

数据标准化

让每一个特征都服从标准正态分布，消除了每个特征分布不同从而导致结果不同的这个因素

数据标准化就是将原来分布范围不同的数据缩放在一个范围之内，一般来说是标准化到均值为 0，标准差为 1 的标准正态分布，均值为 0 是为了让数据中心化，让各个维度的特征都平衡，标准差为 1 是为了让数据在各个维度上的长度都呈现一个单位向量(矩阵)，也就是把原来大小范围分布不通的数据缩放成统一量纲。

数据可以分为有界数据和无界数据。有界数据就是有着明确的数据边界，也就是数据是固定的，无界数据就是没有明确数据边界，也就是说数据是有可能改变的。

对于有界数据的标准化方法，由于数据的边界是固定的，显然会和数据的数量以及最小值、最大值有关系，所以一般会将原数据与数据数量大小、最小值及最大值进行计算操作，从而将数据按照一定的比例缩放在一定的范围内。无界数据是指数据的边界不一定是固定的，也就是数据的边界不确定，随时有可能发生变化，或者不知道数据的边界，这类数据主要是使用均值方差的方法标准化。

一、数据标准化的方法

1.1 最大值标准化

最大值标准化就是让数据中的每个值都除以最大值，把数据缩放到[0,1]之间：

$$x = \frac{x}{\max(x)} \quad (1 - 1)$$

这种方法适合数据都是正数的情况，比如图像像素是[0,255]之间的正数

1.2 绝对最大值标准化

数据中存在负数时，先对数据取绝对值，然后再进行最大值标准化把数据缩放到[0,1]之间：

$$x = \frac{x}{\max(|x|)} \quad (1 - 2)$$

1.3 最大最小值标准化

就是让每一个数据都减去最小值，然后除以最大值减去最小值的结果，将数据缩放到[0,1]之间：

$$x = \frac{x - \min(x)}{\max(x) - \min(x)} \quad (1 - 3)$$

1.4 均值方差标准化

均值方差（标准差）标准化是最为常用的数据标准化方法，操作过程就是让每个数据都减去均值，然后除以标准差： $x = \frac{x - \text{mean}(x)}{\sqrt{x - \text{mean}(x)}^2}$ ，目的是将数据缩放到均值为 0，标准差为 1 的标准正态分布 $N[0,1]$ 。

二、 批量规范化

仅仅对原始输入数据进行标准化是不充分的，因为虽然这种做法可以保证原始输入数据的质量，但它却无法保证隐藏层输入数据的质量。浅层参数的微弱变化经过多层线性变换与激活函数后被放大，改变了每一层的输入分布，造成深层的网络需要不断调整以适应这些分布变化，最终导致模型难以训练收敛。

在 BN 层中，输入 $x : B = \{x_1, \dots, m\}$ ，输出：规范化后的网络响应 $\{y_i = \text{BN}_{\gamma, \beta}(x_i)\}$

2.1 计算数据均值

$$\mu_B = \frac{1}{m} \sum_{i=1}^m x_i \quad (2 - 1)$$

2.2 计算数据方差

$$\sigma_B^2 = \frac{1}{m} \sum_{i=1}^m (x_i - \mu_B)^2 \quad (2 - 2)$$

2.3 规范化

$$\hat{x}_i = \frac{x_i - \mu_B}{\sqrt{\sigma_B^2 + \varepsilon}} \quad (2 - 3)$$

三、 引用

[1] 《7.5. 批量规范化 — 动手学深度学习 2.0.0 documentation》. 见于 2024 年 7 月 9 日. https://zh.d2l.ai/chapter_convolutional-modern/batch-norm.html.

[2] 《[1502.03167] Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift》. 见于 2024 年 7 月 9 日. <https://arxiv.org/abs/1502.03167>. 《批量规范化 (Batch Normalization, BN) 激活函数和批量标准化的区别-CSDN 博客》. 见于 2024 年 7 月 9 日. <https://blog.csdn.net/jgj123321/article/details/105291672>.

[3] 《深度学习之数据标准化方法综述 ai 数据标准化-CSDN 博客》. 见于 2024 年 7 月 9 日. <https://blog.csdn.net/DeepAIedu/article/details/124281964>.

[4] Ioffe, Sergey, 和 Christian Szegedy. 《Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift》. arXiv, 2015 年 3 月 2 日. <https://doi.org/10.48550/arXiv.1502.03167>.

[5] 《Pytorch 归一化(MinMaxScaler、零均值归一化)-CSDN 博客》. 见于 2024 年 7 月 9 日. https://blog.csdn.net/qg_36158230/article/details/120925154.