







感知机目标是找到一个线性可分的超平面:

$$w_1 x_1 + w_2 x_2 + \dots + w_m x_m + b = 0$$

感知机

$$f(x) = sign(w^T x + b)$$

其中输入样本的空间 $X \subseteq \Re^n$, 输出样本空间Y = (+1,-1),

输入样本 $x \in X$, 输出样本 $y \in Y$, w 和 b 是感知模型参数,

符号函数为

$$sign(x) = \begin{cases} +1, & x \ge 0 \\ -1, & x < 0 \end{cases}$$





对于一个正样本 x1, 若

$$w^T x_1 + b \ge 0$$

则被分为+1类

若负样本 x1

$$w^T x_1 + b < 0$$

则被分为-1类





感知器模型损失函数

把满足 $w^T x_1 + b \ge 0$ 输出为+1,满足 $w^T x_1 + b < 0$ 输出为-1。

因此正确分类样本满足 $y(w^Tx+b)>0$

错误分类样本满足 $y(w^Tx+b)<0$





感知器模型损失函数

对于一个错分样本 x_i , 有 $y_i(w^Tx_i+b)<0$

该样本到超平面(分类面)的距离是:

$$y_i \left(w^T x_i + b \right) / \left\| w \right\|_2$$

 $||w||_2 = 1$

所有错分点到超平面的距离之和为

$$-\sum_{x_i \in E} y_i \left(w^T x_i + b \right)$$





感知器模型损失函数

最终的损失函数为

$$L(w,b) = \sum_{x_i \in E} y_i \left(w^T x_i + b \right)$$

损失函数的梯度为

$$\nabla_{w} L(w,b) = \sum_{x_i \in E} y_i x_i$$

$$\nabla_b L(w,b) = \sum_{x_i \in E} y_i$$





感知器参数更新

对于一个错误分类点, (x_i, y_i) 其对 w,b 的更新如下:

$$w \leftarrow w + \eta y_i x_i$$

$$b \leftarrow b + \eta y_i$$

通过参数更新,不断调整分类面,使损失函数不断减小,最终使其对错误分类点,能够正确分类。





算法: 感知器分类算法的原始形式

输入: 训练数据集 $T = \{(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)\}$, 其中 $x_i \in \chi = \Re^n$,

 $y_i \in \Omega = \{-1, +1\}, i = 1, 2, \dots, N; \quad \text{$\not=$} \exists x \eta (0 < \eta \le 1);$

输出: w,b; 感知器模型 $f(x) = sign(w \cdot x + b)$ 。

- (1) 选取初值 w_0, b_0
- (2) 在训练集中选取数据 (x_i, y_i)
- (3) 如果 $y_i(w \cdot x_i + b) \le 0$

$$w \leftarrow w + \eta y_i x_i$$
$$b \leftarrow b + \eta y_i$$

(4) 转至(2),直至训练集中没有误分类点。