

法律声明

□ 本课件包括演示文稿、示例、代码、题库、视频和声音等内容，小象学院和主讲老师拥有完全知识产权的权利；只限于善意学习者在本课程使用，不得在课程范围外向任何第三方散播。任何其他人或机构不得盗版、复制、仿造其中的创意及内容，我们保留一切通过法律手段追究违反者的权利。

□ 课程详情请咨询

■ 微信公众号：小象

■ 新浪微博：ChinaHadoop



决策树和随机森林实践



小象学院
ChinaHadoop.cn

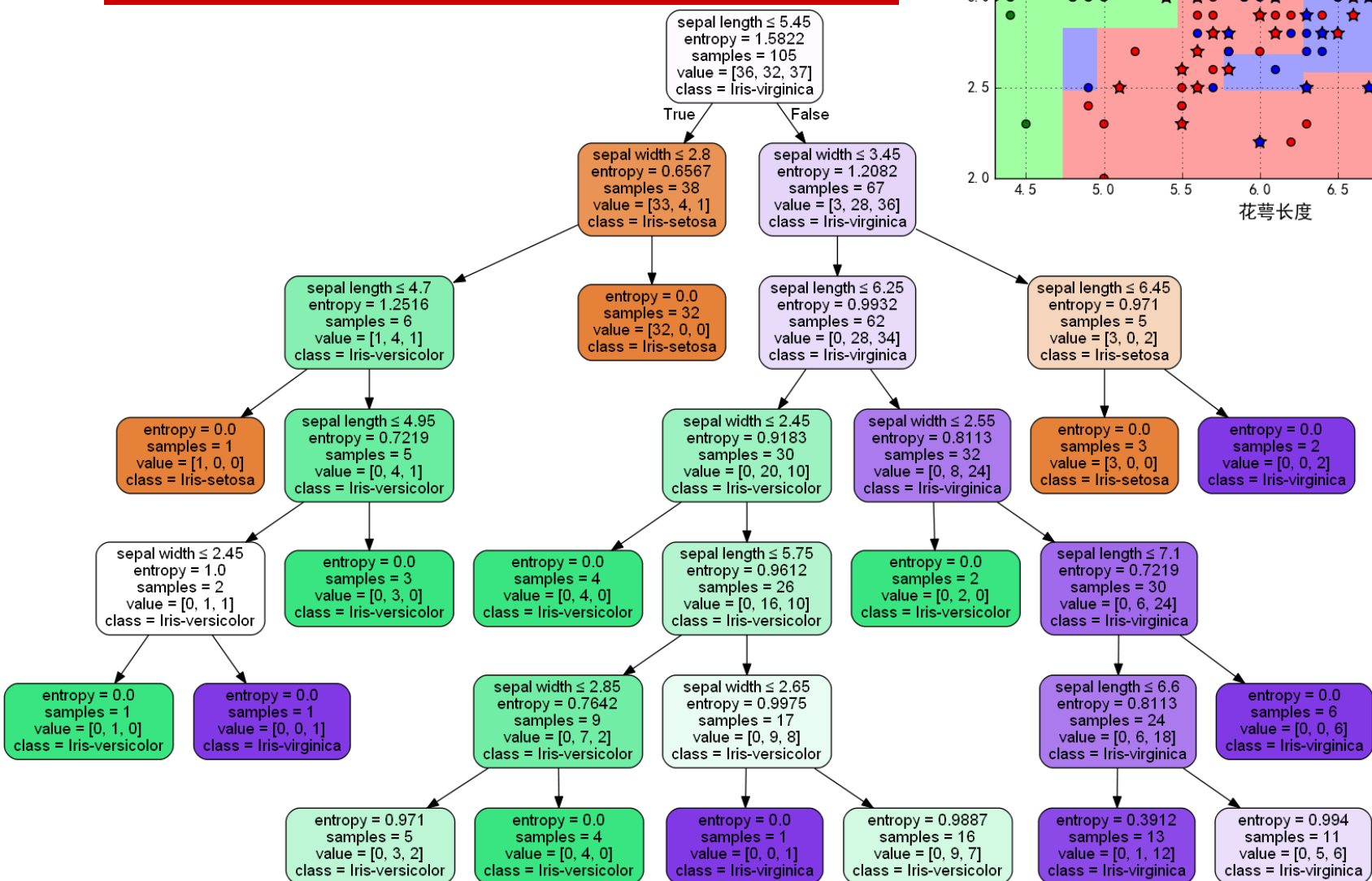
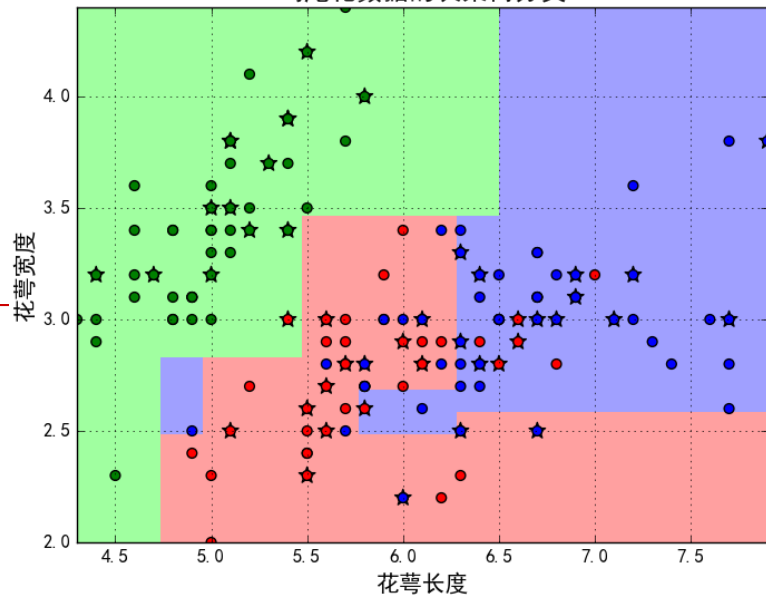
邹博

