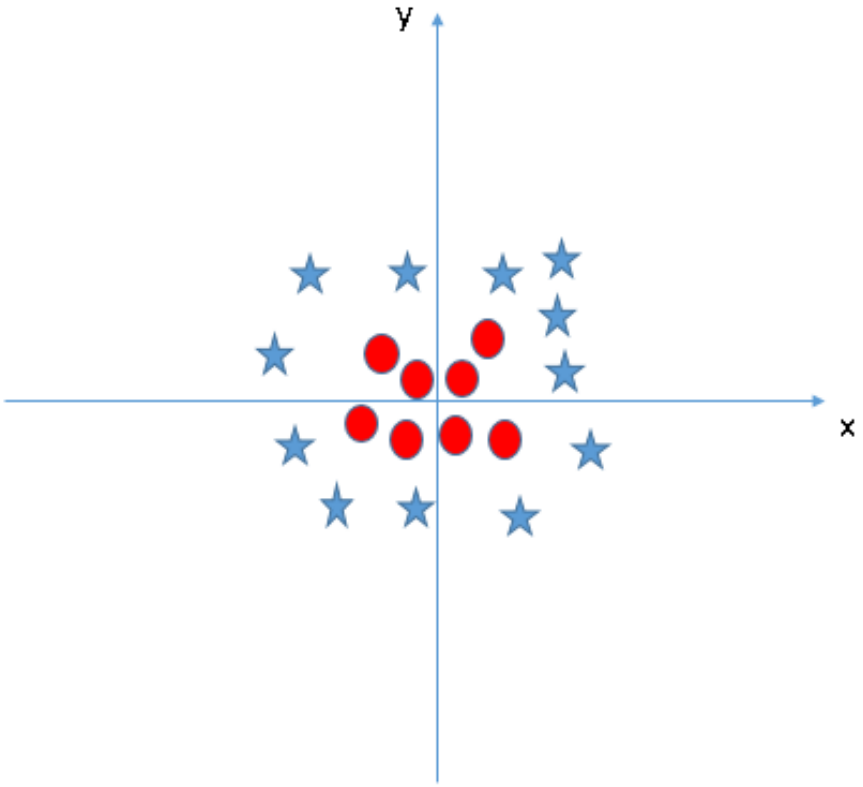
Scenario 4

Support Vector Machine (Concept)



