



华南理工大学
South China University of Technology

汇丰金融科技创新黑客松大赛

金融投资评估系统

商业计划书

项目名称：	金融投资评估系统
项目负责人：	李竑緯

二〇一九年四月

一、技术进步与使用

1.1 现有技术与资源

互联网金融数据中潜藏着未知的知识价值，但也存在着巨量的数据冗余。通过 Python 可以进行互联网金融数据的获取、解析、提取关键信息并进行存储，其方式方法是灵活多样的，可采用的开源库也比较多。使用 Python 时可以依据实际的研究内容来决定需要采集的数据。Python 是互联网数据的获取工具，也是进行互联网金融数据分析的基石。同时，上市公司按时提交的财务报表对该企业运营分析提供了很好的参考。

1.2 本项目技术简述

本项目技术以企业的年度财报相关信息为基础，通过整合相关信息和利用相关金融计算公式得到资产负债比率，五大财务比率，用数据展现企业运营情况，最后分析股票走势，综合考虑来判断该企业是否值得投资。在以往的企业财报中，由于报表中的涉及的项目众多，信息繁杂，往往需要我们花费大量的时间和精力提取相关的信息进行一一分析，在现有的金融分析展示中，结果往往仍然只是诸如“企业负债率”，“企业净利率”等单独的项目，虽然比起直接面对报表已经直观很多，但还是不够清楚的展现各项之间的相关联系，因此我们在传统分析的基础上，对各项数据进行划分，分为“企业获利能力”，“企业偿债能力”，“企业经济效率”，“企业财务结构”四项分析，使各大银行或是投行可以更直观，更清楚地分析企业数据。

在对财报的数据分析中，如果视报表的基础数据为第一级数据，“负债率”，“利润率”等通过简单公式计算的数据称为第二级数据的话，我们的分析意在构建基于“获利”，“偿债”，“经济效率”，“财务结构”的第三级数据。

第三级数据提出的意义在于整合第二级数据，将相关的第二级数据进行整合，从而得到对各企业更为直观，立体，全面，易分析的更高阶数据。

在对第三级数据的搭建上，我们结合了企业的三个核心活动：“经营活动”，“投资活动”，“筹资活动”，以及财报的“资产负债表”，“利润表”和“现金流量表”，提取重要的第二级数据最终划分为了四项。

在“获利能力分析”中，我们选取了营业毛利率，营业净利率，以及所有者权益报酬率，资产报酬率四项作为基础，来反映出企业的成本受否得到控制，业务是否赚钱，以及企业的融资能力和资产配置效率，较为立体的阐述了企业的整体获利能力。

在“偿债能力分析”中，我们选择了流动比率和利息保障倍数两项，以求能够反映出企业的偿还短期债务能力和支付利息的能力，全面的展现企业在面对债务问题上的能力。

在“经济效率分析”中，我们选取了存货周转率，应收账款周转率和总资产周转率三项来反映出企业的“赚钱”效率，市场营销的效率以及整体运营的效率，三方面的结合可以使分析人更容易更快速的把控整个企业的效率问题。

在“财务结构分析”中，我们选取了所有者权益比率和负债比率两项反映股东投资或是自身获利在经营中的比重和作用，以及整个企业的负债情况，大体的展现企业的财务结构。

二、创新方法与实用性

大数据时代，数据分析是十分必要的，人们对某个假设的判定往往是根据数据分析得出来的，数据分析可以帮助人们对得到的结论制定相应的方案。数据采集，分析，模拟的有机结合，使研究人员更容易掌握公司的脉搏和趋向，让投资行业多了一个投资参考，也能让企业本身意识到自己潜在的问题并进一步生成解决方案。在互联网普及，互联网金融广泛的今天，能够搭建平台，让数据搜集，处理和预期一体化，无疑会为更多的企业，个人带来帮助。

三、项目拓展空间

在当前的分析中，我们结合形成了四项第三级数据，并使用四项数据作为标准来评判企业的整体情况，但当前存在的问题是建立四项数据评定企业的标准，因此在接下来的时间内，我们的目的在于即时的提取数据，建立标准，更加进阶的让投资者易于比较，判断和选择不同企业，因此我们将从下面几个方面展开：

2.1 人工智能

机器学习和深度学习在时间序列数据的预测上具有很高的准确率，在金融机构中获得了广泛的应用。有大量的研究来进一步提升金融数据相关模型的准确率，未来可以使用 AlphaAI 项目中的栈式神经网络的架构（stacked neural network architecture）结合财务报表来预测多只股票未来的价格。

在栈式神经网络的架构的基础上，可以利用爬虫技术爬取网络上关于每个公司的文章，使用自然语言处理（NLP）对文章分类成正向性与负向性，再使用此数据加入到神经网络的权重中，使得预测结果更加精准。

2.2 量化

我们拟通过建立分数的方式给予投资者最直观的判断依据，因此接下来要做的便是对四项数据进行“定量”。通过建立公式来建立起获利能力分数，偿债能力分数，经济效率分数和财务结构分数。以及总分数五项。对各项分数的建立中，我们需要对选取的第二级数据赋予合适的权重，在最简单的模式下即为平均分配，而我们需要做的便是寻找合适的比例因子，建立起合适的分配方案，以至于使我们的分析更加真实。

在通过创建公式建立分数之后，我们便可以获得五项分数，因此投资者既可以单方面的了解每一项的分数状况，又可以得到企业总的分数情况。通过对企业分析的数字化和分数化将大大简化对企业的分析，减少对企业分析时产生的成本。

2.3 行业化

众所周知，不同的行业在金融数据的体现上很可能会产生很大的不同，因此单方面以一个标准来反映所有企业的状况是并不现实的。因此在我们的分析中，要更加体现行业的差异性，同行业之间的比较才更加具有说服力。因此在后续的改进中，针对不同的行业建立不同的评分标准以及建立图像反映同行业企业的直观比较也是非常重要的。

四、社会与行业影响

伴随着互联网的发展与移动应用的普及，互联网金融得到了长足的发展，随之而来的是规模巨大的数据，金融从业者要通过数据挖掘、大数据分析获取有价值的信息，完成精准

营销，进而降低营销成本，提高业绩。这时，一项重要的前置工作就是数据的采集，以便于后续的数据分析，尤其是实时数据分析。

数据分析可以让网络营销更加的顺畅，根据网络上出现的数据，分析出人们的需求，从而制定出消费圈子，保证了网络营销的更大利益。所以说，数据分析对人们来说是十分重要的，而数据分析软件的出现给人们带来了不少的便利，根据数据分析软件对现有的数据进行合理准确的分析，从而得到人们想要的结果，这是数据分析的魅力所在。