



常用损失函数速览



Uncle.D

D大叔 来自 Plaid.tech (格子衫们 筹建中)

+ 关注他

损失函数是机器学习和深度学习中非常重要的概念，它用于衡量模型预测结果与真实值之间的差距。不同的任务和模型需要使用不同的损失函数。

1. 均方误差损失函数⁺ (Mean Squared Error, MSE)

均方误差损失函数是最常用的[回归损失函数⁺](#)之一，用于衡量预测值与真实值之间的差异。其公式如下：

$$L(y, \hat{y}) = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

其中， y 是真实值， \hat{y} 是预测值， n 是样本数量。

均方误差损失函数的优点是计算简单，且对异常值具有[鲁棒性⁺](#)。但是，它对大误差的惩罚较大，可能导致模型过于关注大误差样本。

2. 平均绝对误差⁺损失函数 (Mean Absolute Error, MAE)

平均绝对误差损失函数是另一种常用的回归损失函数，用于衡量预测值与真实值之间的差异。其公式如下：

$$L(y, \hat{y}) = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$$

其中， y 是真实值， \hat{y} 是预测值， n 是样本数量。

平均绝对误差损失函数的优点是对大误差的惩罚较小，且计算简单。但是，它对异常值不够鲁棒。

3. 交叉熵损失函数⁺ (Cross-Entropy Loss)

交叉熵损失函数是最常用的分类损失函数之一，用于衡量预测[概率分布⁺](#)与真实分布之间的差异。其公式如下：

$$L(y, \hat{y}) = -\frac{1}{n} \sum_{i=1}^n \sum_{j=1}^m y_{ij} \log(\hat{y}_{ij})$$

其中， y 是真实分布， \hat{y} 是预测分布， n 是样本数量， m 是类别数量。

交叉熵损失函数的优点是对预测概率分布的估计具有较好的鲁棒性。但是，它对类别不平衡问题不够敏感。

4. 二元交叉熵损失函数[★] (Binary Cross-Entropy Loss)

二元交叉熵损失函数是交叉熵损失函数的变体，用于二分类任务。其公式如下：

$$L(y, \hat{y}) = -\frac{1}{n} \sum_{i=1}^n [y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)]$$

其中， y 是真实值， \hat{y} 是预测值， n 是样本数量。

二元交叉熵损失函数的优点是对预测概率分布的估计具有较好的鲁棒性，且计算简单。但是，它对类别不平衡问题不够敏感。

5. 对数损失函数[★] (Log Loss)

对数损失函数是另一种常用的分类损失函数，用于衡量预测概率分布与真实分布之间的差异。其公式如下：

$$L(y, \hat{y}) = -\frac{1}{n} \sum_{i=1}^n y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)$$

其中， y 是真实值， \hat{y} 是预测值， n 是样本数量。

对数损失函数的优点是对预测概率分布的估计具有较好的鲁棒性，且计算简单。但是，它对类别不平衡问题不够敏感。

6. Huber损失函数 (Huber Loss)

Huber损失函数是均方误差损失函数和平均绝对误差损失函数的折中，用于回归任务。其公式如下：

$$L(y, \hat{y}) = \begin{cases} \frac{1}{2}(y_i - \hat{y}_i)^2 & \text{if } |y_i - \hat{y}_i| \leq \delta \\ \delta|y_i - \hat{y}_i| - \frac{1}{2}\delta^2 & \text{otherwise} \end{cases}$$

其中， y 是真实值， \hat{y} 是预测值， δ 是阈值。

Huber损失函数的优点是对大误差的惩罚较小，且计算简单。但是，它对异常值不够鲁棒。

7. 指数损失函数⁺ (Exponential Loss)

指数损失函数是另一种常用的分类损失函数，用于衡量预测概率分布与真实分布之间的差异。其公式如下：

$$L(y, \hat{y}) = \frac{1}{n} \sum_{i=1}^n e^{-y_i \hat{y}_i}$$

其中， y 是真实值， \hat{y} 是预测值， n 是样本数量。

指数损失函数的优点是对预测概率分布的估计具有较好的鲁棒性，且计算简单。但是，它对类别不平衡问题不够敏感。

8. 合页损失函数⁺ (Hinge Loss)

合页损失函数是另一种常用的分类损失函数，用于支持向量机⁺ (SVM) 模型。其公式如下：

$$L(y, \hat{y}) = \frac{1}{n} \sum_{i=1}^n \max(0, 1 - y_i \hat{y}_i)$$

其中， y 是真实值， \hat{y} 是预测值， n 是样本数量。

合页损失函数的优点是对大误差的惩罚较大，且计算简单。但是，它对异常值不够鲁棒。

损失函数的选择取决于任务和模型类型。对于回归任务，常用的损失函数有均方误差损失函数和平均绝对误差损失函数。对于分类任务，常用的损失函数有交叉熵损失函数和二元交叉熵损失函数。对于多标签分类任务，常用的损失函数有二元交叉熵损失函数和二元交叉熵损失函数的变体。