

大学生创新训练项目申请书

项目编号 S201910536032

项目名称 基于大数据技术结合网络舆情对商品销量的预测

项目负责人 李嘉依 联系电话 17375870834

所在学院 数学与统计学院

学 号 201764100112 专业班级 数学与应用数学 1701 班

指导教师 赵康

E-mail 1121306809@qq.com

申请日期 2019 年 4 月 29 日

起止年月 2019 年 5 月至 2021 年 5 月

长沙理工大学

填 写 说 明

1、本申请书所列各项内容均须实事求是，认真填写，表达明确严谨，简明扼要

2、申请人可以是个人，也可为创新团队，首页只填负责人。“项目编号”一栏不填。

3、本申请书为大 16 开本（A4），左侧装订成册。可网上下载、自行复印或加页，但格式、内容、大小均须与原件一致。

4、负责人所在学院认真审核，经初评和答辩，签署意见后，将申请书（一式两份）报送××××大学项目管理办公室。

一、基本情况

项目名称	基于大数据技术结合网络舆情对商品销量的预测						
所属学科	学科一级门:	理学		学科二级类:	统计学类		
申请金额	2万元		起止年月	2019年5月至2021年5月			
负责人姓名	李嘉依	性别	女	民族	汉	出生年月	1999年10月
学号	201764100112	联系电话	宅:		手机: 17375870834		
指导教师	赵康	联系电话	宅:		手机: 13875972975		
负责人曾经参与科研的情况	<p>负责人参与竞赛情况如下:</p> <p>2019年美国大学生数学建模比赛三等奖;</p> <p>长沙理工大学数学与统计学院2018年第一届BP大赛三等奖;</p> <p>长沙理工大学数学与统计学院2018年互联网+大赛三等奖;</p> <p>长沙理工大学数学与统计学院2019年第二届BP大赛二等奖;</p> <p>长沙理工大学数学与统计学院十月诗会诗歌抢答赛三等奖;</p> <p>参加长沙理工大学第九届数学竞赛;</p> <p>参加2019年长沙理工大学第五届“互联网+”大赛</p>						
指导教师承担科研课题情况	<ol style="list-style-type: none"> 长沙理工大学青年教师成长计划项目,有限环上矩阵图的图自同态与矩阵几何,2019.1-2021.12,主持,在研; 湖南省重点实验室自由探索项目,有限环上矩阵图的图自同构与矩阵几何,2018.1-2019.12,主持,在研; 湖南省教育厅科研项目(16C0037),有限环上几类矩阵图的次成分研究,2016.1-2018.12,主持,结题; 国家自然科学基金面上项目(11371072),矩阵空间的图同态与矩阵几何,2014.1-2017.12,结题,参与; 国家自然科学基金面上项目(11371073),两类特殊非线性方程组的算法与理论研究,2014.1-2017.12,结题,参与; 						
指导教师对本项目的支持情况	<p>近三年,赵康老师指导大学生情况如下:</p> <ol style="list-style-type: none"> 数模竞赛:高教杯全国大学生数模竞赛获二等奖一项;湖南省大学生数模竞赛一等奖三等奖各一项;指导湖南省研究生数学建模竞赛二等奖1项;指导美国大学生数学建模竞赛三等奖两项。 指导长沙理工大学2019年“挑战杯”大学生课外学术科技作品竞赛三等奖; 指导本科生发表科研论文一篇; <p>由此可见,赵康老师在指导大学生学科竞赛科研方面有丰富的经验。对于本项目,赵康老师可以在数据分析、论文撰写等方面给出专业的指导。</p>						

	姓名	学号	专业班级	所在学院	项目中的分工
项目组主要成员	敖璇	201747100108	应用统计学 1801 班	数学与统计学院	调研, 数据分析
	周诗雨	201747100115	应用统计学 1801 班	数学与统计学院	数据分析, 撰写论文
	董佳勋	201730010427	交通运输类 1704 班	交通运输学院	数据采集, 数据分析
	张增瑞	201730010511	交通运输类 1705 班	交通运输学院	数据采集, 数据分析

二、 立项依据（可加页）

（一） 项目简介

随着互联网技术与大数据技术的快速发展，大数据、云计算技术在物流领域已进入广泛应用期。目前大数据在需求预测、配送优化、库存控制方面的应用尤为明显。本项目运用大数据、云计算技术对商品历史销售记录信息进行整合分析，提出结合网络舆情建立数学模型，预测未来不同地区不同商品的销售情况与需求。同时考虑不同地区仓库间爆款商品的调配问题，得到调配方案，达到防止货物积压，优化商品库存，提高运输效率与经济收益的目的。

