Applied Machine Learning

Lecture 12
Support Vector Machine

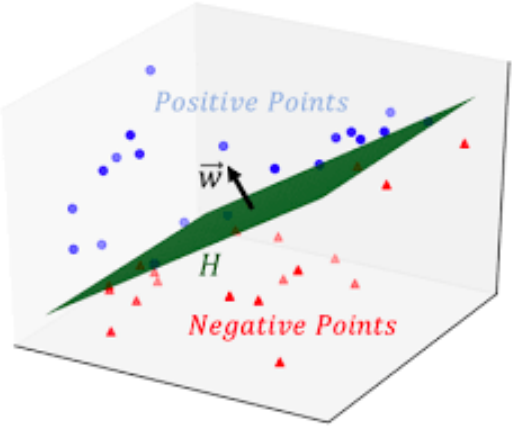
Ekarat Rattagan, Ph.D.

Slides adapted from Andrew's tutorials: http://www.cs.cmu.edu/~awm/tutorials .

Outline

1. Definition
2. Linear classifiers
3. How SVM works?
4. Cost function
5. Optimization

1. Definition



https://waterprogramming.wordpress.com/2019/01/29/intro-to-machine-learning-part-4-support-vector-machines/

Given a training dataset of points, $(\vec{x}_1, y_1), \ldots, (\vec{x}_m, y_m)$,where y_i are either $+1$ or -1, each indicating the class to which the point $\vec{x}_i$ belongs.

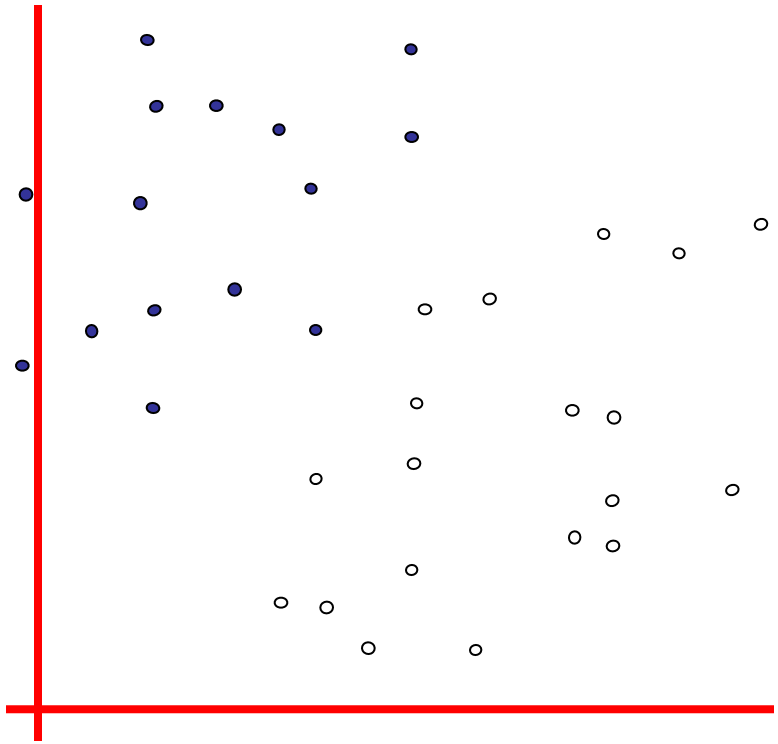
The objective is to find the "maximum-margin hyperplane" that divides the group of points $\vec{x}_i$ for which $y_i = 1$ from the group of points for which $y_i = -1$, so that the distance between the hyperplane and the nearest point $\vec{x}_i$ from either group is maximized.

