

广西科技大学启迪数字学院

《2019 级综合实训》实施方案

实训合作单位： 昆山杰普软件科技有限公司

实训方向： 云计算大数据开发方向

报价：

2022 年 月 日起至 2022 年 月

时段： 日止（待定）

专业（年级） 2019 级

一、实训对象

广西科技大学启迪数字学院
2019 级学生

二、实训周期

2022 年 月 日起至 2022 年 月 日止（根据学校放假时间待定）

三、实训地点

昆山杰普软件科技有限公司
江苏省昆山市巴城镇学院路 828 号昆山浦东软件园 4 号楼 A 座

四、实训费用支付方式

现金或者转账或者就业后付款

五、实训目标:云计算大数据开发工程师

1. 项目实践能力¹

通过多个项目来学习掌握软件开发技术,学习过程中每天需要完成课堂作业和课后作业,每学完一个专题就配有一个项目实战,将累计完成 40000 行以上程序。每日 6 小时上课,3 小时自习,全程上机操作,并有专业技术人员进行辅导,一课一练,每日 300-500 行代码,500 个经典案例教学,关注学员软件开发及设计思想的训练,培养动手能力。项目实战期间以分组形式进行,整个过程中完全采用企业开发流程,模块化管理,学生可以通过版本控制工具 SVN 以及项目管理工具 Maven 等进行团队开发,保证人人都是可以参与到项目中。通过项目实战的方式巩固之前学到的知识并充实自身的项目经验。

2. 职业能力/岗位技能

专注于培养企业当下紧缺的云计算大数据平台运维及开发工程师。学员不仅掌握云计算大数据平台的构建、管理、性能检测及评估,而且掌握在云平台上进行大数据应用系统设计、开发及调优等技术及技巧。学习的核心课程包括Linux运维、KVM虚拟化技术、Linux集群管理、Hadoop离线大数据分析、Hbases数据库、spark&flink等在线实时大数据分析等。

六、实训环境与师资

1. 住宿环境与基本生活费用

2 人间公寓:酒店式公寓(浦东软件园人才公寓):650 元/月(含水、宽带)、电费按园区实际使用结算、热水器、空调、独立卫生间。



4 人间人才公寓：美居客人才公寓：480 元/月、水电费按园区用水电实际使用结算、空调、热水器、独立卫生间。



6 人间人才公寓：美居客人才公寓：380 元/月、水电费按园区用水电实际使用结算、空调、热水器、独立卫生间。



每天的基本生活费用 18 元左右。

2. 实训教学环境（硬件情况、软件配置环境）

24 小时开放，全天 16 个小时学习上机时间，全程项目老师辅导。



2) 专家师资



杰普公司拥有 100 多位全职的来自各行业一线的资深架构师、工程师、咨询人力资源顾问。全面的 IT 技术教育和职业背景教育是杰普 IT 培训的一大亮点。企业一线技术专家（全职授课），10 年经验倾囊相授，每位老师讲授个人技术精髓。

Kevin Ding(丁老师) 复旦大学计算机专业硕士，前任美国 Eclair 集团研发中心(Mainnet-System)(上海)系统架构师

擅长嵌入式物联网系统设计、大数据分析处理，人工智能领域机器学习、深度学习算法与建模。完成项目包括：Eclair 集团 ERP 及 CRM 平台基础架构设计、Web Services 中间件应用构件库研发、受富士通东京研究院委托主持 J2EE 程序自动生成工具研发工作、国家 863 及上海科委多媒体智能信息检索系统研发、中国石化集团油田开发决策支持系统咨询、上海汽车南站智能化系统等。

Robin Liu(刘老师) 杰普金牌讲师

原美国 N.E.T 公司高级软件工程师，具有多年开发及培训经验，精通 Linux/Unix 操作系统、ARM 体系结构、Linux 系统移植、Linux 设备驱动开发、上层应用开发等，曾参与美国 N.E.T 公司为北约组织提供 EMS 和 NMS 系统项目。

Terry Ren(任老师) 资深软件设计师 杰普金牌讲师

曾参与开发法国雪铁龙汽车公司的生产流水线管理调度系统（基于 Unix 平台下的 C/C++ 技术），中国股指期货上市风险控制系统、特别会员结算系统等项目。精通 Java EE 及 C++ 编程, GUI/QT 编程、ARM 硬件结构、arm 汇编指令、熟悉 Linux 核心代码，操作系统内核的裁剪和移植，Linux 设备驱动开发，嵌入式数据库编程等。

Jun Yuan(袁老师) 杰普金牌培训师

曾就职于英国 TravelFusion 公司，资深软件设计师。参与和主持过的项目有 TravelFusion 公司的基于 SOA 架构的旅游信息垂直搜索引擎基础架构设计和开发、杰普公司无线监控系统、台湾某银行信托资产评估系统设计、台湾某银行网银系统开发。

Zhongliang Hu (胡老师) 杰普金牌培训师

高级软件设计师，擅长大数据技术栈，尤其精通离线数据处理。曾参与智慧校园产品群 10 多种产品的设计与研发；主要主持的典型项目如上海 AI 西岸管理系统、知识产权管理系统等大数据相关的项目。

七、实训教学模式与就业保证

1. 教学模式

杰普引进欧美流行的“以研发带动培训”的新型人才孵化模式，与上海各大软件企业合作，进行企业定制的人才培训。

项目驱动培训模式，企业级的真实项目，以项目带动实训，通过多个项目来学习掌握软件开发技术，同时累积至少 1-2 年的软件开发经验，学习过程中每天需要完成课堂作业和课后作业，每学完一个专题就配有一个项目实战，学习过程中学员一人一机，边讲边练，将累计 40000 万行以上的程序。

2. 教学特色

杰普公司凭借多年在 IT 行业和人力资源行业的经验优势、紧贴 IT 行业技术人才的用人需求，推出 2022 年云计算大数据培训课程。该课程以就业为中心，重点关注行业内新技术的流行趋势，对企业开发中需要具备的技术点及时的进行更新和调整，以便根据企业需求，培养市场严重紧缺的具有相当经验的大数据开发人员。

3. 过程管理

杰普独创了“**过程化管理**”体系。通过科学的教学资源配置，标准的质量管理体系，将管理的每个过程细化，确保每位学员的学习质量。

杰普为每位班级组建了专门的管理团队。**管理团队由课程顾问、班主任、技术老师、技术辅导老师、求职顾问组成**。分别负责学员的：住宿管理、班级管理、全天 12 小时技术指导答疑、就业辅导及求职推荐。

在线学习：杰普搭建的“企业级实训平台”，免费开放杰普所有课程视频教程。提供学员在线学习机会。

一课一测：杰普研发的“在线考试系统”，每门课程结束，将进行基本技能、动手能力、重点课程的测评。并根据班级的测评结果，进行阶段性技术串讲及答疑。

每日一练：每天课程结束，学员需要完成技术导师布置的实习任务，巩固当日所学技能，做到活学活用，举一反三。

4. 教学团队

杰普科技有 100 余位中外资深技术和行业专家，技术涵盖在 IT 相关各个行业，资深的技术经验+丰富的教学经验助力学员成功就业。杰普师资，汇聚来自 Sybase、Sun、Eclaire 等国际跨国公司的一线 IT 专家，平均拥有 6 年以上研发经验，是中国内地最具技术实力的专家小组。每位老师具备资深技术实力和企业培训经验，倾囊传授技术精髓，让每位学员成功成材。

5. 实习与就业保证

【“优质就业” 服务理念】

“优质就业”除了要求学员可以找到一份工作，更强调学员的就业速度、质量与职业发展。提高学员进入世界 500 强企业、海外就业、国际知名 IT 企业、国内一流 IT 企业的比例及机会，是优质就业的要求。

