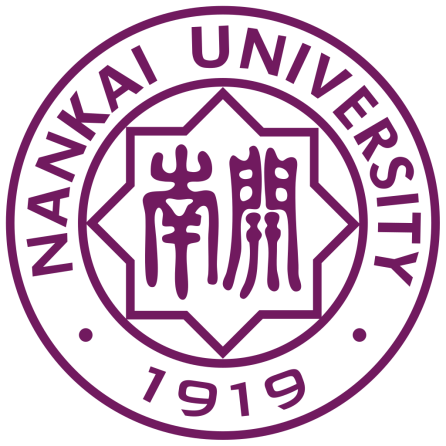


中软国际人才培养解决方案

南开大学软件学院企业综合实训

技术参数及实施方案

v1.0



中软国际教育集团 · 天津

2023.06

**目录**

[一、中软国际公司简介 3](#_Toc24544)

[二、中软国际教育集团简介 5](#_Toc21829)

[三、实训体系四大特性 6](#_Toc31886)

[3.1 实训体系权威性](#_Toc27255)

[3.2 实训体系专业性](#_Toc18140)

[3.3 实训体系实用性](#_Toc2816)

[3.4 实训体系先进性](#_Toc3981)

[四、实训体系五大优势 8](#_Toc14381)

[4.1 秉承“5R”先进的人才培养理念](#_Toc24553)

[4.2 基于“CDIO”国际化工程实践的人才培养规范标准](#_Toc30348)

[4.3 独创的“Project-Driven”项目驱动的授课模式](#_Toc28936)

[4.4 “做中学，学中做”的培训理念](#_Toc20120)

[4.5 企业级标准和工程化的规范](#_Toc30287)

[五、实训总体设计方案 12](#_Toc1957)

[5.1 实训总体时间](#_Toc7457)

[5.2 实训培养目标](#_Toc5992)

[六、实训技术服务内容 13](#_Toc29392)

[6.1项目简介](#_Toc21138)

[6.2实训实施计划](#_Toc11085)

[6.3技术讲座](#_Toc3916)

[七、线上授课保障方案 26](#_Toc22061)

[7.1 开课前一周](#_Toc15610)

[7.2 开课前3天](#_Toc2568)

[7.3 实训过程保障](#_Toc15596)

[7.4 考勤管理](#_Toc17612)

[7.5 成绩评定](#_Toc16246)

[7.6 技术讲座](#_Toc31944)

[7.7 实施资源](#_Toc20567)

[八、配套资源 33](#_Toc14535)

[8.1 师资配比](#_Toc1292)

[8.2 实训环境](#_Toc17075)

[8.3 中软实训项目经理介绍](#_Toc18632)

# 一、中软国际公司简介

中软国际有限公司（以下简称：中软国际）是中国领先的大型综合性软件与信息服务企业，成立于**2000年**，为香港主板上市公司，股票代码**354** 。中软国际总部设于香港，运营总部位于北京，目前公司已在中国北京、上海、深圳、西安等城市、香港特别行政区、美国西雅图、华盛顿、奥斯丁、普林斯顿等城市、英国伦敦、爱尔兰都柏林、匈牙利布达佩斯、日本东京、以及马来西亚吉隆坡等全球**50**余个城市设有**60**多个分支机构和研发中心，员工近**63000**余人。



图1-1：中软国际集团全球交付中心分布图

为适应当今互联网变化，以及**云计算**、**大数据**、**人工智能**带来的挑战，公司将业务划分为技术与专业服务集团（**TPG**）、互联网IT服务集团（**IIG**）和教育科技集团（**ETC**）。

**TPG**主要是为大客户、大行业提供技术和专业服务，目标是成为立足中国、服务全球、行业领先的全链条信息技术服务供应商。公司与华为、微软、汇丰、腾讯、阿里云、中移动等行业巨人有着全面深入的战略合作，深耕金融、电信、互联网、交通、能源等重点行业，持续开拓全球市场，为客户提供全链条、高质量、高效率、具有综合优势和行业特色的技术服务。

**IIG**以“解放号”（JointForce）为核心，并针对政府、制造等典型的长尾市场以互联网组织方式构建核心竞争力。解放号既是一个通过可信人脉组织、由最佳管理实践背书的IT服务众包平台，也是一个海量解决方案研发、实施和集成的工作平台。

**ETC**（中软国际教育集团，简称中软卓越）下设卓越培训中心ETC，是中软国际人才战略的核心组成部分之一，承担着集团发展过程中人才储备和人才培养选拔的任务，是软件及外包业务快速发展的重要人才支撑平台，是集团保持高速增长的动力之一。



图1-2：中软国际集团国内六大研发交付基地

中软国际在全国拥有**六大研发基地**，分别为：北京中软国际总部（**4635**人）、中软国际南京基地（**5463**人）、中软国际西安基地（**10945**人）、中软国际深圳基地（**10439**人）、中软国际上海基地（**3103**人）以及中软国际广州基地（**1553**人）。

