



湖南石油化工职业技术学院  
Hunan Petrochemical Vocational Technology College

# 课程标准

课程名称： 大数据综合实训

课程代码： 31091740

适用专业： 大数据技术与应用

制订时间： 2021年3月

湖南石油化工职业技术学院

## 目录

1 课程概述.....	3
1.1 课程的性质.....	3
1.2 课程定位.....	3
1.3 课程设计思路.....	3
2. 课程基本目标.....	3
2.1 知识目标.....	4
2.2 能力目标.....	4
2.3 职业素养目标.....	4
3、课程教学内容及学时安排.....	4
3.1 课程主要内容说明.....	4
3.2 课程教学内容.....	5
4 教学实施建议.....	6
4.1 教学组织建议.....	6
4.2 教学评价建议.....	7
4.3 教材选用.....	8
4.4 课程主讲教师 and 教学团队要求说明.....	8
4.5 课程思政要求.....	8
4.6 课程教学环境和条件要求.....	9
4.7 教学资源的开发与利用.....	10
4.8 其它.....	10

# 《大数据综合实训》课程标准

课程名称：大数据综合实训

课程代码：31091740

总学时数：24 学时（理论课学时数：0； 实践课学时数：24）

适用专业：大数据技术与应用专业

## 1 课程概述

### 1.1 课程的性质

《大数据综合实训》是一门专业实训课，大数据技术专业课程，为学生搭建起通向“大数据知识空间”的桥梁和纽带，以“构建知识体系、阐明基本原理、引导初级实践、了解相关应用”为原则，为学生在大数据领域“深耕细作”奠定基础、指明方向。

课程将系统讲授大数据的基本概念、大数据处理架构 Hadoop、分布式文件系统 HDFS、分布式数据库 HBase、NoSQL 数据库、云数据库、分布式并行编程模型 MapReduce、流计算、图计算、数据可视化以及大数据在互联网、生物医学和物流等各个领域的应用。在 Hadoop、HDFS、HBase 和 MapReduce 等重要章节，安排了企业级的实践操作，让学生更好地学习和掌握大数据关键技术。

### 1.2 课程定位

通过本课程的学习，使学生学会搭建 Hadoop 完全分布式集群，掌握 HDFS 的原理和基础操作，掌握 MapReduce 原理架构、MapReduce 程序的编写。为将来从事大数据挖掘研究工作以及后续课程的学习奠定基础。

### 1.3 课程设计思路

课程设计围绕大数据开发的基础和方法，考虑到实际企业生产环境中的应用场景，课程中采用实际案例，从案例出发、启动、实施、验证各环节剖析讲解，与真实环境同步，本课程通过大量的案例，引导学生思考问题、分析问题，学生在理解知识点的过程中遇到困难时，可以继续往后学习。可以通过对各学习情景中学习目标、主要内容、授课方式、师生要求等各项内容的描述，来规范课程所要求的内容。真正做到在项目实施过程中各个环节的考察和评价，来完成对课程的评鉴与考核。

## 2. 课程基本目标

## 2.1 知识目标

- (1) 了解分布式文件系统的基本概念、结构和设计需求，掌握 Hadoop 的概念
- (2) 了解布式数据库 HBase 的访问接口、数据模型、实现原理和运行机制
- (3) 了解 NoSQL 数据库与传统的关系数据库的差异、NoSQL 数据库的四大类型以及 NoSQL 数据库的三大基石
- (4) 了解云数据库的概念、基本原理和代表性产品的使用方法
- (5) 掌握大数据相关技术 HDFS、MapReduce
- (6) 掌握 Hadoop 集群及数据批处理
- (7) 掌握 Hive 数据库操作
- (8) 掌握大数据知识体系的系统架构

## 2.2 能力目标

- (1) 能够了解分布式数据库 HBase 的访问接口、数据模型、实现原理和运行机制，并熟练掌握 HBase 的使用方法
- (2) 能够了解 NoSQL 数据库与传统的关系数据库的差异、NoSQL 数据库的四大类型以及 NoSQL 数据库的三大基石
- (3) 基本掌握 Redis、MongoDB 等 NoSQL 数据库的使用方法
- (4) 能够掌握大数据的基本概念
- (5) 能够掌握相关的数据管理、存储、分析计算等的技术基础
- (6) 能够掌握 Hadoop 的相关知识
- (6) 通过对数据库的学习和编程设计，掌握 HBase 的使用方法

## 2.3 职业素养目标

- (1) 培养学生数据分析意识
- (2) 培养学生谦虚、好学的品质
- (3) 培养学生良好的自我表现、与人沟通的能力
- (4) 培养学生的团队精神和服务意识
- (5) 培养学生的自主学习能力

## 3、课程教学内容及学时安排

### 3.1 课程主要内容说明

本课程重点是学习介绍大数据的基本概念和应用领域，并阐述大数据、云计算和物联网的相互关系、介绍大数据处理架构 Hadoop、分布式文件系统 HDFS 的基本原理和使用方法、分布式数据库 HBase 的基本原理和使用方法、NoSQL 数据库的概念和基本原理、云数据库的概念和基本原理、分布式并行编程模型 MapReduce 原理和使用方法、Hadoop 架构再探讨。本课程共 72 学时，其中理论部分占 32 学时，实践部分占 32 学时、和复习考试占学时 8 学时。

