

# 人工神经网络

肖强 张哲槟  
密尚华

March 28, 2018

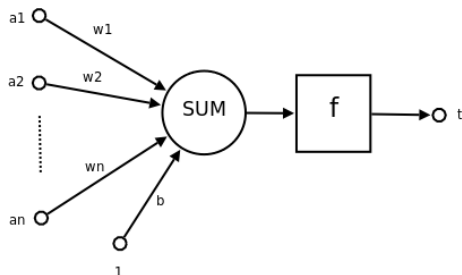
1 人工神经网络简介

2 SVM

# 人工神经网络 (Artificial Neural Network, ANN)

神经网络是一种运算模型，由大量的节点（或称神经元）之间相互联接构成。

网络自身通常都是对自然界某种算法或者函数的逼近，也可能是对一种逻辑策略的表达。

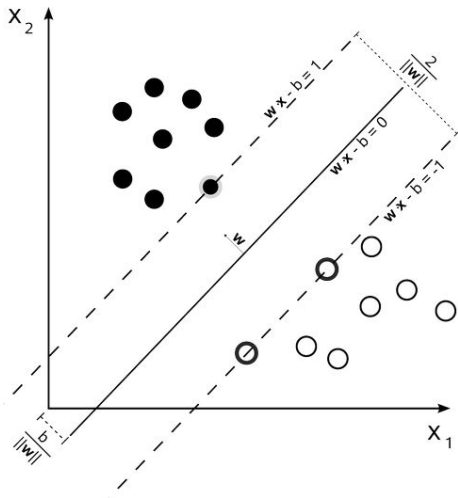


# SVM



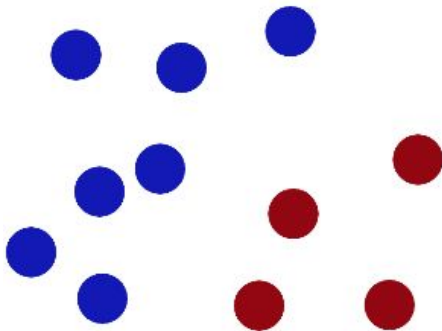
# 支持向量机

- SVM, 俗称支持向量机, 为一种 supervised learning 算法, 属于 classification 的范畴。
- 在数据挖掘的应用中, 与 unsupervised 的 Clustering 相对对应和区别。
- 广泛应用于机器学习 (Machine Learning), 计算机视觉 (Computer Vision) 和数据挖掘 (Data Mining) 当中。