中软国际秉承“**率真存厚，立志有恒，奋斗为本，成就客户，创造分享，共同成长**”的核心价值观，成为客户长期、稳定、可信赖的合作伙伴。

公司凭借完整的生态资源，为客户提供云计算、大数据、人工智能等多领域的技术服务，助力客户应对挑战、构建数字化转型。公司主营业务包含面向大客户、大行业的咨询、解决方案与技术服务；以“解放号”软件众包服务平台为核心的互联网众包线上线下运营服务；以及包含了线下卓越培训中心、卓越体验中心与线上宅客学院相结合的教育培训业务。公司的业务布局真正实现了线上与线下业务相结合的**“双轮驱动”模式**。

# 二、中软国际教育集团简介

**中软国际教育集团**在2008年由中软国际投资成立了，下设卓越培训中心（ETC），是中软国际人才战略的核心组成部分之一，承担着集团发展过程中人才储备和培养的任务，是软件及外包业务快速发展的重要人才支撑平台，是集团保持高速增长的动力之一。

天津开发区中软卓越信息技术有限公司是集团在天津投资的全资子公司。教育集团由中软国际培养事业部发展而成，以“中软总公司计算机培养中心”为基础，随着中软国际集团业务的飞速发展，为了满足各分公司急速增长的用人需求，集团投重金研发了专门针对大学生的准员工**“5R”**IT领域专业人才实训课程体系，并先后在**北京、广州、大连、天津、厦门、沈阳、长沙、重庆、西安、太原、南京、武汉、长春、哈尔滨**等**15**个城市建立了**15**个实训基地，基地总面积超过**50000**平方米，可同时容纳**10000**多名学生，年培训规模逾**50000**人次，并先后与超过500所高校建立稳定的人才培训和实训合作。

毕业学员经过选拔进入中软国际或经推荐进入中软国际业务线及国内知名软件公司、互联网公司实习或就业。各基地通过与高等院校的合作，为广大学生提供企业实习与实训机会，从而大大提高大学毕业生的就业能力和就业质量，并成为**中国顶尖的IT实习实训机构之一**。

# 三、实训体系四大特性

## 3.1 实训体系权威性

依托中软国际集团自身庞大的业务线及战略合作伙伴（华为、百度、腾讯、阿里巴巴、微软）的人才需求，与业务线**共同制定专业人才标准**，细化**岗位技能标准**及**职业能力标准**。对专业人才的**“硬技能+软技能”**建立**六维度专业能力画像模型。**

图3-1：IT专业人才六维度能力画像

与集团业务及合作伙伴线共享**技术资源**、**项目资源**共同制定专业人才培养体系及培训资源建设。共享**人力资源**，产生真实的“**用人需求**”，打通**人才供给渠道**。

## 3.2 实训体系专业性

中软卓越培训中心历经十余年的IT专业人才培养，梳理出**完整的**、**系统的专业知识、技能体系**和**专业人才能力提升体系**。细分技术精进**四个阶段**，即：

* **初级阶段**：掌握并巩固专业基础，提升软件编程思想；
* **中级阶段**：强化领域应用，掌握专业应用的核心编程技术及工程化思想；
* **高级阶段**：企业应用技术，保障参训人员符合企业当前主流技术需求；
* **卓越阶段**：企业前沿技术应用，保障参训人员在行业的可持续性发展。

与此同时，对参训人员的**创新能力**及**综合复杂问题解决能力**也提出了更高的要求。中软卓越培训中心通过**“六维度专业能力画像模型”**，以及科学的人才培养专业知识图谱建设及培训组织形式，对参训人员**六个维度能力**进行提升。

除此之外，对参训人员的**软技能**（职业素质）也提出了一整套解决方案。通过对**沟通**、**表达**、**抗压能力**、**时间管理**以及**团队协作**能力的提升也专门设计一整套**OAC**（职业素质能力课程）。

## 3.3 实训体系实用性

**创新型**及**应用型**专业人才能力培养不断提升实训体系的实用性，同时也满足企业实际用人能力需求。中软卓越ETC通过对**企业真实项目**以及**典型的应用**进行专业人才实训主要资源。目前已形成大量丰富的以**云计算**、**大数据**和**人工智能**为行业应用领域核心的**企业级项目案例库**。