Credit: https://en.wikipedia.org/wiki/Support_vector_machine

2. Linear Classifiers

denotes +1

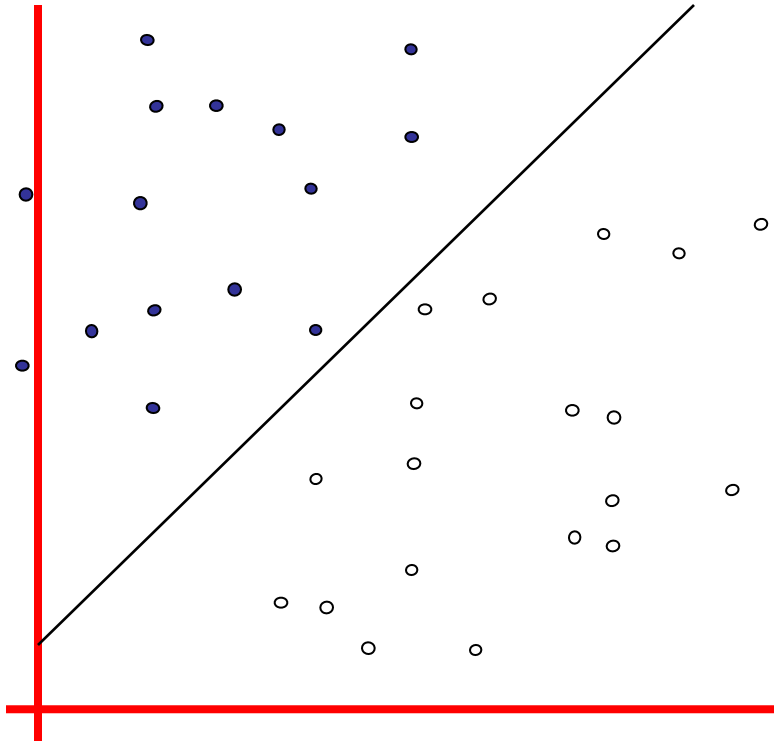
denotes -1



How would you classify this data?

Linear Classifiers

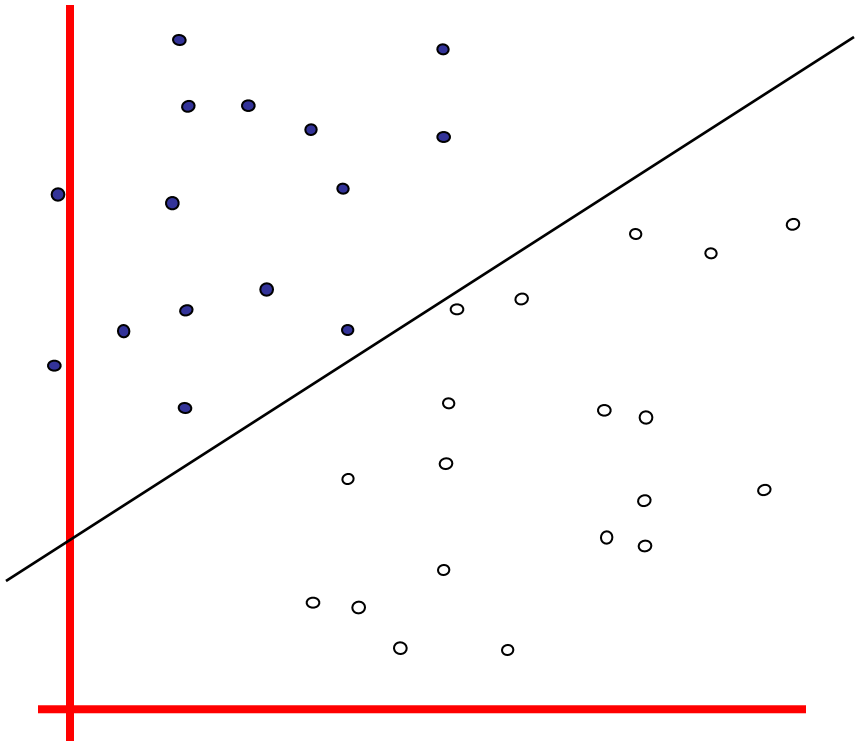
- denotes +1
- denotes -1



How would you classify this data?

