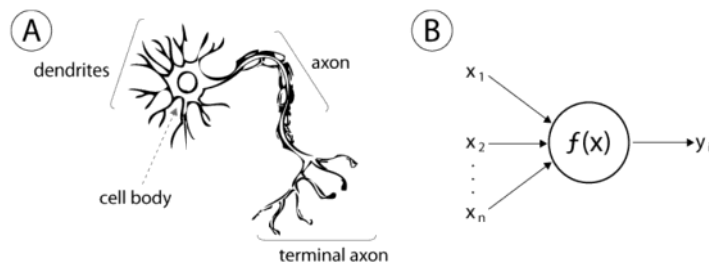


# Class 4 感知机

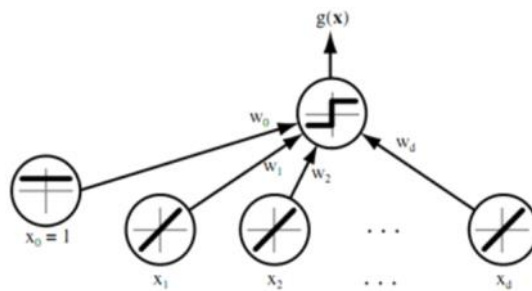
2018年10月19日 10:23

## • 简介:

- 感知机是一种监督分类的算法;
- 感知机是一种线性分类器, 它为人工神经网络 (ANN) 奠定了基础。



## • 模型假设:



$$h_w(x) = \begin{cases} 1 & \text{if } w^T x \geq 0 \\ 0 & \text{if } w^T x < 0 \end{cases}$$

- 上式为二分类问题的模型假设,  $w$  为权重,  $x$  为样本;
- $w^T x \geq 0$  时预测为正类(1),  $w^T x < 0$  时预测为负类(0)。

## • 感知机算法:

### • 成本函数

$$\begin{aligned} J_P(x) &= \sum_{x^{(i)} \in M_0} \omega^T x^{(i)} - \sum_{x^{(j)} \in M_1} \omega^T x^{(j)} \\ &= \sum_{i=1}^N \left( (1 - y^{(i)}) h_w(x^{(i)}) - y^{(i)} (1 - h_w(x^{(i)})) \right) \omega^T x^{(i)} \\ &= \sum_{i=1}^N (h_w(x^{(i)}) - y^{(i)}) \omega^T x^{(i)} \end{aligned}$$

- 上式表示将  $x$  错分的成本之和, 若正确分类则成本为0。
- 第一项表示真实类别为0, 而预测类别为1, 此时  $\omega^T x \geq 0$ ;

- 第二项表示真实类别为1，而预测类别为0，此时 $\omega^T x \leq 0$ ；
- 因此可得， $J_p(x)$ 始终为正值。

#### • 随机梯度下降算法

$$\begin{aligned}\omega &:= \omega + \alpha (y - h_w(x))x && \text{Error} \times \text{Feature} \\ &= \begin{cases} \omega + \alpha x, & \text{if } y = 1 \text{ and } h_w(x) = 0 \\ \omega - \alpha x, & \text{if } y = 0 \text{ and } h_w(x) = 1 \\ \omega, & \text{others} \end{cases}\end{aligned}$$

- Step1: 利用初始  $\omega$  和  $\alpha$  对所有样本进行预测；
- Step2: 随机挑选一个样本来更新参数 $\omega$ ；
- Step3: 利用更新后的参数再次对所有样本进行预测；
- 重复以上步骤。

#### • 多分类感知机：

##### • 简介：

- 多分类感知机是标准感知机的拓展，用于解决多类问题；
- 多分类感知机广泛用于NLP。

##### • 模型假设：

$$C^* = \arg \max_{j=1,\dots,C} \omega_j^T x$$

上式表示，样本的预测类别是  $\omega^T x$  最大哪一个类

##### • 成本函数：

$$J_p(w) = \sum_{k=1}^N \left( \max_{j=1,\dots,C} \omega_j^T x^{(k)} - \omega_{y^{(k)}}^T x^{(k)} \right)$$

上式表示，对每一个样本，预测的类别对应的 $\omega^T x$ 与真实的类别 $\omega^T x$ 之差，即为一个样本的成本，对所有样本的成本求和，即为总成本。

##### • 参数更新规则：

$$\begin{aligned}\omega_j &:= \omega_j - \alpha (1\{j = c^{(k)}\} - 1\{j = y^{(k)}\})x^{(k)} \\ &= \begin{cases} \omega_j - \alpha x^{(k)}, & \text{if } j = c^{(k)} \neq y^{(k)} \\ \omega_j + \alpha x^{(k)}, & \text{if } j = y^{(k)} \neq c^{(k)} \\ \omega_j, & \text{others} \end{cases}\end{aligned}$$

$$\text{where } c^{(k)} = \arg \max_{j=1,\dots,C} \omega_j^T x^{(k)}$$

其中 $c^{(k)}$ 为预测出的样本类别。

#### • 区别与联系：

- **感知机与逻辑回归：**

- 两者都是线性分类器，都可以解决二分类问题；
- 两者的损失函数不同；逻辑回归使用对数损失函数，感知机使用预测值与真实值之差的和作为损失；
- 两者的激活函数不同；逻辑回归的激活函数为Sigmoid函数，感知机的激活函数为阶跃函数。

- **多类感知机与Softmax回归：**

- 二者都是非线性分类器，都用来解决多分类问题；
- 多类感知机是标准感知机的扩展，Softmax回归是Logistics回归的扩展。