2022_Winter_ML

代码连接：

https://github.com/Buzzy0423/2022_Winter_ML

LinearRegression

基本公式$Y = W^TX + B$, 损失函数$L = \frac{1}{2}(Y` - Y)^2$,权重直接由损失函数求导得出

即$\frac{dL}{dW} = (W^TX - Y)X = 0$（bias算作w0）, 推导出$W = (X^TX)^{-1}(X^TY)$

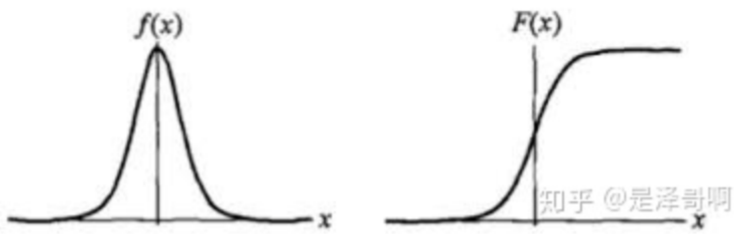
整个过程无需训练

Logistic Regression

Logistic分布

分布函数如下

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



其中，当$\mu = 0$,$\gamma = 1$ 时，即是sigmoid函数。

Logistic函数的一个良好性质为：$\ln \frac{y}{1-y} = \omega^T x + b$

将 y 视为 x 为正例的概率，则 1-y 为 x 为其反例的概率。两者的比值称为**几率（odds）**，指该事件发生与不发生的概率比值，若事件发生的**概率**为 p。则对数几率：$ln(odd) = ln\frac{y}{1-y}$

Logistic Regression的正向传播就是在LinearRegression后加了一个sigmoid激活函数。

Logistic分布的似然函数为$L(w) = \Pi[p(x_i)^{y_i}][1 - p(x_i)]^{1-y_i} = \Sigma[y_i(w^Tx) - ln(1 + e^{w^Tx_i})]$

我们定义Logistic Regression的Loss为$J(w) = -\frac{1}{N}lnL(w)$

对其求导得$g_i = \frac{dJ(w)}{dw_i} = (p(x_i) - y_i)x_i$

最后每轮根据学习率更新$w_i^{k+1} = w_i^k - \mu g_i$

NN

原理见人工智能导论课件