Linear Classifiers

- denotes +1
- denotes -1

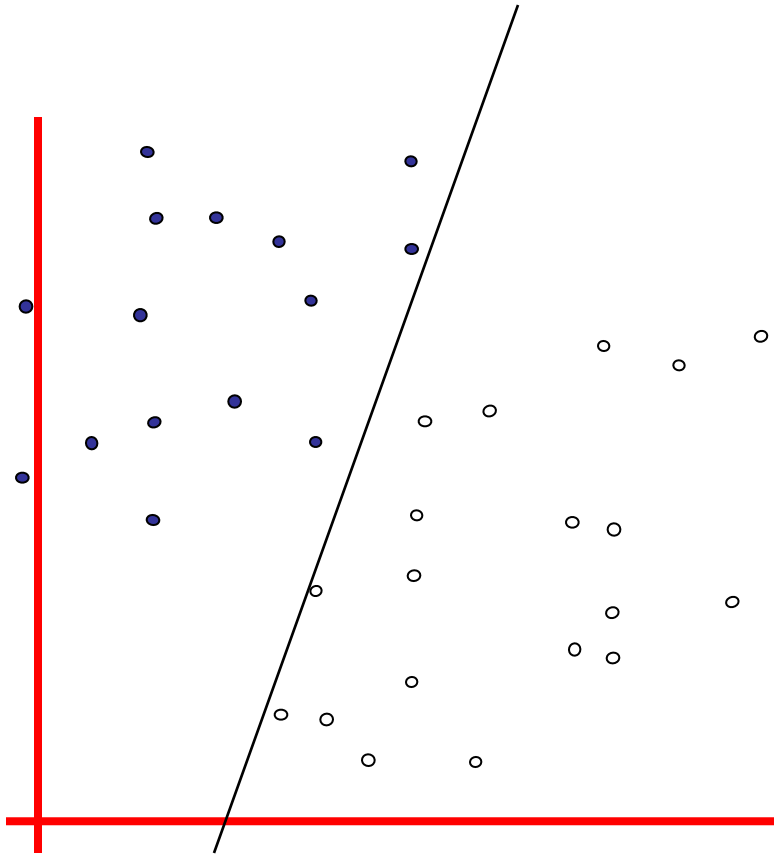


How would you classify this data?