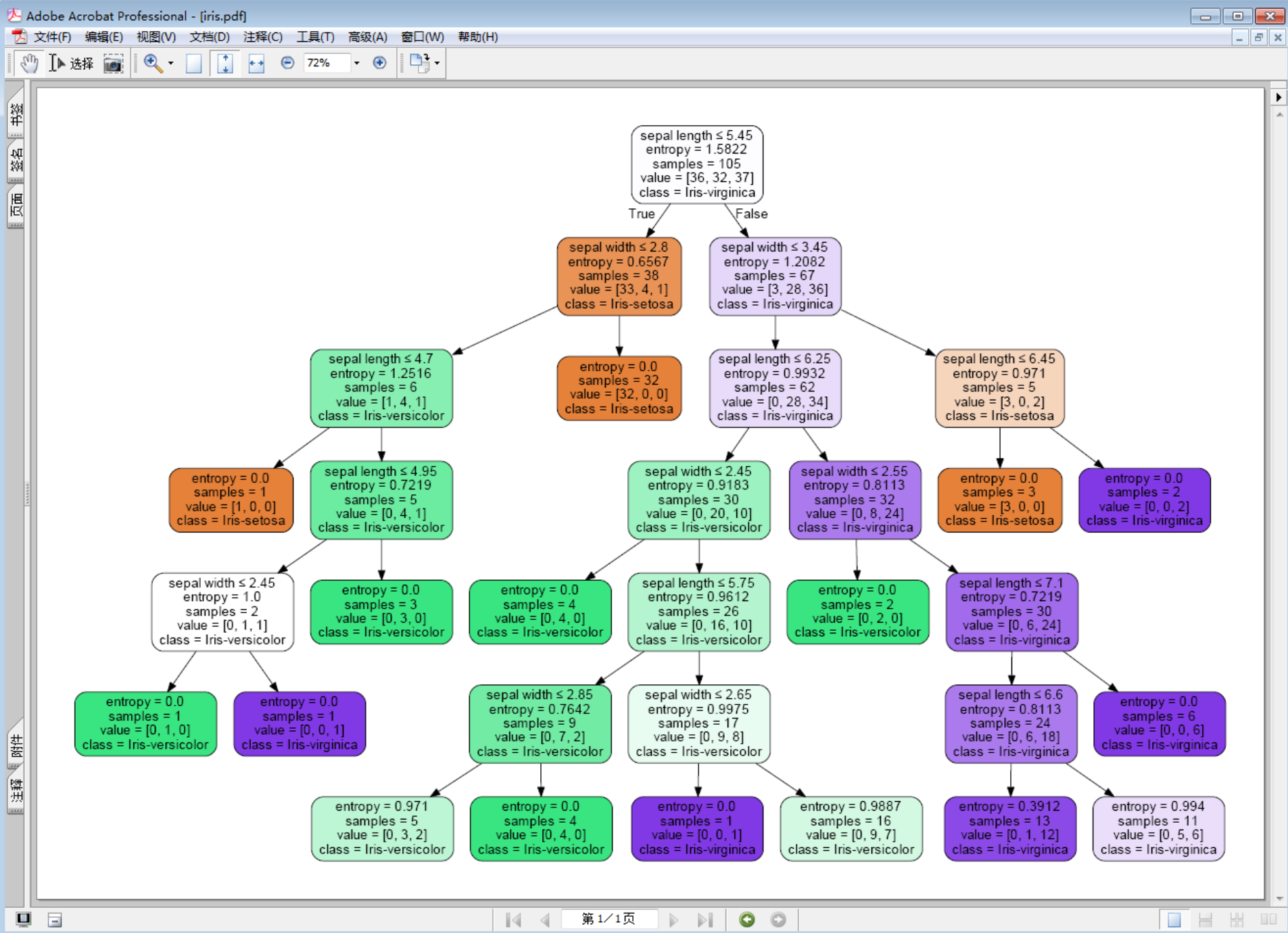
主要内容

- 决策树分类鸢尾花数据
- 随机森林分类鸢尾花数据
