**第十一章 线性分类器实践大作业**

# 实验目的

1. 深入理解SVM算法；
2. 掌握基于SVM/SVR进行分类/回归分析的基本步骤；

# 背景简介

Dosenbach等人与2010年在国际上率先报道了他们基于脑图像数据，结合SVM/SVR方法进行“脑年龄”分类及回归的结果，研究成果发表在著名期刊Science。在该研究中，Dosenbach等人基于238个样本(儿童及成人，7~35岁)的脑图像数据，每个样本提取12720个特征，结合SVM对儿童（7-11岁）及成人（24-30岁）进行分类，用SVR对个体脑年龄进行预测。实验结果表明，基于脑图像特征可以很好地预测个体脑年龄。

继Dosenbach等人的研究之后，脑年龄评估近年来受到国际研究人员的广泛关注，因为很多精神疾病的发生发展都与脑的发育延迟（脑年龄较实际年龄偏低）或者过早老化（脑年龄较实际年龄偏高）有关，因而脑年龄评估有望推进对脑疾病的精准诊断。

# 实验要求

本实验要求基于脑图像数据提取的特征，结合本章学习的SVM/SVR方法，对个体脑年龄进行分类/回归。具体要求如下：

1. 给定数据是高维小样本的，在分类/回归之前需要先对数据进行降维，降维方法自行选定；
2. SVM工具可自行下载，或者用matlab自带工具包；
3. 核函数及惩罚因子等参数自行选择；
4. Leave one out或者leave two out方法进行交叉验证；
5. 分类结果以预测值和真实值的相关系数、MSE两个参数给出正确率，画出ROC曲线；回归结果描述；
6. 尝试改变参数（特征数、核函数、惩罚因子、高斯核参数等），看参数对结果有什么影响；
7. 要求写出实验报告，包括伪代码（或流程图）、源代码（含注释）、实验结果及分析。

# 数据说明

* AALCorrArray\_BJ\_63.mat为用于回归的特征数据，包含63个样本，每个样本6670个特征；
* Age\_BJ\_63.mat为用于回归的标签数据，即63个样本的年龄；
* AALCorrArray\_BJ\_76\_Class.mat为用于分类的特征数据，包含76个样本，每个样本6670个特征；
* Age\_BJ\_76\_Class.mat为用于分类的76个样本的标签数据

# 参考文献

参见课程平台的两个pdf文档

Dosenbach\_Sci\_10\_prediction of maturity.pdf

Dosenbach\_Sci\_10\_Supplimentary material.pdf