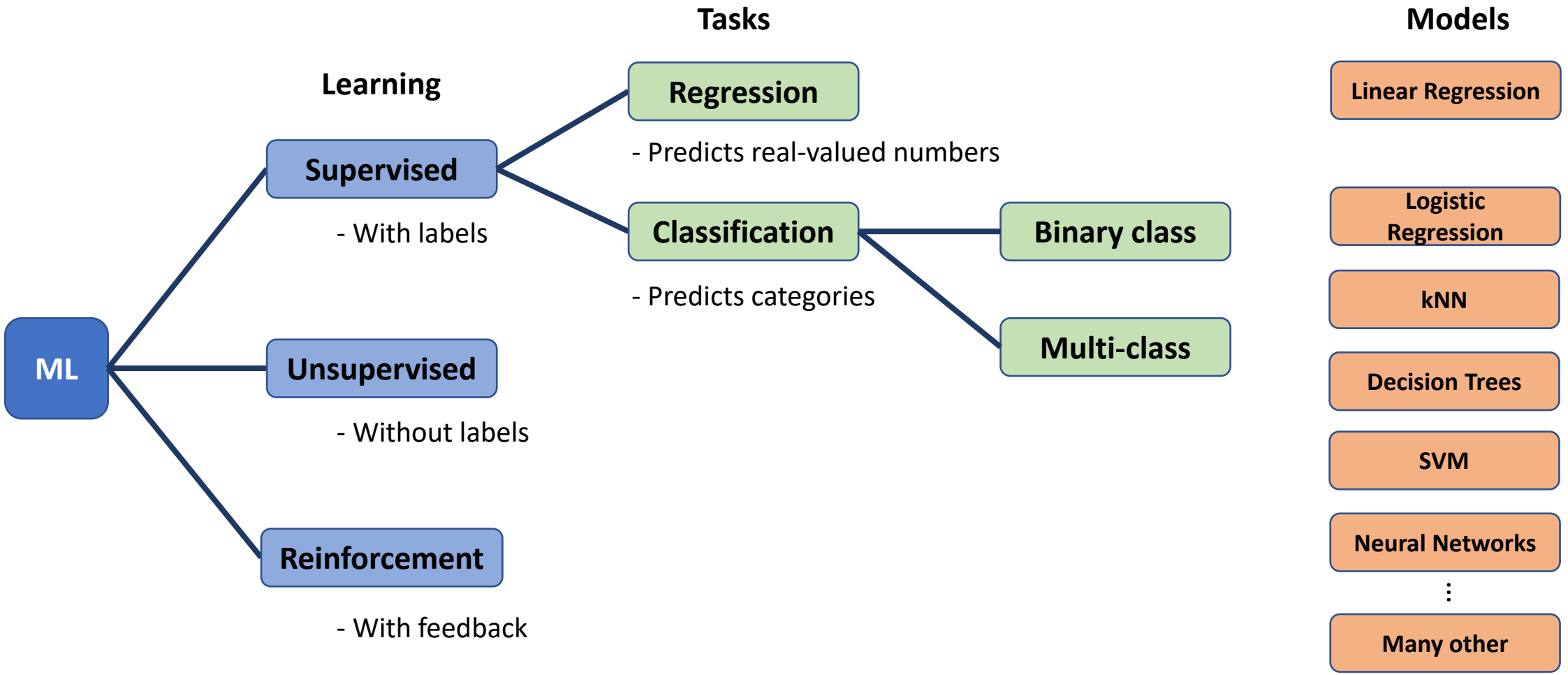
Support Vector Machine

Review: Types of machine learning problems



Review: Types of machine learning problems

Models	Hyperparameters	Parameters	Loss or Criteria
Linear Regression	X	$y = w_1 x_1 + w_2 x_2 + w_0$	MSE / RSS
Logistic Regression	X	$z = \sum w_i x_i$ $\sigma(z)$	BCE
kNN	k	X	
Decision Trees			
SVM			
Neural Networks			

Review: Types of machine learning problems

Models	Hyperparameters	Parameters	Loss or Criteria
Linear Regression	X	$y = w_1 x_1 + w_2 x_2 + w_0$	MSE / RSS
Logistic Regression	X	$z = \sum w_i x_i$ $\sigma(z)$	BCE
kNN	k	X	
Decision Trees	max depth α # tree lr	X	gini entropy MSE / RSS – regression
SVM	C	X	→
Neural Networks	✓	✓	✓

Support Vector Machine

- Use hyperplane
- Use kernel
- Performs well
- Regression, Classification (binary class)

Review: Binary Classification

Yes or No problem

- Creditcard Default
- Fradulant Insurance Claim
- Spam Filtering
- Medical Diagnosis
- Survival Prediction
- Customer Retention
- Image Recognition
- Sentiment analysis

