

损失函数：衡量模型输出与真实标签的差异

$$Loss = f(\hat{y}, y) \quad \text{计算一个样本}$$

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

$$Obj: Cost + Regularization.$$

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