（二） 研究目的

1. 结合网络舆情，运用大数据技术预测未来一段时间不同地区不同商品的销售情况和销量需求；
2. 根据其销量需求，考虑仓库间的商品调配问题，达到缩短运输时间，提高运输效率。

（三） 研究内容

1. 运用大数据技术对商品历史销售记录信息进行收集、整合和分析，结合网络舆情，预测未来一段时间不同地区不同商品的销售情况和核心城市各片区的主流单品的销量需求；

2. 从海量商品中选取针对不同地区的爆款商品并预测其销售量，据此考虑不同地区仓库间的爆款商品的调配问题，通过此来缩短运输时间，防止货物积压，优化库存，提高运输效率。

（四）国、内外研究现状和发展动态

1. 国内外研究现状：大数据即广义的数据，它是一切可以用作数据资料的统称，即可以被收集、测量、分析，能进行可视化处理的信息或知识，包含产生于自然和人类世界中的文字、图片、音频、视频等，以及各种物质材料。传统数据是手工记录、整理的数字，主要服务于数字记录与统计。大数据是智能计算机运算、分析的数据。它一方面指数据产生规模扩大、种类多样化，海量数据自动产生；另一方面指人工智能可以学习和分析大数据，代替人进行预判。

大数据建构了数据化世界理念，一切需通过数据呈现。同时，人类的自主判断会减少，而更依附于算法对大数据的分析和预判。比如：在总统选举中，媒体利用大数据分析和描述总统和选民属性，预测选举结果；在体育赛事中，利用大数据分析球队的进球指数、运动员平时训练和比赛的竞技状态数据，预判冠军球队等等。在算法为王的时代，一切都可以被纳入算法中，算法是否具有价值偏向、算法能否得出客观公正的结论、大数据是否一定是完备的数据等，这些都会左右算法的结果，进而影响人的观念和行为。

大数据技术是在“互联网+”背景下生成的一种新型产物，通过利用新的思维方式进行数据运用，随着互联网技术的不断发展，各种各样的数据也在不断的增加，因此，如何进行有效数据处理也成为一个很重要的问题。

随着大数据技术的广泛应用，其对数据进行有效地采集与整理，然后进行科学的分析，从而将一些有效数据进行收集，为企业发展提供数据支持。目前大数据技术中对数据采集工作来说，主要分为两种方式：集中采集与分布采集。因为互联网数据存在的形式种类非常多，如何有效的进行采集与处理非常重要。

互联网技术的快速发展使得传统的数据存储模式受到限制，其自身储存数据量较小，不能够有效的对大量数据进行分析，导致其不能符合企业的需求。对于不断产生的庞大数据举行有效的计算也是非常重要的。云计算的发展使其能够将不同的数据进行分类存储，并通过对这些数据进行有效的分析。因此，大数据、云计算技术已经是社会发展的一个趋势。

大数据、云计算技术在物流领域已进入广泛应用期，国内外都在不断提高数据的处理能力，研究大数据与信息化对接。目前基于大数据在物流产业上的研究

范围已经非常广泛，比如智慧物流模式的构建、物流配送优化选址、物流成本预测等方面。与国内相比，国外基于大数据、云计算技术的运用和研究更加成熟，且深入到了物流产业的各个环节，使得其物流行业已经进入成熟阶段^[2]。

2018年7月，中国产业信息网发布了2018年中国物流供应链的管理现状，并对其行业发展趋势进行了分析。2018年1-5月，全国物流总额为105.3万亿元，同比增长7.1%^[3]。由此可知，我国物流发展情况良好，仍有巨大的潜力，但是我国现代物流和供应链管理行业仍处在初级发展阶段，提供的增值服务比较少，而增值服务十分重要，已经成为主流趋势，并且大数据技术、云计算技术将成为新的价值点。大数据、云计算技术的逐渐成熟为物流行业提高服务效率、产业结构升级提供了新的思路。由此，如何利用大数据技术对物流服务进行升级，以及目前企业应用大数据的现状成为有价值的研究方向。

2. 发展动态：

百世汇通曾运用大数据管理网点的运营行为，通过数据分析模型预测了几次网点的异动，将问题消灭在了萌芽时期。可见，准确预测市场行情和消费需求在很大程度上能够避免重大问题的出现。如今已经有将GM(1,1)模型、非线性三次指数平滑预测模型和二次回归预测模型相结合组成的组合预测模型、霍尔特指数平滑预测等产品销量预测方法。

