

《2024Fall 机器学习与数据挖掘》——作业 1

问题:

根据提供的数据，训练一个采用在不同的核函数的支持向量机 SVM 的 2 分类器，并验证其在测试数据集上的性能。数据下载 FTP 地址: <ftp://172.18.167.164/Assignment1/material> (建议使用 FTP 客户端链接, 用户名与密码均为 student)

要求:

- 1) 考虑两种不同的核函数: i) 线性核函数; ii) 高斯核函数
- 2) 可以直接调用现成 SVM 软件包来实现
- 3) 手动实现采用 hinge loss 和 cross-entropy loss 的线性分类模型, 并比较它们的优劣

实验报告需包含 (但不限于):

- 1) SVM 模型的一般理论
- 2) 采用不同核函数的模型和性能比较及分析
- 3) 采用 hinge loss 的线性分类模型和 SVM 模型之间的关系
- 4) 采用 hinge loss 线性分类模型和 cross-entropy loss 线性分类模型比较
- 5) 训练过程 (包括初始化方法、超参数参数选择、用到的训练技巧等)
- 6) 实验结果、分析及讨论

将实验报告 (.doc 或 .pdf) 和代码 (不要包含数据) 打包成一个压缩包, 设置其文件名为: 学号+姓名.tar, 然后将其上传到课程 FTP: <ftp://172.18.167.164/Assignment1/report>

Due: 2024.11.10