- 决策树的可视化
- 决策树回归
- 多输出决策树

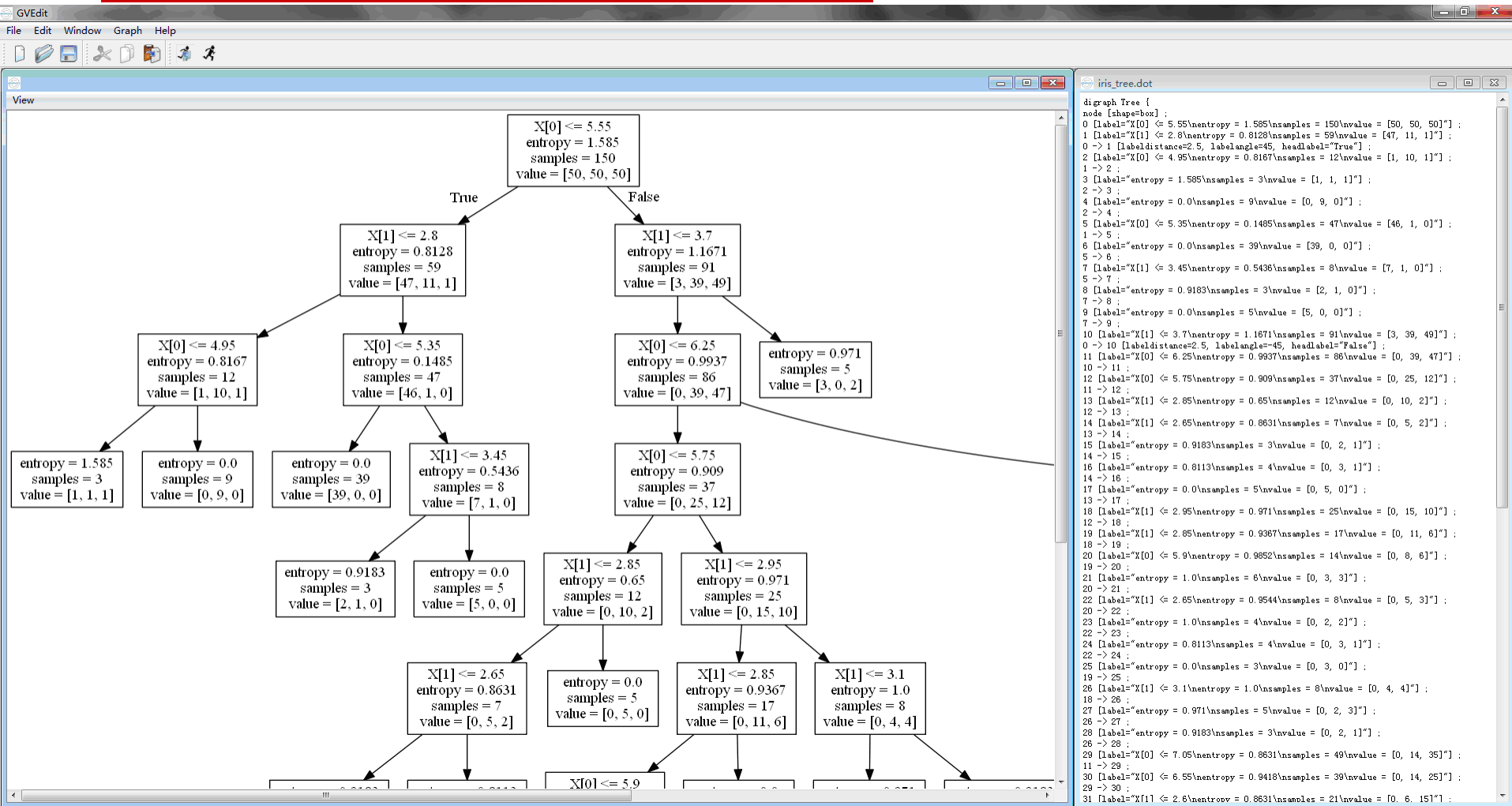