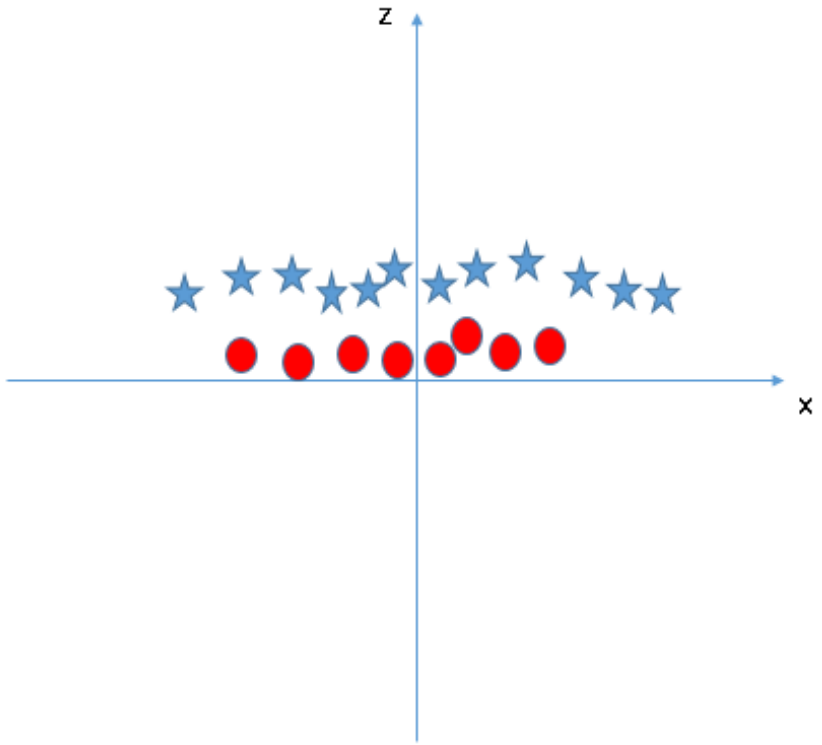


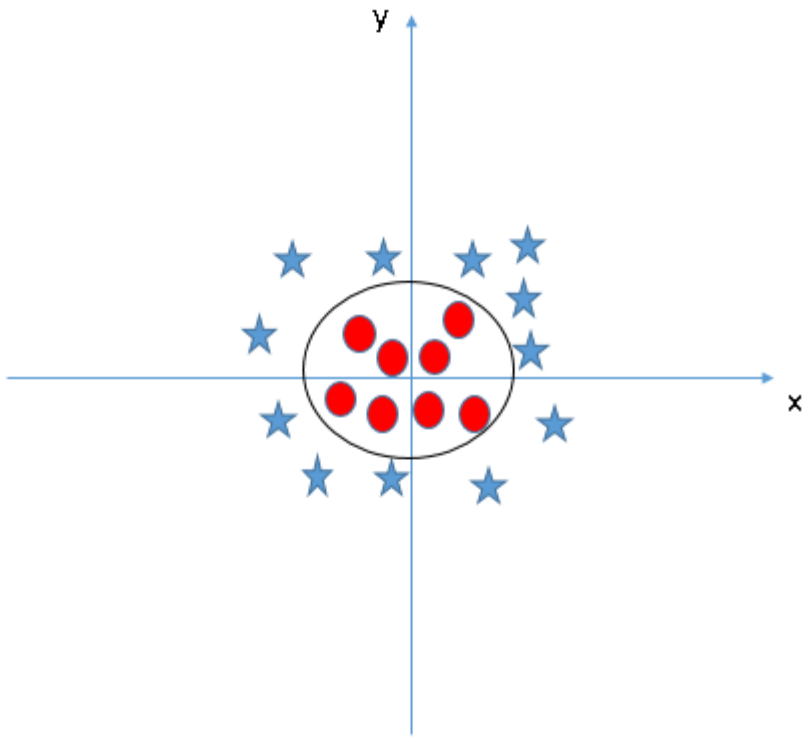
Support Vector Machine (Concept)



Scenario 5

Support Vector Machine (Concept)



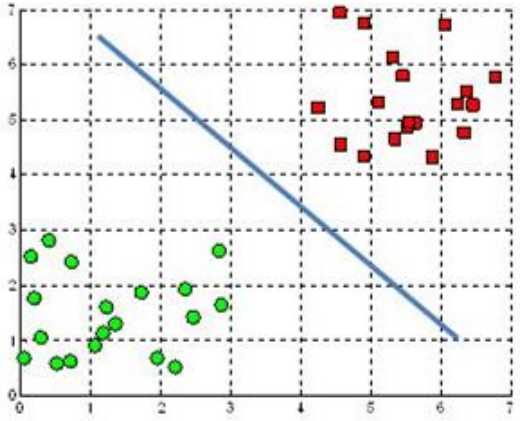


Scenario 6

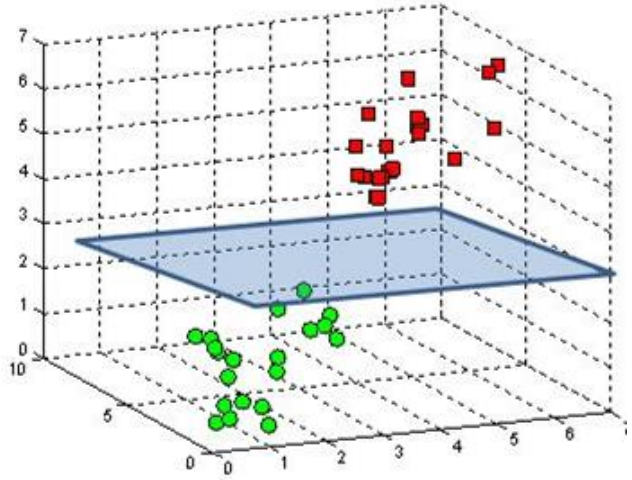
Support Vector Machine (Concept)

- What is Decision Surface?
- Hyperplane is a linear decision surface that splits the space into two parts
- Hyperplane is a binary classifier