以京东为例。京东在2004年正式涉足电商领域，截至2017年3月，京东集团有超过12万员工，主要涉及电商、金融和物流三大块。传统供应链主要面临计划管理、业务监控、成本控制、客户服务等挑战。计划管理要求适合地备货与调拨，业务监控要求将各个环节的数据打通及时跟踪，成本控制主要降低库存成本，客户服务需要在适当的时间，以合理的价格提供给客户最需要的产品。京东利用大数据、云计算技术构建了包括好计划系统、好商品系统、好价格系统等，有效地提高了供应链整体效率。最近，京东发布了智慧供应链战略，以数据挖掘、人工智能、流程再造和技术驱动为源动力，形成覆盖商品、价格、计划、库存、协同五大领域的智慧供应链解决方案。京东目前服务超过2亿活跃用户，超过1万家活跃的供应商，每月产生的对外采购订单超过200万单。仅京东大数据平台数据总量已经突破200PB（拍字节），每天新增超过1.5PB数据，每天约有20万个报表分析作业运行，日处理达到150亿行数据量。京东通过大数据极大的提高了库

存管理能力，用数据预测某款商品未来 28 天在每个仓的销量，从而助力采销提前且高效地进行库存管理，能让商品现货率保持 90%以上^[1]。

目前，对于商品效率预测都是基于商品历史销量进行的预测，无法对由网络舆论造成的短时间内商品销量的大幅度改变进行预测。此外，企业对应用大数据技术进行供应链物流服务升级的需求是比较明显的，其中应用大数据进行需求预测、配送优化、库存控制的需求尤为明显，应用大数据进行采购管理、成本分析的需求更加迫切。

如今互联网发展迅速，微博、微信等社交软件的使用大量普及，网络舆论逐渐发展成影响商品销售量的一大因素，因此，我们团队要综合商品历史销量和网络舆论两大因素来预测商品的销量。

参考文献：

- [1]. 马洁，郑彩云，应用大数据进行供应链物流服务升级的现状和需求分析[J], 物流商论, 2019 (4): 12-13.
- [2] 丰佳栋，云计算视角下的第三方物流服务质量创新模型[J], 中国流通经济, 2015, 29 (2): 33-38.
- [3] 张娜，马燕，崔恒睿，刘逗逗，大数据技术进展与发展趋势，通信设计与应用[J], 2019, 26 (2): 78-79。

(五) 创新点与项目特色

创新点：

1. 将大数据技术与物流问题结合。
2. 结合互联网发展迅速现状，考虑网络舆情因素对商品销售情况的影响。

项目特色：

1. 综合商品历史销量与网络舆情得到优化的商品销量预测模型。
2. 采用大数据的数据处理与分析技术实现对海量数据的处理与分析与分析结果可视化。

(六) 技术路线、拟解决的问题及预期成果

1. 技术路线（见图 1）

- (1) 数据来源：爬虫获取互联网媒体舆情信息；大数据共享平台。
- (2) 数据处理：分布式储存文件系统 HDFS HBase；分布式计算框架, MapReduce；统计分析软件 SPSS SAS；图计算平台 Pregel。
- (3) 数数据应用：数学计算软件 MATLAB。