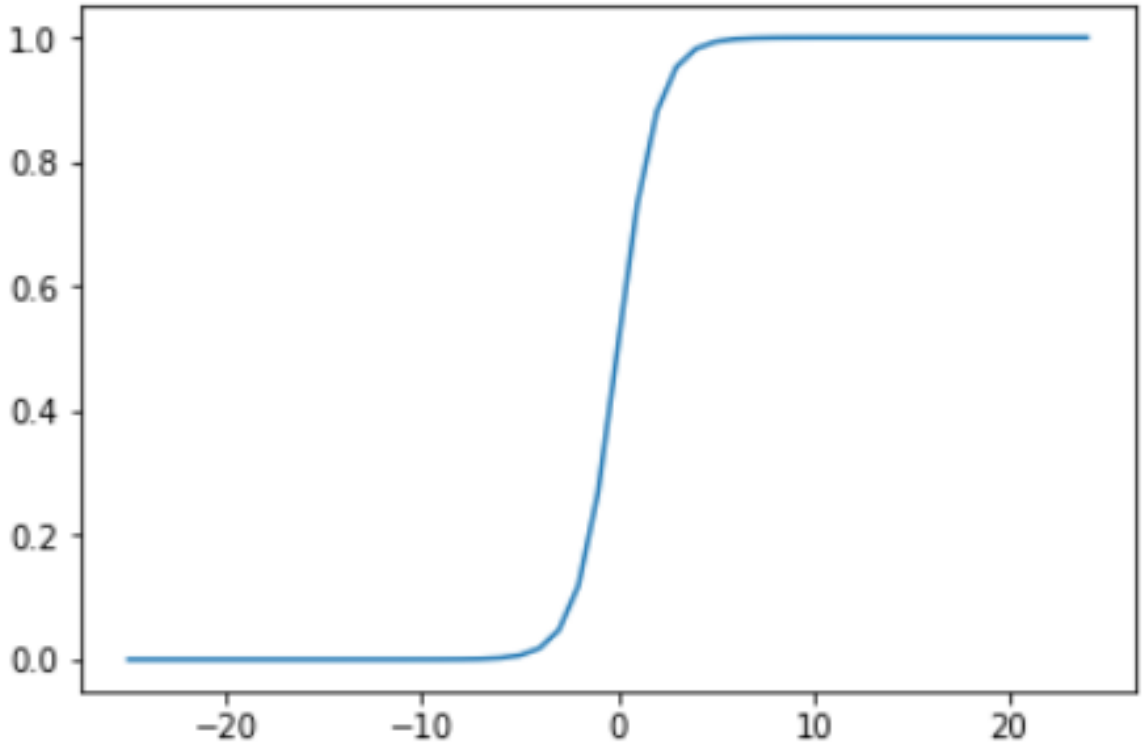
Review: Logistic Function

$$P^{(i)} \in \mathbb{R}[0, 1]$$

$$P^{(i)} = \sigma(z^{(i)})$$

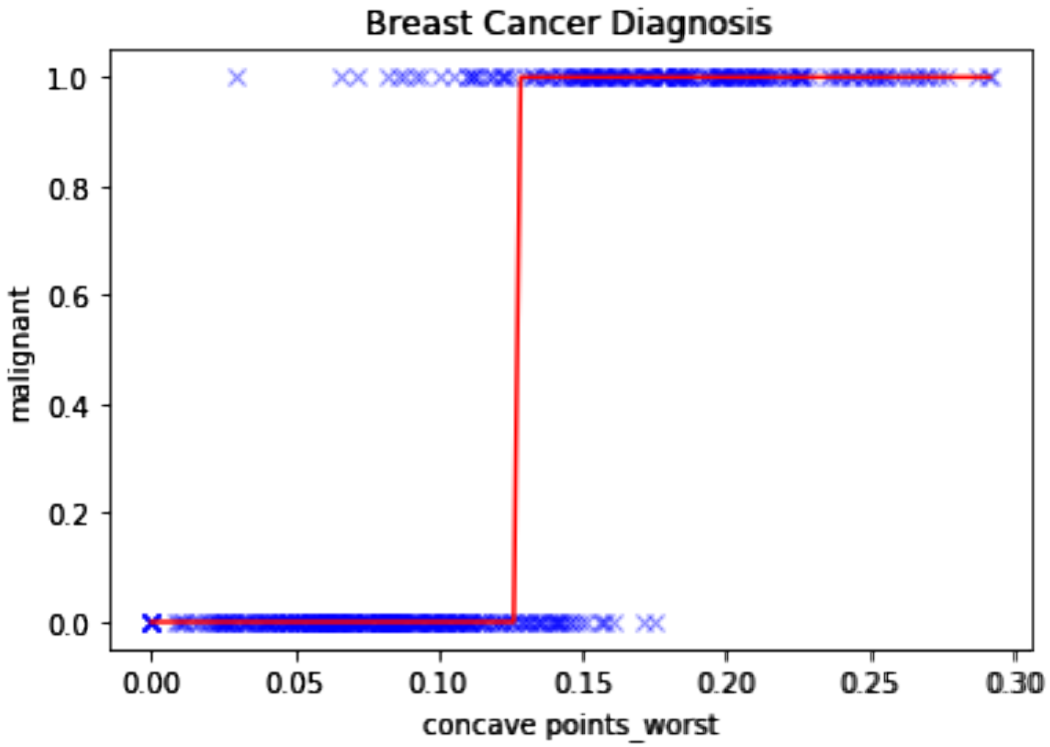
$$\sigma(z) = \frac{1}{1 + e^{-z}}$$

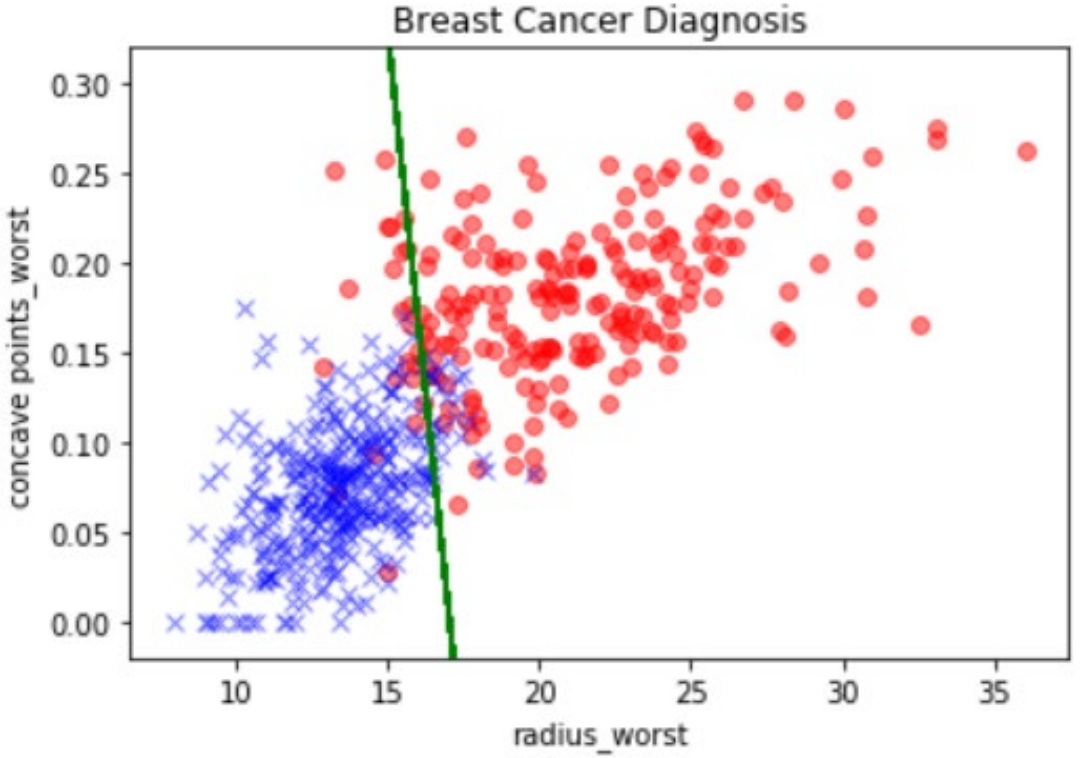
$$z^{(i)} = \boldsymbol{W} \cdot \boldsymbol{X} + b$$



Called “logit” and is related to the decision boundary

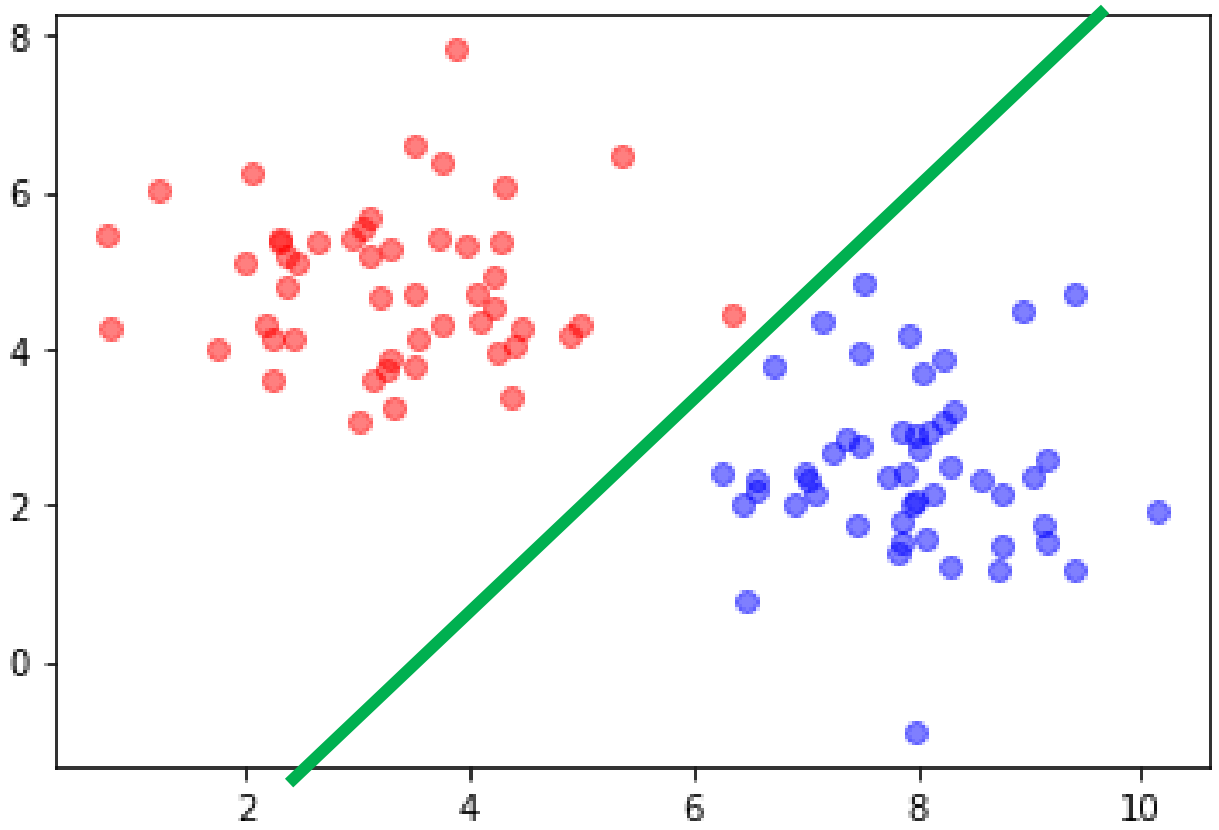
Review: Logistic Regression Decision Boundary