【就业承诺】

培训合格学员毕业后 100%推荐就业，许多学员通过企业定制的方式进入知名企业！合格学员专享杰普资深职业规划师、猎头顾问提供的就业推荐和专业的求职指导服务。

1、已毕业的学生 100%高质量就业，成功是可以复制的。

- 2、 签订就业协议，不就业 100%退还学费。
- 3、 杰普百家企业联盟。
- 4、 2022-2023 年校园招聘及人才培训订单，确保求职无忧。

【就业专家—杰普 IT 人才服务部】

杰普成立专门的人才服务部门，为企业组建 IT 团队，为学员提供专业的就业推荐服务。确保就业品质。

- 职业顾问，猎头专家，技术顾问组成专业团队
- 提供专业化求职训练指导及个性化就业推荐服务
- 广泛的企业合作，帮助学员求职无忧
- 学员寄语：技术决定未来成就的高度
- 职业拓展能力训练课程
- IT 企业类型介绍与职位需求分析
- 如何制定个人职业发展的 5 年计划
- 有效的简历制作与投递
- 面试技巧—如何打动主考官
- Face to face 模拟
- IT 求职英语集训
- 职场新人宝典

就业推荐：





上市公司：鼎捷软件

500强：上海银行



上市公司：熙菱软件

国有企业：宝信、欧冶

500强：捷咨咨询



上市公司：安硕

上市公司：艾融

上市公司：润和

上市公司：亚信



上市公司：汉得

重点软件企业：群硕

上市公司：软通动力

上市公司：金在线

部分就业展示：





永辉云创 (500强企业)



爱奇艺 (互联网公司)



今日头条 (互联网公司)



航信德利 (外资企业)



恒生电子 (上市公司)



天正信息 (知名企业)



唯品会 (互联网公司)



茂轮科技 (互联网公司)



新浪微博 (互联网企业)



京东 (互联网公司)



永辉云创 (500强企业)



挖财网 (互联网公司)



同程旅游 (互联网公司)



复星国际 (500强企业)



聚光科技 (上市公司)



欧冶云商 (国有企业)



用友政务 (上市公司)



联鼎软件 (上市公司)



知微瑞驰 (知名企业)



国际航空 (国企)



亚信科技 (上市公司)



朗驹信息 (高新技术企业)



通用汽车 (500强)



凯捷咨询 (500强)



东软集团 (上市公司)



汉得信息 (上市公司)



博科资讯 (行业龙头)



帮你玩 (高新技术企业)



用友软件 (500强)



亚信集团 (上市公司)



德明通讯 (知名企业)



复星国际 (上市公司)



跨越信息 (知名企业)



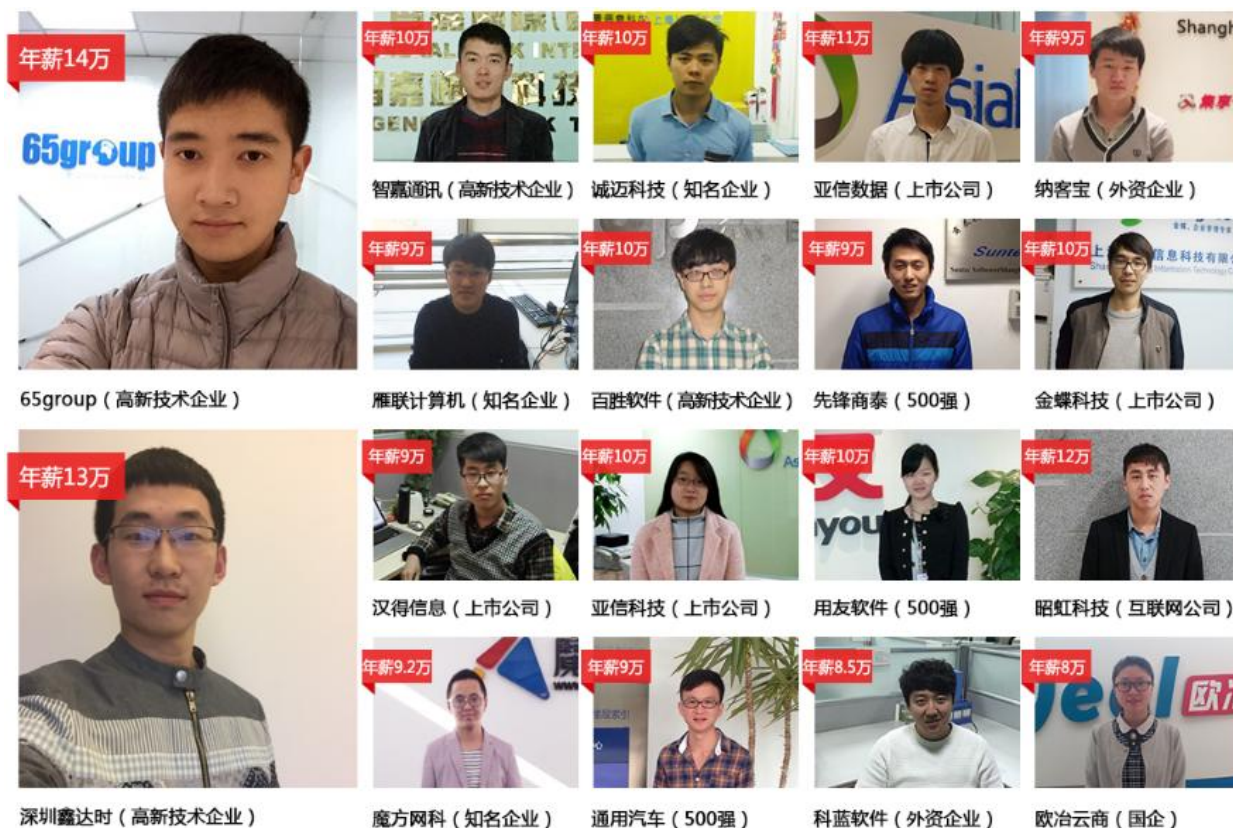
汉得信息 (上市公司)



HP (500强)



宝信软件 (国企)



学员就业在百度公司-9万



学员就业在淘宝-年薪10万



学员就业在京东



学员就业在爱奇艺-年薪20万



学员就业在新浪公司-年薪10.4万



学员就业在500强



学员就业在国有企业



学员就业在外资公司



学员就业在上市公司



八、实训纪律与成绩评定

1) 课堂纪律:

学员上课时间不得从事任何与教学无关的活动, 严禁玩游戏。违反规定三次以上者, 取消就业推荐资格, 学员上课时间不得擅自离开教室。如携带手机, 须处于关机或无声状态, 不得影响课堂秩序。学员只能在学籍所在班级听课, 不得擅自串班, 如有发生, 取消培训资格。

2) 考勤制度:

教学管理员每天检查和记录学员出勤状况。

迟到: 学员在规定的上课时间, 未按时到教室者, 视为迟到。学员迟到次数累计三次, 警告批评; 迟到次数累计超过五次者, 取消证书申办及就业推荐资格! 迟到超过一小时者, 视为旷课一次。

旷课: 学员未经请假或请假未准, 擅自缺勤, 视为旷课。旷课次数累计超过三次者, 取消证书申办及就业推荐资格。

病假: 学员因身体不适, 不能坚持正常学习者, 须到医院就诊并向班主任出示相关病假证明。

事假: 学员在公司学习期间(包括课程阶段和项目阶段)如无特殊情况不得请事假。若遇特殊情况, 须严格履行请假手续, 由本人提出申请, 报由班主任批准同意。

3) 教室管理制度:

学员须严格遵守相关的教室管理制度。未经 MIS 部同意, 学员不得自行使用、拔插、移动或调换相应设备。违者按设备价值的 10—100 倍罚款, 并写出书面检讨, 通报全班。因违反操作章程而造成设备的损坏, 视情节轻重, 由责任人负责一部分或全部赔偿, 必要时报请公司领导给予处分。

