作业一:简单神经网络

一、目标

搭建基本的多层神经网络,并在给定测试集上进行精度测试。

注 1: 不使用深度学习框架完成网络搭建。

注 2: 不限制编程语言,推荐使用 python 进行神经网络搭建,允许使用 numpy 等工具包。

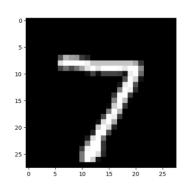
注 3: 使用给定的训练集和测试集,可使用提供的代码模板(bp_template.py)并在其基础上进行修改,也可以重新进行编写。

二、数据集

输入数据为取值 0~9 的手写数字集,以.csv 格式保存。其中训练集包含 60000 张数字图像,测试集包含 10000 张数字图像,每张图像均为 28*28 大小。

csv 文件中每一行都是一个图像记录,每个记录共有 28*28+1=785 个数值,其间以逗号(;')分隔。每个记录的第一个取值为该记录的正确 label,剩余 784 个取值为该图像数据的像素值 (0~255)。

示例:如图所示的手写数字7(显示的是灰度处理后的图片)



其记录结构为(实际上是一行,这里为了方便显示进行了换行):

可以看到这个记录中第一个数值为7,其余数值为它的像素值。

三、网络要求

- 1. 网络输入: 784 个输入节点 (每个节点对应图片的一个像素)
- 2. 网络输出: 10 个输出节点(分别代表 0~9 这 10 个数字)
- 3. 网络深度建议为3至5层即可,如果太深则需要太长运行时间。
- **4.** 使用给定训练集(mnist_train.csv)进行权重训练,使用测试集(mnist_test.csv)测试并给出测试精度。(不对精度做特别的要求,只需在合理范围内即可)

四、作业提交

- 1. 将训练得到的权重数据用 numpy 保存为.npy 格式的文件,让结果可以复现。
- 2. 需撰写结果报告,包括但不限于网络结构介绍、测试精度结果截图等。报告命名格式: 学号+姓名,如 XXXXXXXXX+张三+作业一报告。(更多细节参考"实验报告撰写格式")
- 3. 将代码、权重数据、报告打包成压缩文件,命名格式: 学号+姓名,如 XXXXXXXXX+张三+作业一。

五、注意细节

本次作业给出了 6000 张图片进行训练,如果在 CPU 上训练时间过长,可以考虑采用 GPU,或者减少训练集的大小(但会降低训练后的模型精度)。由于本次作业并不让大家采用深度学习框架,而是采用 numpy 等包,所以要想进行 GPU 加速则可以采用 CuPy,具体可以参考:

https://towardsdatascience.com/heres-how-to-use-cupy-to-make-numpy-700x-faster-4b920dda1f56

https://cupy.dev/

https://github.com/cupy/cupy