$z = 0.443\ x1 + 2.76\ x2 - 7.57 = 0$

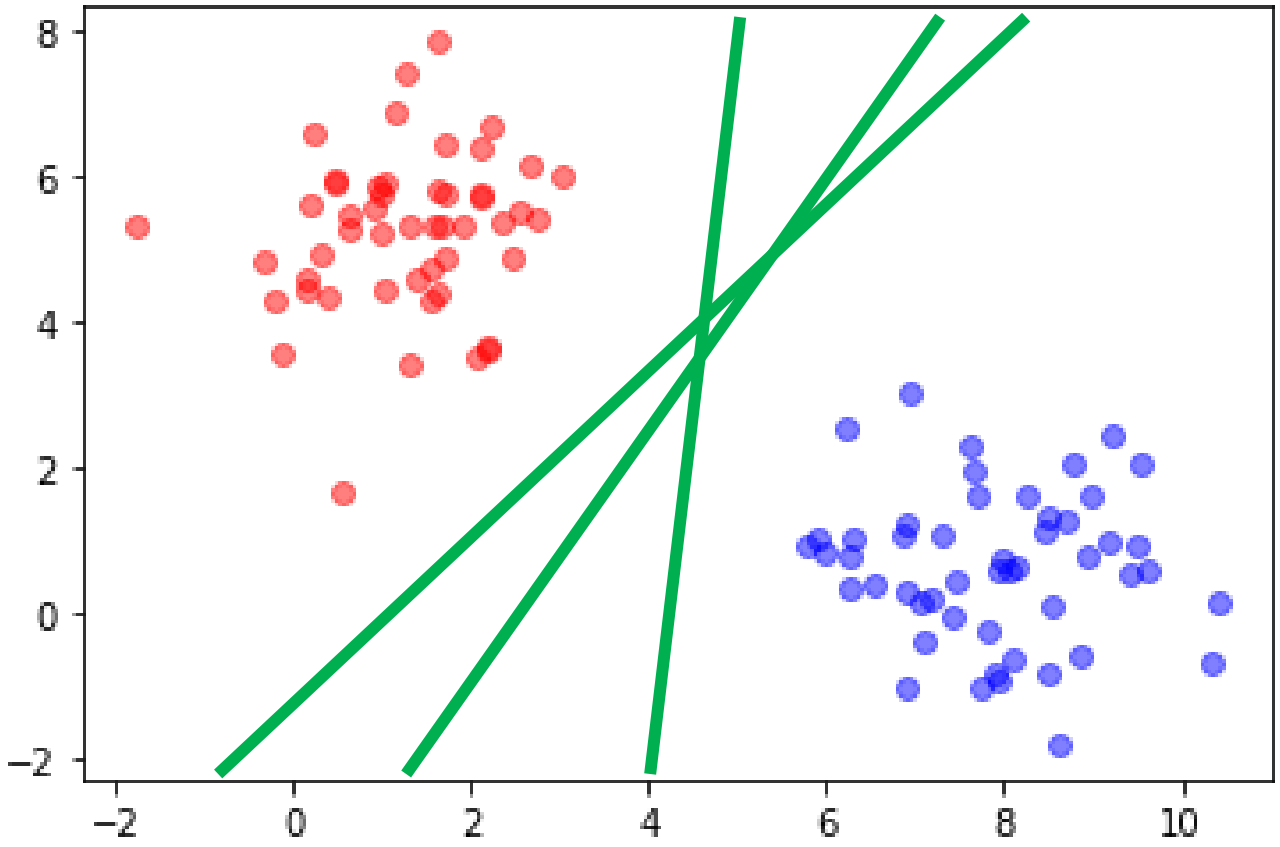
Hyperplane as a Decision Boundary



We can separate the two classes using a hyper plane!

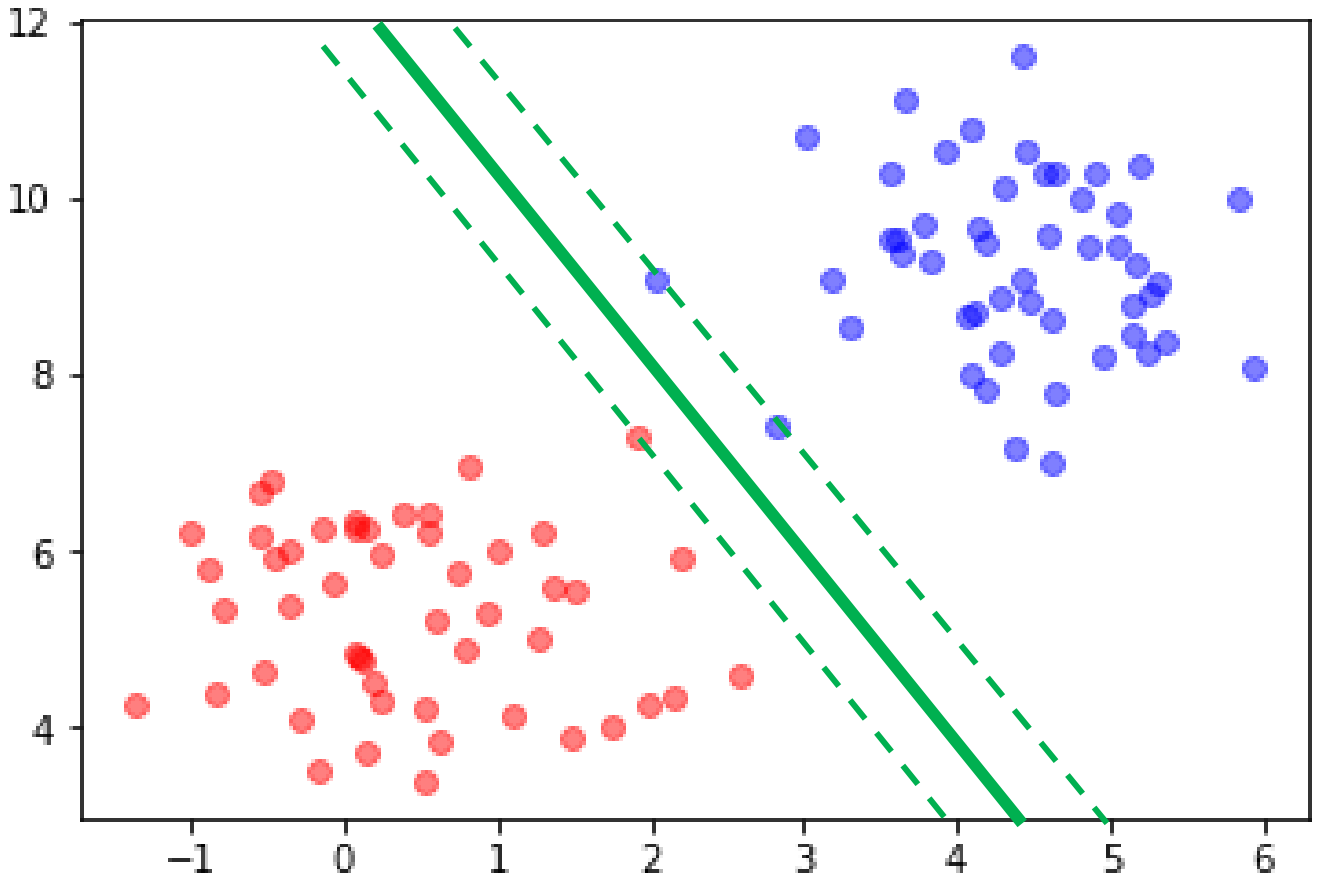
This hyperplane is called "separating hyperplane"

Hyperplane as a Decision Boundary



But which hyperplane should we choose?