公司有权调整教室的安排和使用。学员须在公司指定的教室学习和练习。教室对学员免费开放。课余时间学员不得在教室内做与学习无关的事情。违反规定三次以上者, 上报学院。

未经 MIS 部同意, 学员不得下载、安装和使用与教学无关的软件。否则由此引起的一切后果(包括但不限于病毒传播, 数据丢失等), 由责任人承担。

学员须保持教室安静、整洁。严禁在教室内进食、抽烟、随地吐痰、乱丢杂物等。学员须使用有盖水容器, 以免液体倾漏。

学员在当天学习结束后, 须确保教室的灯和空调都关闭后锁门再离开教室, 养成良好的维护习惯。

4) 实训成绩的评定标准与规则

考核是检验教学效果、保证教学质量的重要手段, 其目的在于指导和督促学员系统地复习和巩固所学知识和技能, 检验其理解程度和灵活运用能力, 调动学员学习的主动性和积极性, 培养学员的创新精神和创

新思维。对学员的考核由任课教师负责。成绩不合格或缺失考核成绩的, 取消就业推荐资格。考试方式主要包含以下三种形式:日常课堂测验、单门课程考试、项目实战考评。

九、实训课程

1. 专业课程

专注于培养企业当下紧缺的云计算大数据平台运维及开发工程师。

1) 实训内容简介:

在整个实训阶段, 需要深入理解 Java 虚拟机原理及程序调优方法; 熟练掌握 Java 程序设计方法; 熟练掌握面向对象程序设计方法及程序鲁棒性设计原则及方法; Java 8 中的函数式程序设计方法; MySQL 数据库体系结构; 掌握 MySQL 数据库集群建设及管理方法; 具备数据模型设计、数据库构建、开发及管理的能力。深入理解虚拟化原理、云计算与虚拟化应用场景; 掌握 Linux 平台重要虚拟化技术 Qemu+KVM+Libvirt 的建立、虚拟机动态迁移、网络及存储管理、虚拟机克隆等方法; 掌握虚拟化集群的远程管理方法。深入理解 Hadoop 体系结构及产品生态环境构成; 熟练掌握 HDFS 分布式文件系统建立、监控方法; 熟练掌握用 Java 语言对 HDFS 的输入输出操作; 掌握 Hadoop 集群的构建方法; 深入掌握 MR 编程方法。

2) 实训环境:

实训环境采用杰普自主搭建的云计算大数据技术实验平台。

3) 实训方法:

一) 全程企业实战演练、项目经理手把手传授实战经验

全程模拟企业工作场景, 采取一人一机, 边讲边练, 企业项目经理全程指导。从项目的行业介绍、需求分析展开, 将项目实训按实际开发流程划分为多个阶段, 并设立各阶段的实训目标及项目验收方案, 将具体的知识点、技术要义、软件开发规范、各角色的工作内容、开发经验以项目案例的形式, 传授给学生。平均每日学生需完成 400-500 行代码, 全程 4-5 万行代码的强化训练, 由浅入深展开实训。

二) “准职业人”的企业化人才培养

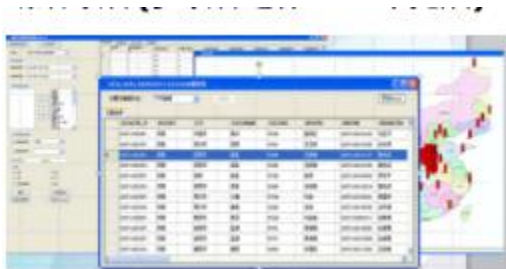
按企业的工作方式及行为模式, 结合学校的相关制度, 培养学生, 锻炼学生的职场技能及职业素养。

- 项目开发职场化: 项目阶段, 为每名学生开通工作邮箱, 采取团队开发的形式, 为学生划分团队角色(各阶段角色轮换): Team Leader、工程师、技术支持、QA 测试人员, 要求学生根据自己的角色及工作内容, 每日汇报工作报告(工作报告参照软件企业格式要求)。通过这样的学习形式, 让学生在实训中学习团队合作、团队沟通、职场规范等技巧。

- 成绩评定职业化: 每阶段项目结束后, 采取答辩以及考试等方式, 验收学生学习成果。锻炼学生的沟通表达、演讲能力。实训结束, 模拟企业的员工考核制度, 对学生进行人事+技术评测, 对学生的整体实训成果全面测评。并评选“优秀员工”, 予以奖励。

2. 实训项目

综合项目(多个不同项目中选择 1-2 个)



项目实战 1:电子商务系统-商品推荐子系统

生活中你一定见过很多的推荐系统，不管是亚马逊的推荐还是淘宝的猜你喜欢，又或是其他站点的智能个性化推荐。很多时候你会疑惑“他是怎么知道我正需要这个商品的”，或“他是怎么知道我喜欢这个风格的音乐的”，又或是“他是怎么拥有‘智慧’知道我的喜欢偏好的？”，该案例正好解答你的疑惑，同时也为你打开人工智能这个新世界大门。

推荐系统是大数据人工智能、机器学习中最常见和最容易理解的应用之一，如亚马逊的推荐引擎其为浏览网站的用户提供个性化的内容。大数据智能推荐系统就是根据电商网站的历史数据，智能化的、个性化的帮助用户决定应该购买什么产品，模拟销售人员帮助客户完成购买过程。

大多数推荐引擎工作有两种办法。它们都可以依靠每个用户喜欢的商品的属性来发现用户也可能喜欢的其他商品，或者推荐引擎可以依靠喜好和其他用户的愿望来计算用户之间的相似性指数并据此推荐项目给他们。

