**企业实习中期报告**

**信息与软件工程学院**

学 号：2013220402019

姓 名： 陈秋远

专业方向： 大型主机

企业名称：华西期货

实习岗位名称：金融软件工程师

企业指导教师：杜野

院内代管教师：刘峤



目 录

[1. 企业实习目标任务 1](#_Toc449007504)

[1.1. 实习单位情况和岗位职责 1](#_Toc449007505)

[1.2. 实习目标、任务与复杂工程问题 1](#_Toc449007506)

[2. 实施方案和可行性研究 1](#_Toc449007507)

[3. 知识技能学习情况 1](#_Toc449007508)

[4. 前期任务完成度与后续实施计划 1](#_Toc449007509)

[参考文献 2](#_Toc449007510)

**说明:**

1. **报告要求6000字以上。**
2. **本模板仅为内容基本参考，请各位同学根据个人情况进行目录结构扩展。**
3. **封面和目录单面打印，从正文第1页开始双面打印。**

## 企业实习目标任务

### 实习单位情况和岗位职责

华西期货有限责任公司是华西证券公司的全资控股子公司，其前身为成都大业期货经纪有限公司，是中国最早成立的期货经纪公司之一。华西期货公司在秉承华西证券稳健运营理念的基础上，已进一步完善了风险控制和内部管理制度。重组后的华西期货公司注册资金为1.2亿人民币，成为西南地区为数不多的具有券商背景的期货公司。

我的岗位是IT部门的金融软件开发工程师。我主要的责任是进行数据整理、数据可视化，以及进行一些量化交易的研究和测试。

### 实习目标、任务与复杂工程问题

开发一套软件，帮助决策者进行从选股到行情判断的决策以及交易。同时学习到金融相关的知识。作为一个全栈工程师独立解决一套软件从需求到开发以及维护的过程。其中包括了对于一门新语言的掌握，对python社区的熟悉，对于python库的熟悉，对于python源代码的阅读理解。以及软件的架构选择，从网络爬取数据，存储数据以及数据库的设计和维护，还有因为是自己设计的数据库，需要解决远程连接的问题。

## 实施方案和可行性研究

使用敏捷开发，由于只有我一个人，暂时使用迭代的办法，一个版本一个版本来加功能，看最终能够做到什么程度。刚开始觉得特别复杂的东西在慢慢一步一步地做的过程中被慢慢分解并解决掉。在边学习边写代码的过程中，从无到有，最终还是出来了一个暂时能够使用的框架。然后在这个框架上加一些东西就能够满足各种各样的需求。在开发这套软件的时候，我下意识地使用了软件工程的知识，对于文档的管理，版本的迭代都有了很详细的记录，也每天在写工作日志，保证进度的推进，代码在去除保密部分后上传到了github上面，方便版本的管理和迭代。由于每过一段时间会对于旧的代码有所遗忘，所以我会定期阅读老的代码，每次温习都会有新的收获，具体来说就是对现在搭建起来的框架有了更深层次的理解和修正改良。以便于进行更好地拓展。

## 知识技能学习情况

刚开始入职的时候在老师的建议下选择了python作为工具语言进行一系列的操作，具体使用的是python2.7.因为在python平台上的科学计算和网络支持。Python也有很多的库能够调用。另外还管理了一个大小约为4T的数据库，使用的sqlserver。在实习中大量使用了库，但是因为很多库的文档写的不是很全，所以解决问题的时候还是要读源码，在读源码的时候碰到了很多的困难，但是最终还是解决了。因为从事的是金融相关的工作，还需要很多金融和数学的知识，为此我读了很多金融相关的书籍，也在准备考取期货从业资格证。数学知识方面我在下班的时候重新复习了微积分和概率论，但是发现学校学习的数学知识并不够，所以还需要具备一些统计学知识，这些学习任务都在我下班之后进行。我希望以后从事量化交易的工作，所以在做好IT之余还是需要尽量掌握更多的金融知识。

## 前期任务完成度与后续实施计划

在工作中首先打交道的是excel，我花了一些时间搞清楚了excel的一些基本知识，也熟悉了VBA的一些操作，然后试用了python对于excel的操作库，最后使用了一个名为openpyxl作为工作库，将这个库的源码看了很多，让我能够用代码自动生成符合需要的excel。然后由于要为其他人使用，需要做一个界面，我开始使用的是python自带的tkinter库，但是美观程度不够，于是在迭代到第二个版本的时候，使用了python的qt库。感觉还行。以上是代码的细节处理部分。在收集了数据之后，遇到了新的需求，那就是一般的数据作为excel表存放已经非常冗余了，于是有了存放数据库的必要。这时候我考虑了很多的数据库，传统的关系型数据库是首要考虑的，比如mysql或者sqlserver。但是问题在于：首先，我存放的数据不是重型数据，不会像我维护的TB级别数据库，所以关系型数据库的高性能体现不出来。其次，我要存放的数据不是规定死了的格式，会有各种各样的样子。于是我开始考虑nosql数据库了，redis是内存数据库，键值对让它很方便，但是问题在于不是很好迁移，所以最终我选择了mongodb，我考虑到这个数据库很火，换句话说支持比较多，遇到问题能够很快找到解决办法，而且python也对它有很好的支持。于是最终选择了mongodb。除了使用各种工具外，原生代码的训练一般涉及字符串处理、日期处理。以及将数学公式转化成代码的过程也比较锻炼人。在某些情况下python有现成的数学公式库，但是我发现这并不是给软件用的，他的输入格式太简单，不适合软件中遇到的复杂的数据结构，所以解决办法还是将代码看懂自己重写一遍以便于使用。

在之后的工作中我主要要完成的是将一个交易思路进行建模，并使用数据可视化的方法体现在图中，以便于进行决策。这个涉及一些数学建模的问题，我查阅了很多的资料，但是这个不能一蹴而就，我需要反复修正才能达到我的目的。另外的计划就是将现有的软件进行重构，做到更好的健壮性。

## 参考文献

1. Openpyxl. http://openpyxl.readthedocs.io/en/default/
2. matplotlib. <http://matplotlib.org/>
3. 《机器学习实战》,Peter Harrington
4. 《统计学方法》， 李航
5. Pyqt4,https://sourceforge.net/projects/pyqt/