Maximum margin classifier

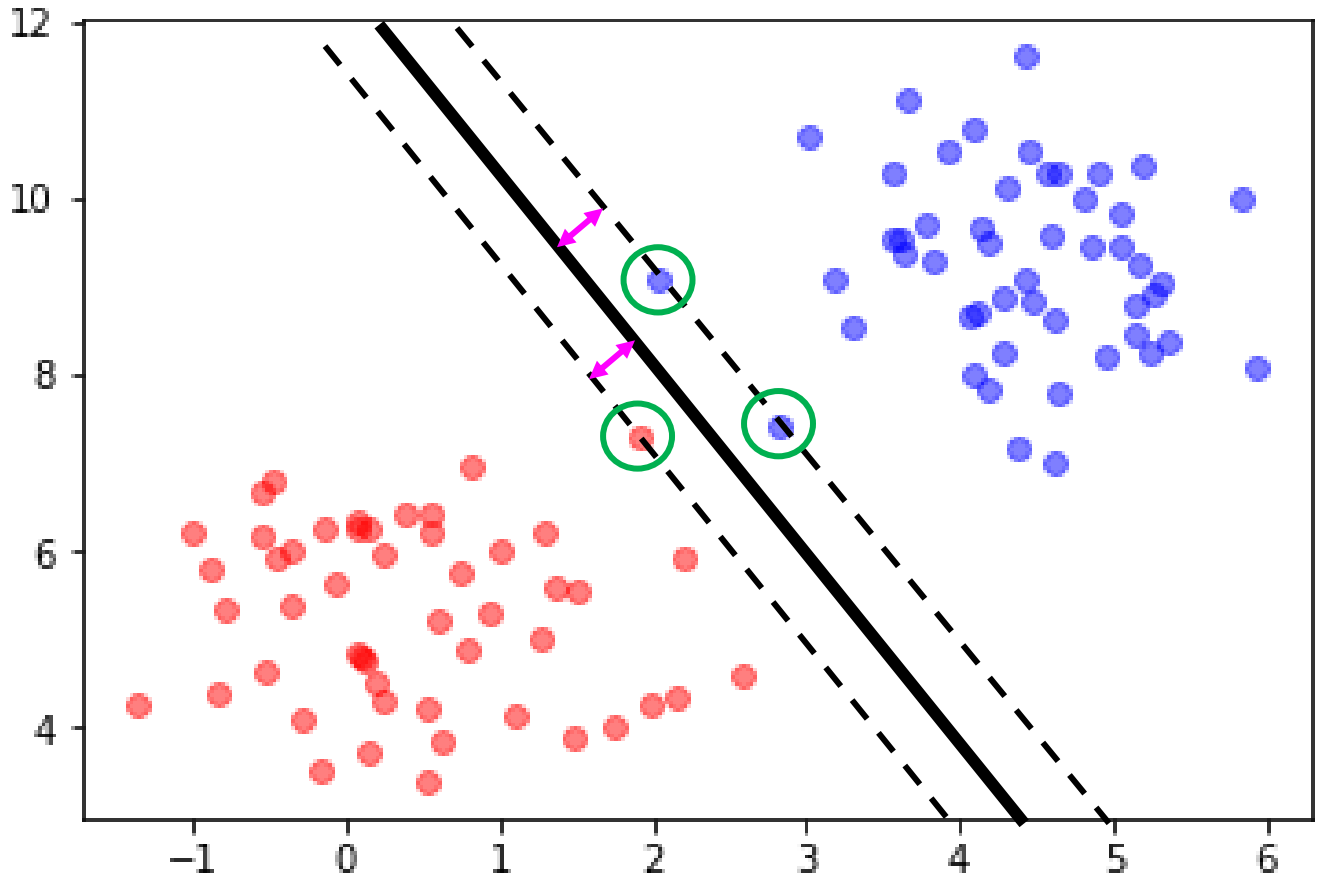


Which hyperplane should we choose?

