

海鲜价格预测系统的设计与实现

开题答辩

答辩人: XXX

指导老师: XX

2019 年 3 月 25 日

研究意义

基本内容

研究途径

可行性分析

研究意义

研究背景

随着我国现代化建设的逐步完善，大部分领域都实现了现代化和智能化，唯独农业在现代化建设方面还有所落后，十九大报告指出：我国是一个农业大国，没有实现农业的现代化和繁荣富强，国家现代化是不完整不全面和不牢固的，建设农村现代化关系到建设社会主义现代化强国的大局问题。习主席指出：“加快推进农村农业现代化”，解决好“三农”问题，要走中国特特色现代化道路。现代社会的大数据和人工智能发展飞速，数据分析和智能算法在人们日常生活中的各个方面都有渗透，从大型智能家电的智能管控和用户的个性化推荐都有所渗透，是当代人们生活工作中不可或缺的重要组成部分。

国内外研究现状

机器学习的回归算法

机器学习及其算法在国内外备受瞩目，许多优秀的科研及技术人员致力于此。一些机器学习相关的期刊及会议应运而生，集中收录了大批高质量的论文，汇聚了一些创新性前瞻性的思想。尤其是回归算法，有着非常多的表现优秀的算法，例如：SVR 回归、最小二乘法回归、神经网络回归、决策树回归、xgboost 回归、catboost 回等。

机器学习在价格预测方面的表现

在价格预测方面，国内也有许多优秀的论文和作品，例如：国内外黄金价格的预测、房地产价格的预测、发病预测、大豆期货价格预测等。是一个十分成熟的领域。

发展水平和趋势

机器学习已经形成一种不可逆转的历史性趋势，我们需要立足于此考量如何进行跨部门日常事务处理并将自身业务与市场整体经济状况加以结合。在 39 年的发展历程中，众多企业一直在努力消化、采用并从机器学习技术的发展进步与相关最佳实践中获益。已经有企业能够真正将其转化为自身业务优势。机器学习同时也将成为实现“大数据”的重要途径。同时衍生出许多非常优秀的机器学习框架，例如：TensorFlow, H2O, Ai-One 等。

基本内容

课题基本内容

本课题主要是通过大量的不同地区的海鲜价格的基础数据 (数据集) 和通过机器学习的相关算法, 预测当地海鲜价格, 同时在传入新的数据之后能够实时的反馈出最新的合理价格, 也能存储历史数据, 图表展示价格走向, 用两种或两种以上算法同时预测, 对比分析, 得出最合理的结果。

- * 主要工作

- 数据集的工作, 数据清洗和数据集的扩充

- * 算法选择

- 拟准备用神经网络拟合算法和另一种算法对数据进行预测处理, 对比分析

- 神经网络算法

- 包括搭建和训练神经网络的过程, 包括参数的选取, 网络层数的选取, 喂入数据集, 训练, 迭代调整参数, 得到最优模型

- 其他机器学习预测算法做比较

- * 保留历史数据, 做出价格走势图

- * 留出接口给其他模块调用

可能遇到的困难

1. 数据集的获取，目前没有开源的国内海鲜市场数据集可以使用.
2. 数据集量需求大，至少一万条的数据，工作量巨大
3. 神经网络可能欠拟合
4. 神经网络可能过拟合

解决办法

1. 在爬虫获取数据的同时搜集更多公开的数据.
2. 考虑伪数据集，用均值或者其他方法来扩充数据
3. 欠拟合问题
 - 添加其他特征项
 - 添加多项式特征
 - 减少正则化参数
4. 过拟合问题
 - 重新清洗数据
 - 增大数据的训练量
 - 采用正则化方法

研究途径

- 爬虫搜集数据（大约一万条），数据集去重，数据集降维，去除噪音等问题。
- 算法选择，拟准备用神经网络拟合算法和另一种算法对数据进行预测处理，对比分析。
- 神经网络算法包括搭建和训练神经网络的过程，包括参数的选取，网络层数的选取，喂入数据集，训练，迭代调整参数，得到最优模型
- 保存历史数据，做出价格走势图留出接口以便其他模块进行调用

可行性分析

可行性分析

技术可行性

TensorFlow 是一个开放源代码软件库，用于进行高性能数值计算。借助其灵活的架构，用户可以轻松地将计算工作部署到多种平台（CPU、GPU、TPU）和设备（桌面设备、服务器集群、移动设备、边缘设备等）TensorFlow 最初是由 Google Brain 团队（隶属于 Google 的 AI 部门）中的研究人员和工程师开发的，可为机器学习和深度学习提供强力支持，并且其灵活的数值计算核心广泛应用于许多其他科学领域。

法律可行性

这个系统的设计是在相关法律法规下实施完成的，所以不存在任何法律问题。因此，在法律上是完全可行的，只要人物做的不要违规就没有问题，还有数据的来源比较正规，不用做商业用途，TensorFlow 是提供给机器学习开发人员的引擎，这个设计没相关的专利或者商业问题，在法律上是完全可行的

经济可行性

TensorFlow 是一个免费的框架，实现上并不需要太多的资源，本系统所需要的经济很少，需要的成本很低，经济上是完全可行的，从目前来看，基本可以由一个人来实现，所以经济上是没有任何问题的。

That's All, Thanks.