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



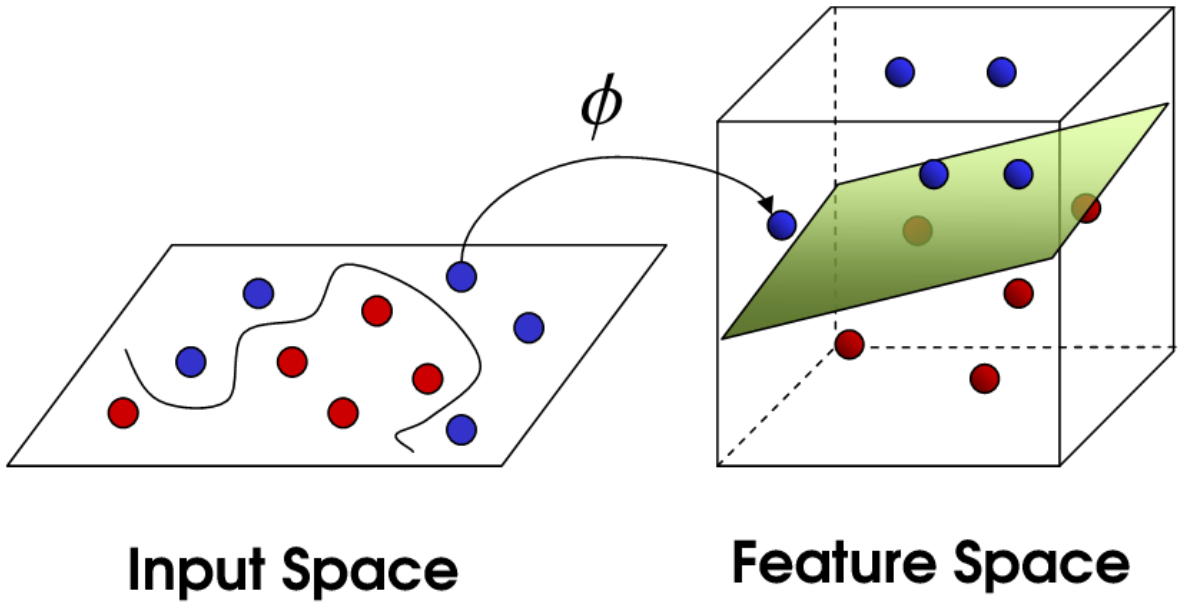
A hyperplane in $\mathbb{R}^3$ is a plane



A hyperplane in $\mathbb{R}^n$ is an n-1 dimensional subspace

Support Vector Machine (Concept)

- Sometimes our data is linearly separable, But sometimes it is not!



Support Vector Machine (Concept)

- What is SVM Kernel?
- Kernel: Mapping function that transforms a given space into some other (usually very high dimensional space)
- They can provide both linear and nonlinear model

