

一、专业实践训练整体情况

| | | |
|--|------------------------|---|
| 实践单位名称 | 杭州美平米科技有限公司 | |
| 实践单位地点 | 浙江省杭州市上城区高德置地 A3 座 9 楼 | |
| 实践岗位名称 | 高级测试工程师 | |
| 专业实践训练时间 | 集中进行 | 2021 年 06 月 01 日开始 至 2022 年 02 月 28 日结束 |
| | | 专业实践训练累计 272 天（单位考核前），其中项目研究天数 105 天（单位考核前） |
| <p>（1）基本概况（含实践单位简介、实习实践内容等）</p> <p>实践单位简介： 本人所在的实践单位是杭州美平米科技有限公司，位于杭州市上城区高德置地中心 1 号楼 9 楼。公司经营范围包括：电子商务平台技术、计算机软硬件、计算机网络技术的技术开发、技术服务、技术咨询、成果转让，计算机系统集成，国内广告的设计、制作、代理、发布，经营性互联网文化服务、网站建设、网页设计、市场营销策划，批发（含网上销售）计算机及辅助设备、通讯设备等。公司建设现代数字技术与供应链融合模式，打通供应链上的商流、信息流、资金流、物流，实现供应链可视化管理。具有即时、可视、可感知、可调节的能力，满足个性化订单、库存、智能生产之间的平衡，建立柔性供应链，提升企业的效率和效益，打造全生命周期产业互联网供应链服务。</p> <p>实习实践内容： 我的实践内容有以下几点</p> <ol style="list-style-type: none"> 1、配合产品经理进行需求分析，参与需求评审，提出建设性意见； 2、主导负责公司软件产品的业务测试、自动化测试、性能测试并编写相关技术文档； 3、参与数据清洗、数据分析、模型训练的相关工作。 | | |

(2) 项目研究概述（含项目名称、项目来源、项目经费、主要研究目标和技术难点等）

项目名称：电商平台数据分析

项目来源：实践单位

项目经费：20 万左右

主要研究目标：女装行业是目前服装行业的最大类目，女装有更新速度快、受众广、品类多样的特点，而女装行业中销量占比较高的几个品类依次是：连衣裙、T 恤、休闲裤、衬衫及针织衫。于是针对女装零售及批发行业，从我们的电商平台获得月销售量高于 100 笔的女装商品信息，每个商品包含自身的一些属性、交易信息、商品所属店铺的信息以及评价信息及购买商品的个人信息。通过这些信息，我们对畅销商品的共性做了进一步分析，即畅销商品的性质、销量的相关因素。由此来把握整个市场的宏观特征，发掘高销量服装共性，分析影响销量的显著因素，并且对这些因素做量化。从而指导商家更合理有效的捕捉流行趋势、定位市场、合理管理库存以及更改营销策略。

技术难点：

- 1、一般从平台获取的数据不够规整，如获取大段的文字信息，字符处理需要尽量全面的考虑，也要细致进行数据清洗和核对结果。
- 2、根据描述性分析、探索性分析及建模需求，对数据进行合理的转化和处理，使得数据更具解释性。
- 3、绘制词云图，找到关键标签。

(3) 项目开展情况（含项目研究内容、研究方案及技术路线，研究团队分工、本人承担任务及完成情况，存在问题与改进建议等，不少于 500 字。）

研究内容、方案及技术路线：

项目研究的电商平台分为用户端和商家端和商家运营后台。用户端运用先进的互联网技术以及大数据云计算，为用户提供方便快捷的购买体验，针对每个用户的不同喜好和属性推荐用户可能会喜欢和需要购买的商品，对用户来说更便捷安全放心有保障。商家端致力于打造潮流女装批发和零售市场，为入驻的商家提供了具有：简单开店，快速上新产品，方便管理订单，全面统计经营数据，通过数字化运营为商家提供更精准的营销模式、同时提供直播售卖商品等特色功能的店铺管理软件。商家运营后台为商家运营提供数据统计、数据分析的功能，以数据驱动运营，为商家精准营销提供数据支撑。