Linear Classifiers

denotes +1

denotes -1

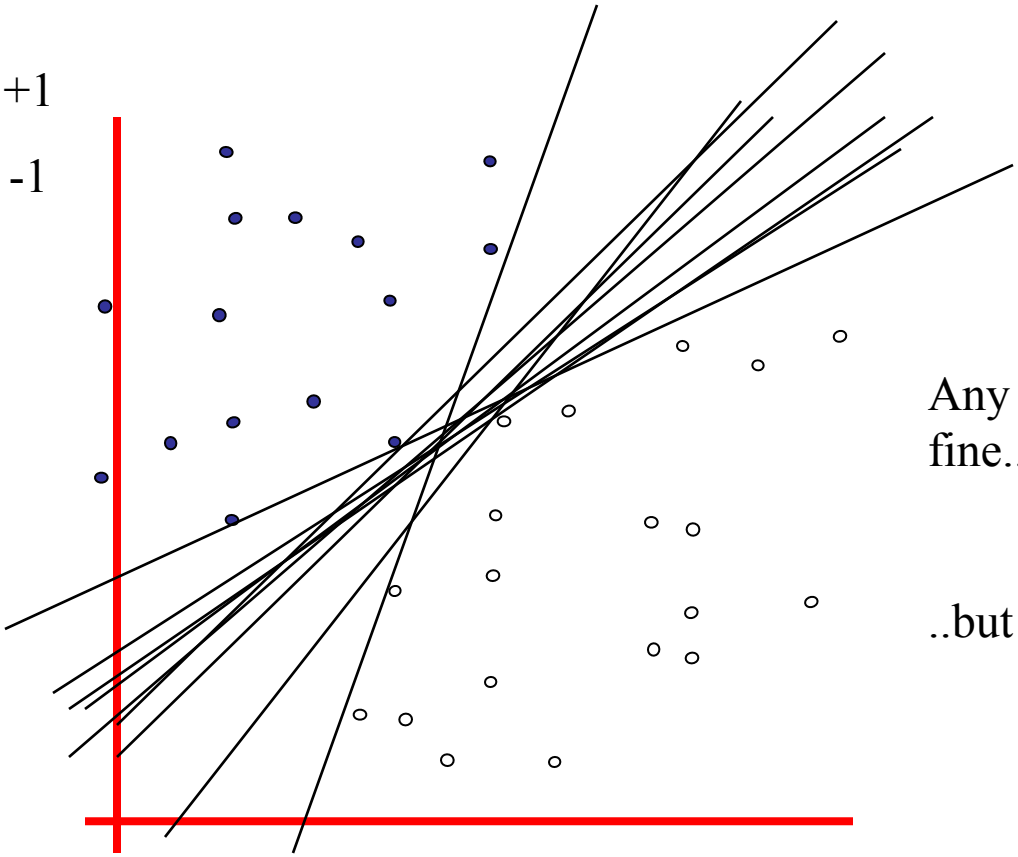


How would you classify this data?

Linear Classifiers

denotes +1

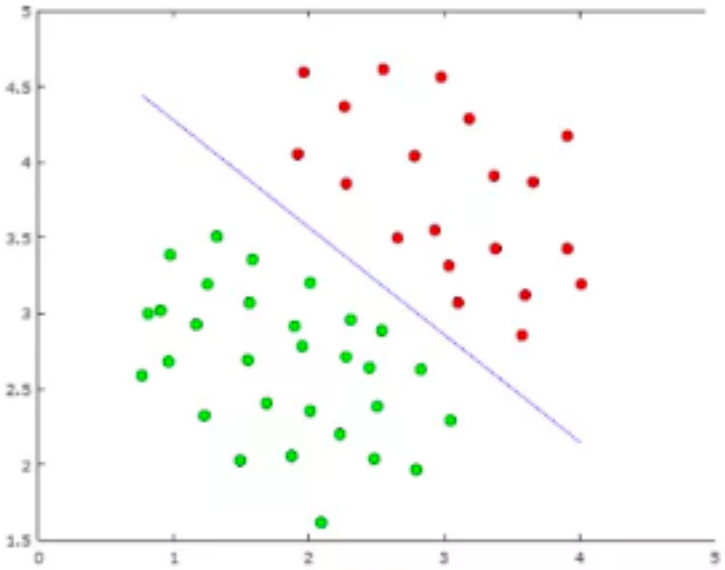
denotes -1



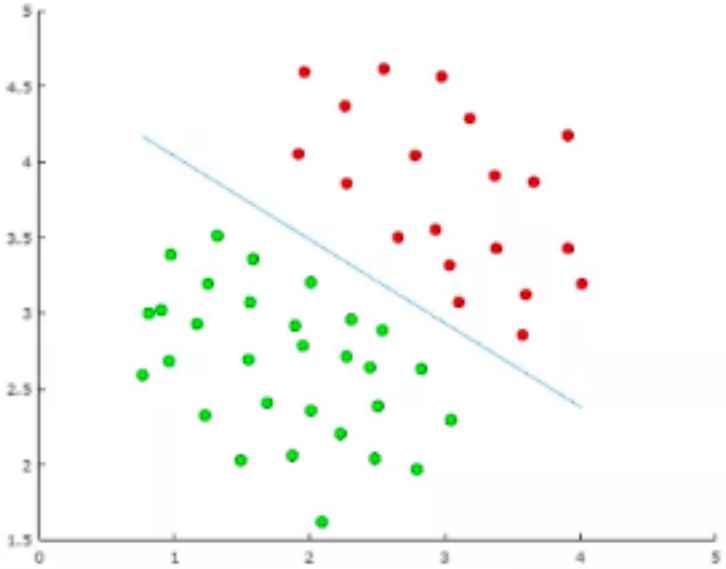
Any of these would be fine..

..but which is best?

