

20220523-机器学习

1.学习内容

1.1 机器学习

批量规范化

为什么需要批量规范化

如何使用批量规范化

2.结果描述

1.学习内容

1.1 机器学习

批量规范化

为什么需要批量规范化

- 数据预处理的方式对于最终结果会产生巨大影响。在使用数据时，我们的第一步通常是标准化输入特征，使其平均值为0，方差为1。
- 神经网络中间层的模型参数随着训练更新变幻莫测。这些变量分布中的这种偏移可能会阻碍网络的收敛。
- 深层网络很复杂，容易过拟合。

如何使用批量规范化

批量规范化可以应用于单个层，也可以应用到所有层。在每次训练迭代中，首先通过规范化输入（即减去均值并除以其标准差，均基于当前小批量处理），之后应用比例系数和比例偏移。

$$\text{BN}(\mathbf{x}) = \gamma \odot \frac{\mathbf{x} - \hat{\mu}_{\mathcal{B}}}{\hat{\sigma}_{\mathcal{B}}} + \beta.$$

$$\hat{\mu}_B = \frac{1}{|B|} \sum_{\mathbf{x} \in B} \mathbf{x},$$

$$\hat{\sigma}_B^2 = \frac{1}{|B|} \sum_{\mathbf{x} \in B} (\mathbf{x} - \hat{\mu}_B)^2 + \epsilon.$$

γ 为拉伸参数， β 为偏移参数，它们的形状与 \mathbf{x} 相同，需要与其它模型参数一起学习。其中计算方差在最后面添加了一个小的大于零的常量，以确保不会除以零。

- 对于全连接层，可以将批量规范化层置于仿射变换和激活函数之间：

$$\mathbf{h} = \phi(\text{BN}(\mathbf{W}\mathbf{x} + \mathbf{b})).$$

- 对于卷积层，可以在卷积层之后和非线性激活函数之前应用批量规范化。当卷积有多个输出通道时，需要对这些通道的每个输出执行批量规范化，每个通道都有自己的拉伸和偏移参数，这两个参数都是标量。

2.结果描述

这两天有点不在状态，顾着看msi比赛，基本没把心思放在学习上。今天只稍微看了点动手深度学习中的批量正则化。要收收心了。