- 思考：树的深度与过拟合

鸢尾花数据决策树



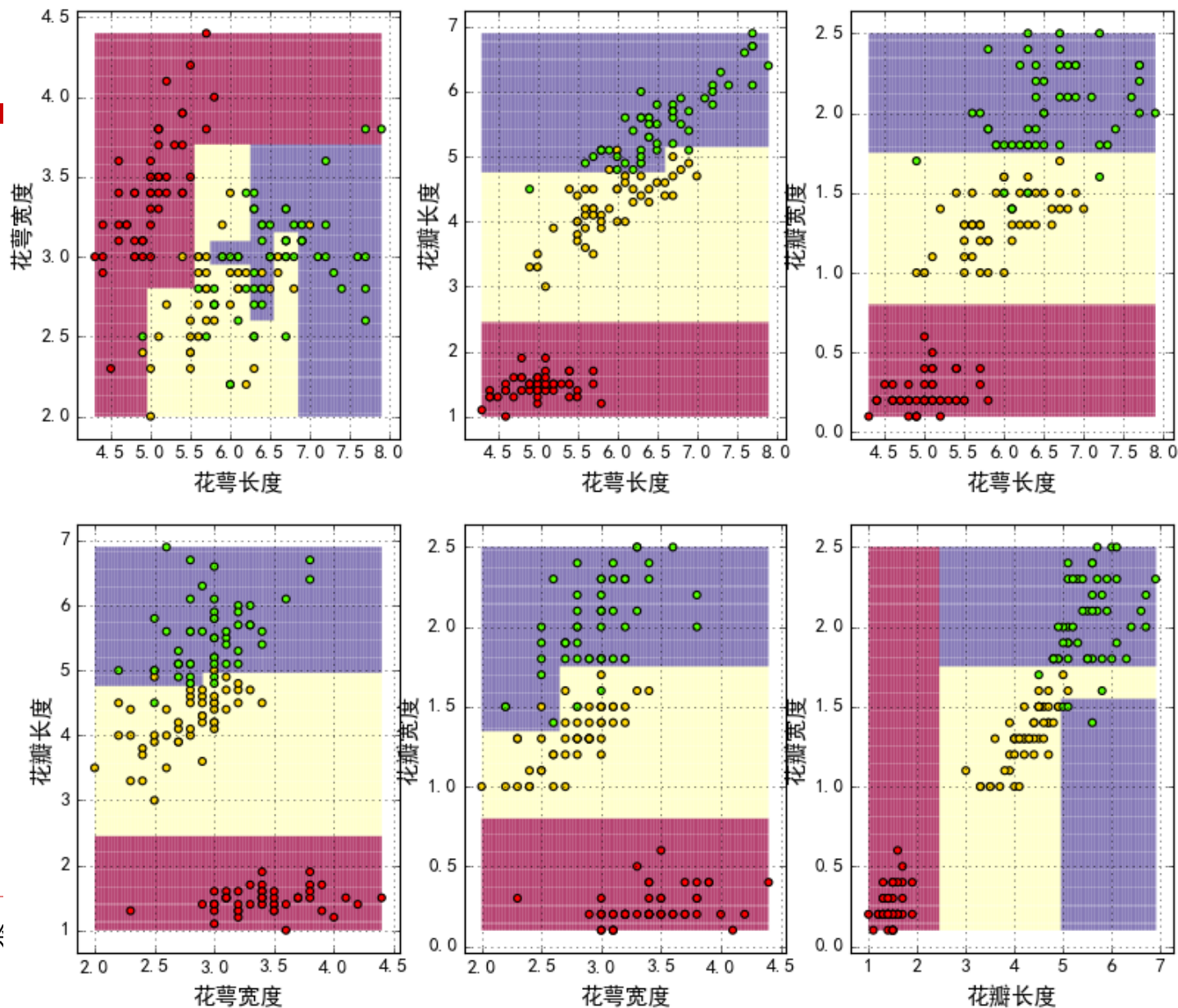


Graphviz