Linear Kernel

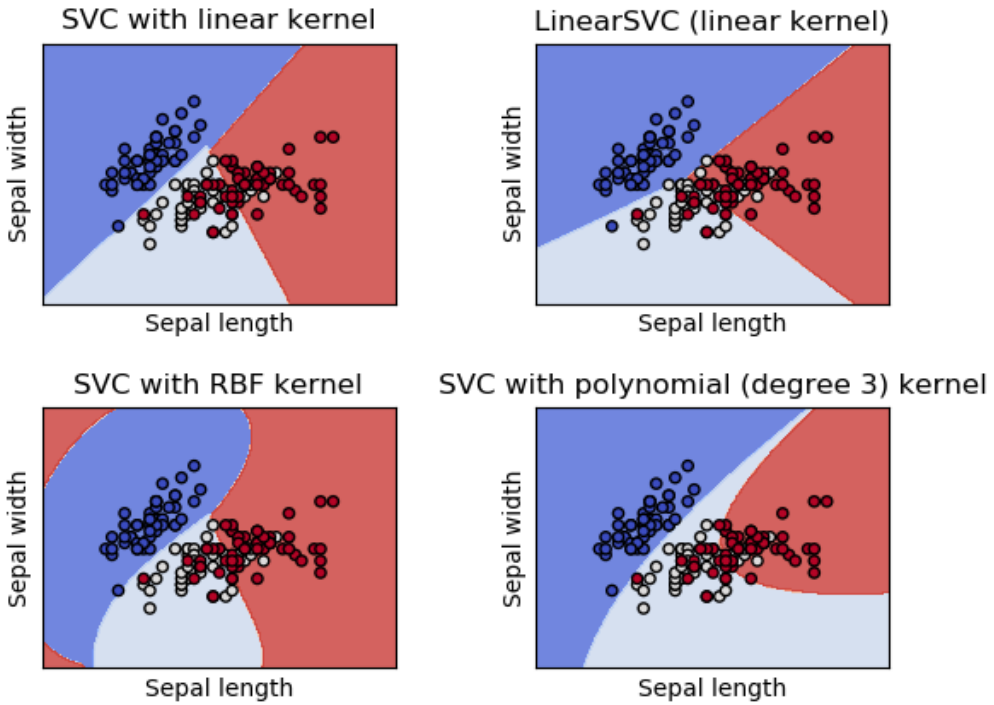
$$K(x_i, x_j) = x_i \cdot x_j$$

Polynomial Kernel

$$K(x_i, x_j) = (x_i \cdot x_j + c)^d$$

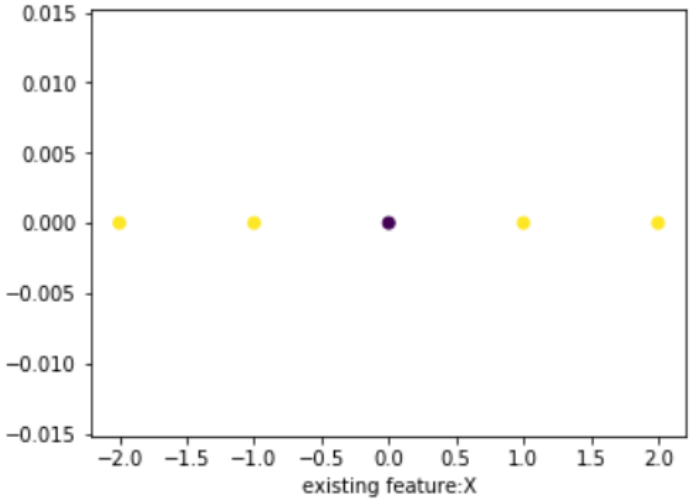
RBF Kernel

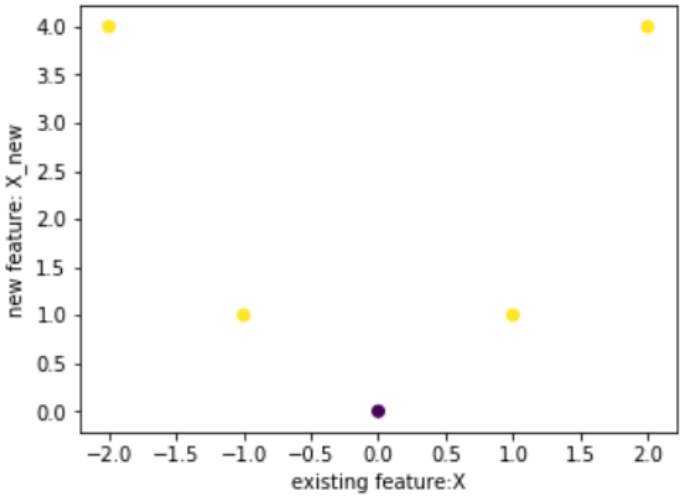
$$K(x_i, x_j) = exp(-\gamma ||x_i - x_j||^2)$$



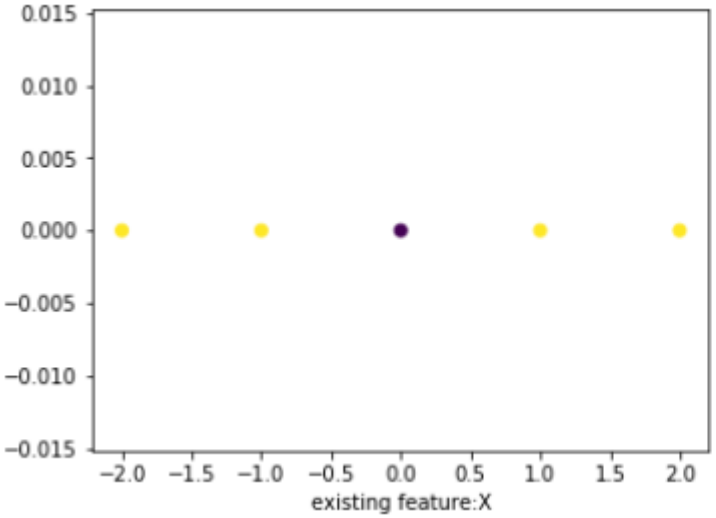
Support Vector Machine (Concept)