Linear Classifiers (SVM VS Logistic Regression)

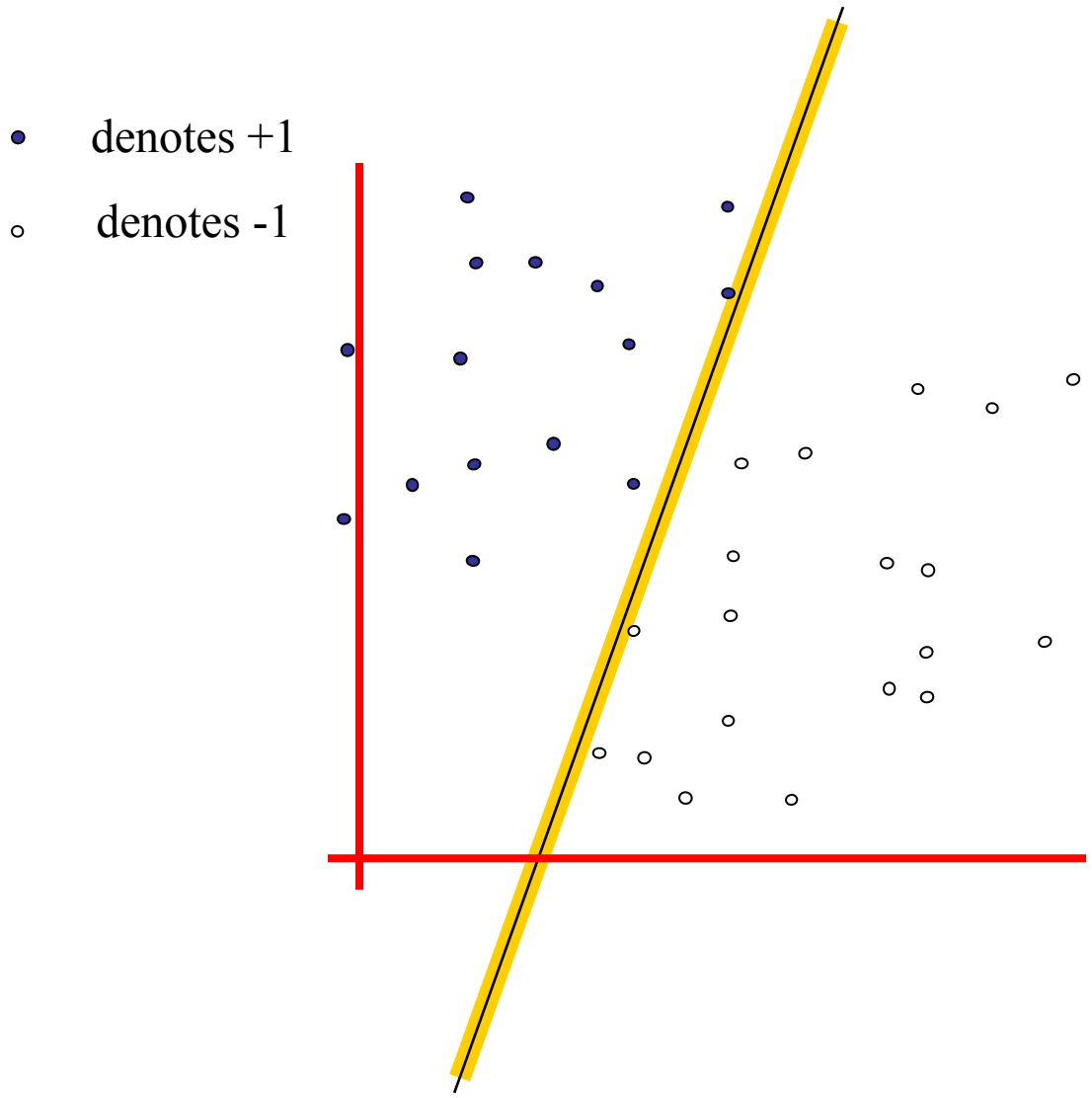


SVM



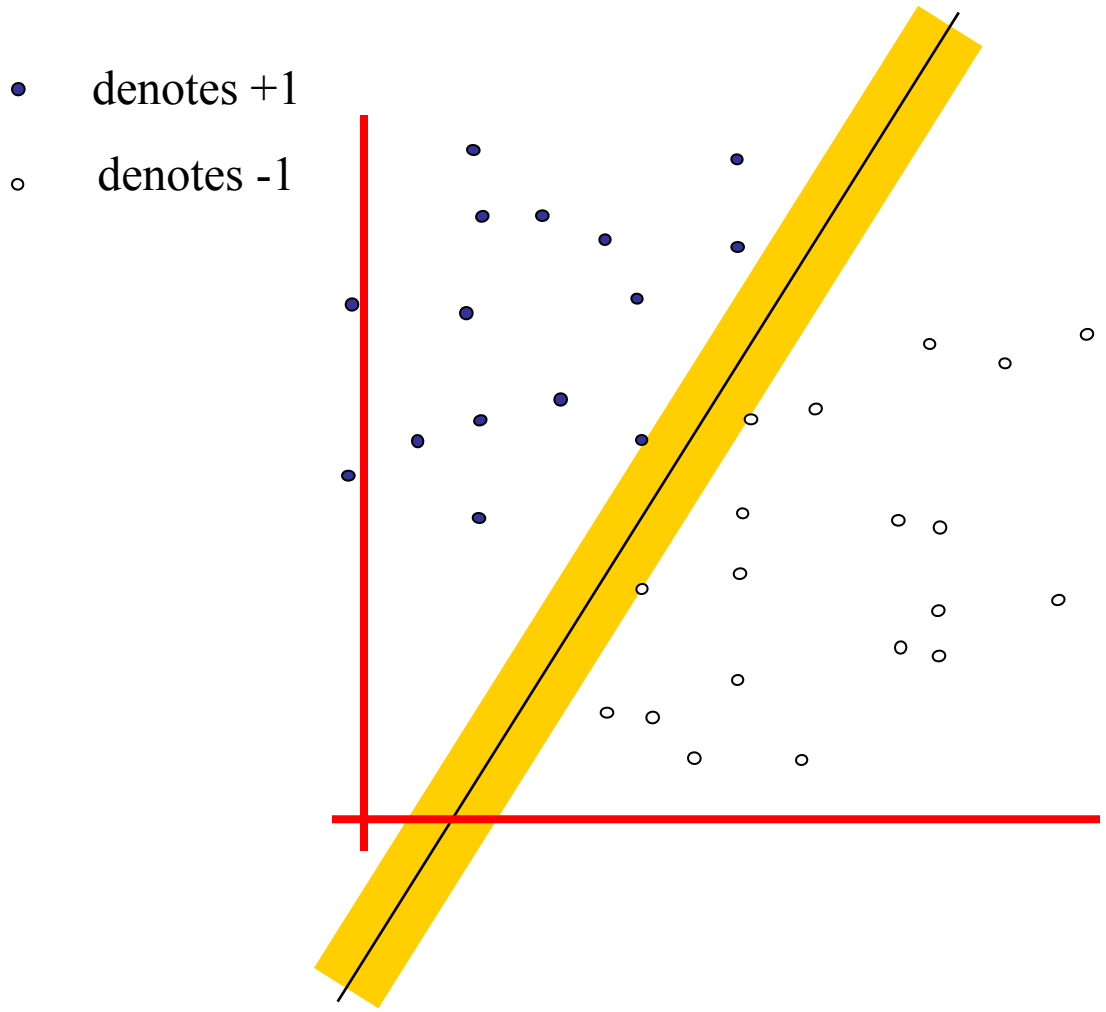
Logistic Regression

Classifier Margin



Define the margin of a linear classifier as the width that the boundary could be increased by before hitting a datapoint.

Maximum Margin



The maximum margin linear classifier is the linear classifier with the maximum margin.

This is the simplest kind of SVM (Called an Linear SVM)