团队分工、本人承担任务及完成情况：

整个项目团队共 7 人：产品 1 人、测试 1 人、前端开发 2 人、后端开发 3 人，我负责的是整个项目的测试和部分数据分析工作。

按功能来分，用户端主要包括首页商品及活动展示、用户推荐、购物车、订单模块、售后模块，商户端主要包括商品售卖模块、活动管理、订单模块、直播模块、售后模

块、在线客服等模块，这些模块的测试工作我都按时完成，并及时提出缺陷和跟进问题的解决；另外承担了部分数据清洗和数据分析的工作，为模型训练提供数据准备，保证了整个项目于 2022 年 1 月 20 日按时上线运营。

问题与改进建议：

数据质量是数据分析项目的障碍性问题。良好的数据质量会带来更准确的结果和更好的预测，而糟糕的数据质量会产生误导性的预测，进而可能导致公司采取有误导性的行动。解决数据质量的问题，可以使用机器学习的转换、插补和预测的组合；另外在源数据中进行手动数据清洗和细致核对结果，防止数据不准确的发生。

二、专业实践训练收获

(一) 围绕考核评价指标体系，举例说明以下收获（不少于 800 字）

知识掌握：

结合团队成员以往的工作经历和项目敏捷开发的需要，项目后台架构采用了微服务架构，在测试后台的过程中对微服务架构有了一定的了解，挑战了测试的难度，微服务各服务之间通过接口来交互，当接口有改变的时候，对所有的调用方都有影响，这时自动化测试就显得非常重要了，如果要靠人工一个个接口去测试，工作量繁重，整个项目实现了接口的链路自动化，大大提高了测试效率。

数据清洗过程用到的技术：用 RStudio 读入多个数据文件并合并，通过提取文件名并运用 apply 函数族简单快速的读取多个数据文件到一个列表中；在数据清洗中对缺失值进行处理，从字符串中提取有用信息，采用函数对变量进行计算生成新的变量，实现因子变量标签转换。

数据分析用到的技术：数据描述性分析方法，对调查的所有变量进行统计性描述，描述内容包括数据的集中趋势分析、数据的离散程度分析、数据的分布，并利用图表形式，对总体变量进行展现。

通过本项目的实施，我积累了数据分析的项目经验，加深了对数据分析的理解，具备了基本数据分析的能力。

在模型训练方面，主要工作都是由大数据团队完成，我参与了解了模型训练的整个过程，从数据收集、数据清洗、数据分析，到模型训练、模型测试，以及最终的模型上线和迭代优化，积累了实际的工程经验，这为我的论文写作打下了基础。

能力提升：

在实际项目中接触 Spring Cloud 微服务架构，掌握了测试微服务架构的方法，进一步丰富了我的技术栈。

数据分析过程中用到了数据描述性分析的方法，通过数据描述性分析过程可以很直观的得到一些基本的结论，并发现数据之间的内在关系。

在团队从需求确定到项目上线，整个项目进程中，团队内外的不断交流沟通，提升了我的表达能力和沟通技巧。

素质养成：

学无止境的理念，身处这个新兴技术层出不穷的年代，越来越感受到学习的重要性。作为一个有着近十年工作经验的技术人员，掌握了不同的开发语言 Java\Python 和不同的工具如，UI 自动化测试 SONIC 云测平台、接口自动化及性能测试工具 Jmeter、数据分析工具 Rstudio 等；还有正在学习的新技术：数据挖掘、神经网络，我深刻的认识到，只有不断提升自己，才能在竞争日益激烈的社会中更好的实现自己的价值。

其他：实践过程中，校内、校外导师都给予了我很多指导和帮助，非常感激两位老师的栽培，谢谢你们的耐心教导。

(二) 取得成效

在电商平台发展过程中，传统的营销模式已经难以适应市场发展变化的需求。市场营销是电商企业应对市场竞争的有效手段，但是传统营销方式主要是以大量的营销广告为主，同时配合大量的线下人工宣传，进而导致营销成本高居不下。精准营销可以充分借助大量的数据分析技术，对电商企业发展过程中的市场需求情况进行准确分析，根据其商品的特点在市场中寻找准客户，确保商品销售的精准性和了解客户的需求，有针对性地进行营销，进而可以为商户节约营销成本，提升商户商品的销量、提高商户整体收益。

