**《深度学习与大数据智能》实验报告二**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **年级、专业、班级** | | **20计算机科学与技术（卓越）01** | | **姓名** | | **黄昊** | **学号** | **20204205** |
| **实验题目** | **FashionMnist分类任务正则化对比实验** | | | | | | | |
| **实验时间** | **2023/05/07** | | **实验地点** | | **DS3302** | | | |
| **学年学期** | **2021-2021（2）** | | **实验性质** | | **□验证性 □设计性 ■综合性** | | | |
| 一、实验目的  1. 进一步理解卷积神经网络的基本构成。  2. 理解并掌握正则化技术在卷积神经网络中的作用。 | | | | | | | | |
| 二、实验项目内容  1. 基本内容：构建基于Fashion-MNIST数据集的分类卷积神经网络模型，该模型中需使用相应的正则化技术。  2. 进阶内容：  1）基于基本实验内容，进行L1/L2/Dropout/BN正则化前后对比实验。  2） 阅读相关文献，使用上述1）中未提到的一种正则化技术进行实验，并分析实验结果。  3）分析实验中使用的正则化技术的特点及其使用范围。 | | | | | | | | |
| 三、实验的过程和主要源程序  ①原理介绍  在模型训练过程中，常常会碰到过拟合问题。过拟合问题表现为，在训练集上的测试性能良好，但在训练集上的表现非常差，即反映了模型的泛化性能较差。而为了解决过拟合问题，我们在模型训练过程中常常会使用各种各样的正则化技术，以缓解模型的过拟合问题。下面介绍一些常见的正则化技术：  A.L1正则化与L2正则化  L1正则化和L2正则化的思想类似，都是在原先的损失函数上面加上一个正则化项目，对于L1，有：  L2，有：  其中*L0*是原来的损失函数，*w*是模型权重。这两个正则化技术的思想类似，认为权重过大有向训练集过拟合的趋势，于是希望模型权重不那么大，以缓解过拟合。其中*λ*是超参数，需要根据模型不断调节/  B.BatchNorm正则化  首先，对于一个模型的输入，我们常常会作正则化处理：  其中*μ*为均值，*σ*为输入的标准差。因此，我们也想在模型中加入一个BatchNorm层，这一层的作用是对上一层的输出作正则化处理。因此，本层需要学习两个参数：*μ*和*σ。*  C.Dropout正则化  Dropout的作用是在每个训练批次中，通过忽略一定比例的节点（让这些节点输出恒定为0），使得减少过拟合现象，减少各个节点之间的相互作用，以提高泛化能力。  D.Early Stop  根据相关文献指出，在模型训练过程中，刚开始训练的时候，模型通常处于欠拟合状态。而到达一定时刻后，模型开始有过拟合的倾向。而Early Stopping的思想是关注测试集上的一定指标，当这个指标开始变差时，就停止训练，而不是训练到指定的训练次数。  ②模型搭建  结合本次实验内容，本次实验搭建了四种模型：不带任何正则化层的模型，带Dropout层的模型，带BatchNorm层的模型和带BatchNorm和Dropout层的模型，如下从左到右依次如下所示：    ③主要源程序 | | | | | | | | |
| 四、实验结果及分析和（或）源程序调试过程、实验总结与体会  1、实验运行结果截图  2、程序编写过程中遇到的问题及解决思路、方法等  本次实验遇到的问题主要是对mindspore框架的callback回调函数的不熟，主要是mindspore的训练过程直接封装到train函数里面去了，然后对mindspore的设计思路不熟。最后查阅文档，发现mindspore对于每个模型设计了一个metric类，metric类可以传入常见的Accuracy和Loss作为度量标准进行计算，最后callback函数里面可以读取metric来打印中间的loss和accuracy变化过程。最后主要的问题还是来自于对mindspore设计的不熟，这只能多查文档，多查资料一步步解决。  另外，在实验时，发现L1和L2正则化使用后，分类效果急剧下降，接近于随机：    通过分析公式并思考，正则项是权重大小的绝对值或绝对值平方之和，此时正则化项的超参数系数如果设置得过大，模型将会倾向于使模型的所有权重趋于0，导致欠拟合。因此，正则化项的超参数选取十分重要。  3、实验总结和体会  ①正则化效果前后对比  ②Early Stop效果及分析  ③正则化技术的特点及使用范围  ④最终总结与体会 | | | | | | | | |

实验报告填写说明：

1、第一、二部分由老师提供；

2、第三部分填写源程序和算法，源程序要符合程序编写风格（缩进、注释等）；

3、第四部分主要填写程序结果（截图）、解决问题的方法、总结和体会等；

4、报告规范：包含报告页眉、报告的排版、内容是否填写，命名是否规范等。

5、源程序和实验报告命名：学号姓名序号.py 学号姓名序号.docx，例如学号20181234的张三同学，**他的第一次实验命名为：20181234张三1.py和20181234张三1.docx**

**6、向蓝墨云提交两个文件，一个实验报告，一个程序文件（代码合并）**