



感知机

🕒 Created	@April 7, 2022 4:30 PM
▼ Class	统计学习方法
▼ Type	work
🔗 Materials	
☑ Reviewed	<input type="checkbox"/>

基本概念

感知机是二类分类的线性分类模型，其输入为实例的特征向量，输出为实例的类别，取+1和-1二值。感知机对应于输入空间（特征空间）中将实例划分为正负两类的分离超平面，属于判别模型。感知机学习旨在求出将训练数据进行线性划分的分离超平面。

$$f(x) = \text{sign}(w * x + b)$$

感知机有如下几何解释：线性方程

$$(wx + b) = 0$$

对应于特征空间中的一个超平面S，其中w是超平面的法向量，b是超平面的截距。超平面将特征空间分为两个部分，位于两部分的点（特征向量）分别被分为正、负两类，因此超平面S称为分离超平面

感知机学习策略

感知机学习的目标是求得一个能够将训练集正实例点和负实例点完全正确分开的分离超平面，需要确定感知机模型参数w，b，因此需要定义损失函数并且将损失函数极小化

感知机的损失函数采用了误分类点到超平面S的总距离，首先输入空间中任一点x0到超平面S的距离：

$$\frac{1}{\|w\|} |w \cdot x_0 + b|$$

对于误分类的数据 (x_i, y_i) 来说

$$-y_i(w \cdot x_i + b) > 0$$

因此误分类点 x_i 到超平面 S 的距离是：

$$-\frac{1}{\|w\|} y_i(w \cdot x_i + b)$$

所有误分类点到超平面 S 的总距离为

$$-\frac{1}{\|w\|} \sum_{x_i \in M} y_i(w \cdot x_i + b)$$

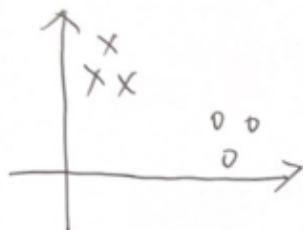
因此感知机的损失函数定义为

$$L(w, b) = - \sum_{x_i \in M} y_i(w \cdot x_i + b)$$

其中 M 为误分类点的集合，这个损失函数就是感知机学习的经验风险函数

线性分类 { 硬分类 { LDA
感知机
软分类 { 生成式: 高斯判别分析
判别式: LR.
朴素贝叶斯.

感知机:



模型:

$$f(x) = \text{sign}(w^T x + b)$$

$$\begin{cases} 1 & w^T x + b \geq 0 \\ 0 & w^T x + b < 0 \end{cases}$$

损失函数:

$$L(w) = \sum_{i=1}^N I(y_i(w^T x_i + b))$$

$I()$ 为当 $y_i w^T x_i > 0$ 时, $= 0$.

当 $y_i w^T x_i < 0$ 时, $= y_i w^T x_i$.

但此 $L(w)$ 不可导, 不连续.

因此改为

$$L(w) = \sum_{i=1}^N \sum_{x \in D} [-y_i(w^T x + b)] \quad D \text{ 为全体数据点集合.}$$

学习算法: 随机梯度下降. SGD.

$$\nabla_w L = w - \lambda(-y_i x_i) \quad \lambda \text{ 为学习率.}$$

感知机算法. 不同的初值会产生不同的结果.