The one with the least likely to misclassify the test data

= The one with the biggest margin

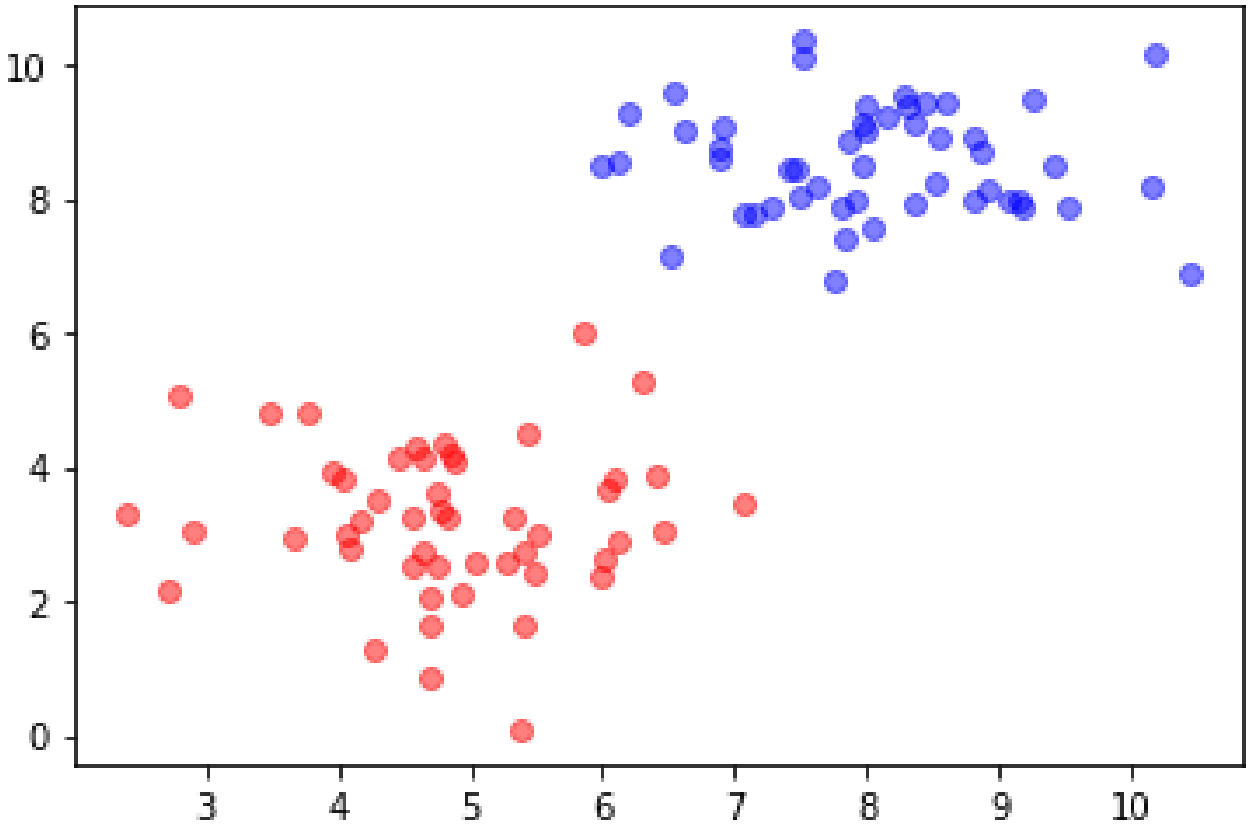
Maximum margin classifier



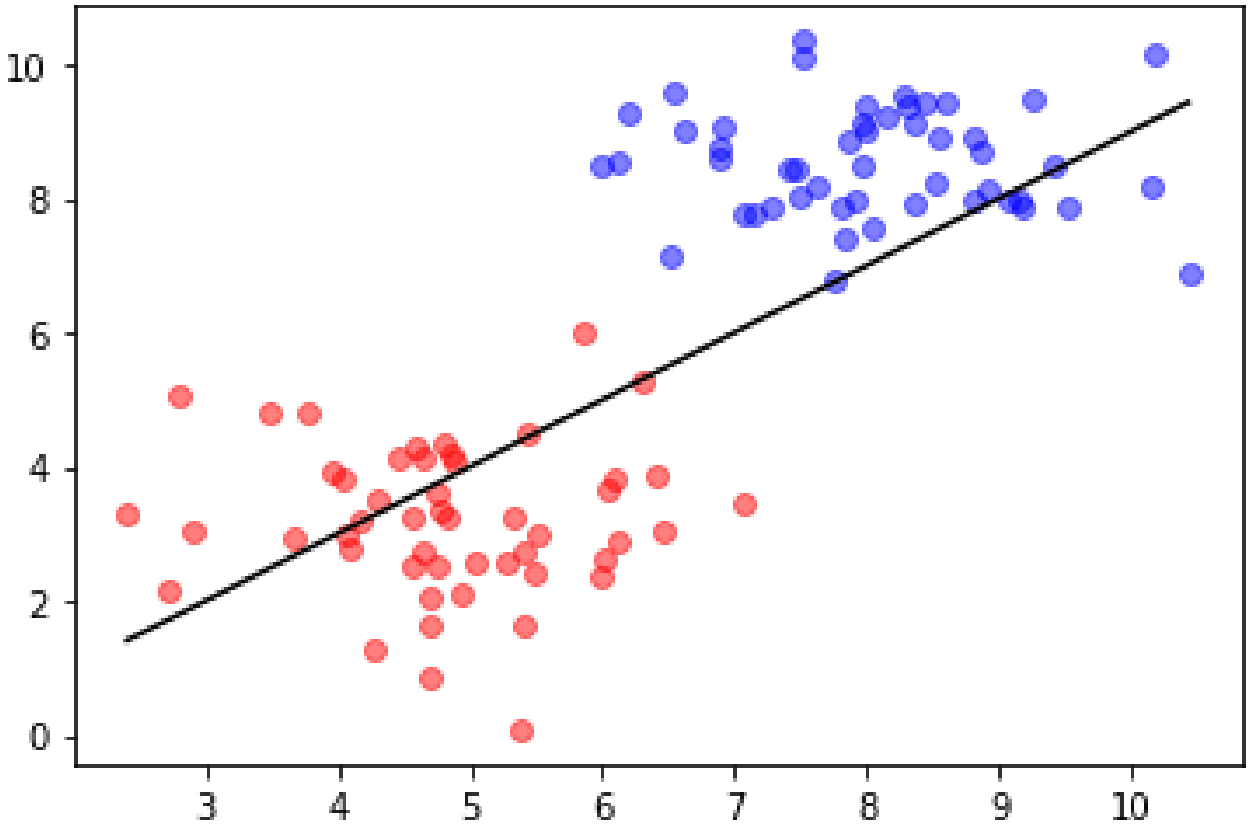
Support

Margin

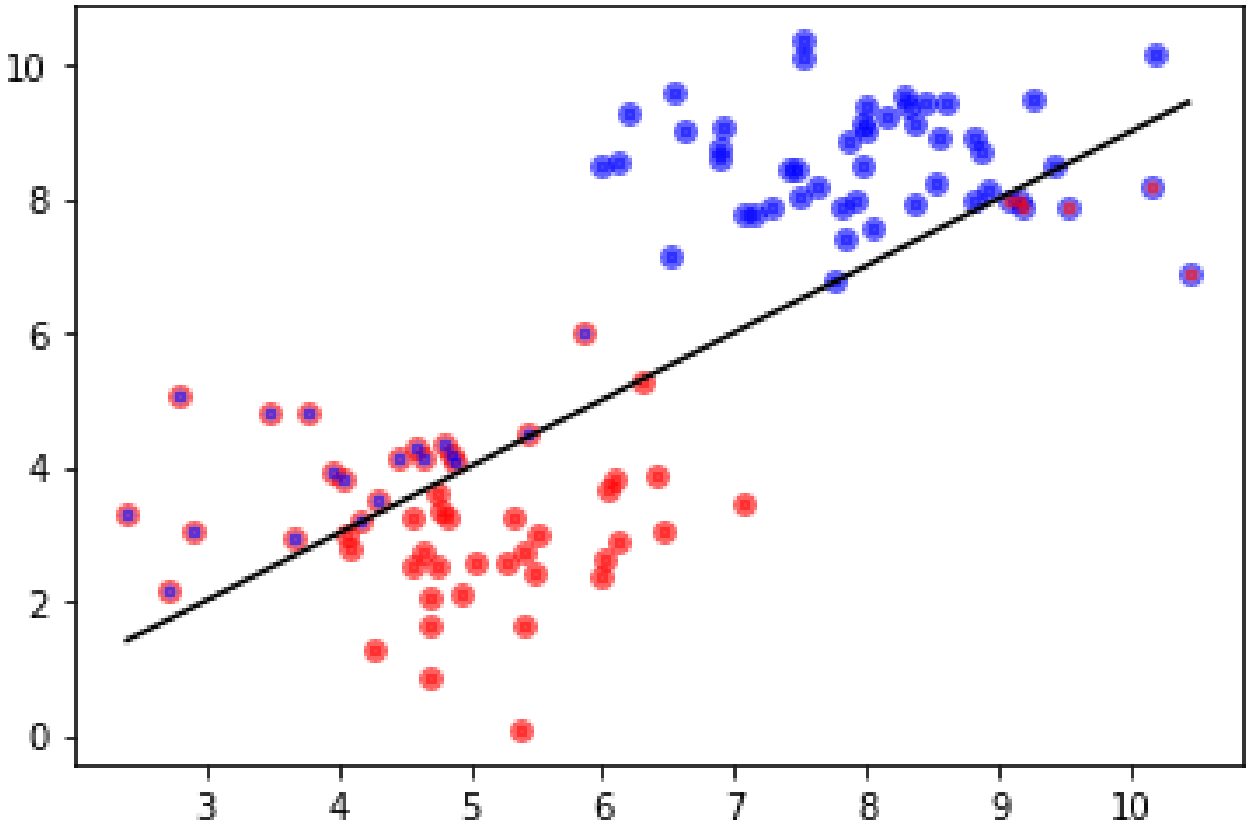
How to find the maximal margin hyperplane?



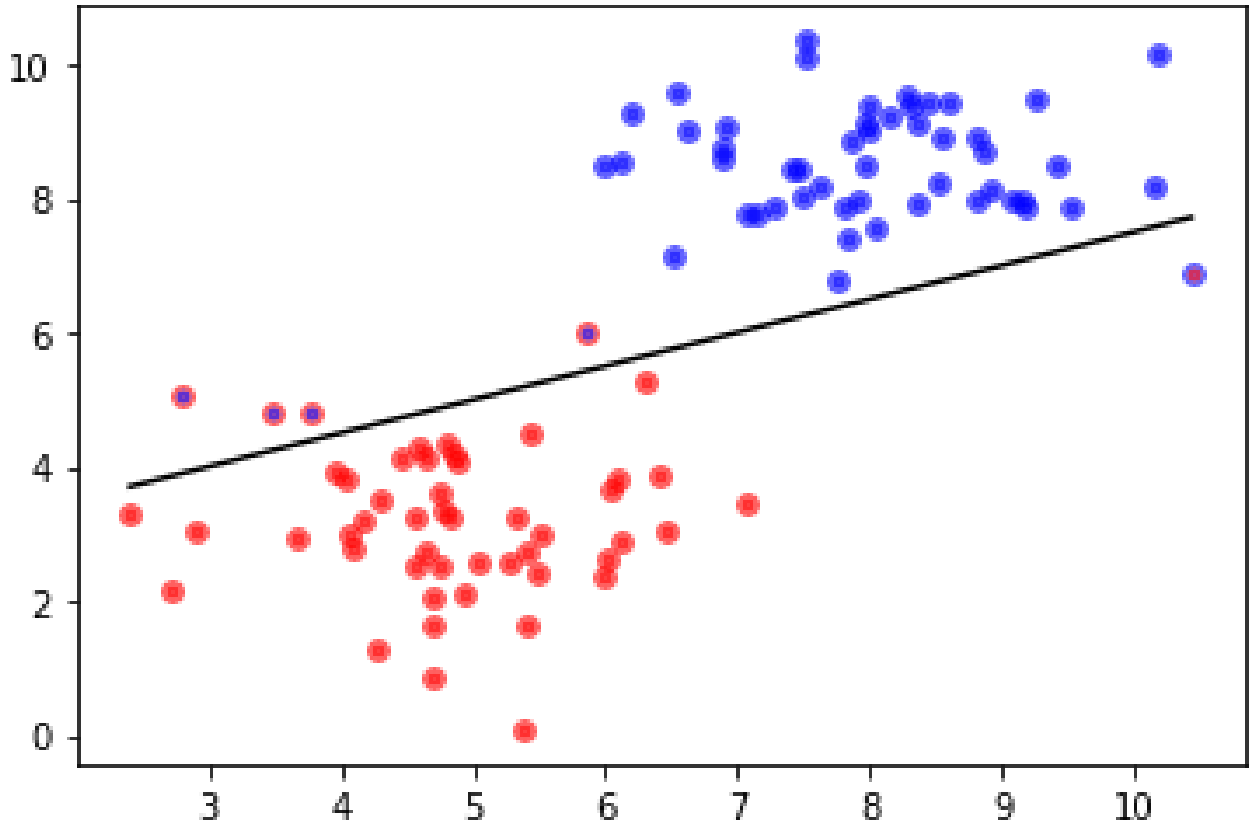
How to find the maximal margin hyperplane?