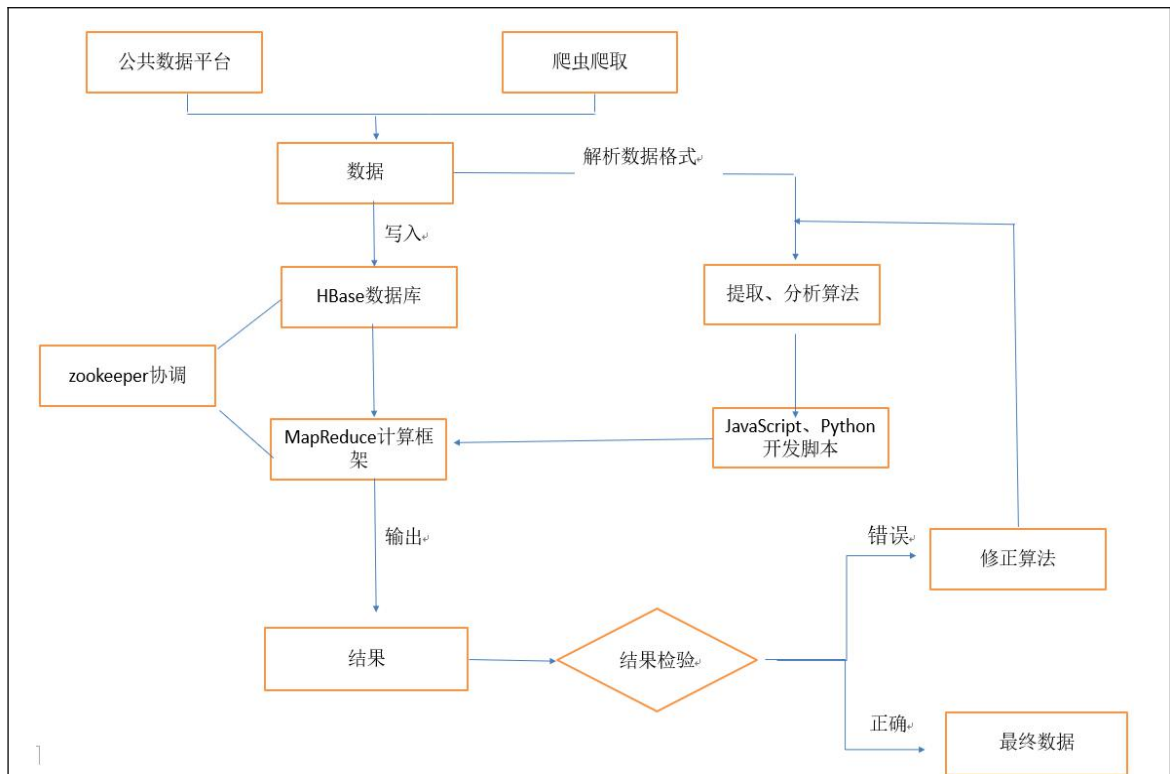


图 1 项目技术路线图

2. 拟解决的问题

(1) 商品存量不足、商品积压、商品调配不当、运输效率低等造成经济效益损失的问题。

(2) 短期内商品需求的预测，并根据预测进行商品采购、调配等问题。

3. 预期成果

(1) 获取与互联网媒体舆情信息有关的商品；得到商品在不同地区的销售情况；得到网络舆情对商品销量的影响情况；得到商品的预测销售情况；得到各地区仓库间商品调配方案。

(2) 研究报告一份；发表论文 1-2 篇；

(七) 项目研究进度安排

1. 2019.5-2019.12: 构建结合网络舆论与历史销量对各商品销量的预测模型

工作内容: 网络舆情分类考察、商品销售记录数据收集、商品销量预测模型构建、检验预测模型准确度

2. 2020.1-2020.12: 构建仓库间商品调配模型。工作内容: 商品仓库信息收集、商品调配模型的建立与检验、系统整理研究成果, 撰写论文。

3. 2021.1-2021.5: 整理前面工作成果, 完成研究报告。

(八) 已有基础

1. 与本项目有关的研究积累和已取得的成绩

项目组成员成绩优秀，基础扎实，积极参加学科竞赛，并多次获奖。其中：部分成员参与了长沙理工大学 BIM 技术大赛，‘基于 BIM+GIS 技术协同在海滨高速公路运行于维护应用示范’的项目研究，并获得二等奖；项目组成员组队参加了 2018 年高教杯大学生数学建模比赛以及 2019 年美国大学生数学建模比赛，为本项目的开展打下了坚实的基础。

2. 已具备的条件，尚缺少的条件及解决方法

长沙理工大学数学与统计学院学术交流非常活跃，平均每周一次的专家学术报告为学生创造了很好的学术氛围。另外，学院建有完善的实验室与资料室，为学生参与各类学科竞赛提供了平台。

三、 经费预算

开支科目	预算经费 (元)	主要用途	阶段下达经费计划(元)	
			前半阶段	后半阶段
预算经费总额	20000		10000	10000
1. 业务费	14000		6000	8000
(1) 计算、分析、测试费	0		0	0
(2) 能源动力费	0		0	0
(3) 会议、差旅费	6000	参加会议比赛等的会务费、培训费、参赛费等	3000	3000
(4) 文献检索费	0		0	0
(5) 论文出版费	8000	论文版面费	3000	5000
2. 仪器设备购置费	2000	购置 1 个移动硬盘与 5 个大容量优盘	2000	0
3. 实验装置试制费	0		0	0
4. 材料费	4000	打/复印装订费，购买专业图书费	2000	2000
学校批准经费	20000			

四、 指导教师意见

该项目基于大数据技术对商品销售量进行预测，研究内容具有很强的应用背景，研究目的明确，研究方法合理。该项目符合大学生创新训练项目的要求，能充分锻炼项目组成员解决问题的能力。

特此推荐，希望予以支持！

导师（签章）：

年 月 日

五、 院系大学生创新创业训练计划专家组意见

专家组组长（签章）：
年 月 日

六、 学校大学生创新创业训练计划专家组意见

负责人（签章）：
年 月 日

七、 大学生创新创业训练计划领导小组审批意见

负责人（签章）：
年 月 日