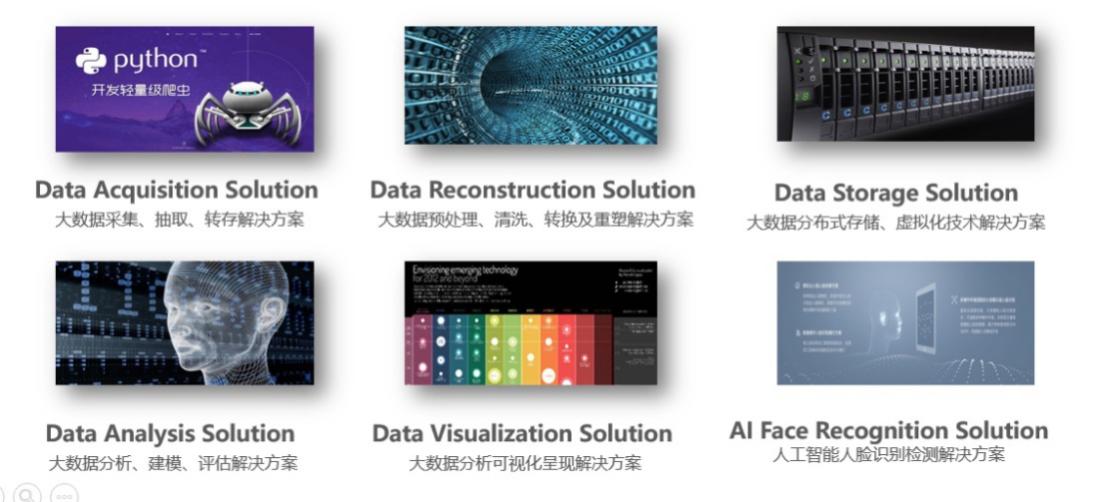


图3-2：实训体系项目案例资源库

## 3.4 实训体系先进性

为了更好保证参训人员符合企业实际应用技术的需求，实训体系的课程内容始终保持**闭环**、**持续性资源迭代**，达到与企业实际用人需求共成长。

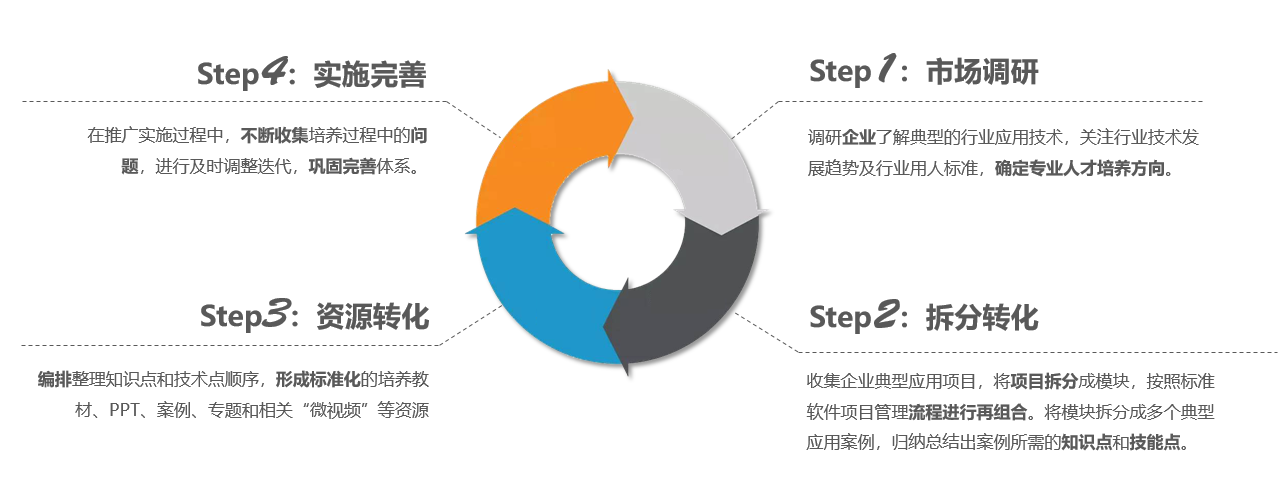


图3-3：课程体系闭环迭代研发流程

中软卓越ETC凭借雄厚的专业课程资源研发能力，将企业真实的项目及应用拆分成**知识域**、**知识点**和**技能点**，同时转化成**标准化**、**立体化**的培训教学资源。

# 四、实训体系五大优势

## 4.1 秉承“5R”先进的人才培养理念

**5R人才培养理念**是中软国际教育科技集团在业界首创及独创的专业技术人才培养理念，同时也是十余年的专业技术人才培养经验的结晶及核心指导原则。**“5R”**指的是**“5个真实”**，即**真实的企业环境**、**真实的项目**、**真实的项目经理**、**真实的压力**以及**真实的就业**。