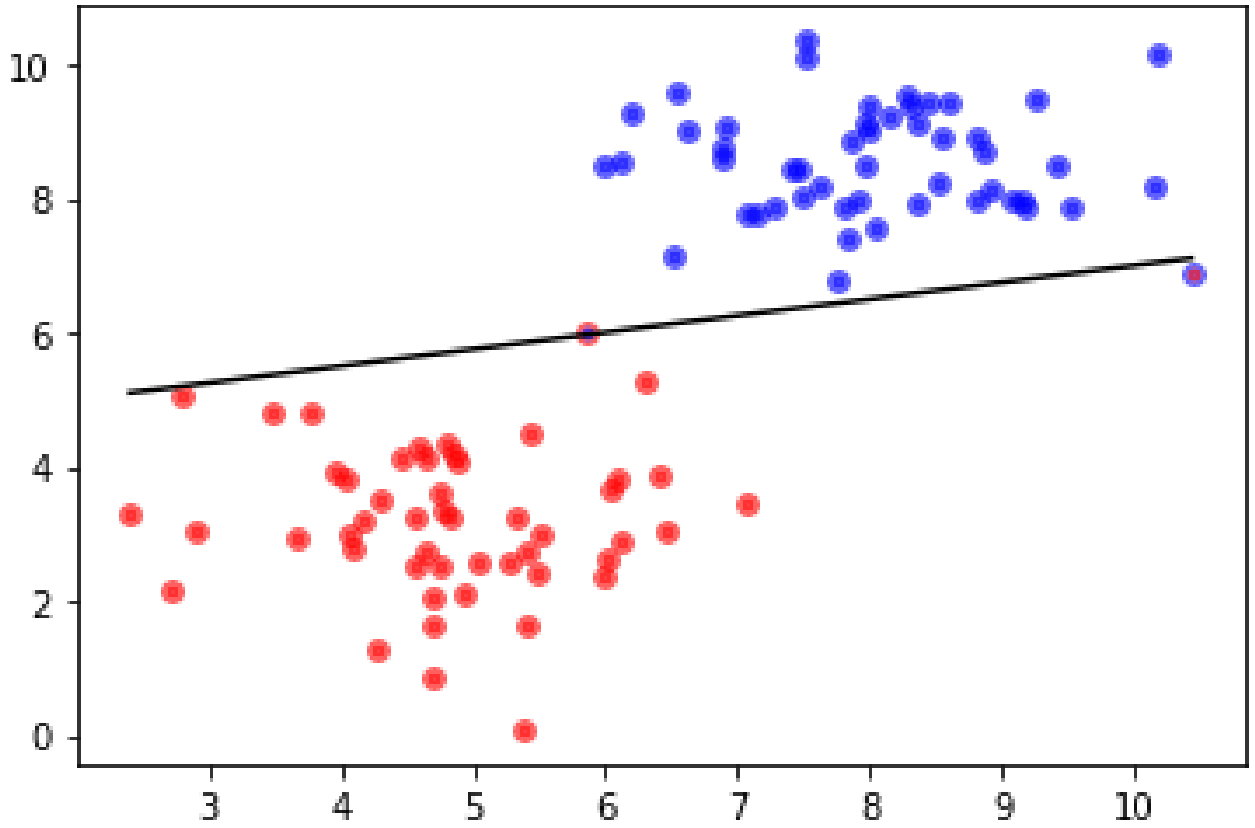
How to find the maximal margin hyperplane?



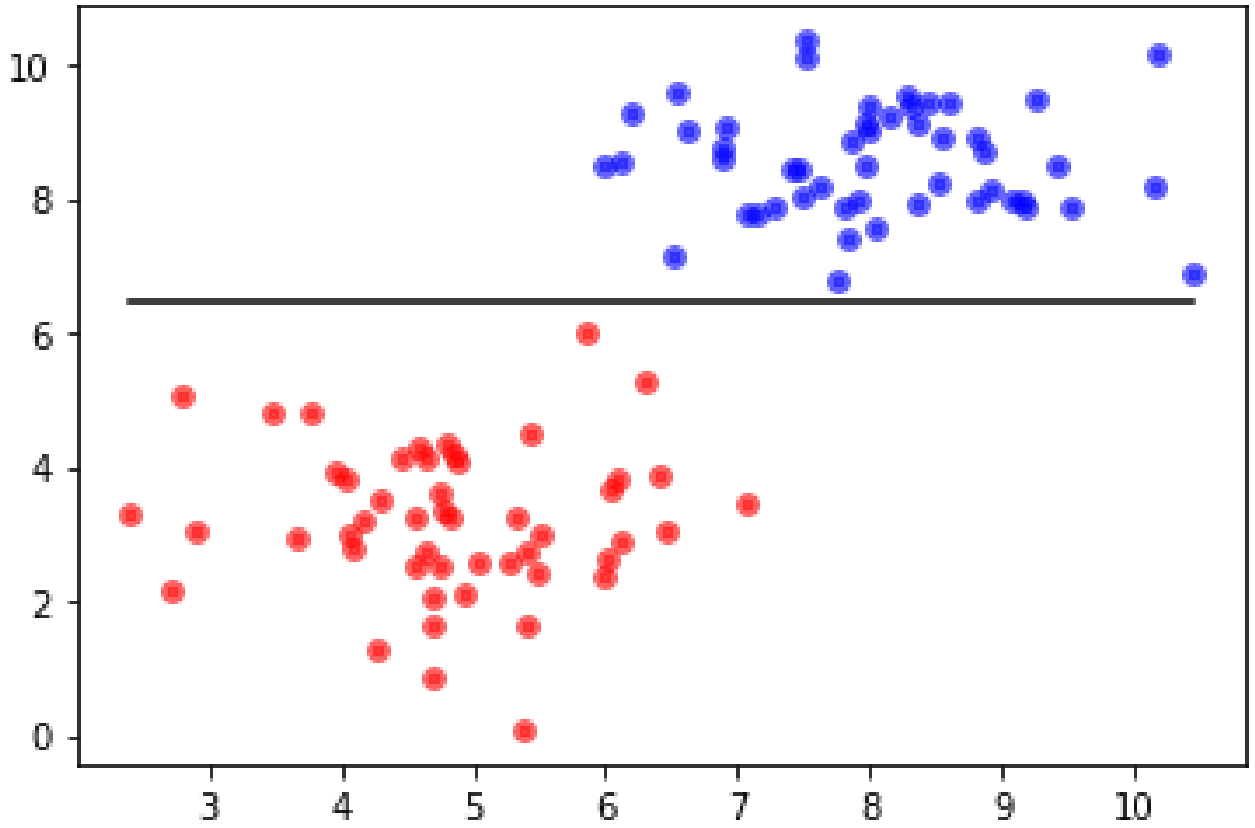
How to find the maximal margin hyperplane?



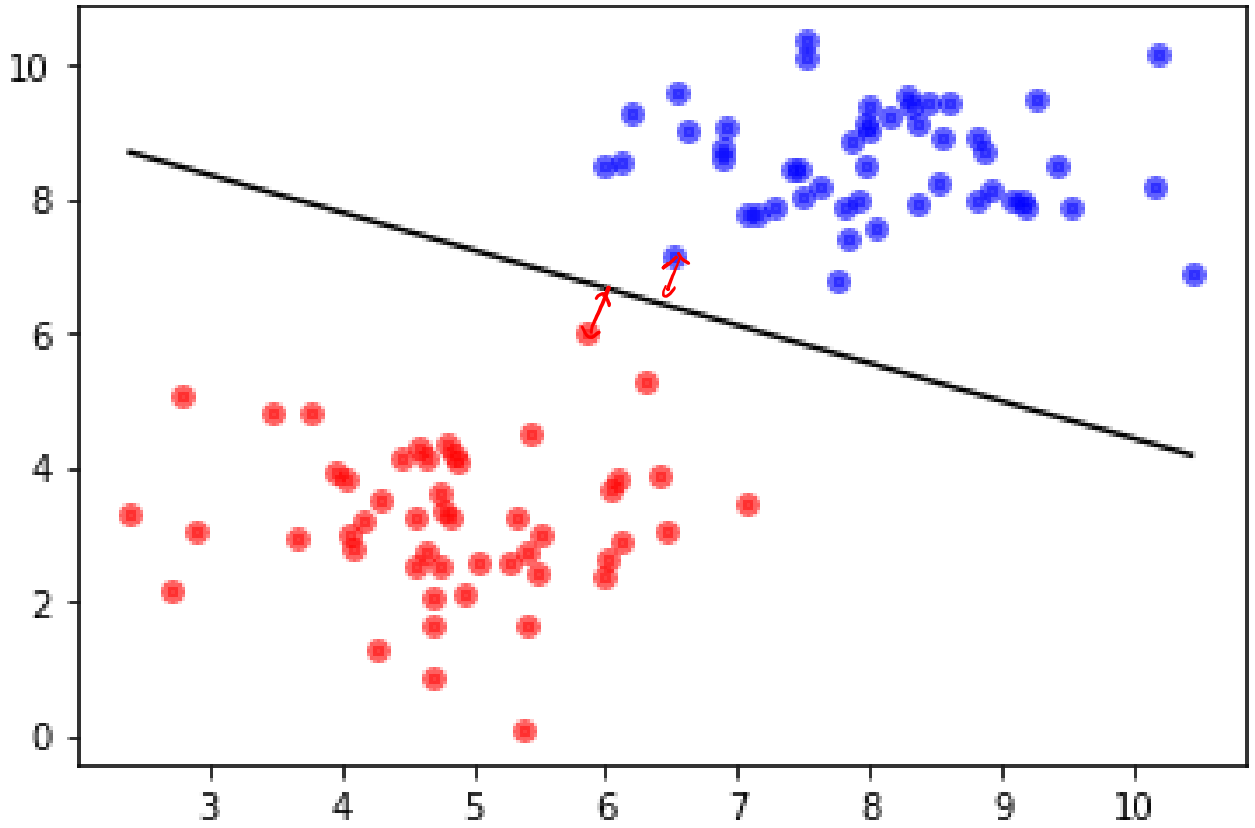
How to find the maximal margin hyperplane?