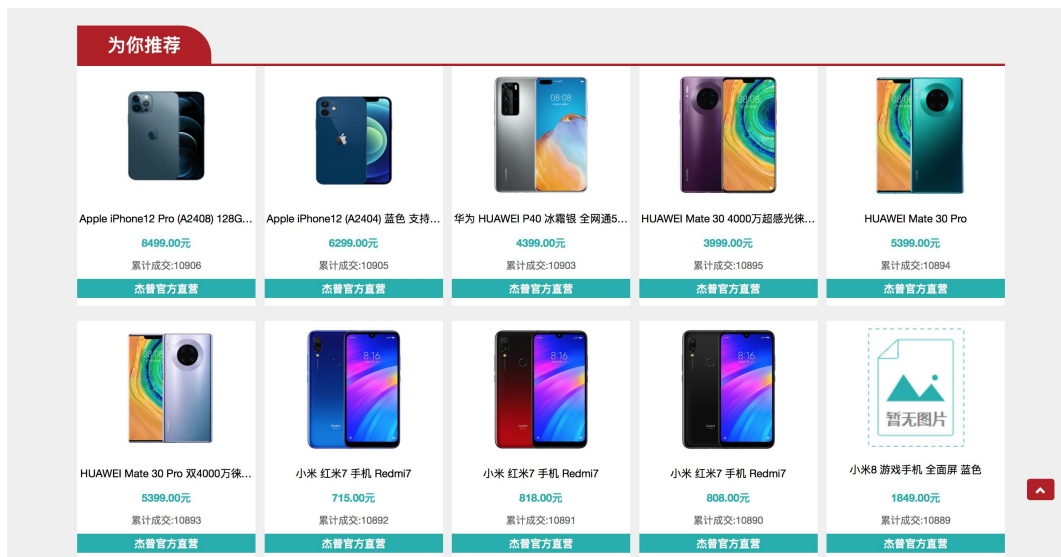
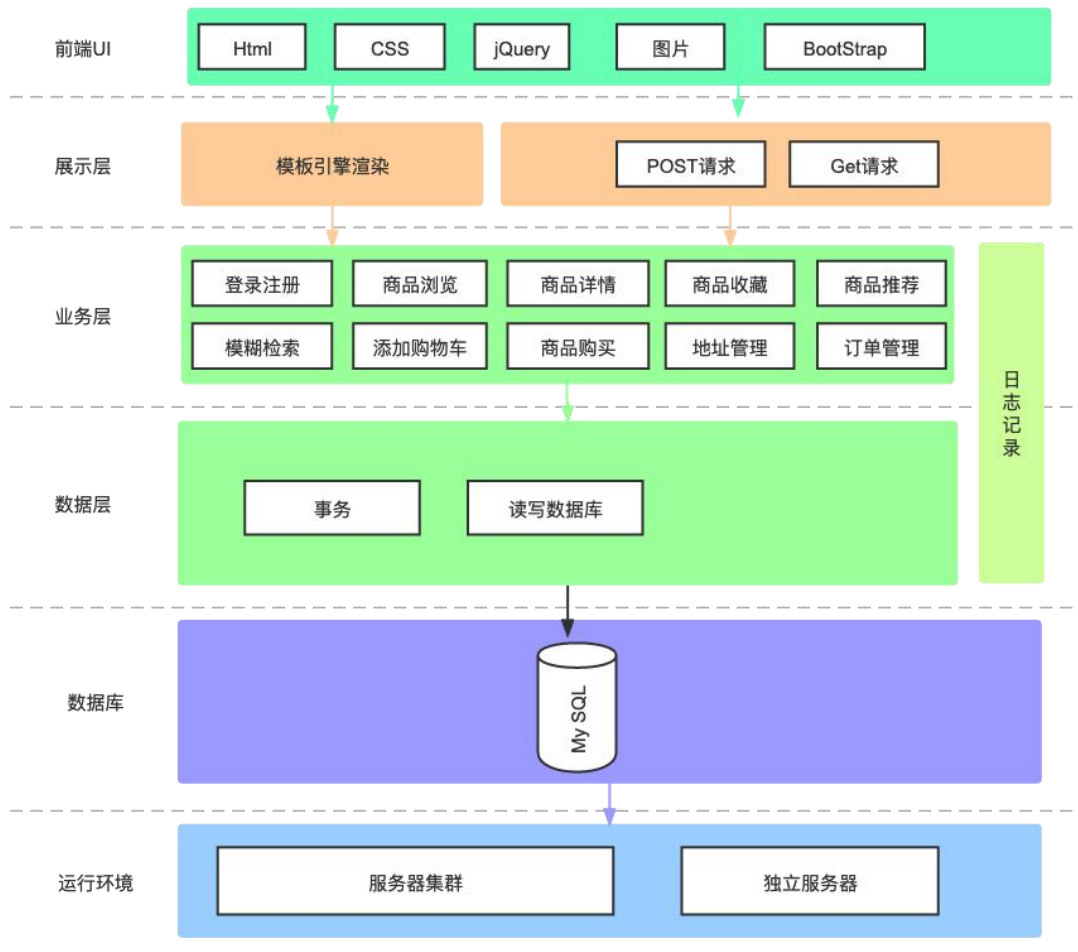
该案例使用协同过滤算法来实现推荐引擎，所以该系统的四个主要功能模块为用户画像建模，商品特征建模、用户行为分析、以及商品推荐。

1、用户画像建模用于计算用户之间的相似性，两个用户相似意味着他们拥有相同喜好，或者相同的商品偏好。两个偏好相同的用户其中一个购买了某个商品，意味着另一个用户也有可能喜欢。从而把这个商品推荐给另一个用户

2、商品特征建模用于计算商品之间的关联关系或者相似性，商品的关联是可以用来提高销售额，例如某用户购买了网球拍，系统可以推荐网球给这个用户。而商品相似意味着商品之间可以相互替代。这可以改善用户体验。

3、用户行为分析可用于分析的数据非常多，整个日志数据都可以用来分析，其中登录日志可以分析用户购物时间，关键词可以回馈推荐系统的推荐效率等，为了达到教学目的，该案例分析用户的购买记录，转发以及评价评分。这些数据可以用于用户对物品偏好度的降噪，甚至直接用于用户画像建模或者物品特征建模等。

4、商品推荐这个功能模块就是根据用户画像或者商品特征导入推荐算法计算推荐结果，从而最终将推荐商品推荐给用户。



项目实战 2:电信业务运营支撑平台-用户画像子系统

大数据资产变现是国内外电信运营商普遍关注的问题，从分析通信运营商的数据来源及特点入手，从内对外两个方面，分多个场景对数据资产的价值挖掘及应用模式进行了阐述，全面规划了电信运营商的大

数据发函策略，并就实施路径给出了具体建议。

用户画像可以使电信产品的服务对象更加聚焦，更加的专注。本子系统分别从用户的通话数据，短信数据以及流量数据等多个维度构建用户画像模型，实现最终精细化的数据展示及精准的套餐推荐服务。

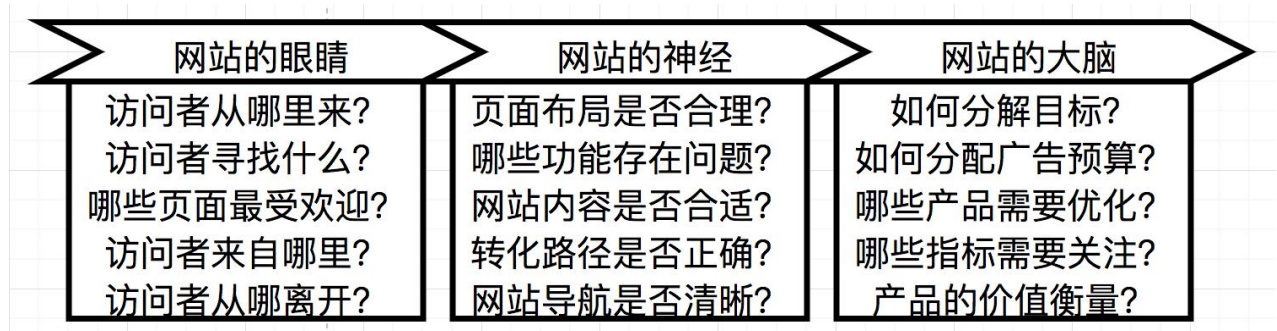


项目 3：电子商务系统-网站统计分析子系统

项目简介：

网站分析，可以帮助网站管理员、运营人员、推广人员等实时获取网站流量信息，并从流量来源、网站内容、网站访客特性等多方面提供网站分析的数据依据。从而帮助提高网站流量，提升网站用户体验，让访客更多的沉淀下来变成会员或客户，通过更少的投入获取最大化的收入。

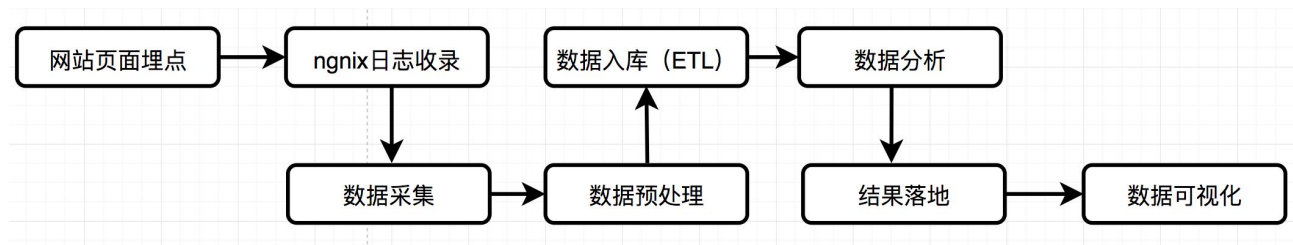
事实上网站分析设计的内容非常广泛，由很多部分组成。每一部分都可以单独作为一个分析项目，如下所示：