图4-1：中软卓越5R专业人才培养理念

**真实的环境**：中软卓越培训中心提供完全真实的项目研发办公区域，标准的项目管理模式。环境对于参训人员很重要，环境和企业的管理模式能够让参训人员以一个全新的状态投身到开发工作中。

**真实的项目**：项目均来自中软国际7条核心业务线或行业典型领域应用，只有真实的、典型的项目才会有真实的功能需求和技术需求，才能够提升参训人员对于项目的真实理解。

**真实的压力**：在项目实训过程中，我们通过目标管理、时间管理、压力管理提升参训人员的整体职业素质以及高强度持续工作能力。在实训过程中我们加强阶段审查和考核的次数和频率，以便于提升参训人员抗压力和团队管理能力，突出目标管理意识。

**真实的项目经理**：在培训过程中，我们的培训师均具备3年的软件项目开发经验以及1.5年以上的项目管理经验。在培训过程中突显**实际场景还原**、**突发事件处理**，解决实际开发项目中出现的各种“1+1不等于2”的问题。

**真实的就业**：只要经历以上四个真实并达到要求和标准，我们相信参训人员基本具备为了作为一个软件及互联网行业职业人的基本标准，从而“水到渠成”，自然而然提升其就业竞争力。

**4.2 基于“CDIO”国际化工程实践的人才培养规范标准**

中软实训的不断提升参训学生的**构思**（Conceive）、**设计**（Design）、**实现**（Implement）和**运作**（Operate）的四个方面能力 ，以实训项目产品研发到产品运行的生命周期为载体 ,让学生以主动与实训项目之间通过有机联系的方式学习工程化理念。

CDIO培养大纲将工程类参训学生的能力分为工程基础知识、 个人能力、 人际团队能力和工程系统能力四个层面，实训最终结果要求以综合的培养方式使学生在这四个层面达到预定目标。

从实训的整体安排上可以看出，在实训过程中从初期的项目立项，到前期的项目原型设计和数据库设计，再到中期的项目业务整合及技术攻坚，再到最终的项目结项收尾。软件工程贯穿整个项目实训实施。

过程中的各种项目组织过程资产的收集以及项目活动的组织，也凸显出中软卓越培训中心在**企业级**实训的**工程化程度**、**标准化程度**和**专业化程度**。

## 4.3 独创的“Project-Driven”项目驱动的授课模式

项目驱动模式有别于传统高等教育的知识点驱动模式，企业级培训的核心目标是解决问题的能力及良好的职业素质，这与高等教育之间是一个良好的**“互补模式”**，同时也建立起高校、企业、学生三者之间的**“共赢模式”**。

项目驱动模式，强调问题的解决，以问题作为出发点，如何运用所学的知识能够解决问题，同时培养参训人员对技术点和技能点的综合应用能力。在实训过程中通过整体目标、分解目标及阶段里程碑的确定，提升参训人员的效率、提升参训人员的抗压力，不断提升参训人员灵活运用所学的各种专业知识及技术技能解决复杂综合实际问题的能力。



图4-2：中软卓越“Project-Driven”项目驱动专业人才培养模式

## 4.4 “做中学，学中做”的培训理念

任何一个行业应用领域所涉及的技术、知识都非常庞大。完全的掌握是依靠时间的积累和经验的总结。中软卓越培训中心依照“2-8”原则，强调实际动手能力和解决问题的能力的培养。因此在高校课堂教学的基础之上，在项目实训知识点讲授过程中“源于课程高于课程”，同时通过一系列的实际问题，让参训人员不断体验和了解并掌握所学知识的实际应用技巧，经过总结和归纳就最终变为自己的专业能力。所以在项目实训过程中“没有什么都能解决的程序员，只有在实践中不断学习成长的开发工程师”。

## 4.5 企业级标准和工程化的规范

标准和规范是一个大型软件企业的生命线，中软卓越实训体系的最主要的一个特色是严格按照CMMI5及软件工程标准对项目的全流程进行监控管理，严格按照软件项目管理“启动->规划->监控->实施->收尾”五大核心流程，不断提升参训人员**企业级**开发的**流程规范**、**技术规范**、**行业领域业务规范**、**编码规范**、**注释规范**以及**团队管理规范**。而这些规范恰恰是在校学生最缺乏的东西，而这又恰恰是企业最看重的能力。

与此同时，通过实训过程中的不断考核，让参训人员了解和掌握规范，在各种实训活动中，让参训人员体验学习规范。这样在今后就业过程中会更加有助于提高参训人员的就业竞争力。

