

QG工作室2025寒假训练营培训内容

Tips

- 遇到问题一定要先自己想办法解决
 - 灵活运用搜索引擎
 - chatgpt: 学会如何提问, 如何使用, 但是不能依赖
- 每次完成项目实战和学习新知识后, 需要上传代码和笔记文件到github仓库中
- 可以借鉴别人的代码, 但是一定要理解, 不能直接照抄
- 学习资料很多且不唯一, 可以使用参考资料也可以自行检索, 选择适合自己的进行学习

整体安排

时间	内容
1.15-1.16	相关工具的安装与使用
1.17-2.5	python学习
2.6-2.19	机器学习入门

第一阶段：相关工具的安装与使用

- 阶段目标: 安装会使用到的相关工具, 并掌握基本使用方法。

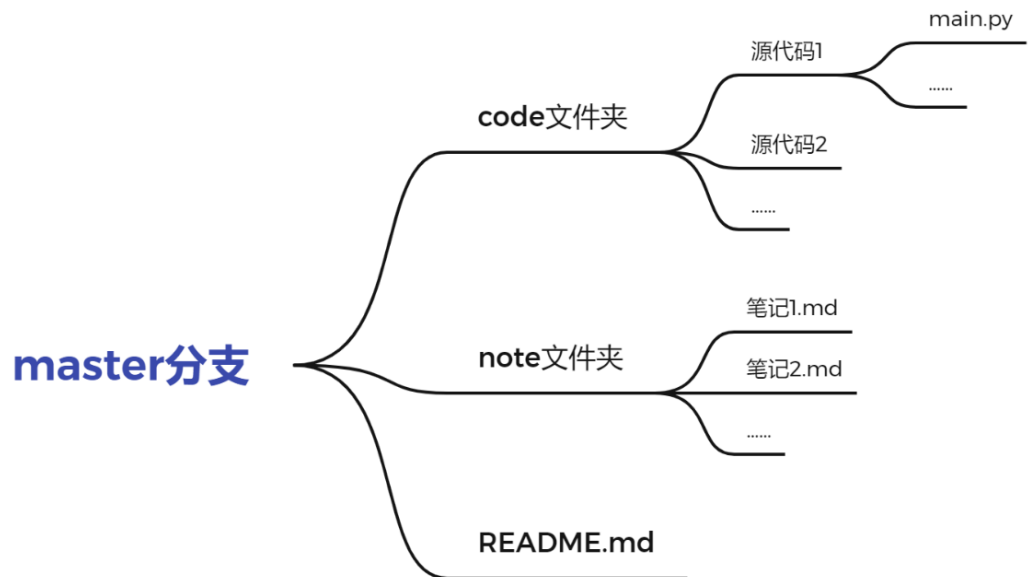
Day1 (1.15)

- anaconda, pycharm, jupyter的安装与使用
- anaconda: 官网下载
- pycharm: 账号可以使用学生邮箱免费申请免费, 或者直接上淘宝购买
 - jupyter: 可以在anaconda下载, 轻量级的网页编程, 便于调试分析
- 简单的markdown语法
 - 编辑器安装 (推荐使用 typora)
 - markdown语法学习: 可以尝试照着案例敲一遍, 后继需要再继续查
 - 图床配置 (PicGo+阿里云OSS): 便于将图片添加到markdown文件中

配置环境会很折磨人, 但配置环境是第一步, 不要被劝退! 工欲善其事, 必先利其器。

Day2 (1.16)

- 拥有自己的 github 仓库
 - **科学上网** (有需要的私聊导师)
 - GitHub desktop 或 Git 的使用 (都是上传文件到GitHub的工具, 选择其中适合自己的一个熟练使用即可)
 - GitHub仓库结构参考:



各部分要求说明如下:

- code文件夹: 用于存放学习实操的python源代码, 不同项目代码源文件新建文件夹分开进行存放, 文件夹按照项目内容命名。
- note文件夹: 用于存放学习笔记, 要求使用markdown进行书写, 命名按照笔记内容决定, 对笔记内容不做进一步要求。
- README.md文件: 书写对仓库的内容、结构等的说明。
- 仓库创建好后, 及时将地址私发给导师, 方便导师进行检查
- 参考资料
 - [关于 GitHub Desktop - GitHub 文档](#)
 - [Git 教程 | 菜鸟教程 \(runoob.com\)](#)
 - [廖雪峰git学习笔记](#)
 - [Git开发流程](#)

第二阶段: python学习

- 阶段目标:
 - 掌握python的基础语法, 数据类型, 数据结构
 - 掌握条件控制以及循环控制, 推导式, 错误和异常捕获, map, lambda, filter函数
 - 理解运用面向对象 (类与实例化, 类变量以及方法, 私有变量和私有方法, 封装, 继承, 多态等)

