python与数据挖掘学习规划

**研究方向：**推荐系统（图神经网络、知识图谱+推荐）

**学习目标：**

（1）熟悉python编程语言，基本掌握pycharm编译工具的使用方法。

（2）在python语言中掌握像其它语言一样的面向对象思想以及方便、常用的设计模式。

（3）熟练使用python语言完成针对小问题的程序编写。

（4）提高对应研究方向的代码阅读能力。

（5）学习课上老师要求掌握的科研期间常用的技术。

**学习内容：**

（1）pycharm基础教学。

（2）python中的数据类型大致可分为6类：

①Number(数字) ②String(字符串) ③Tuple (元组) ④List(列表)

⑤Dictionary (字典) ⑥Set(集合)

（bool布尔类型、int整型、float浮点型、complex复数等都归为Number数字类型）  
这6种数据类型中不可变类型为：number、string、tuple

可变类型为：list、dict、set

（3）\*args和\*\*kwargs的使用，传参时这两种可变值参数比较万能。

（4）采用类工厂模式进行随机数生成的类封装，各个实现的类分别写在不同的py文件里，通过模块调用（import...）的方式使用其它模块内的方法，实现python的类工厂设计模式。

（5）学习了函数修饰器与类修饰器，了解到类修饰器相较于函数修饰器可以作为接口使用，功能优于函数修饰器。此外还实现了修饰器ACC（正确率统计）和MCC（马修相关系数计算），使其能够分别修饰不同参数类型函数的数据，使得最终获取到经修饰器修饰后的结果。

（6）学习了XML可扩展标记语言，使用lxml第三方库，实现大型XML文件解析修饰类，同时编写此类的调用演示，例如天气数据，其中使用修饰类修饰做一个XML元素处理的函数，这个函数能够打印每一个元素的子树结构。