# 五、实训总体设计方案

本方案根据南开大学软件工程专业的实际需求量身定制，结合前期的专业课程,以基于hadoop的电影行业分析可视化平台实战项目展开。通过课程实践提高学生对JavaEE技术的深入理解，了解并掌握基于Web的设计思想及高性能企业级架构实现，通过具体的业务流程了解分布式开发架构和基于Hadoop生态圈的大数据开发，HDFS的分布式文件存储，MapReduce的数据分析，相关数据模型的设计，以及SpringBoot等框架的整合等，各项技术在实际应用过程中的使用技巧和经验。同时强调Java语言Web开发领域的应用，包括但不限于HTML5、CSS、JavaScript、jQuery等应用开发技术，对于大数据的可视化技术进入相关研究，并根据项目情况给出合理的数据呈现方案，最终实现完整的企业级JavaEE+大数据全流程开发能力。

课程以实际动手为主题思想，提高学生发现问题解决问题的能力，同时提升参训人员对目前主流的应用程序设计以及高性能、高可靠性架构模式有更深层次的认知和理解，为今后的工作和学习打下坚实的理论基础和实际的动手能力。同时注重项目管理内容，让参训学员了解体验项目开发的全生命周期，实训以企业真实项目管理方式开展，项目开发+项目管理方式让学员完成实训后就具备进项目组工作的能力。

## 5.1 实训总体时间

|  |  |  |
| --- | --- | --- |
| 培训信息 | 具体内容 | 备注 |
| 培训主题 | 基于hadoop的电影行业分析 | 暂无 |
| 培训时长 | 5周 | 暂无 |
| 实施地点 | 校内实训基地 | 暂无 |

## 实训培养目标

1. 了解H5/CSS3/JS(ES6)/jQuery等大前端web开发技术；
2. 掌握JSP/Servlet/ajax/jQuery等web开发能力；
3. 了解分布式开发原理及远程调用RPC框架应用；
4. 了解Hadoop生态体系，掌握平台的集群环境搭建；
5. 掌握HDFS原理及应用，基于JavaAPI的开发；
6. 掌握MapReduce原理及应用，基于JavaAPI的开发；
7. 了解高性能高可靠性企业级程序的设计思路；
8. 掌握springboot等JavaEE应用程序的应用开发能力；
9. 提高学生的开发调试能力及系统架构思想；
10. 了解项目全生命周期，掌握项目各阶段提交物及标准。

# 六、实训技术服务内容

## 6.1 项目简介

项目1：《基于hadoop的电影行业分析》

**项目概述：**

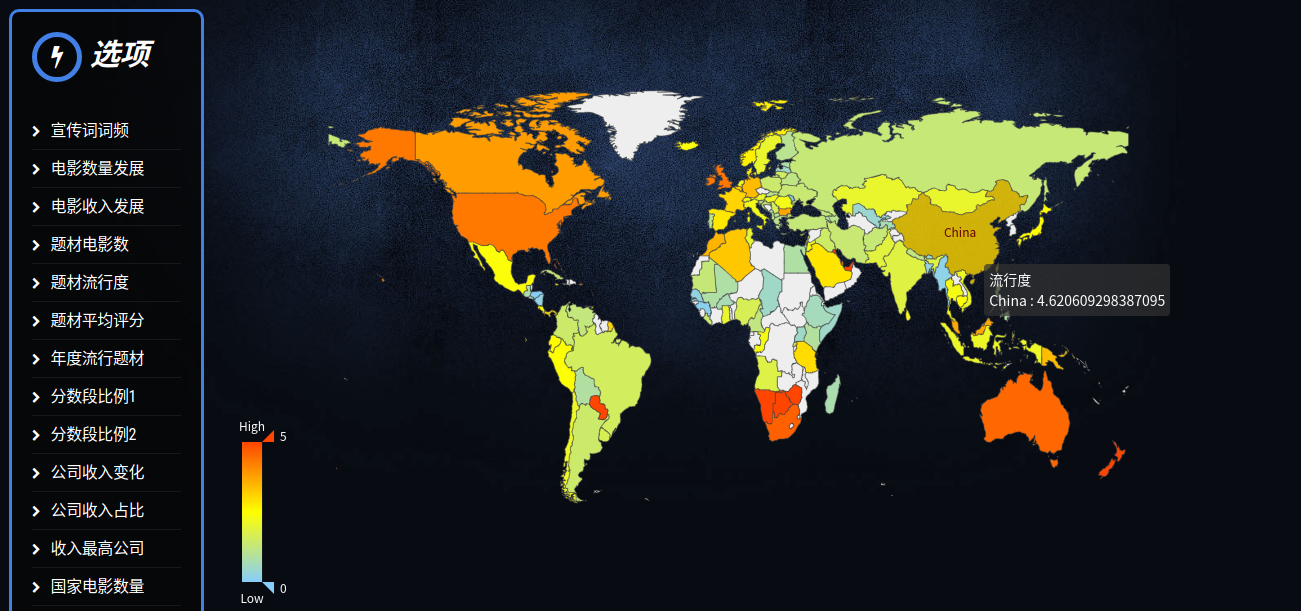
本系统是一个基于hadoop的大数据分析可视化项目，可以将收集来的电影信息，进行分析得到结果后，将结果已图表的形式进行可视化展示让用户更加直观看到大数据分析出来的结果。

