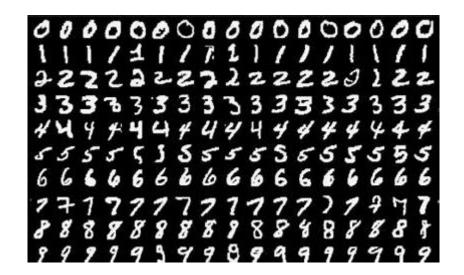
实验一: SVM 图像分类

实验目的

- 1.手动实现一个 SVM 分类器, 完成 MNIST 数据分类任务
- 2.掌握 SVM 的基本原理以及优化流程
- 3.对 SVM 分类器训练的超参数(包括收敛终止条件, 学习率等)进行调优

实验数据

MNIST 数据库是由 Yann 提供的手写数字数据库文件 , 下载地址为: MNIST handwritten digit database, Yann LeCun, Corinna Cortes and Chris Burges。(其中包含数据集格式等详情), 主要包含了 60000 张的训练图像和 10000 张的测试图像



实验内容

基础内容(100%):

1.SVM 的实现

熟悉 SVM 的原理与优化求解 SVM 分类器的算法的过程。

2.MNIST 数据集分类

利用 MNIST 训练集训练、验证,在测试集测试,报告你在测试集上的准确率,并使用现成的库实例化的一个 SVM 分类器(python 下如 sklearn),并同样在 MNIST 上进行测试和训练,对比这个分类器与你设计的 SVM 分类器的结果,如果你设计的分类器效果相较于这个分类器的准确率较低的话,请分析一下原因。

3.对 SVM 进行超参数的搜索

对 SVM 分类器进行超参数的搜索(比如收敛终止条件, 学习率等), 选取你的最优的超参数组合下的 SVM 分类器的优化结果

额外加分项(bonus 50%)

1. 构建使用 kernel 方法的 SVM 分类器

构建你的使用 kernel 方法的 SVM 分类器,以及相关的训练优化流程,利用 MNIST 训练集训练、验证,在测试集测试,报告你在测试集上的准确率

2. 对比不同 SVM 方法

对比一下未使用 kernel 方法的 SVM 分类器和不同 kernel 方法下的 SVM 分类器,分析一下不同方法间的区别.(比如可视化降维后的数据在不同方法下的分类界面)

注意:在设计你自己的 SVM 分类器时,以及在优化你的 SVM 分类器的流程中,请尽可能使用常见的科学计算库(python 中如 numpy)和基本库来构建和优化你的 SVM 分类器。如果使用现成的库来直接构建并优化你的 SVM 分类器将会被酌情扣除相应的分数。如需要进行可视化的分析可以使用任意的可视化库以及工具

作业提交时间:

2022.10.23, 周一, 23:59

提交内容: 报告 (pdf) 和代码 (zip),请确保你的代码清晰可读、可复现、无 bug、无 特殊环境依赖,无法复现 的代码会极大影响你的得分。

评分依据:

实验报告占整个实验分数的 60%,其中实验报告中的实验分析占整个实验分数的 30%; 代码部分占整个实验分数的 40%

很好的完成基础内容部分的工作即可以拿到满分。

最终的实验得分计算方法为:

Min(基础内容得分 + bonus 得分, 100%)

本次实验占课程总成绩的 6 分

代码语言: 不限