How to find the maximal margin hyperplane?

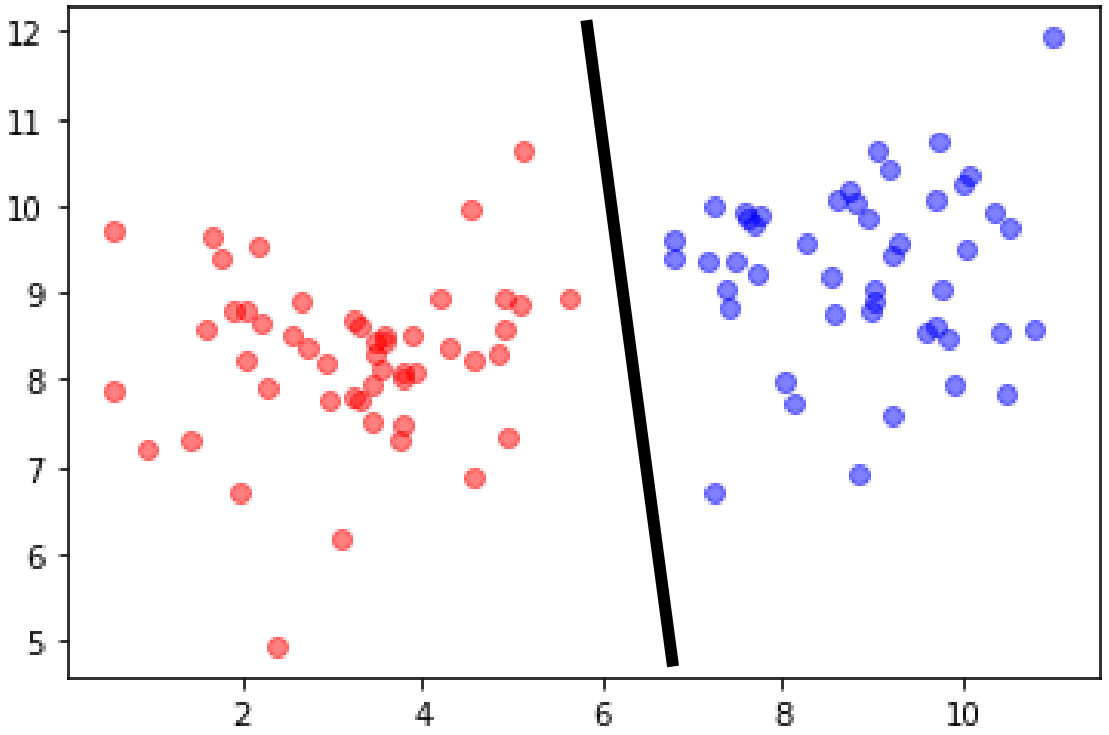


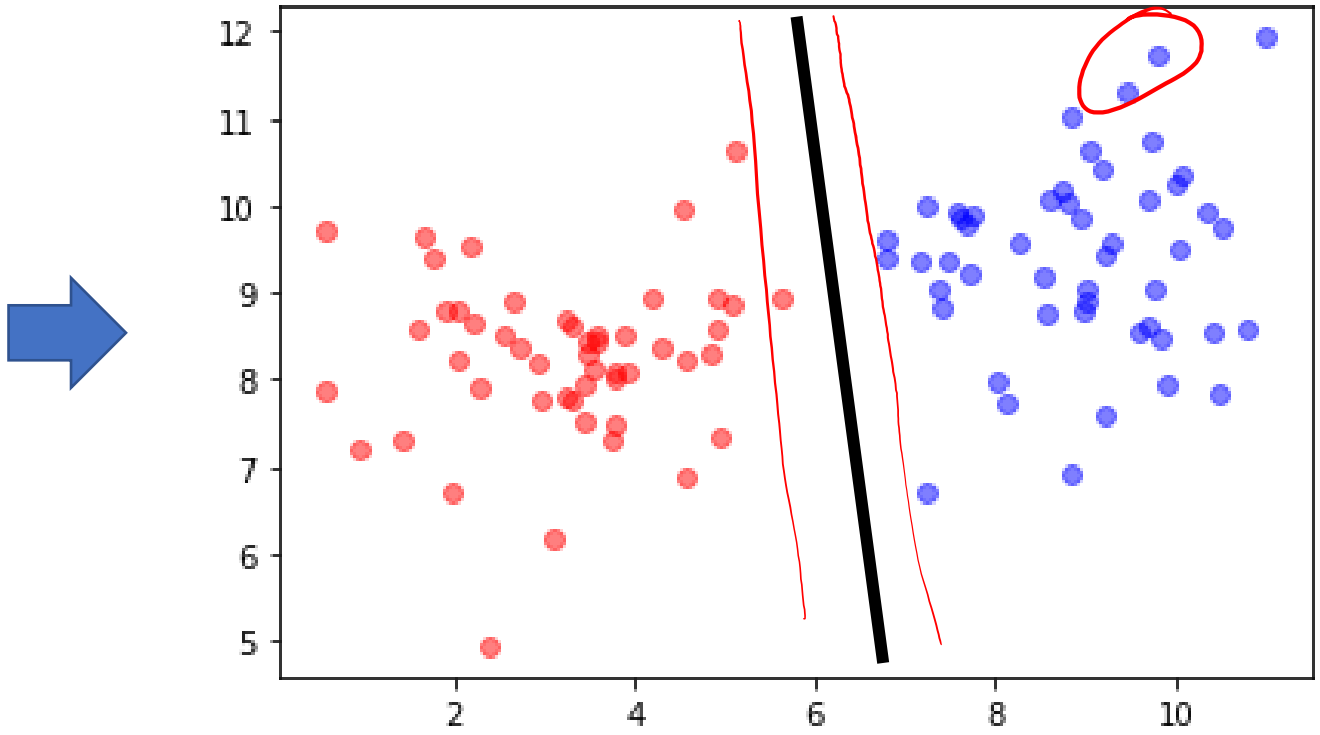
How to find the maximal margin hyperplane?



