

大数据分析挖掘能力提升集训营

授课内容

| 日期 | 课程内容 | 课时数 | | | 备注 |
|------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------|---------|----------------------------|----|
| | | 课堂教学学时 | 是否多媒体教学 | 时间 | |
| Day1 2021. 3. 20 | 数据挖掘概论 Python 编程基础 1.1 掌握 Python 基础语法 1.2 变量常量与注释 1.3 字符串的常用方法 Python 数据结构与流程控制语句 2.1 列表, 字典, 元组, 集合增删改查 2.2 流程控制语句判断语句 2.3 流程控制语句循环语句 2.4 列表推导式与字典推导式 Python 编程基础函数编程 3.1 函数初识, 默认参数与动态参数 3.2 函数的返回值, 可变长参数 3.3 函数作用域与高阶函数 3.4 内置函数和生成器及推导式 3.5 闭包, 生成器, 迭代器, 装饰器 | 6 | √ | 09:00-12:00 14:30-17:30 | |
| Day2 2021. 3. 21 | 文件基础 4.1 文件与字符编码 4.2 读取 txt 文件中的数据 4.3 保存数据为 csv 格式文件 4.4 二进制数据读取 模块和内置模块 5.1 模块和包 5.2 模块和包的安装 5.3 内置模块 time 和 datetime 模块 5.4 json 模块 5.5 os 模块和 sys 模块 5.6 random 模块 | 6 | √ | 09:00-12:00 14:30-17:30 | |
| Day3 2021. 3. 27 | Python 数据分析之 NumPy 6.1 Numpy 概述 6.2 Numpy 数组的创建、变换、本操作 6.3 Numpy 数学计算 Python 数据分析之数据可视化 | 6 | √ | 09:00-12:00 14:30-17:30 | |

| | | | | | |
|---------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---|---|----------------------------|--|
| | <p>7.1 Matplotlib Matplotlib 概述、基础语法 Matplotlib 常见图形绘制</p> <p>7.2 PyEcharts 基本图形绘制、局部及全局设置 高级图表绘制（地图，水球图，词云图等）</p> | | | | |
| <p>Day4 2021. 3. 28</p> | <p>实战案例一：金融贷款数据可视化 数据载入与预处理 1.1 数据导入 1.2 数据清洗 Matplotlib 绘图 2.1 柱状图 2.2 叠加柱状图 2.3 饼图 2.4 折线图 2.5 多面板绘图 Seaborn 绘图 3.1 散点图 3.2 点图 3.3 箱线图 3.4 多面板绘图 3.5 小提琴图 3.6 直方图 3.7 计数图 3.8 热力图</p> <p>实战案例二：国家发展经济情况可视化 数据集概述 可视化总收支变动情况 调整收支情况折线图 绘制多条曲线 探索 GDP 与收支情况变动</p> | 6 | √ | 09:00-12:00 14:30-17:30 | |
| <p>Day5 2021. 4. 10</p> | <p>Python 数据分析之 Pandas 实战案例三：金融贷款数据的清洗 1.数据读取 2.数据清洗 2.1 缺失值的查看 2.2 删除缺失值 2.3 简易填补缺失值 2.4 向前向后与插值法进行缺失值的填补</p> | 6 | √ | 09:00-12:00 14:30-17:30 | |

| | | | | | |
|------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---|---|----------------------------|--|
| | 2.5 异常值处理 2.6 重复值处理 数据保存 3.1 Python 自带文件写入函数的存储 3.2 Pandas 中函数进行文件的存储 | | | | |
| Day6 2021. 4. 11 | 机器学习与 Python 实践 1 8.1 Scikit-learn 简介 8.2 数据预处理 1. 缺失值处理 2. 数据规范化 3. 主成分分析 8.3 线性回归 1. 一元线性回归 2. 多元线性回归 3. Python 线性回归应用 机器学习与 Python 实践 2 9.1 股票价格涨跌趋势预测 9.2 案例目标及实现思路 9.3 指标计算 9.4 预测模型的构建 9.5 预测结果分析 9.6 量化投资策略设计与分析 | 6 | √ | 09:00-12:00 14:30-17:30 | |
| Day7 2021. 4. 17 | 机器学习与 Python 实践 3 10.1 K 最近邻(kNN,k-NearestNeighbor)分类算法 10.2 手动实现一个 KNN 算法 10.3 Sklearn 中实现 KNN 10.4 模型选择与训练 10.5 网格搜索 Grid-search 10.6 KNN 案例 1: 鸢尾花的分类 10.7 KNN 案例 2: 手写数字的识别 | 6 | | 09:00-12:00 14:30-17:30 | |
| Day8 2021. 4. 18 | 机器学习与 Python 实践 4 11.1 关联规则 11.2 一对一关联规则挖掘及 Python 实现 11.3 多对一关联规则挖掘及 Python 实现 11.4 关联规则挖掘应用: 国际股票指数关联分析 1. 问题描述 2. 数据预处理 3. 关联规则挖掘 4. 问题扩展 5. 结果分析 | 6 | √ | 09:00-12:00 14:30-17:30 | |
| Day9 | 一、金融大数据教学实训基础案例 1.1 众包任务特征指标的计算 | 6 | √ | 09:00-12:00 14:30-17:30 | |

| | | | | | |
|----------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---|---|----------------------------|--|
| 2021. 4. 24 | 1.2 股票价格指数周收益率和月收益率的计算 1.3 上市公司净利润增长率的计算 1.4 股票价、量走势图的绘制 | | | | |
| Day10 2021. 4. 25 | 一、金融大数据教学实训综合案例 1.1 总体设计 1 案例背景 2 案例目标及实现思路 3 基于总体规模与投资效率指标的综合评价 4 基于成长与价值指标的综合评价 1.2 基本设计步骤 1) 指标选择 2) 数据处理 3) 主成分分析 4) 综合排名 5) 收益率计算 6) 量化投资策略设计与分析 往届全国大学生“泰迪杯”数据挖掘挑战赛指导讲解 数据挖掘流程介绍 论文写作与注意事项 往届赛题解读 | 6 | √ | 09:00-12:00 14:30-17:30 | |
| 学员通过本次学习和考试可获取工信部“大数据分析师（高级）”证书，工信部考试中心官网可查询 | | | | | |

