

基于Python的机器学习

金融案例分享

主讲人：何芳芳

2017.07.13

获取更多资料



微信扫一扫
“上海数据分
析”公众号



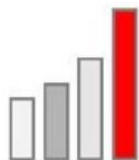
微信扫一扫
“助教老师”
获取更多资料

1 数据与人才

2 机器学习金融业应用场景

3 信用评级：Python的实现

- 新资产



量大



高效性



多样性



价值洼地

- 人才



“我认为，对数据分析的基本理解对下一代年轻人来说非常重要，这就是你们要进入的世界。”

- 知识：专业知识（统计、机器学习）、行业知识
- 工具：SPSS、SAS、R、PYTHON
- 流程：



1 数据与人才

2 机器学习金融业应用场景

3 信用评级：Python的实现

- Numpy –python 科学计算库
- Pandas – python 数据分析处理库
- Scikit_learn –python 机器学习库

• 常用场景



信用卡申请授信



风险评估



反欺诈分析



客户细分



客户流失预警



金融产品推荐

• 案例一：信用评级

• 数据：



Microsoft Excel
Comma Separated Values Fi

1 数据与人才

2 金融业应用场景

3 信用评级：Python的实现机器学习

- 分析处理数据

- 读取
- 数据类型
- 数据概览：样本数、特征数、特征值范围
- 缺失值处理
- 异常值处理
- 转换处理

- 模型建立

- ✓ 训练数据
- ✓ 测试数据
- ✓ 过拟合

- 模型评估

- ✓ 评估指标：准确率、召回率、F1 Score
- ✓ 评估图：ROC

- 应用：预判信用级别（分类）
- 决策树算法