Quiz

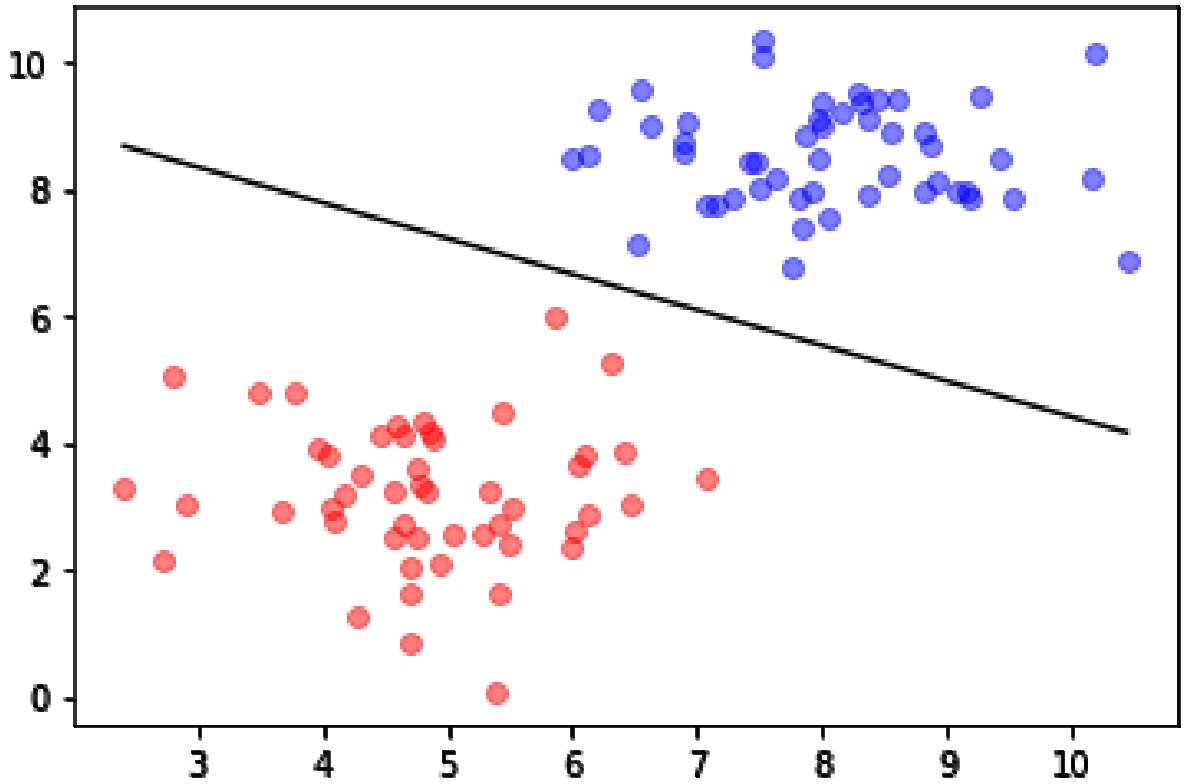
What happens to the separating hyperplane when adding new train data points?

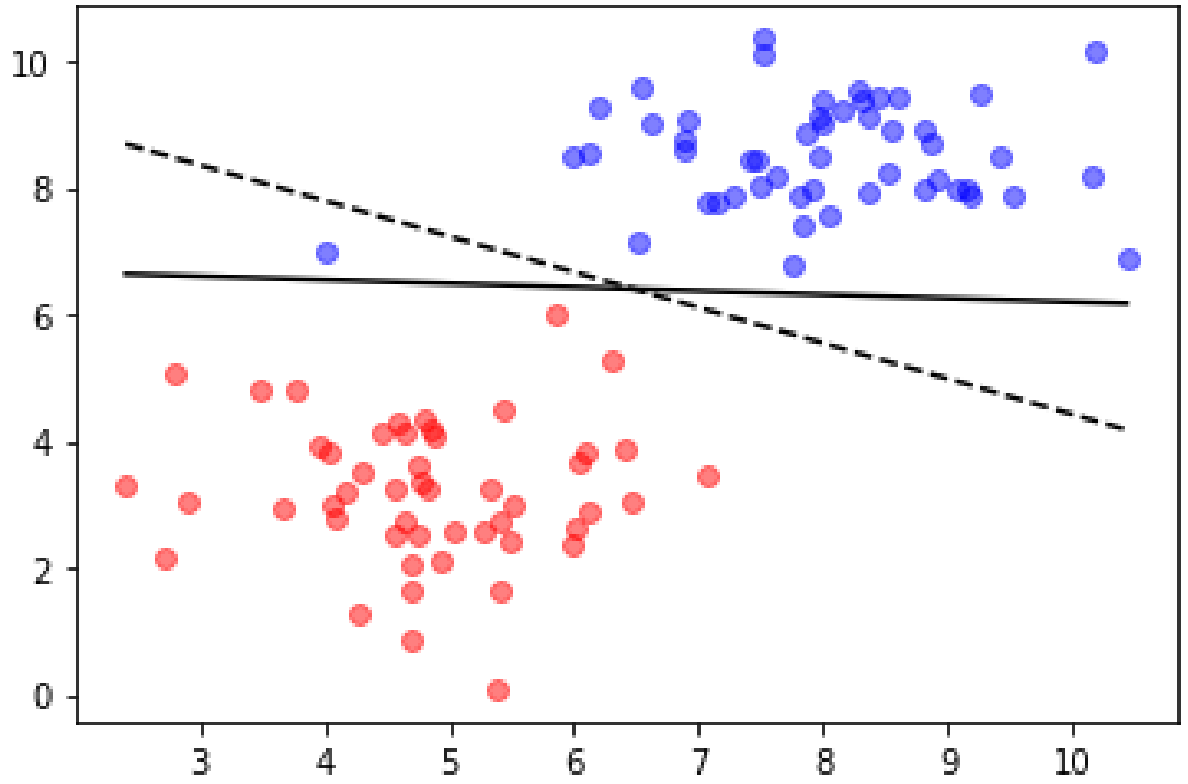




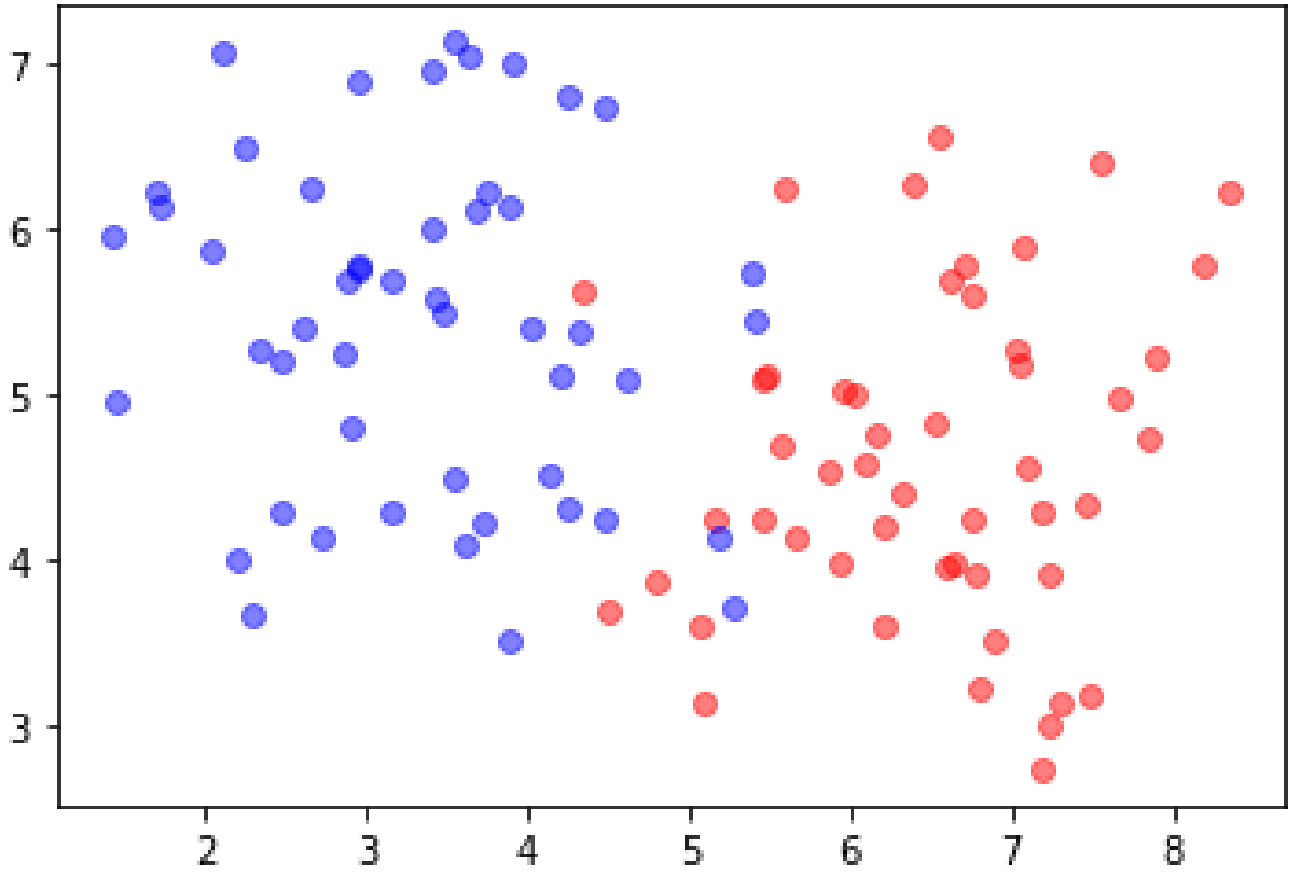
Drawback of the hard margin

What happens to the separating hyperplane?





The impossible case…



Can you separate this with a hyperplane?