编程练习.md

# 研一新生的学习任务

## 1. 学习python基础

第一步：

* Python基础（手敲代码！不允许直接粘贴复制，那样什么都学不到！）  
  地址：https://www.runoob.com/python3/python3-tutorial.html))
* python进阶（有时间翻翻就可以）  
  地址：https://python3-cookbook.readthedocs.io/zh\_CN/latest/index.html

第二步：

数据采集  
 京东评论、或者是豆瓣影评  
 *要求：数据量大于1万（清洗后）*

*学习时长建议：****2-3****周，不建议花费很长时间。* 数据采集掌握方法即可，不必要花费太长时间钻难点，在以后工程需要的时候查询即可。

## 2. 机器学习

机器学习资料

* 书籍：
  + 统计学习方法
  + Python机器学习基础教程ilovepython
  + **极力**推荐用这个 “Python机器学习基础教程ilovepython”，一方面书写的好，另一方面有完整可运行的配套代码。
* 代码库：
  + introduction*to*ml*with*python-master（在Python机器学习基础教程下属文件夹）
* 视频
  + 吴恩达 bilibili搜索即可
  + 李宏毅 bilibili搜索即可  
    编程任务：  
    sklearn 全家桶实现所爬取的数据，对爬取的数据做3分类。 可以将1、2分归为一类，3分归为一类，4，5分归为一类。要求输出F1值。

机器学习建议：本部分建议花费几天开天，看一下基础的概念，之后就可进入深度学习阶段，如果遇到概念问题可以返回来再看相应的概念。

学习时长概念：3-4周，了解基础概念即可。开题之前不太需要纠结公式，只需要会用就好。

## 3. 深度学习

框架**优先**次序：

* [pytorch](https://pytorch.org/tutorials/beginner/pytorch_with_examples.html#warm-up-numpy)
* [tensorflow](http://c.biancheng.net/view/1881.html)
* 飞桨

编程任务：（二选一）

* 使用Word Average Model模型进行三分类
  + RNN分类

任务要求：  
有问题可以先同级内部讨论解决，之后有需要再和师兄师姐讨论。

学习建议：有问题多看官方文档手册，可以解决你可以遇到的90%的问题。同样不必看公式推导、反向传播，只需要会出结果就行，足够做工程和毕业。