具体功能：宣传词词云图，电影数量变化图，电影收入变化图，各年份各电影题材数量变化图，各题材流行度，各年份最流行题材top5，各年份不同分数段电影数据和占比图，各题材不同分数段电影数据和占比图，各公司收入随年份变化，等图表。

项目功能：

1. 宣传词中关键词的词频率。
2. 逐年全球各地的电影发行量的变化。
3. 逐渐全球发行电影题材的变化。
4. 逐年电影利润的变化。
5. 逐年电影题材的流行变化。
6. 各种题材的平均评分排行。
7. 各电影公司的逐年输入变化。
8. 分析影响电影收益的原因占比。
9. 根据往年数据预测未来电影的流行走向。
10. 根据往年数据预测未来各种电影题材的收益走向。

部分截图：

****

## 6.2实训实施计划

|  |  |  |  |
| --- | --- | --- | --- |
| 天次 | 项目实施阶段 | 阶段任务及提交 | 备注 |
| 第一天 | 项目立项、定义及设计 | 项目启动  项目开发讲解  项目组成立  确定项目开发小组  项目组汇报  提交物：《项目分组信息表》、《项目小组汇报演示稿》  项目讲解  大数据图形可视化项目开发现状  大数据图形可视化项目主流标准  大数据图形可视化项目应用领域和行业分析  各小组项目立项  立项汇报  提交物：《项目立项申请书》、《项目开题演示稿》  项目定义与设计  项目全生命周期的基本管理  项目阶段标准及提交  项目管理工具的使用（Project）  提交物：《WBS任务分解表》、《项目开发周期表》、《任务分配矩阵表》  需求分析  需求分析的意义和方法  需求分析工具的使用  需求分析报告的撰写  项目需求分析验收  提交物：《项目需求规格说明书》 |  |
| 第二天 | 基于大前端技术的静态原型实现技术 | 项目实现：  《基于hadoop的电影行业分析及可视化的图表静态原型图表的实现》  上午讲授涉及知识点：   1. echarts的介绍 2. echarts的基础用法 3. echarts的十大常用图表。 4. echarts的自定义图表。 |  |
| 第三天 | 基于大前端技术的静态原型实现技术 | 项目实现：  《基于hadoop的电影行业分析分析及可视化的图表静态原型图表的实现》  上午讲授涉及知识点：   1. Hcharts的介绍 2. Hcharts的基础语法。 3. Hcharts的常用图标。 4. .Hcharts的自定义图表。 |  |
| 第四天 | 企业级应用框架 | 项目实现：  《基于hadoop的电影行业分析分析及可视化的后台数据读取实现》  上午讲授涉及知识点：  1.SSM框架回顾串讲  2.springboot简介  3.springboot的第一个程序  4.springboot的环境搭建  5.springboot的配置 |  |
| 第五天 | 企业级应用框架 | 项目实现：  《基于hadoop的电影行业分析及可视化的后台数据读取实现》  上午讲授涉及知识点：  1.springboot restful风格 CRUD  2.Linux下安装jdk发布springboot项目  3.Springboot与数据库交互  4.结合数据库进行restful风格数据读取。 |  |
| 第六天 | Linux系统 | 项目实现：  《基于hadoop的电影行业分析及可视化的大数据平台的搭建》  上午讲授涉及知识点：  1. Linux入门概述  2. VM安装相关  3. Linux目录结构  4. VI/VIM编辑器 |  |
| 第七天 | 分布式处理架构 Hadoop | 项目实现：  《基于hadoop的电影行业分析及可视化的大数据平台的搭建》  上午讲授涉及知识点：  1.单机版 Hadoop 部署（安装、配置、验证）  2.伪分布式 Hadoop 部署（安装、配置、验证）  3.完全分布式 Hadoop 部署（安装、配置、验证） |  |
| 第八天 | 分布式处理架构 Hadoop | 项目实现：  《基于hadoop的电影行业分析及可视化的大数据平台的搭建》  上午讲授涉及知识点  1. Hadoop核心架构  2. Hadoop与高效能和网络计算区别  3. Hadoop发展现状  4. Hadoop运行模式详解  5. Hadoop目录结构 |  |
| 第九天 | 分布式处理架构 Hadoop | 项目实现：  《基于hadoop的电影行业分析及可视化的大数据平台的高可用及运维维护》  上午讲授涉及知识点：  1.联邦集群  2.Hadoop HA  3.故障的切换与规避  4.Hadoop 运行中的排错 |  |
| 第十天 | HDFS组件 | 项目实现：  《基于hadoop的电影行业分析及可视化的大数据平台的数据存储的实现》  上午讲授涉及知识点：  1.HDFS命令行接口（文件基本操作）  2.Java API（Configuration类、FileSystem抽象类、Path类、FSDataInputStream类、FSDataOutputStream类等）接口操作  3.基于 HDFS 的 I/O 操作  4.选讲：数据流，压缩  案例演示：  1.命令行操作数据集（豆瓣，IMDB电影等数据文件的操作）  2.客户端程序（基于 Java 语言）与 HDFS 数据流实现交互  实战演练：  1.使用 Java API 操作文本格式的豆瓣电影数据集文件  2.使用 Java 编程实现将爬取的数据清洗后保存到 HDFS 文件系统中 |  |
| 第十一天 | MapReduce组件 | 项目实现：  《基于hadoop的电影行业分析及可视化的大数据平台的数据分析的实现》  上午讲授涉及知识点：  1.MapReduce简介  2.在 Hadoop 集群环境中分析数据  3.理解map、reduce计算模型  4.MapReduce的工作流程  5.Yarn运行模型的剖析和调度  案例演示：  1.经典案例：WordCount 1.0 版本  实战：  1.借鉴 WordCount 1.0 程序框架实现电影数据集当中的词频统计分析  2.编程实现电影评分的平均分、最高分、最低分统计 |  |