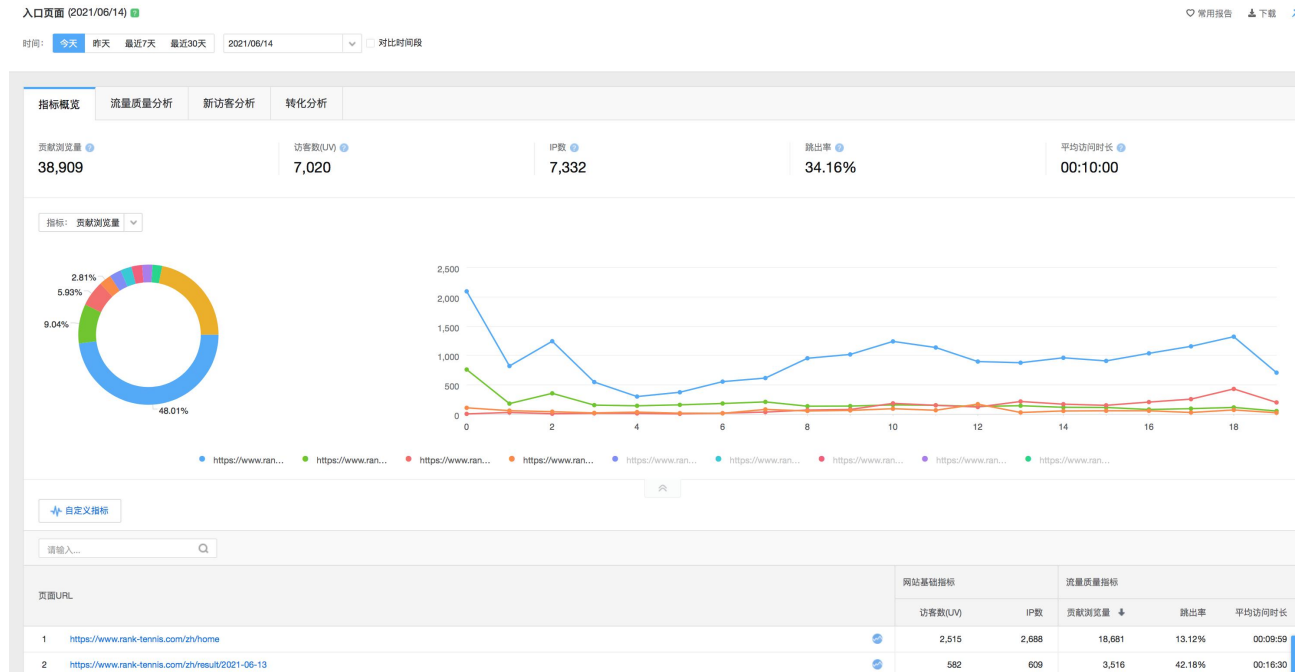


1、网站分析是网站的眼睛。是从网站的营销角度看到的网站分析。在这部分中，网站分析的主要对象是访问者，访问者在网站中的行为以及不同流量之间的关系

2、网站分析是整个网站的神经系统。这是从产品和架构的角度看到的网站分析。在这部分中，网站分析的主要对象是网站的逻辑和结构，网站的导航结构是否合理，注册购买流程的逻辑是否顺畅

3、网站分析是网站的大脑，在这部门中，网站分析的主要分析对象是投资回报率（ROI）。也就是说在现有的情况下，如何合理的分配预算和资源以完成网站的目标

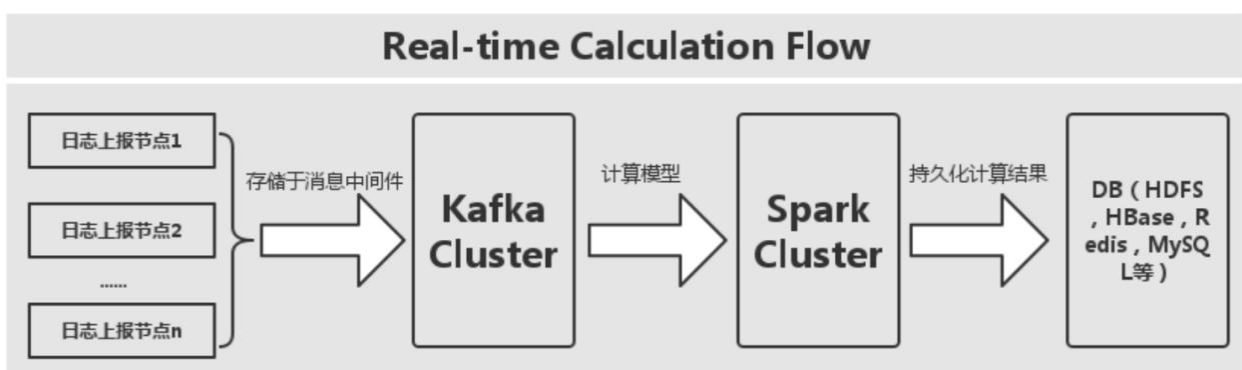


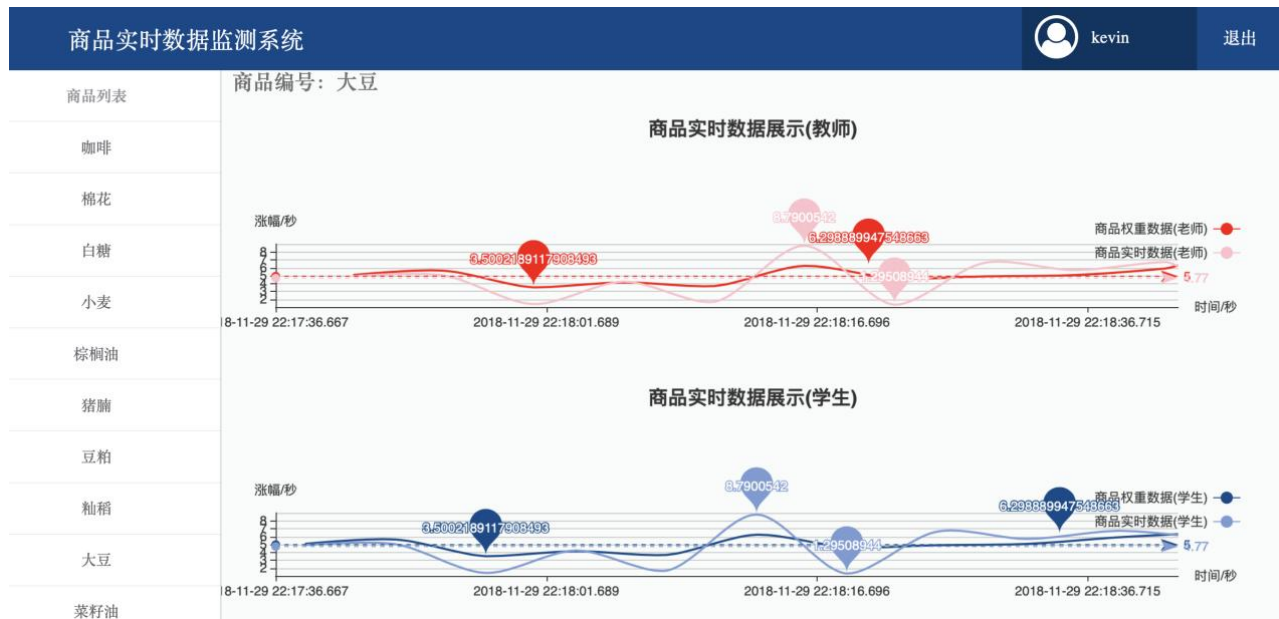


项目 4：期货实时数据动态分析系统

项目介绍：

本项目对大数据平台中引用服务器一方产生的商品数据进行实时分析，便于我们更好的观察到商品的价格变动等相关信息，以便更快的进行商品买卖，在股票交易系统、期货交易中占重要的地位的子系统。整个实时商品数据分析系统可以分为“数据采集-数据缓冲-数据处理-数据存储”四个步骤。Apache 项目下的 Kafka 框架可以很好的从多源目标收集流式数据数据，由于采集数据与处理数据的速度可能会出现不一致，该系统也使用消息中间件 Kafka 来作为缓冲。对流式数据的处理，采用 spark streaming 做实时的数据分析；最后使用 HBase 和 MySQL 作为处理数据的存储工具。