我们研发的项目于 2022 年 1 月 20 日正式上线运营，项目上线后平稳运行，没有出现技术性缺陷。并以我们公司自营店为试点商户，在商户运营方面自取得了不错的效果，通过一个月的营销数据统计，销量提升了 5%左右、客户留存率提升了 7%左右，库存积压率降低了 10%左右，并为其他商家运营提供了更精准的营销数据。

例如，通过对商品属性的特定特征描述性分析，我们得到了如下几个结论：

1. 月销量 100-200 笔的商品数量较多，且随着商品单价的上升有下降趋势；
2. 年龄层对销量影响有差异，其中在 25-29 岁之间的年龄层消费的商品数量较多；
3. 从商品属性中提取关键标签，发现“显瘦”、“显高”、“气质”、“新款”、“中长款”、“韩版”等特征整体对销量都有不同程度的积极作用，并且不同年龄层对销量的影响不同。

由这些结论，使得我们平台根据商家运营的不同品类商品给出了年龄段、商品档位、款式特征等相关推荐，为商家运营提供了数据依据，从而达到了更加精准和精细化运营的效果，提高了生产效益，降低了库存积压带来的损失，给商户带来了实质性的经济利益，并受到了一些商家的好评。

与学位论文撰写的相关程度：

我的学位论文是服装搭配推荐算法相关，需要使用测试数据集训练模型，在项目中我不仅熟悉了数据清洗、数据分析和模型训练的整个流程，而且实际项目中使用了数据描述性分析的方法挖掘出了一些数据之间的关系，最终都取得了不错的效果，这对我后期的论文的撰写有很大的帮助。

3. 在校期间主要研究成果【含产品与样机、专利（含申请）、著作、软件著作权、论文、标准、获奖、成果转化等】

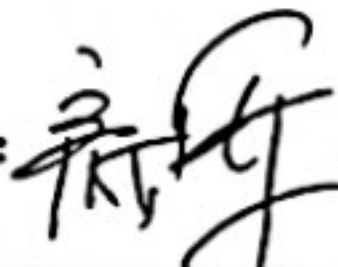

| 成果名称 | 类别含产品与样机、专利（含申请）、著作、软件著作权、论文、标准、获奖、成果转化等] | 发表时间/授权或申请时间等 | 刊物名称/专利授权或申请号等 | 本人排名/总人数 | 学校排名/总参与单位数 |
|------|---|---------------|----------------|----------|-------------|
|------|---|---------------|----------------|----------|-------------|