### 3.2 课程教学内容

课程模块	课程主题	主要内容	课时
模块一	大数据概述了解大数据发展历史以及未来发展趋势	<ul style="list-style-type: none"> <li>◆ 传统大规模系统存在的问题</li> <li>◆ 了解大数据相关概念</li> </ul>	6
模块二	大数据处理框架 Spark	<ul style="list-style-type: none"> <li>◆ 掌握大数据系统架构掌握 Spark 技术</li> <li>◆ 掌握 Spark 发展历程、版本演变、生态系统</li> <li>◆ 掌握大数据系统架构掌握 Spark 的安装和使用方法</li> </ul>	4
模块三	Hadoop 分布式架构	<ul style="list-style-type: none"> <li>◆ Hadoop HDFS 基本结构</li> <li>◆ 掌握 HDFS 简介、相关概念、体系结构、存储原理、读写过程</li> </ul>	4
模块四	数据库握 NoSQL 数据库展、四大类型、三大基石	<ul style="list-style-type: none"> <li>◆ NoSQL 数据库的四大类型、三大基石、NewSQL 数据库</li> </ul>	4
模块五	Spark 集群规划	<ul style="list-style-type: none"> <li>◆ 集群和网络拓扑要求</li> <li>◆ 集群软件的端口配置</li> </ul>	4
模块六	集成 Hadoop 到现有 workflow 及 Hadoop API 深入探讨	<ul style="list-style-type: none"> <li>◆ 存储系统</li> <li>◆ 利用 Sqoop 从关系型数据库系统中导入数据到 Hadoop</li> <li>◆ 利用 Flume 导入实时数据到 Hadoop</li> <li>◆ ToolRunner 介绍、使用 MRUnit 进行测试</li> <li>◆ 使用 Configure 和 Close 方法来进行 Map/Reduce 设置和关闭</li> </ul>	4

模块七	使用 Hive 和 Pig 开发及技巧	<ul style="list-style-type: none"> <li>◆ Hive 和 Pig 基础</li> <li>◆ Hive 的作用和原理说明</li> <li>◆ Hadoop 仓库和传统数据仓库的协作关系</li> <li>◆ Hadoop/Hive 仓库数据数据流</li> <li>◆ Hive 部署和安装</li> <li>◆ Hive Cli 的基本用法</li> <li>◆ HQL 基本语法</li> <li>◆ 运用 Pig 过滤用户数据</li> </ul>	6
模块八	Hbase 安装和使用	<ul style="list-style-type: none"> <li>◆ Hbase 安装部署</li> <li>◆ Hbase 原理和结构</li> <li>◆ Hbase 运维和管理</li> </ul>	4
模块九	Spark 的存储	<ul style="list-style-type: none"> <li>◆ Spark 的存储级别;Cache 介绍;Spark 的容错原理</li> </ul>	4
模块十	Spark 案例	<ul style="list-style-type: none"> <li>◆ Lineage 容错;Checkpoint 容错;RDD 的创建</li> <li>◆ 案例—统计单词的个数</li> </ul>	4
模块十一	RedHadoop 企业版本	<ul style="list-style-type: none"> <li>◆ 运用 RedHadoop 快速构建服务集群</li> <li>◆ 运用 RedHadoop DW 构建数据仓库</li> </ul>	4
模块十二	Spark 原理	<ul style="list-style-type: none"> <li>◆ Spark 原理; Spark 的架构图; Spark 运行模式介绍</li> <li>◆ —local; —standalone; —messos; —yarn; Spark 的 RDD 什么是 RDD; RDD 的种类; —Tranformation; —Action</li> </ul>	8
模块十五	复习	<ul style="list-style-type: none"> <li>◆ 大数据相关理论知识的复习与编程实践的复习</li> </ul>	4
模块十六	考试	<ul style="list-style-type: none"> <li>◆ 教师出卷笔试</li> </ul>	2
<b>合计</b>			72

## 4 教学实施建议

### 4.1 教学组织建议

- 建议选择运行稳定教学软件版本教学;
- 教师要在开展教学前应有一个已准备好的项目为教学依据, 围绕着开展教学;
- 项目教学要尽量采用分组教学法;
- 要注重学生自主编程的培养, 避免养成抄写代码的习惯;

- 适当组织些活动，化解抽象枯燥的教学；
- 使用通俗易懂的例证，方便学生简单直观理解。

## 4.2 教学评价建议

### 4.2.1 课程内容评价要点

序号	单元（模块）	考核标准	权重比例%
1	课堂学习	包括出勤、课堂表现及课堂积极回答问题等	15
2	课后作业	是否按时、按质、按量完成教师布置的课后练习	15
3	课堂实训	能否实操出课堂练习	20
4	期末考试	由教师评定的笔试成绩	50

