|  |  |  |
| --- | --- | --- |
| 第二章 神经网络设计实验 | | |
| 姓名： 苏云鹤 | 学号：21373007 | 班级： 212114 |
| **注意：课程实验代码和报告需独立完成，严禁抄袭!如发现有抄袭情况（同年级或者抄袭高年级），相应的实验或报告部分直接0分。** 实验2.1基于三层神经网络实现手写数字分类(60’’) **1.介绍一下“手写数字分类任务”和本实验中使用的数据集。**  实验所采用的数据集是 MNIST 手写数字库。这是一个常用的手写数字图像数据集，它包含了大量的手写数字图像，用于机器学习和图像处理领域的训练和测试。MNIST数据集由60,000个训练图像和10,000个测试图像组成，其中每张图像的大小为28x28像素。图像的像素值按行优先顺序存放，取值范围为[0, 255]，其中0表示黑色，255表示白色。  **2.本实验的神经网络都包含哪些层？各个层的功能是什么？**  本实验的神经网络包含3个全连接层、2个ReLU激活函数层以及1个Softmax层。它们的功能分别是：   1. 全连接层：将前一层的输出向量转换为当前层的输入向量，将多个输入特征进行线性组合并得到新的特征表示。具体而言，它以一维向量作为输入，输入在与权重相乘后再与偏置相加得到输出向量。其中，权重和偏置项都可以在学习的过程中被更新，从而更精确地反映特征。 2. ReLU激活函数层：作用是对输入信号进行非线性映射，引入非线性因素，从而增强模型的表达能力并提高其分类效果。具体来说，ReLU函数定义为，即当输入为正数时，输出等于输入；当输入为负数时，输出为0。ReLU函数的特点是简单、快速，并且能够有效地解决梯度消失问题，使得网络能够更加深层、更加复杂，提高了神经网络的表达能力。 3. Softmax损失层：将神经网络输出的原始分数（也称为logits）转化为概率分布，使得模型能够对不同类别进行概率预测，并计算预测结果与真实标签之间的损失。   **3.网络的训练(training)过程和网络的推理(inference)过程有什么区别？**  训练过程的目的是获得更加准确的网络模型。为此，过程中需要使用专门的训练数据集，通过神经网络的前向传播和反向传播不断更新修正模型的参数。  推理过程的目的是使用已经训练好的神经网络进行预测等工作。为此，过程中需要加载训练好的模型参数，读入测试样本，进行网络的前向传播，并输出结果。  **4.参数和超参数的区别是什么？本实验中都使用了哪些超参数？**  参数指的是模型中可学习的变量，这些变量的值会通过训练数据进行自动调整以最小化损失函数。对于参数，模型可以根据训练数据进行学习和调整，使得模型能够适应特定的任务和数据分布。参数的最终值是由训练过程中的优化算法决定的，目标是通过不断迭代来寻找损失函数的最小值，从而使模型的预测结果更加准确。  相比之下，超参数是在模型训练之前需要手动设置的一组参数。超参数不会通过训练数据进行学习和调整，而是由人工或经验确定的。超参数包括但不限于学习率、批大小、迭代次数、正则化参数等。超参数的选择往往需要经验和实验来确定，因为不同的超参数组合可能会对模型的性能产生显著影响。  本次实验中使用到的超参数有：  batch\_size=100, input\_size=784, hidden1=158, hidden2=64, out\_classes=10, lr=0.01, max\_epoch=1, print\_iter=100  **5.什么是正则化？有哪些正则化技术？**  正则化是机器学习中用于减轻或避免模型过拟合的技术。常见的正则化技术有L2正则化、L1正则化、Dropout正则化等。 实验2.2 基于 DLP 平台实现手写数字分类(30’’) **1.什么是量化？量化会带来什么好处/坏处？**  量化是指为了减轻过多网络参数对算力、存储空间带来的压力，将权重、偏置等参数等用低精度的数据表示。  好处：在不对精度产生较大影响的前提下显著加快网络的训练和推理速度；减小存储空间。  坏处：量化后数据位宽太低，则会对结果的精度产生影响；对于简单算子(如 add)，量化之后对速度提升不大，反而会降低精度。  **2.pycnml、CNML库以及DLP之间的关系是什么？**  pycnml 作为接口将 CNML 模型加载到 DLP 上并执行推理任务。具体来说：   1. pycnml 是一个 Python 绑定的库，用于与 CNML (Cambricon Neural Network Model) 相关的操作，提供了方便的接口来加载、运行和管理 CNML 模型。 2. CNML 是神经网络模型描述语言，可以描述和定义神经网络模型的结构和参数。CNML 包含了一系列的操作符和模块，可以实现各种深度学习任务。 3. DLP (Deep Learning Processor) 是 Cambricon 公司基于硬件设计的深度学习处理器，用于高效地执行深度神经网络推理任务。DLP 能够利用 CNML 描述的模型结构和参数进行硬件加速，提供高性能、低功耗的深度学习推理能力。   **3.使用DLP相对于CPU有哪些优势和劣势?**  优势：性能更高，因此计算速度更快；架构专门针对深度学习任务设计，对于 CNN 和 RNN 等常见模型，DLP 的数据传输和计算方式也进行了专门的优化；包含多个计算单元和内存控制器，在执行深度学习任务时可以利用更多的并行计算资源，从而提高计算效率。  劣势：成本较高，限制较多，开发和调试复杂。 | | |