| 第十二天 | MapReduce组件 | 项目实现：  《基于hadoop的电影行业分析及可视化的大数据平台的数据分析的实现》  上午讲授涉及知识点：  1.MapReduce 优化策略总结  2.MapReduce 和 JavaBean 交互  3.MapReduce 和 DB 交互  案例演示：  1.使用 MapReduce 读取和分析 MySQL 数据表中数据  2.使用 MapReduce 分析提取含有关联关系的数据集文件的特性  3.实用MapReduce实现k-means算法。 |  |
| 第十三天 | Hive组件 | 项目实现  《基于hadoop的电影行业分析及可视化的大数据平台的数据分析》  上午讲授涉及知识点：  1.数据类型  2.表、表分区（静态分区、动态分区）、分桶概念及操作  3.Hive客户端种类以及选用  4.使用 CLI 完成数据连接  5.使用 CLI 进行表操作  6.使用 CLI 完成数据导入  7.使用 CLI 完成数据查询  8.使用 Java JDBC 实现对 Hive 数据集的 CRUD  案例演示:  1.电影数据集分区、分桶的实现  2.使用 HQL操作电影数据集  实战：  1.使用 HQL 操作 电影数据集 |  |
| 第十四天 | sqoop组件 | 项目实现  《基于hadoop的电影行业分析及可视化的大数据平台的数据的导入导出》  上午讲授涉及知识点：  1.sqoop的安装与使用  4.Sqoop导入导出  5.使用Sqoop讲数据导入到mysql当中供可视化系统提取进行可视化。 |  |
| 第十五天 | ZooKeeper组件 | 项目实现：  《基于hadoop的电影行业分析及可视化的大数据平台的高可用模式搭建》  上午讲授涉及知识点：  1.理解分布式系统构建的基础架构  2.ZooKeeper体系结构  3.ZooKeeper集群安装  4.ZooKeeper的数据模型  5.ZooKeeper的常用操作  6.ZooKeeper命令行操作(ZK服务命令、ZK客户端命令、查询命令等)  7.借助Java API使用ZooKeeper  案例演示：  1.ZooKeeper的安装和使用  2.ZK服务的启动、查看、停止、重启  3.借助ZK客户端（命令行）完成根目录显示、文件的创建、获取|修改文件内容、删除文件等  4.使用常见的四字命令: stat、ruok、dump、kill、conf、cons、wchs、wchc、wchp等  5.使用Java API完成ZK的节点操作 |  |
| 第十六-十七天 | Spark组件 | 项目实现：  《基于hadoop的电影行业分析及可视化的大数据平台的数据分析》  上午讲授涉及知识点：  1.Spark设计与运行原理  2.Spark简介与定义  3.Spark运行架构  4.Spark的部署模式  5.Spark应用程序开发  6.RDD的设计与运行原理  7.RDD编程实践  8.RDD键值对实现  9.编写Spark应用程序  10.编写RDD应用程序  11. Spark的安装与使用 |  |
| 第十八—十九天 | Spark组件 | 项目实现：  《基于hadoop的电影行业分析及可视化的大数据平台的数据分析》  上午讲授涉及知识点：  1. 使用Java编程开发疫情数据病例数据的TOP N获取  2. 基于Java编程开发实现疫情数据的二次排序  3. 基于Java编程完成分组Top N的实现  4. 使用RDD完成一维数组的MapReduce(使用Scala演示)  5. Shark和Spark SQL  6. Spark SQL原理与实践  7. Spark SQL简介  8. RDD与DataFrame |  |
|  |
| 第二十天 | Spark组件 | 项目实现：  《基于hadoop的电影行业分析及可视化的大数据平台的数据分析》  上午讲授涉及知识点：  1.数据读取和保存(parquet方式、JDBC方式)  2. Parquet方式查询购物者年龄大于指定数值的购物者信息  3.Spark SQL借助JDBC方式读取传统数据库数据放入DataFram  实战:  1.启动Spark集群  2.计算各省疫情病例总数  3.计算所有每月各省疫情病例增加量  4.统计所有各省疫情病例的top十名  5.使用JDBC读入MySQL数据库的疫情数据表放入DataFrame |  |
| 第二十一天 | Flink组件 | 项目实现：  《基于hadoop的电影行业分析平台的大数据平台的数据分析》  涉及内容：  1.什么是flink  2.Flink 的重要特点  3.什么是有界流和无界流  4.flink和sparkStreaming 对比  5. spark DAG 如何划分stage  6. flink 运行的组件  7.Flink的安装及配置。 | 第四十三天 |
| 第二十二~~二十三天 | Flink组件 | 项目实现：  《基于hadoop的电影行业分析平台的大数据平台的数据分析》  涉及内容：  1. jobmanager 的作用  2. jobmanager接收到的应用程序包含  3. taskManager 的作用  4. flink的资源管理器  5. Dispatcher的作用  6. 任务提交流程 | 第四十四天 |
| 第二十四~~二十五天 | Flink组件 | 项目实现：  《基于hadoop的电影行业分析平台的大数据平台的数据分析》  涉及内容：  1. 任务提交  2. Source源  3. Transformation  4. Sink  5. Window  6. Window Function  7. Flink Window State  8. Flink Dataset  9. Flink SQL与案例 | 第四十五天 |
| 第二十六~三十天 | 项目研发提交及验收 | 1. 各小组研发项目 2. 项目收尾   项目整合及调优  项目测试  代码检查  生产环境配置  项目打包及发布   1. 项目验收   各小组项目汇报  优秀项目展示  提交：《项目最终提交件综合包》   1. 实训总结   各小组可选题目：  智慧城市交通流量分析  哔哩哔哩弹幕流量分析  网易云音乐数据分析  短视频推荐平台 |  |

