

损失函数: 衡量模型输出与真实标签的差异

$Loss = f(\hat{y}, y)$ 计算一个样本

$Cost = \frac{1}{N} \sum_i f(\hat{y}_i, y_i)$ 计算所有样本

$Obj = Cost + Regularization.$

交叉熵 = 信息熵 + 相对熵

熵(信息熵) \rightarrow 事件发生的不确定性大小

相对熵 \rightarrow 衡量两个分布的距离

`nn.CrossEntropyLoss(weight, size_average, ignore_index, reduce, reduction)`

weight: 各类型的 loss 权重

ignore_index: 忽略某个类别

reduction: 计算模式 none/sum/mean

`nn.NLLLoss(...)`

负对数似然函数(对样本值取负号)

nn.BCELoss(...)

二分类交叉熵

nn.BCEWithLogitsLoss(...)

自动为 input 取 sigmoid