决策树

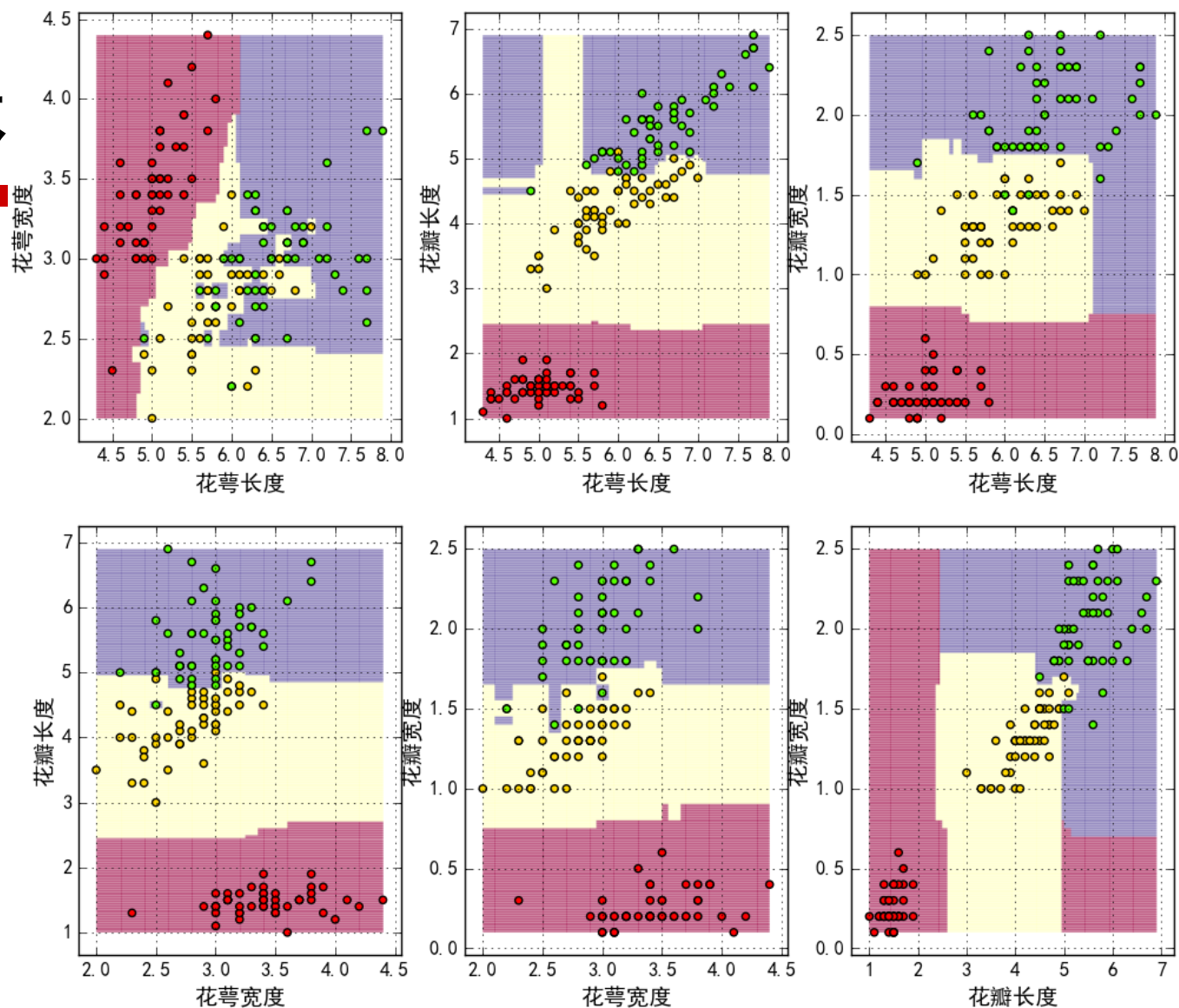
鸢尾花数据二特征决策树结果