3. 职业素养与职业发展

杰普公司凭借多年在IT行业 and 人力资源行业的经验优势、紧贴IT行业技术人才的用人需求,推出2022年云计算大数据培训课程。该课程以就业为中心,注重行业内新技术的流行趋势,对企业开发中需要具备的技术点及时的进行更新和调整,以便根据企业需求,培养市场严重紧缺的开发人员。另外,通过礼仪培训、就业讲座、模拟面试培养学生的职业素质。最终使其成为满足企业需求的大数据开发人员。

4. 用人单位需求:

企业需求: 来自 51job

<input type="checkbox"/> Hadoop运维架构师	游族网络	上海-徐汇区	30-40万/年	11-16
<input type="checkbox"/> hadoop工程师	好买基金网	上海-虹口区	25000-29999/月	11-16
<input type="checkbox"/> Hadoop高级工程师	Tukmob (土壳移动)	上海-普陀区	15000-19999/月	11-16
<input type="checkbox"/> 大数据平台HADOOP工程师(赴恒...	上海微创软件股份有限公司	上海	6000-7999/月	11-16
<input type="checkbox"/> 大数据开发工程师/hadoop开发...	华尔特人才咨询(上海)有限公司	上海	40000-49999/月	11-16
<input type="checkbox"/> 资深Hadoop开发架构师	平安付智能技术有限公司	上海-徐汇区		11-16
<input type="checkbox"/> 数据工程师(Hadoop)	上海浦东软件园汇智软件发展有限公...	上海	8000-12000/月	11-16
<input type="checkbox"/> Hadoop开发工程师/Hadoop开发	上海智子信息科技股份有限公司	上海	10000-14999/月	11-16
<input type="checkbox"/> hadoop售前工程师	上海音智达信息技术有限公司	上海		11-16
<input type="checkbox"/> Hadoop软件工程师(金融业务)...	上海威士顿信息技术股份有限公司	上海	8000-9999/月	11-16
<input type="checkbox"/> 高级大数据架构师HADOOP/SPAR...	悦岚(上海)数据服务有限公司	上海-杨浦区	25000-29999/月	11-16
<input type="checkbox"/> hadoop研发工程师	上海图聚信息技术有限公司	上海	15000-19999/月	11-15
<input type="checkbox"/> hadoop工程师	深圳万里通网络信息技术有限公司	上海-徐汇区	10000-14999/月	11-16
<input type="checkbox"/> Hadoop高级开发工程师	猫酷	上海-闸北区	10000-15000/月	11-16

<input type="checkbox"/> Hadoop开发工程师(数据分析力...	旦动通大图信息技术有限公司	上海	10-15万/年	11-16
<input type="checkbox"/> 初级数据挖掘工程师 (Hadoop) ...	新新贷(上海)金融信息服务有限公...	上海	8000-9999/月	11-16
<input type="checkbox"/> Hadoop开发工程师	银联智惠信息服务(上海)有限公司...	上海-浦东新区	10000-14999/月	11-16
<input type="checkbox"/> Hadoop工程师	小沃科技有限公司	上海	15000-19999/月	11-16
<input type="checkbox"/> Hadoop工程师	上海迅时通信设备有限公司	上海-徐汇区		11-16
<input type="checkbox"/> 资深hadoop工程师	上海高欣计算机系统有限公司	上海-浦东新区	15000-25000/月	11-16
<input type="checkbox"/> Hadoop运维工程师	上海谦问万答吧云计算科技有限公司...	上海-闵行区	15000-19999/月	11-16
<input type="checkbox"/> 大数据工程师 (Hadoop)	软通动力信息技术(集团)有限公司...	上海	10000-14999/月	11-16
<input type="checkbox"/> Hadoop架构师	上海巨人网络科技有限公司	上海		11-16
<input type="checkbox"/> Hadoop开发工程师 / 数据挖掘...	上海久懿网络科技有限公司	上海	10000-18000/月	11-16
<input type="checkbox"/> hadoop 大数据工程师	上海视畅信息科技有限公司	上海-浦东新区	10000-14999/月	11-16
<input type="checkbox"/> Hadoop开发工程师 / 数据挖掘...	上海我快网络科技有限公司	上海-虹口区	15000-19999/月	11-16
<input type="checkbox"/> Hadoop开发工程师	上海朝阳永续信息技术股份有限公司...	上海-浦东新区	10000-14999/月	11-16
<input type="checkbox"/> Hadoop兼职讲师	上海信息化培训中心	上海		11-16
<input type="checkbox"/> Hadoop开发工程师	车风网	上海-杨浦区	8000-9999/月	11-16

已选条件: spark(职位名称) + 上海

<input type="checkbox"/> 全选	<input type="button" value="申请职位"/>	<input type="button" value="收藏职位"/>	<input type="button" value="智能排序"/>	<input type="button" value="发布时间"/>	< 1 / 1 >	共6条职位
职位名称	公司名	工作地点	薪资	更新时间		
<input type="checkbox"/> 高级大数据架构师HADOOP/SPAR...	悦岚(上海)数据服务有限公司	上海-杨浦区	25000-29999/月	11-16		
<input type="checkbox"/> 大数据研发 (Spark) 工程师	上海乾特计算机技术有限公司	上海-浦东新区	10000-20000/月	11-16		
<input type="checkbox"/> 数据挖掘算法工程师(hadoop、...	数云信息技术(上海)有限公司	上海	10000-14999/月	11-16		
<input type="checkbox"/> 大数据研发 (Spark) 工程师	上海亿通国际股份有限公司	上海	10000-14999/月	11-13		
<input type="checkbox"/> [和讯大数据] Spark &Had...	和讯信息科技有限公司-大数据	上海-浦东新区	15000-19999/月	11-06		
<input type="checkbox"/> HADOOP SPARK STORM HBASE大数...	上海创业接力企业服务有限公司	上海		10-23		

☐ 全选

共1页, 到第 1 页

以上数据来自<http://www.51job.com>

十、实训进度安排

主讲教师：丁长明

分布式计算技术专家、复旦大学理学硕士，20年IT行业从业经验。

主要完成项目有美国Eclair集团ERP及CRM平台基础架构设计、Web Services中间件应用构件库研发、受富士通东京研究院委托主持J2EE程序自动生成工具研发工作、国家863及上海科委多媒体智能信息检索系统研发、中国石化集团上海海洋石油研究院油田开发决策支持系统、上海汽车南站智能化系统方案设计等。

主讲教师：胡忠良

Java EE及大数据开发项目经理，8年IT行业从业经验。

参与了杰普公司与上海某公司联合开发的“在线校园影视平台”，主持了“企业综合信息发布系统”的研发工作。参与了上海某房地产公司的房产评估系统设计、上海某高校的科研管理系统开发，精通

Java EE 体系架构、目前正在关注大数据、云计算领域开发,并主持和参与了相关平台的设计与研发工作。

第一阶段：面向对象项目编程实战

该阶段学生需要学习 Linux 操作系统、Java 核心技术基础以及 Java 面向对象编程等课程,重点需要掌握 Java 核心技术基础和面向对象编程。该阶段以“物联网环境数据监控中心”为驱动项目,通过监控烟雾传感器、温度传感器、湿度传感器、光亮度传感器等物联网设备传输的数据,利用 Java 高级编程知识开发一款具有分析数据、监控数据、实时警报、日志管理功能的综合应用程序。

提供成果:

在该阶段中,学习完每一天的内容,需要提交当天的作业;完成该阶段项目后,学生需要提交完整的项目需求分析、项目代码以及演示 PPT。

考核方法:

在此阶段学习完成后,包含在该阶段的课程采用考试的方式进行考核,包括 Linux 操作系统基础、Java 核心基础以及 Java 面向对象的内容;学生需要通过简单的答辩检测项目是否完成,以及检测项目的完成度。

第二阶段：大数据离线数据分析实战及 Nosql 数据库编程实战

大数据离线数据分析实战使用基于协同过滤算法的推荐系统为实战项目。在电商应用中,有一种推荐业务,根据电商网站的历史数据,智能化的、个性化的帮助用户决定应该购买什么产品,模拟销售人员帮助客户完成购买过程。该案例主要包含以下几个功能:商品特征建模、用户行为分析、以及商品推荐。本项目中使用 Hadoop HDFS 作为数据存储,利用 MapReduce 作为分布式计算框架进行相关算法计算,将结果利用 MR 的多数据格式输出技术输出到 Mysql 数据库,对接前台电子商城,从而在客户登陆以后给予合适的商品推荐。