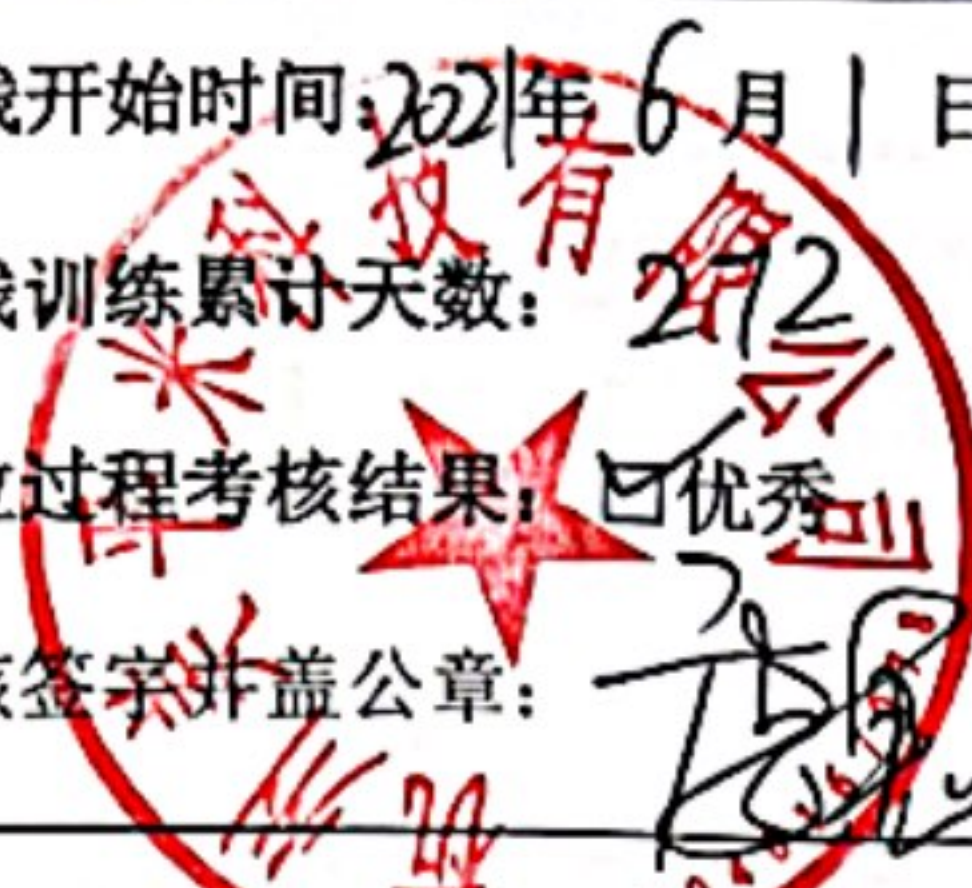
本人承诺

在专业实践训练及考核报告撰写过程中，如实提供材料，严守
学术道德、遵循学术规范。

签字： 王庆娅 2022年6月1日

三、考核评价

| | |
|--|---|
| <p>校外合作 导师(或现 场导师) 评价</p> | <p>重点对研究生项目研究开展情况、职业素养、行业知识掌握、环境和岗位适应能力、工程实践能力、团队协作能力，以及通过技术创新、成果转化、解决工程实际问题等取得的经济和社会效益等方面的评价：</p> <p>王庆娅同学工作积极严谨，认真负责，注重团队合作，知识面广，专业基础扎实，能灵活的将专业知识应用到工作任务中，取得了很好的业绩，具有一定的开拓和创新精神。</p> <p>校外合作导师（或现场导师）签字： 2022年6月1日</p> |
| <p>校内导师 评价</p> | <p>重点对研究生科学素质、基础及专业知识掌握、技术应用创新能力、取得的研究成果、项目研究与学位论文撰写的相关程度等方面的评价：</p> <p>王庆娅同学认真完成专业实践，将专业知识应用到实际业务领域中并取得了较好的成果，为后续学位论文撰写积累了一定的理论基础，同时也培养了研究能力。</p> <p>校内导师签字： 2022年6月1日</p> |

| | |
|-----------------------------|---|
| <p>实践单位 过程考核 意见</p> | <p>实际实践开始时间:2022年6月1日 实际实践结束时间:2022年2月28日 专业实践训练累计天数:272 其中项目研究天数:105 实践单位过程考核结果: <input checked="" type="checkbox"/>优秀 <input type="checkbox"/>良好 <input type="checkbox"/>合格 <input type="checkbox"/>不合格 审核签字并盖公章:  2022年6月1日</p> |
| <p>最终考核 结果审核 备案</p> | <p>考核总成绩 (由现场答辩考核成绩90%+单位过程考核成绩10%组成): 是否重修: <input type="checkbox"/>是 <input type="checkbox"/>否 教学管理部 (或相关分院) 审核签字 (公章): _____ 年 月 日</p> |

四、相关支撑材料

在校期间主要研究成果【含产品与样机、专利（含申请）、著作、软件著作权、论文、标准、获奖、成果转化等】证明材料原件扫描件，具体提交要求如下：

1. 产品与样机扫描件包含企业证明材料（含产品与样机功能及创新性介绍、社会经济效益、个人贡献说明及相关照片等）。

2. 授权专利扫描件包含专利证书授权页；未授权专利扫描件包含专利受理书扫描件和专利请求书扫描件需加盖事务所公章或发明专利申请页（有二维码）。

3. 著作扫描件包含封面、封底和版权页。

4. 软件著作权扫描件包含著作权证书和事务所出具著作权人排序证明。

5. 论文扫描件包含封面、封底、目录和论文全文（含收录证明）。

6. 标准扫描件包含封面、版权页、发布公告、前言和目次。

7. 获奖扫描件包含显示单位和个人排名的获奖证书。

8. 成果转化扫描件包含企业证明材料（含成果技术说明、社会经济效益、个人贡献说明及相关照片等）。