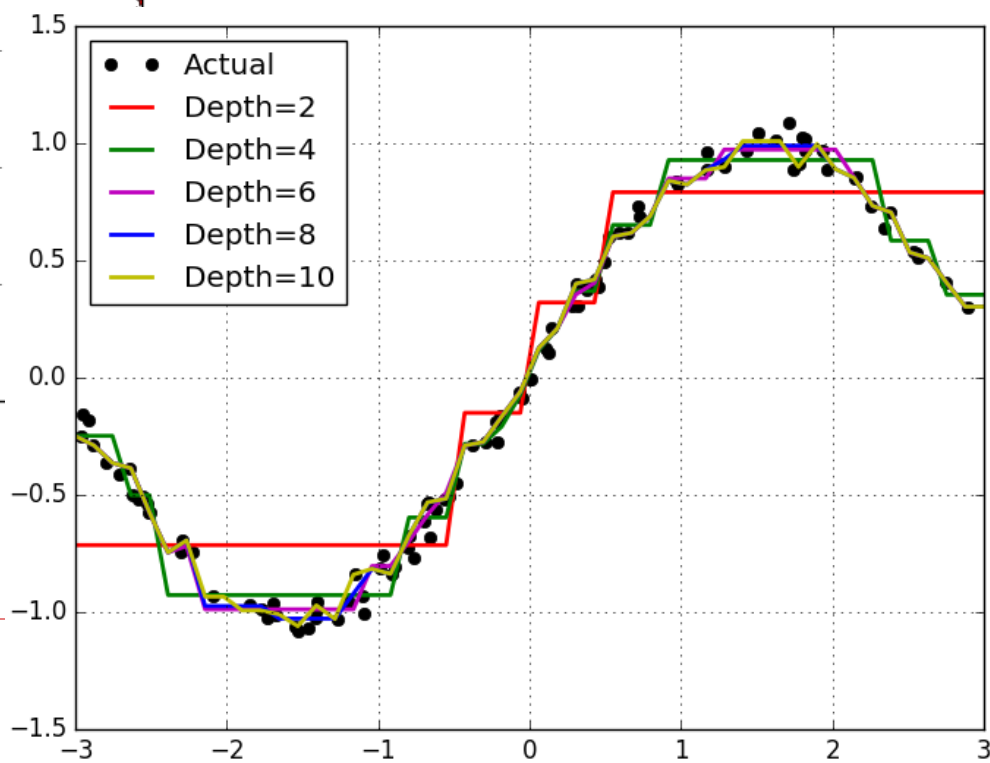
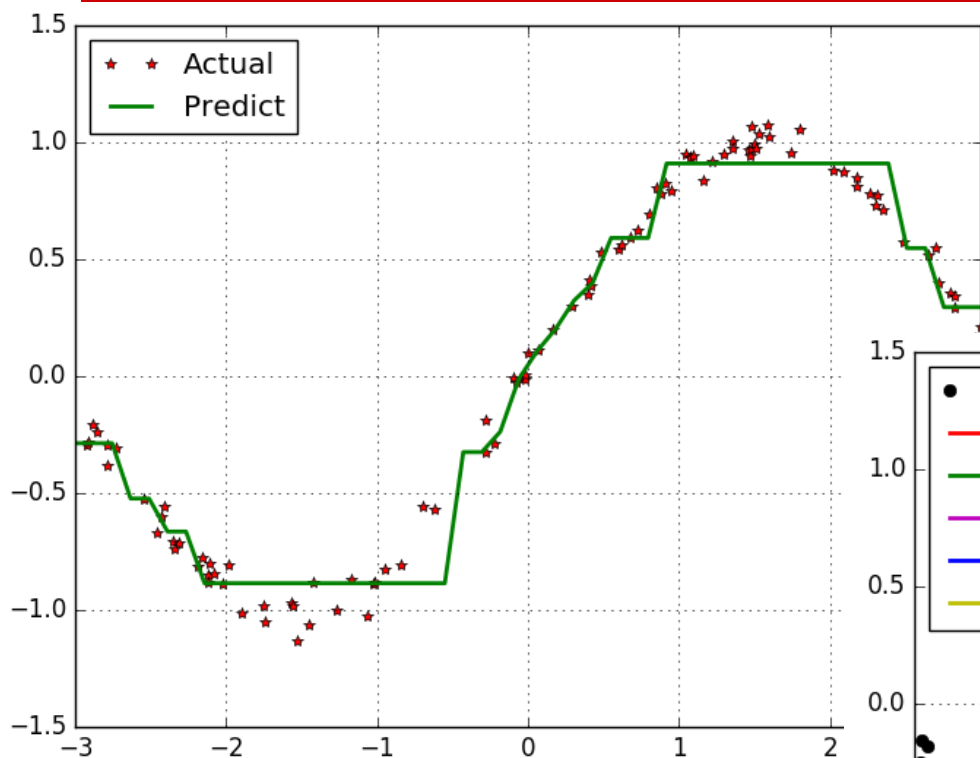
随机森林