Nosql 数据库编程实战以基于 Pagerank 算法的搜索引擎项目为驱动,该搜索引擎主要利用 Nutch 爬虫框架,在互联网中抓取网页信息、网页间关联关系,利用 Xml 配置文件对接到 Hbase 数据库。Hbase 作为一款 Nosql 数据库擅长存储倒置索引结构,并能做到毫秒级访问,数据存储之后利用基于 hbase 的 MapReduce 程序实现 Pagerank 算法,用来计算网页间的重要性排名,最后构建倒置索引表存储于 Hbase 中。在前台程序中将传统 Dao 层替换为能对接 Hbase 的 Api,从而连接 Hbase 数据库进行查询。该项目在互联网的使用中非常普遍,也能使学生更加理解和熟练使用 NoSql 数据库。

提供成果:

在该阶段中,学习完每一天的内容,需要提交当天的作业;完成该阶段项目后,学生需要提交完整的项目需求分析、项目代码以及演示 PPT。

考核方法:

在此阶段学习完成后,包含在该阶段的课程采用考试方式进行考核,包括 MySQL 数据库技术、JDBC 操作数据库等内容;学生需要通过简单的答辩检测项目是否完成,以及检测项目的完成度。

第三阶段：大数据实时数据分析实战

本阶段以 100 年新生儿数据分析和基于实时流式处理的商品价格分析引擎为项目,利用 Spark 生态圈中的模块完成不同的功能。在新生儿数据分析的例子中,以 Zeppelin 作为可视化工具,将新生儿信息进行各种数据分析,如观测新生儿名字多样性趋势,进行可视化后可以更好的理解 Spark-Sql 模块的核心理念。基于实时流式处理的商品价格分析引擎项目,利用 Spark-Streaming 模块完成实时商品价格分析,对接 kafka 分布式消息订阅发布系统,可以模拟动态商品折线图以及相关数据的计算处理,该项目的原理在股票交易系统、期货交易系统中均有广泛的应用优势。能够熟练使用 JAVA 操作 redis 数据库。最终完成大型交友网站的相关模块功能。

提供成果:

在该阶段中,学习完每一天的内容,需要提交当天的作业;完成该阶段项目后,学生需要提交完整的项目需求分析、项目代码以及演示 PPT。

考核方法：

在此阶段学习完成后，包含在该阶段的课程采用闭卷考试方式进行考核，包括 hadoop、hbase 等技术内容；学生需要通过简单的答辩检测项目是否完成，以及检测项目的完成度。

第四阶段：流式数据处理分析实战

通过 flume、kafka、Flink 等整合开发。通过网站流量统计分析，可以帮助网站管理员、运营人员、推广人员等实时获取网站流量信息，并从流量来源、网站内容、网站访客特性等多方面提供网站分析的数据依据。从而帮助提高网站流量，提升网站用户体验。让访客更多的沉淀下来变成会员或客户，通过更少的投入获取最大化的收入

提供成果：

在该阶段中，学习完每一天的内容，需要提交当天的作业；完成该阶段项目后，学生需要提交完整的项目需求分析、项目代码以及演示 PPT。

考核方法：

在此阶段学习完成后，包含在该阶段的课程采用操作实践的方式进行考核，实践考核包括利用 Flume 进行数据采集、kafka 进行数据缓存，flink 完成项目案例的实时计算。学生需要通过简单的答辩检测项目是否完成，以及检测项目的完成度。

实训支持资源：

建立了“数字化、共享型”专业教学资源。为了向学生提供更好的学习平台支持，满足不同层次学生的学习需求，杰普提供全套的实训资源，包括：每门课程的授课课件、教学视频、文档模板、练习题及答案、随堂笔记、项目案例等。学生可随时学习巩固或跟随项目导师完成最新的公司项目。

具体内容如下：

时间		实训内容
注：该计划可根据学生的基础、接受程度和教学互动情况而适当调整		
第 一 个 月	第 1/2 周	<p>Linux & KVM</p> <p>理解 Linux 操作系统体系结构，掌握 Linux 文件系统管理、进程管理、任务管理、用户管理、权限管理、网络设备配置管理、分布式文件系统 NFS 管理、远程访问、操作系统性能检测等方法。深入理解虚拟化核心概念，虚拟化方法。掌握以 KVM 为主的虚拟化原理、相关系统安装配置、虚拟化网络管理、虚拟机构建访问配置、虚拟存储管理、虚拟域管理及虚拟机迁移。</p> <p>Java 核心基础</p> <p>掌握 Java 编程语言的核心基础，掌握 Java 面向对象编程。</p>

	第 3/4 周	<h2>Java 高级编程 Java 8 程序设计</h2> <p>深入理解 Java 虚拟机原理及程序调优方法，熟练掌握 Java 程序设计方法，熟练掌握 OO 程序设计方法及程序鲁棒性设计原则及方法，能综合运用图形程序、并发程序、网络程序开发；初步掌握并行程序设计方法。</p> <h2>Java 8 程序设计</h2> <p>深入理解函数式程序设计的理论，熟练掌握 java 8 中的函数式程序设计方法，熟练掌握 java 8 中基于函数式程序设计方法构建的集合框架；</p>
第二个月	第 5/6 周	<h2>数据库与数据访问技术</h2> <p>从数据的结构上分类，商业数据类型分为结构化、半结构化和无结构化。在大数据分析领域经常需要将结构化数据从传统的关系型数据库导入大数据平台，为后期大数据分析做准备工作。另 web 上存在大量用 xml 描述的半结构化数据也是大数据领域中特别关注的一类数据。因此是否掌握结构化及半结构化数据的处理分析方法直接关系到大数据分析能否正常实施。</p> <h2>MySQL 数据库</h2> <p>理解 MySQL 数据库体系结构；掌握 MySQL 数据库集群建设及管理方法；具备数据模型设计、数据库构建、开发及管理的能力。</p> <h2>JDBC&XML</h2> <p>理解 JDBC 体系结构，掌握 Java 操作关系型数据库的方法，对数据访问事务有深刻认识，具备分层应用设计的能力。掌握大数据领域中半结构化数据模型建立方法，掌握 Java 处理半结构化数据的方法和</p>

		<p>技巧，理解分布式计算领域中 XML 的作用。</p> <p>项目管理及编程规范（SVN、Maven）</p> <p>掌握软件开发编程规范。熟练掌握 svn 版本控制器的搭建和使用，熟练掌握使用 Maven 构建项目</p>
	第 7/8 周	<p>Hadoop</p> <p>掌握大数据领域中半结构化数据模型建立方法；掌握 Java 处理半结构化数据的方法和技巧；理解分布式计算领域中 XML 的作用。理解 JDBC 体系结构；掌握 Java 操作关系型数据库的方法；对数据访问事务有深刻认识；具备分层应用设计的能力。</p> <p>hbase</p> <p>理解 HBase 体系结构及数据模型，掌握 Hbase 集群构建与管理、HBase 的大数据存储模型设计、HBase 系统的 Java 连接及编程、Hbase 与 MapReduce 的连接技术、信息汇总的方法、Hbase 的表设计、协处理器扩展、Hbase 集群性能检测及评估。</p>
第三个月	第 9/10 周	<p>zookeeper</p> <p>深入理解 Zookeeper 集群体系结构及分布式系统构建的核心控制因素和 Zookeeper 的应用场景，掌握 Zookeeper 集群构建及优化、Zookeeper Cli 常用操作、Zookeeper 异步程序设计及开发、基于 Zookeeper 的分布式锁开发及集群配置模块开发。</p> <p>Hive</p> <p>理解 Hive 体系结构及设计原理，熟练掌握 Hive 系统构建方法、Hive 数据处理方法及常见大数据算法实现、Hive 系统原理及系统优</p>