- 参考资料:
 - [【尚学堂】2023最新版Python400集零基础入门学习教程Python教程含爬虫教程python400集初学者教程PYTHON编程开发python全栈](#)
 - [Python 教程 — Python 3.12.1 文档](#)
 - [practical-tutorials/project-based-learning: Curated list of project-based tutorials](#)

Day3 (1.17)

- 基础语法与缩进
- 数字类型
 - int float bool complex
- 数据结构
 - 字符串 列表 字典 元组 集合 切片

Day4 (1.18)

- 条件控制以及循环控制
- 推导式
- 错误和异常捕获
- map, lambda, filter函数的使用

Day5 (1.19)

- 面向对象
 - 类与实例化 类变量以及方法 私有变量和私有方法 封装 继承 多态

Day6 (1.20)

- 复习前面学习的内容，查漏补缺，上传一份总结

Day7 (1.21)

数学基础

- 线性代数
 - 向量 矩阵乘法 行列式 矩阵转置以及对应的性质 矩阵的逆以及对应的性质

tip: 由于时间关系，不建议对线性代数进行系统的完全学习，可以只简单过一下上面这些概念，其他理论问题遇到再查，也可以买本下学期的书简单看一下。

- 参考资料

- [【线性代数的本质】合集-转载于3Blue1Brown官方双语】](#)
- [麻省理工学院 - MIT - 线性代数](#)
- 两个视频建议结合起来看

Day8-9 (1.22-1.23)

- numpy库
 - 基于列表构建矩阵 特殊矩阵构建 矩阵乘法 矩阵广播机制 矩阵转置 矩阵的逆 矩阵存取
 - pandas库
 - 学习读取数据 掌握简单的数据清洗
 - matplotlib库学习
 - 基础绘图语法（比如散点图、曲线图、条形图、扇形图等）
- 不同绘图之前差异较大，不建议广度地学习，不需要学太深,需要时再学习查阅即可，熟能生巧

Day10 (1.24)

- 复习前面学习的内容，查漏补缺，上传一份总结

Day11 (1.25)

- 泰坦尼克号数据集可视化
 - 分别设计对各指标的生还数（率）进行对比可视化
 - 分析可视化结果
- 参考资料
 - [Titanic数据集](#)
 - [泰坦尼克号数据可视化分析](#)

新年快乐！春节期间不安排学习任务

第三阶段：机器学习入门

Day12 (2.5)

数学基础

- 多元微分学的学习
 - **偏导的概念** 微分 **方向导数和梯度** **多元函数极值与偏导的关系** 条件极值
- 复习熟练矩阵内容
- 参考资料

- B站高数课程简单学习对应知识点即可，高数下课本会进行系统的学习

Day13-15 (2.6-2.8)

- 机器学习概念入门
 - 有监督学习&无监督学习，训练集&验证集&测试集，过拟合&欠拟合，损失函数，模型评价指标，常见数据清洗，特征工程
- 参考资料
 - [2022吴恩达机器学习Deeplearning.ai课程](#)
 - 台大李宏毅教授系列

Day16 (2.9)

- 多元线性回归解析解
 - 最小二乘法构建损失函数
 - 求矩阵与向量偏导并手推公式
- 多元线性回归梯度下降法求解
 - 梯度下降法
- 参考资料
 - [可能是全网最详细的线性回归原理讲解!!!](#)
 - [\[5分钟深度学习\] #01 梯度下降算法](#)
 - 两个视频最好都看下

Day17-18 (2.10-2.11)

- 完成多元线性回归的类封装
 - 代码实现解析求解和梯度下降法求解
 - 参考 sklearn 接口形式
 - 仅可使用 numpy，不允许调用机器学习库
 - 以 [波士顿房价数据集](#) 为例，划分训练集，测试集，并完成数据分析和预处理
 - 完成模型训练并予以评价分析结果

元宵节快乐! 休息一天

Day19-20 (2.13-2.14)

- 学习K-means算法
 - 熟练掌握K-means算法的流程
- 完成K-means的类封装
 - 参考 sklearn 接口形式

- 仅可使用 numpy，不允许调用机器学习库
- 以 鸢尾花数据集 为例，划分训练集，测试集，完成模型训练，并予以评价分析结果

Day21-25 (2.15-2.19)

- 提前学习数据结构，为开学后的考核做准备
 - 链表
 - 栈
 - 队列
 - 树
 - 图
- 尝试用c或c++实现，不建议用python
- 参考资料
 - [数据结构与算法基础 \(青岛大学-王卓\)](#)

训练营结束！ 寒假快乐！