

Author: Liu Jian

Time: 2020-06-03

机器学习9-逻辑斯谛回归

1 逻辑斯谛回归/对数几率回归

2 附录 - 逻辑斯谛分布 (logistic distribution)

机器学习9-逻辑斯谛回归

1 逻辑斯谛回归/对数几率回归

逻辑斯谛回归源自逻辑斯谛分布，是一种**对数线性模型**，虽然叫“回归”，但却是一个二分类模型。

模型：

$$\ln \frac{p(y=1|\mathbf{x})}{p(y=0|\mathbf{x})} = \mathbf{w}^T \mathbf{x} + b$$

即：

$$p(y=1|\mathbf{x}) = \frac{1}{1 + e^{-(\mathbf{w}^T \mathbf{x} + b)}}$$
$$p(y=0|\mathbf{x}) = \frac{1}{1 + e^{\mathbf{w}^T \mathbf{x} + b}}$$

逻辑斯谛回归比较两个条件概率值的大小，将实例 \mathbf{x} 分到概率值较大的那一类。一个事件的几率 (odds) 是指该事件发生的概率 p 与不发生的概率 $1-p$ 的比值 $\frac{p}{1-p}$ ，对数几率 $\ln \frac{p}{1-p}$ 。

参数估计方法：极大似然估计，即最大化对数似然函数（等价于最小化交叉熵）。

优化问题求解方法：梯度下降法、牛顿法等。

进一步，推广到多分类问题即得**多项逻辑斯谛回归模型** (multi-nominal logistic regression model)：

$$\mathbb{P}_r(y=k|\mathbf{x}) = \frac{e^{\mathbf{w}_k^T \mathbf{x} + b_k}}{1 + \sum_{i=1}^{K-1} e^{\mathbf{w}_i^T \mathbf{x} + b_i}}, \quad k=1, 2, \dots, K-1$$
$$\mathbb{P}_r(y=K|\mathbf{x}) = \frac{1}{1 + \sum_{i=1}^{K-1} e^{\mathbf{w}_i^T \mathbf{x} + b_i}}$$

2 附录 - 逻辑斯谛分布 (logistic distribution)

定义：设 X 是连续型随机变量，其分布函数 $F(x)$ 和密度函数 $f(x)$ 如下：

$$F(x) = \mathbb{P}_r(X \leq x) = \frac{1}{1 + e^{-(x-\mu)/\gamma}}$$
$$f(x) = F'(x) = \frac{e^{-(x-\mu)/\gamma}}{\gamma(1 + e^{-(x-\mu)/\gamma})^2}$$

其中， μ 为位置参数， γ 为形状参数。

分布函数的图形是一条 S 形曲线 (sigmoid curve), 该曲线以 $(\mu, \frac{1}{2})$ 为中心对称, 形状参数 γ 的值越小, 曲线在中心附近增长得越快。

Sigmoid 函数:

$$y = \frac{1}{1 + e^{-x}}$$