		化方法。
	第 11/12 周	<p>Scala 程序设计</p> <p>熟悉掌握 Scala 编程核心规则以及规范、熟悉掌握 Scala 集合库，熟悉 Scala 模式匹配、理解 Scala 面向对象编程、掌握 Scala 与 java 集成、熟悉掌握 Scala 函数式编程，理解 Scala 的隐式转换以及 Scala 类型系统、理解 Scala Actor</p> <p>Spark Core</p> <p>理解 Spark 体系结构及原理、Spark 程序设计模型及并程序运行机制，掌握各种 Spark 集群的部署，深入理解 Spark 概念和编程模型 RDD、理解 Spark 的容错机制以及优化操作，熟悉掌握 Pair RDD 的数据连接，聚合，分组，排序的处理；掌握 Spark 操作数据的方式</p> <p>Spark SQL</p> <p>掌握 Kafka 上的 java 编程技术。理解 Spark SQL 架构和原理、掌握 DataFrame、DataSet DSL 和 SQL 开发、Spark 多数据源整合（txt、CSV、Json、parquet、JDBC、Hive），掌握 Spark SQL 执行计划原理、Spark SQL 性能调优</p>
第 四 个 月	第 13 周	<p>Structured Streaming</p> <p>熟练的掌握操作数据源（Parquet、Json、CSV、Hive、以及 JDBC 连接数据库等</p> <p>Flume&Kafka</p> <p>理解 Flume 架构及设计概念，掌握 Flume 通道与接收器、Flume 源与通道选择器、Flume 拦截器、ETL 与路由、Flume 的规划、部署</p>

		与监控，理解 Kafka 体系结构及设计原理，掌握 Kafka 的配置管理、Kafka 存储、分区操作、Kafka 发布订阅、Flume、Kafka 集成及与其他组件对接操作。
	第 14 周	<p>Redis</p> <p>了解 Redis 在 java 程序开发中的作用。掌握 redis 的使用，熟练应用 redis 的命令。熟悉 Redis 在项目中的集成开发</p> <p>Oozie 工作流</p> <p>理解工作流的概念，掌握 oozie 工作流引擎的安装、掌握 oozie 配置运行方式、掌握 oozie java 编程运行方式、掌握多 mapreduce 任务工作流传参、掌握 EL 常数和函数。</p>
	第 15 周	<p>ELK</p> <p>掌握物联网中传感数据实时处理方法。理解 ELK 的架构和原理，熟练的掌握 Elasticsearch 开发、Logstash 数据采集、Kibana 数据可视化。</p>
	第 16 周	<p>Flink</p> <p>掌握 Flink DataStream 的使用，及 Kafka 和 Flink 的结合，掌握 Flink 性能监控、Flink 调优和 Flink SQL 执行计划，深入理解通过 Watermark、Checkpoint、任务调度与负载均衡、状态管理进行任务调优，熟练的掌握 flink 中的 sql 开发，及 Flink SQL 和 hive 的结合</p>
第五个月	第 17/18 周	<p>综合项目(多个不同项目中选择 1-2 个)</p> <p>项目实战 1:电子商务系统-商品推荐子系统</p> <p>本项目针对电商的销售记录数据构建离线分析平台及采用 MR</p>

并行程序设计方法构建面向用户的分布式个性化商品推荐引擎。该系统采用机器学习方法对用户特征进行画像，深度挖掘用户在商品购买方面的个性化爱好，从而让电商能实现精准的商品推荐。本项目构建的引擎亦可适用于音乐分享推荐，电影爱好推荐等领域。

项目实战 2:电信业务运营支撑平台-用户画像子系统

大数据资产变现是国内外电信运营商普遍关注的问题，从分析通信运营商的数据来源及特点入手，从内对外两个方面，分多个场景对数据资产的价值挖掘及应用模式进行了阐述，全面规划了电信运营商的大数据发函策略，并就实施路径给出了具体建议。

用户画像可以使电信产品的服务对象更加聚焦，更加的专注。本子系统分别从用户的通话数据，短信数据以及流量数据等多个维度构建用户画像模型，实现最终精细化的数据展示及精准的套餐推荐服务。

项目 3：电子商务系统-网站统计分析子系统

项目简介：

网站分析，可以帮助网站管理员、运营人员、推广人员等实时获取网站流量信息，并从流量来源、网站内容、网站访客特性等多方面提供网站分析的数据依据。从而帮助提高网站流量，提升网站用户体验，让访客更多的沉淀下来变成会员或客户，通过更少的投入获取最大化的收入。

项目 4：期货实时数据动态分析系统

项目介绍：

本项目对大数据平台中引用服务器一方产生的商品数据进行实

		<p>时分析，便于我们更好的观察到商品的价格变动等相关信息，以便更快的进行商品买卖，在股票交易系统、期货交易中占重要的地位的子系统。整个实时商品数据分析系统可以分为“数据采集-数据缓冲-数据处理-数据存储”四个步骤。Apache 项目下的 Kafka 框架可以很好的从多源目标收集流式数据数据，由于采集数据与处理数据的速度可能会出现不一致，该系统也使用消息中间件 Kafka 来作为缓冲。对流式数据的处理，采用 spark streaming 做实时的数据分析；最后使用 HBase 和 MySQL 作为处理数据的存储工具</p>
	第 19/20 周	<p>就业指导</p> <p>求职指导、简历撰写、面试技巧等。</p>

十一、毕业设计指导

毕业设计环节制定如下实施方案：

1. 组织机构

1. 成立毕业生毕业设计答辩委员会，组成人员如下：

主 任：罗成果

副主任：胡忠良

成 员：朱凯、袁军、李楠、甄磊、胡莉、王亚娟、郭苗

2. 成立毕业设计指导小组，组成人员如下：

组长：郭苗

指导教师：郭苗、崔略、沈维妮、安海龙、王羽、李光辉、白骏程

2. 论文撰写与指导工作要求：

毕业生论文撰写要本着实事求是的原则，以能够解决工作中的实际问题为出发点，引导学生从现实工作中组织素材，不能生搬硬套，敷衍了事；实习指导老师要本着以身作则、讲究效率和效果、主动跟踪指导的原则，妥善完成论文指导任务。

3. 工作内容

- 1) 毕业设计选题
- 2) 毕业设计开题报告
- 3) 毕业设计任务书
- 4) 毕业论文撰写
- 5) 毕业设计中期检查
- 6) 毕业设计集中当面指导及初稿审定

7) 毕业设计初稿打印 2 份/人及再次修订

8) 毕业设计答辩

4. 毕业设计题目及选题预案:

学生通过以上各个阶段项目的实训,可掌握云计算大数据体系的基本技术架构和实现细节,在此基础上,结合实训过程中掌握了解的软件开发的基本流程和训练的团队协作,项目管理,项目协调,实现和实施的各技能,结合杰普公司毕业设计项目库以及当前正在实施的合同项目,特制定如下毕业设计选题预案供毕业设计工作开展参考。

基于 MapReduce 框架的不一致数据发现算法	基于 Docker 的虚拟化运维平台
基于 VMware 的集群云渲染实验平台分析与设计	海洋环境数据的可视分析
气候数据的可视化和可视分析	基于图像的人体区域检测
基于图像的人体手部识别	遗传规划算法性能与特征关系分析
遗传算法性能与特征关系分析	基于神经网络的网络病毒预测模型的实现
基于大数据的教学硬件资源使用效益分析系统的设计与实现	基于大数据的疾病种类统计系统的设计与实现
基于云计算的教学辅助系统的设计与实现	电影推荐系统的设计与实现
微博情感分析	大数据流处理技术及系统部署——以 spark streaming 为例
大数据流处理技术及系统部署——以 storm 为例	大数据流处理平台性能对比研究
基于 Spark 的教育大数据处理系统设计	Openstack 平台上 Hadoop 的部署与维护