### 4.2.2 课程评价方法和内容

评价类型	评价方法	评价内容
职业素养 (10%)	过程性评价 (10%)	到课考勤，学习及工作态度、安全意识、质量观念、合作精神、敬业精神等纳入职业素养考核，在具体考核指标中体现。
理论知识 (50%)	过程性评价 (20%)	主要是课堂提问、平时作业、单元测验、期中测验等。
	终结性评价 (30%)	主要是期末考试，评价综合专业理论知识掌握和运用能力，由计算机随机命题或人工命题组成标准试卷，尽量与国家临床医学检验技师职称资格考试接轨。
职业技能 (40%)	过程性评价 (20%)	实训报告、实际操作过程评价。

终结性评价 (20%)	建议考核核心技能项目 参照技能考核标准与要求，编制核心技能项目的评分标准，评分标准应涵盖操作规范性、结果准确性、人文关怀、沟通交流、操作安全等。
----------------	---

### 4.3 教材选用

#### 1、教材选用建议

为了让学生掌握职业岗位工作所需的技术知识，顺利实施职业技能训练，授课承担部门应选用近几年出版的全国优秀的高职规划教材，并且采用项目驱动式的编写思路为宜。

#### 2、教材编写建议

为了使教材适合高职教育以及现代技术发展快、创新多的特点，突出强调理论教学与实践操作紧密结合的一体化教学模式，自编教材应以“项目导向，任务驱动”为主线。

### 4.4 课程主讲教师和教学团队要求说明

本课程要求任课教师首先牢固树立中国特色社会主义理想信念，践行社会主义核心价值观，自觉增强立德树人、教书育人的荣誉感和责任感，学为人师，行为世范。最好由具有双师型素质的高学历的教程承担。要求教师具有扎实的专业知识和丰富的相关行业实际工作经验，具有一定职业教学教学能力，能够开展课程教学改革和科学研究。

### 4.5 课程思政要求

全面推进课程思政建设，发挥好专业课程的育人作用。专业课程教学过程以专业知识和技能为载体，加强思想政治教育，充分发挥课堂主渠道功能，努力发掘课程中立德树人的要素，与思想政治理论课同向同行，形成协同效应。本专业课程思政具体要求如下。

#### 1、课程教学与爱国主义教育相结合

通过选择优秀典型的行业企业案例、视频题材等重要思政教育内容，激发爱国热情，培养家国情怀。在专业教师引导之下，通过我国 IT 行业和大数据技术应用发展成就和实力的展示，开展爱国主义教育、中国梦教育，增强学生的国家

认同感与民族自豪感。

## 2、课程教学与团队合作精神相结合

专业核心课程实训教学过程中，以实训任务为载体，以工作小组为单元，引导学生将企业本职工作经历融入学习过程，调动学习积极性，重点强调项目成员团队合作的原动力和凝聚力，树立了正确的集体观，培养团队合作精神。

## 3、课程教学与职业素养培养相结合

通过实践教学环节和企业经历，结合企业生产实际和行业人才素养需求，引入企业对优秀员工必备素质和基本规范的要求，引导学生自觉实践相关行业的职业精神和职业规范，增强职业责任感，培养学生良好的职业品德、职业纪律及职业责任心，教育学生爱岗敬业、讲究诚信、精益求精，在潜移默化中提高了学生未来岗位的适应能力。

## 4、课程教学与高职学生学情相结合

高职院校学生普遍基础薄弱、学习主动性不强，在这样的学情下，课程教学中教师应实时自我反思和自我总结，不断完善教学手段，增强学生的学习兴趣，提升学生的信心，提高学生的专业能力。

## 5、课程教学与实际项目案例相结合

教学中，引入实际企业或公司案例，通过理论课程教学结合实际项目案例的教学模式，引导学生提升自我意识、养成良好的职业精神和职业规范，在实际项目案例中不断总结自己、提升自我、提升团队作战意识和团队协作能力。

### 4.6 课程教学环境和条件要求

主要能够满足正常的课程教学、实习实训所需的专业教室、实训室。

#### 1、专业教室基本条件

配备交互智能教育平板、黑（白板）、多媒体计算机、投影设备、音响设备，互联网接入或 WIFI 环境，并具有网络安全防护措施。安装应急照明装置并保持良好状态，符合紧急疏散要求、标志明显、保持逃生通道畅通无阻。

序号	教学场地	设施配置	功能
1	投影室	投影仪、相关软件等	公共课程教学
2	多媒体机房	电脑、投影仪、相关软件等	专业课理实一体化教学

## 4.7 教学资源开发与利用

### 1、常规教学文件

常规教学文件应包括：授课计划、教案、讲稿、教学课件等资料。

### 2、教学资源

应建立适合教师教学的《教学案例库》和适合学生自主学习的《导学手册》和《习题集》。

### 3、网络资源

有条件的情况下，可以建立本课程相关学习网站，放置教师的授课视频和学习手册、在线测试等教学资源，方便学生在线学习、下载资源、学习讨论。

## 4.8 其它