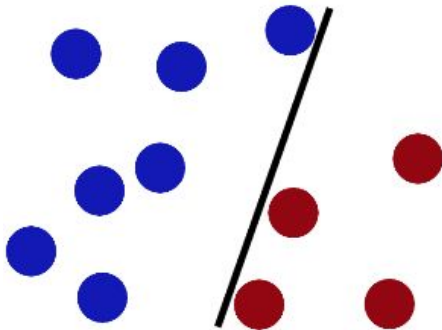
# 一个游戏

现在桌子上有两种颜色的球，现在要把他们分开。



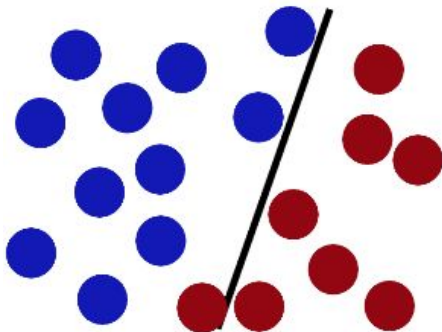
## 一个游戏

我们把一根棍子放在中间，看上去干得不错。



## 一个游戏

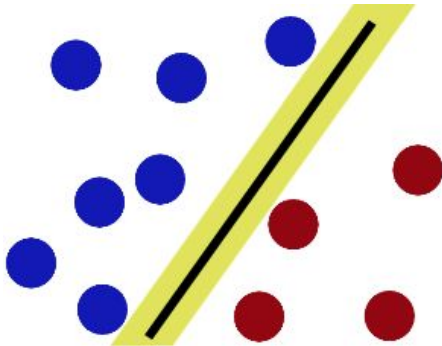
有些人又往桌子上放了一些球，大部分都分对了，但是出现了一个错分的。





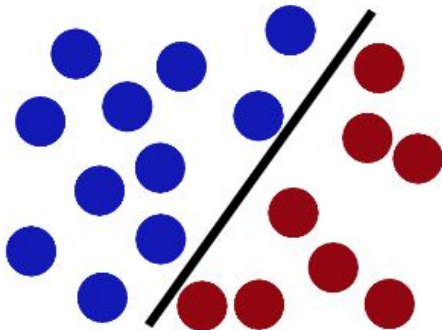
## 一个游戏

SVM 就是试图把棍放在最佳位置，好让在棍的两边有尽可能大的间隙。



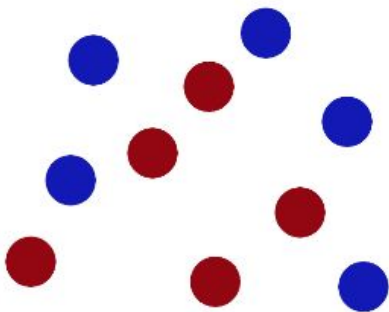
# 一个游戏

现在即使放了更多的球，棍仍然是一个好的分界线



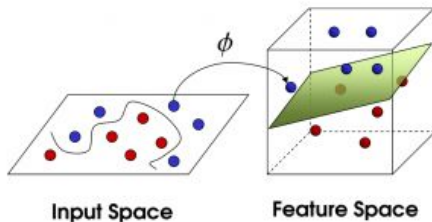
# 一个游戏

我们已经学会了一个 trick，然后，在 SVM 工具箱中有另一个更加重要的 trick，于是又有一个新的挑战。



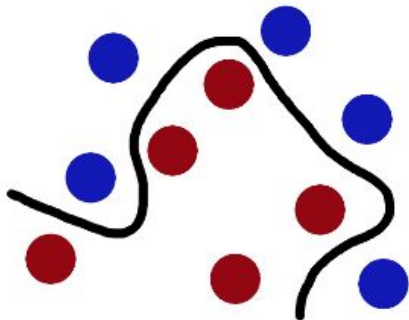
# 一个游戏

现在，没有棍可以很好地分开两种球了，现在怎么办呢？我们可以一拍桌子，球飞到空中。然后，抓起一张纸，插到了两种球的中间。



## 一个游戏

现在，这些球看起来像是被一条曲线分开了。

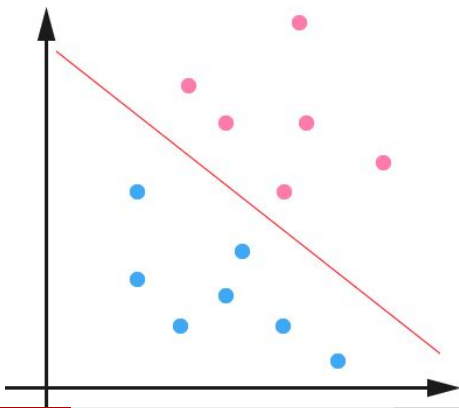


# 怎么求解 SVM?

给定训练样本集  $D = \{(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)\}$  如果可以用一个线性超平面将其完全分开，那么这个超平面可以表示为：

$$\mathbf{w}^T \mathbf{x} + b = 0$$

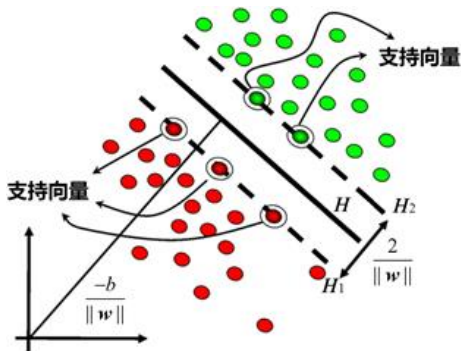
其中  $\mathbf{w}$  决定了平面的方向，而  $b$  决定了平面与原点之间的距离。



# 间隔与支持向量

空间中任意点  $\mathbf{x}$  实际上是一个向量，其到超平面的距离为：

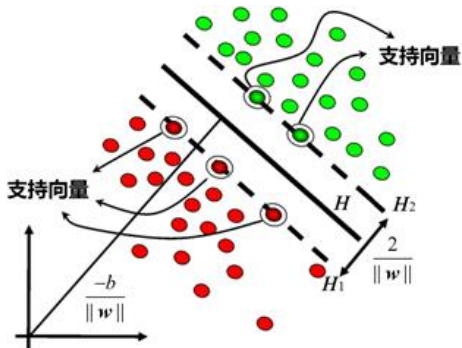
$$r = \frac{\mathbf{w}^T \mathbf{x} + b}{\|\mathbf{w}\|}$$



# 间隔与支持向量

那么，如果这个超平面分类成功，我们令

$$\begin{cases} \mathbf{w}^T \mathbf{x} + b \geq +1, & y_i = +1; \\ \mathbf{w}^T \mathbf{x} + b \leq -1, & y_i = -1. \end{cases}$$





# 间隔与支持向量

这样，那些使等号成立的点，也即距离超平面最近的点称作支持向量 (support vector)，两个类到超平面距离之和也即两类之间的间隔 (margin) 是

$$\gamma = \frac{2}{\|\mathbf{w}\|}$$

训练支持向量机，实际上是希望找到具有最大间隔的划分超平面，即训练目标是在满足分类任务的情况下最大化  $\gamma$ ，也即最大化  $\|\mathbf{w}\|^{-1}$ ，也即最小化  $\|\mathbf{w}\|^2$ 。SVM 的基本型就是：

$$\begin{aligned} \min_{\mathbf{w}, b} \quad & \frac{1}{2} \|\mathbf{w}\|^2 \\ \text{s.t.} \quad & y_i(\mathbf{w}^T \mathbf{x}_i + b) \geq 1, \quad i = 1, 2, \dots, m \end{aligned}$$