逻辑回归

1. 概念

1.1. 概念：逻辑回归是一个分类模型，通过给样本数据打分进行分类，大于阀值归一类，反之亦然

导入：from sklearn.linear_model import LogisticRegreession

引入 sigmoid 函数：sigmoid 函数是由伯努利分布推导出来的，在一个取值连续的变量上(逻辑回归的预测值 Y)，可以将其映射到 0~1 的区间上，也就是概率，当概率大于 0.5 时归为一类，反之亦然，这样就对自变量 x(特征)起到了分类作用

注：求出的概率是归为正例的概率

2. 逻辑回归的损失函数由来和求解

2.1 损失函数的由来

逻辑回归损失函数.

1. 背景：损失函数的目的在于求出联合密度最大的w值.

2. 联合概率的密度函数.

$$L(w)=\prod_{i=1}^{m} P(y^{(i)}|x^{(i)};w)$$

$$=\prod_{i=1}^{m} S(z^{(i)})^{y^{(i)}}\left(1-S(z^{(i)})^{1-y^{(i)}}\right)$$

取对数. 乘积变加和

$$\ln L(w)=\ln\left[\prod_{i=1}^{m} S(z^{(i)})^{y^{(i)}}\left(1-S(z^{(i)})\right)^{1-y^{(i)}}\right)$$

$$=\sum_{i=1}^{m}\left[y^{(i)}\ln S(z^{(i)})+(1-y^{(i)})\ln(1-S(z^{(i)}))\right.$$

3. 引入损失概念. 变为负数.

$$J(w)=-\sum_{i=1}^{m}\left[y^{(i)}\ln S(z^{(i)})+(1-y^{(i)})\ln(1-S(z^{(i)}))\right]$$

$$\frac{\partial \ln A}{\partial w_j}=\frac{\partial \ln A}{\partial A}\cdot$$

损失函数的求解

逻辑回归损失函数求导.

$$J(w) = -\sum_{i=1}^{m}\left[y^{(i)}\ln S(z^{(i)}) + (1-y^{(i)})\ln(1-S(z^{(i)}))\right]$$

w_j表示对其中某一项求导，加和就是全部

$$\frac{\partial J(w)}{\partial w_j} = \frac{\partial}{\partial w_j}\left(-y^{(i)}\ln S(z^{(i)}) + (1-y^{(i)})\ln(1-S(z^{(i)}))\right)]$$

∵ 将$S(z^{(i)})$、$1-S(z^{(i)})$看成一个整体求导、$\frac{\partial \ln A}{\partial w_j} = \frac{\partial \ln A}{\partial A}\cdot\frac{\partial A}{\partial w_j}$

$$= -\left(\frac{y^{(i)}}{S(z^{(i)})} - \frac{1-y^{(i)}}{1-S(z^{(i)})}\right)\frac{\partial S(z^{(i)})}{\partial w_j}$$

$$= -\left(\frac{y^{(i)}}{S(z^{(i)})} - \frac{1-y^{(i)}}{1-S(z^{(i)})}\right)\frac{\partial S(z^{(i)})}{\partial z}\cdot\frac{\partial z}{\partial w_j}$$

$$= -\left(\frac{y^{(i)}}{S(z^{(i)})} - \frac{1-y^{(i)}}{1-S(z^{(i)})}\right)\cdot\left[S(z)(1-S(z))\right]\cdot\frac{\partial z}{\partial w_j}$$

$$= -\left(\frac{y^{(i)}}{S(z^{(i)})} - \frac{1-y^{(i)}}{1-S(z^{(i)})}\right)S(z)(1-S(z))\cdot x_j$$

$$= -\left[y^{(i)}(1-S(z)) - (1-y^{(i)})S(z)\right]\cdot x_j$$

$$= -\left[y^{(i)} - y^{(i)}S(z) - S(z) + y^{(i)}S(z)\right]\cdot x_j$$

$$= -(y^{(i)} - S(z))x_j^{(i)}$$

sigmoid求导.

$$S'(z) = \left(\frac{1}{1+e^{-z}}\right)'$$

令$-z = k$

$$= -(1+e^{-z})^{-2}\cdot(e^{k})'\cdot(k)'$$

$$= -(1+e^{-z})^{-2}\cdot e^{-z}\cdot(-1)$$

$$= -(1+e^{-z})^{-2}\cdot(-e^{-z})$$

变换形式

$$= \frac{1}{1+e^{-z}}\cdot\frac{e^{-z}}{1+e^{-z}}$$

$$= \left(\frac{1}{1+e^{-z}}\right)\left(\frac{1+e^{-z}}{1+e^{-z}} - \frac{e^{-z}}{1+e^{-z}}\right)$$

$$= \left(\frac{1}{1+e^{-z}}\right)\left(1 - \frac{1}{1+e^{-z}}\right)$$

$$= S(z)(1-S(z))$$

$$\frac{\partial z}{\partial w_j} = \frac{\partial}{\partial w_j}(w_1x_1 + w_2x_2 + w_3x_3 + \cdots w_nx_n + b)$$

$$= \frac{\partial w_jx_j}{\partial w_j}$$ 只有$w_j x_j$对w_j有关系

$$= x_j$$

3. 逻辑回归为什么使用 sigmoid 函数

3.1 sigmoid 函数推导

Sigmoid函数推导

背景：由伯努力分布可以推导出Sigmoid函数

伯努力分布的质量函数

$$f(x) = p^x \cdot (1-p)^{1-x}$$

$$f(x) = p^x \cdot (1-p)^{1-x}$$

$$= e^{\ln[p^x \cdot (1-p)^{1-x}]}$$

$$= e^{x \cdot \ln p + (1-x)\ln(1-p)}$$

$$\cancel{= e^{x \ln p + \ln(1-p) - x\ln(1-p)}}$$

$$= e^{x \cdot \ln p + \ln(1-p) - x\ln(1-p)}$$

$$= e^{x \cdot (\ln p - \ln(1-p)) + \ln(1-p)}$$

$$= e^{x \cdot \ln\frac{p}{1-p} + \ln(1-p)}$$ 常数不关心

令 $\eta = \ln\frac{p}{1-p}$

$$e^{\eta} = \frac{p}{1-p}$$

$$e^{\eta}(1-p) = p$$

$$p = \frac{e^{\eta}}{1+e^{\eta}} = \frac{1}{1+e^{-\eta}}$$