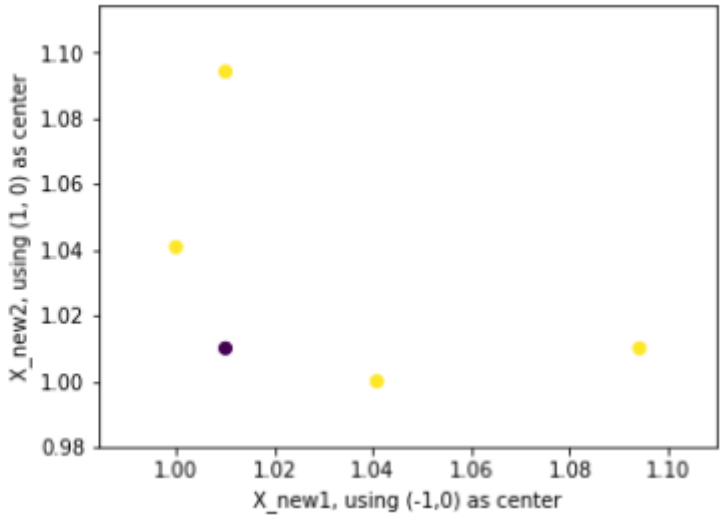
Polynomial Kernel



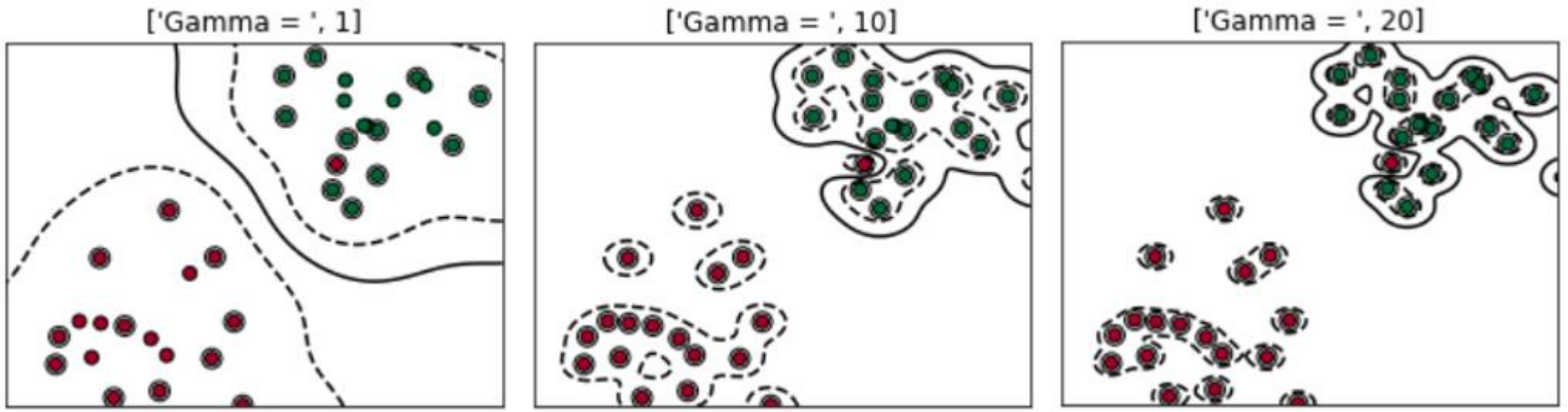


RBF Kernel



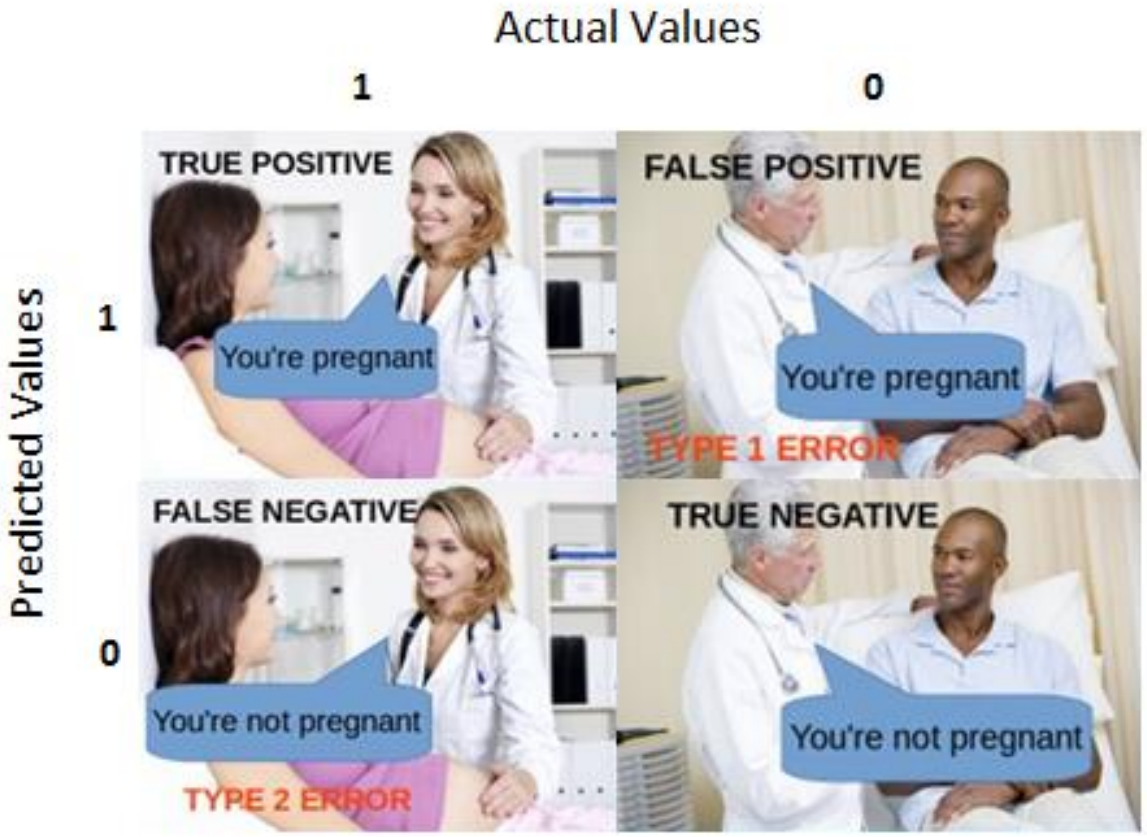


Support Vector Machine (Concept)

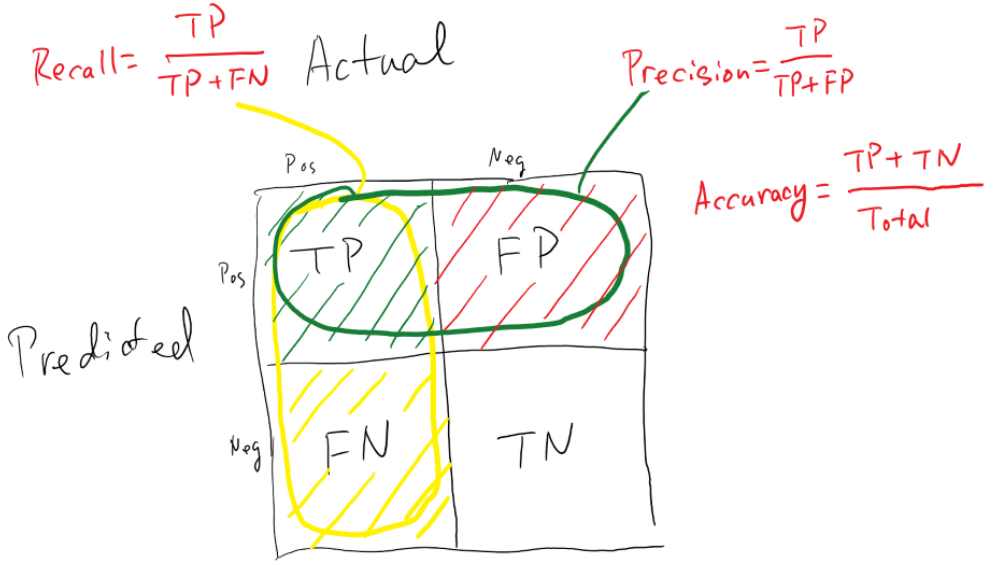


RBF Kernel

Confusion Matrix



Confusion Matrix		Target			
		Positive	Negative		
Model	Positive	a	b	*Positive Predictive Value*	a/(a+b)
	Negative	c	d	*Negative Predictive Value*	d/(c+d)
		Sensitivity	*Specificity*	**Accuracy** = (a+d)/(a+b+c+d)	
		a/(a+c)	d/(b+d)		



Thanks!