□ 50

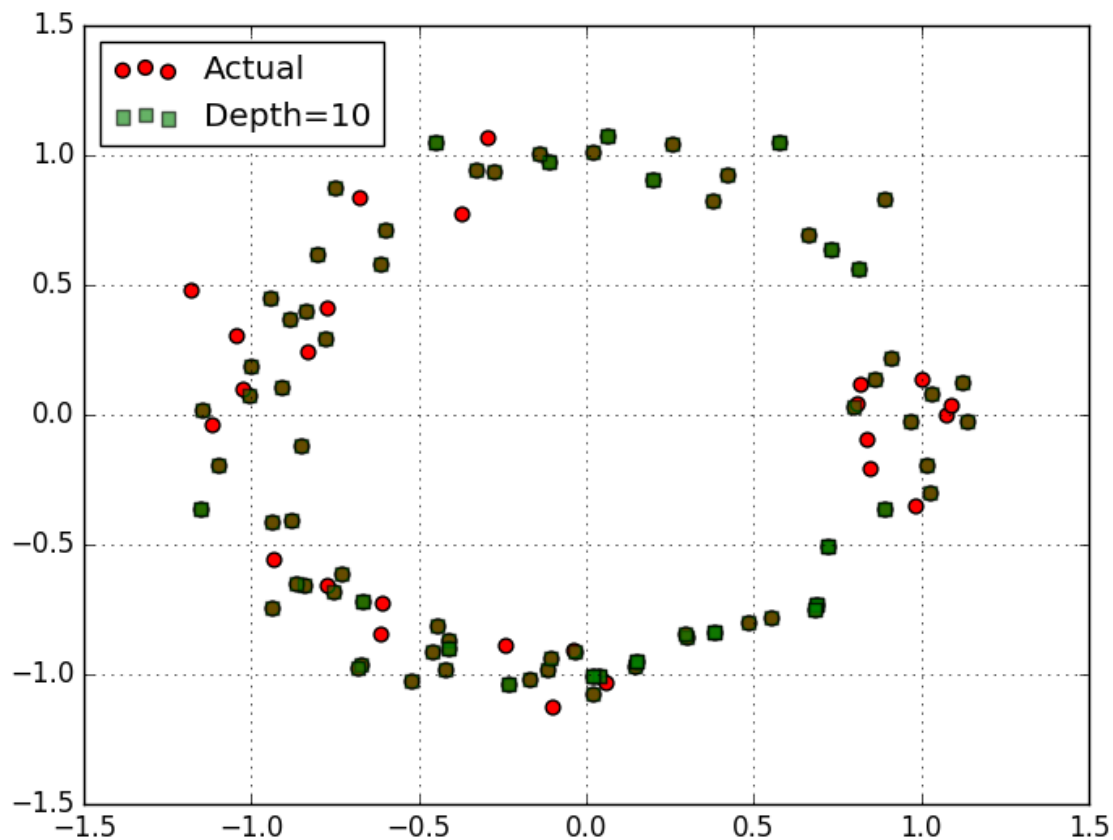
鸢尾花数据二特征随机森林结果



决策树用于拟合



多输出的决策树回归



作业

- 使用决策树做任意数据集的分类。
 - 离散变量
- 使用随机森林做数据回归。
 - 连续变量

我们在这里

□ <http://wenda.ChinaHadoop.cn>

■ 视频/课程/社区

□ 微博

■ @ChinaHadoop

■ @邹博_机器学习

□ 微信公众号

■ 小象

■ 大数据分析挖掘



感谢大家！

恳请大家批评指正！