感知器

chuangketie

感知器是用于二分类的最基础线性模型。该模型的核心式子如下

$$f(x) = sign(w \cdot x + b)$$

损失函数

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

该式子表示所有误分类点到分类超平面的距离

点到直线距离公式推导

$$d = \frac{|(x_1 - x) \cdot w^{'}|}{\|w'\|}$$

$$= \frac{|x_1 \cdot w^{'} - x \cdot w|}{\|w'\|}$$

$$x \cdot w^{'} = 0$$

$$= \frac{|x_1 \cdot w + b|}{\|w'\|}$$

$$distance = \frac{|x_1 \cdot w + b|}{\|w'\|}$$

由 sign 函数的性质可以知道，对于所有误分类的点满足 $y(w^T + b) < 0$,同时发现 □$\|w'\|$ 是一个固定值。因此可以得到损失函数和所有误分类点到分类超平面的距离和等价。

学习算法:

我们可以将这个问题转化成一个等价的最优化问题。

目标函数可以如下表示:

$$minL(w,b) = min_{w,b} \sum_{x_i \in M} -y_i(w^T x_i + b)$$

我们可以利用梯度下降的方法来求解这个问题。损失函数的梯度如下:

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

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

在训练的过程中，一旦遇见误分类的点就对 w，b 进行更新。

$$w = w + \alpha y_i x_i$$

$$b = b + \alpha y_i$$

算法可行性:

证明:

收敛性：即经过有限次迭代后可以得到一个将训练集完全正确划分的分离超平面以及对应的感知器模型。

定义 **0.1** *Novikoff* 定理：设训练集 $T = (x_1, y_1), (x_2, y_2), ..., (x_n, y_n)$，是线性可分得，其中 $x_i \in \chi = R^n, y_i \in Y = \{-1, +1\}, i = 1, 2, 3, .., N$ ，则
(1) 存在满足条件的 $\|W'_{opt}\|$ 的超平面，$W'_{opt} \cdot x' = W_{opt} \cdot x + b_{opt} = 0$ 将数据完全分开，且存在 $\gamma > 0$ ，对所有 *i = 1, 2, 3, ..., N* 满足：

$$y_i(W'_{opt} \cdot x_{opt} + b_{opt}) \geq \gamma$$

(2) 令 $R = max\|x'_i\|$ ，对于感知器误分类次数 K，满足：

$$k \leq (\frac{R}{\gamma})^2$$