

Weekly Report (8)

---By Chen Junlin at UESTC

chenjunlin@std.uestc.edu.cn

Week: 2019-2020-1学期 第14周

Time: 2019/12/2 ~ 2019/12/8

本周完成内容

本周集中火力,

- 完成了 Assignment 1 的 Softmax 部分,
- 观看了 Lecture 4 (Introduction to Neural Network) 视频,
- 开始进行 Two-layer Neural Network 部分的作业。目前已经实现了2层神经网络的前向传播、后向传播、小批量SGD训练, 正在对超参数进行优化。

困难和挑战

- 对于 Softmax 和本周新增的神经网络(包括ReLU激活函数)反向传播的梯度推导, 依然是一个难点。网上找到的很多资料, 在表达式上各有不同, 在推导过程中, 中间省略了很多推导环节, 导致读不懂.....
- 找到一个关于 向量微积分 的[资料](#), 里面介绍关于向量、矩阵的求导, 从最基础的标量求导开始介绍, 非常详细。但自己尝试推导梯度表达式时, 发现其中还有很多坑, 其中的细节也相当麻烦, 甚至一度无法进行下去。最终看网上资料求导出来的结果, 表达式却非常 elegant, 并且易于实现。所以困惑: 究竟是需要从最基础的“标量微积分→向量微积分→梯度”开始推导, 还是直接用推导出的结果编程实现即可?

下周计划

- 期末考试周快要到了, 后面的时间会越来越紧张, 考试周期间(可能从16周左右开始)可能会暂停学习, 全力以赴备战期末。不过个人感觉已经开始慢慢摸到一点门道和套路了, 相信之后的学习过程中会越来越顺手。
- 神经网络的部分在之前有接触, 但仅停留在基础概念和大体感觉。而这一轮学习需要从数学原理上真正理解神经网络的运行机制。后续学习到使用现成的机器学习框架搭建神经网络时, 也许会比现在顺利一些吧。