## 6.3技术讲座

中软国际会根据实训进度，邀请基地高级工程师、项目经理或合作企业高级技术主管，为学生举行行业技术讲座和职业规划讲座。

# 七、配套资源

## 8.1 师资配比

为了更好地实施和保障项目实训的顺利进行以及各阶段的顺利展开，我们在项目实施过程中将配备以下角色和职位的人员参与：

|  |  |  |
| --- | --- | --- |
| 角色 | 人数 | 备注说明 |
| 项目经理（PM） | 1 | 根据实际情况确定，基本配比1：40 |
| 项目经理助理（PMA） | 1 | 根据实际情况确定，基本配比1：80 |
| 项目助理（PA） | 若干 >2 | 根据实际情况确定，基本配比1：60 |
| 组织协调（PO） | 1 | 根据实际情况确定 |
| 注：以上根据项目具体安排和实施人数最终进行确定和调整。 | | |

## 8.2 实训环境

* 硬件环境

|  |  |  |
| --- | --- | --- |
| 设备 | 要求 | 备注 |
| PC机 | 传统PC机即可 |  |
| CPU | 64位支持硬件虚拟化双核4线程以上CPU |  |
| 内存 | 8G以上物理内存 |  |
| 硬盘 | 500G以上硬盘 |  |
| 网络 | 千M网卡 |  |

* 软件环境

|  |  |  |  |
| --- | --- | --- | --- |
| 开发工具 | 工具名称 | 版本号 | 备注 |
| 开发机操作系统要求 | Win7 64位系统或Linux 64位发行版本 | | 暂无 |
| Java开发环境 | JDK | 8 | 暂无 |
| 中间件服务器 | Apache Tomcat | 9 | 暂无 |
| 数据库 | MySQL | 5.1 | 暂无 |
| IDE开发平台 | Eclipse | eclipse-jee-2019-03 | 暂无 |
| 大数据支撑平台 | Apache Hadoop 2.6以上 | | 暂无 |
| 服务发布环境 | Linux | Centos7 | 暂无 |
| 虚拟机 | VirtualBox | 15.1.0 | 暂无 |
| 开发框架 | Spring/MyBatis | 4.2/3.4 | 暂无 |
| 第三方插件 | mysql-connector | 5.1.47 | 暂无 |

联系人：赵秀杰 联系电话：13011371386 公司官网：www.chinasofti.com