

# 计算机类专业（大数据方向）人才培养方案

## 一、招生对象及学习年限

(一) **招生对象:** 全日制普通中学高中毕业生, 招生方式为普通高考招生。

(二) **学习年限:** 基本学制三年, 实行弹性学制, 学生在校时间原则上不能少于两年, 总在校时间(含休学)不得超过六年。

## 二、培养目标

本专业培养面向中国特色社会主义建设, 对接广州及珠三角地区战略性主导产业和战略性新兴产业中大数据或数据挖掘的工程重点领域的人才需求, 具有良好的职业道德和职业精神, 能从事计算机软硬件产品及大数据或数据挖掘的工程性开发与实现、在计算机与互联网企业中从事系统集成或售后服务、数据处理与分析、在政府部门或企事业单位从事信息系统的建设、管理、运行、维护的技术工作, 具备“一技之长+综合素质”的德、智、体、美等方面全面发展的高素质和技术应用性人才。

## 三、就业岗位与就业范围

就业岗位	就业范围	主要业务工作
大数据研发工程师	互联网、金融、IT、制造业、零售企业	大数据产品建设与开发、大数据项目需求分析、设计、业务建模。
大数据算法工程师	互联网、金融、IT、制造业、零售企业	大数据产品建设与开发、大数据项目需求分析、设计、业务建模。
大数据产品工程师	互联网、金融、IT、制造业、零售企业	大数据产品建设与开发、大数据项目需求分析、设计、业务建模。
大数据售前工程师	互联网、金融、IT、制造业、零售企业	数据集成工作的开发、测试与调优、大数据产品测试, 测试报告编写。
大数据运维工程师	互联网、金融、IT、制造业、零售企业	大数据平台搭建, 维护, 调优, 管理, 监控。
数据挖掘工程师	互联网、金融、IT、制造业、零售企业	常规数据报告的制定与信息挖掘、根据公司战略需要进行数据建模。
数据分析师	互联网、金融、IT、制造业、零售企业	数据采集及数据处理工作、对数据进行整理规划, 编写数据说明文档、明确客户方的业务体系。

## 四、人才培养规格

### (一) 综合素质

1. **思想政治素质:** 掌握马克思主义科学的世界观、人生观和价值观。有坚定跟着共产党走中国特色社会主义道路的信心和决心, 有热爱祖国、服务人民的理想信念。具有社会责任感, 能积极践行社会主义核心价值观, 拥有能够支撑职业和人生发展的思想政治素质。

2. **职业素质:** 具有良好的职业态度和职业道德修养, 具有正确的择业观和创业观。坚持职业

操守，爱岗敬业、诚实守信、办事公道、服务群众、奉献社会；具备从事职业活动所必需的基本能力和管理素质；脚踏实地、严谨求实、勇于创新。

3. 人文素养与科学素质：具有融合传统文化精华、当代中西文化潮流的宽阔视野；文理交融的科学思维能力和科学精神；具有健康、高雅、勤勉的生活工作情趣；具有适应社会核心价值体系的审美立场和方法能力；奠定个性鲜明、善于合作的个人成长成才的素质基础。

4. 身心素质：具有一定的体育运动和生理卫生知识，养成良好的锻炼身体、讲究卫生的习惯，掌握一定的运动技能，达到国家规定的体育健康标准；具有坚韧不拔的毅力、积极乐观的态度、良好的人际关系、健全的人格品质。

## （二）职业能力

本专业主要学习常用数据挖掘编程语言（R 语言和 Python）、数据库应用技术、Java 程序设计、Linux 操作系统、Hadoop 大数据存储与运算、Hadoop 大数据存储与运算、大数据查询与处理、Spark 大数据快速运算、Oozie 大数据工作流，具备大数据开发与数据挖掘的实际工作能力，具有创新意识及进一步发展专业技能的良好基础。

本专业毕业生应获得以下几方面的知识和技能：

1. 具有比较扎实的计算机基本理论与技术、数据挖掘的常用算法；
2. 掌握当前使用最广泛的两门数据挖掘编程语言 R 语言和 python；
3. 熟悉 SQL 的计算与存储过程调优，并具有严密的逻辑分析能力；
4. 具备数据的处理、抽取、清洗、转换等能力；
5. 掌握主流的 Hadoop 处理技术，包括 MapReduce、Hive、Hbase 等；
6. 逻辑思维能力强，具备较强的文档编写和良好的沟通表达能力。

## （三）职业拓展能力

1. 对大数据基础架构和平台有深刻理解；
2. 熟悉 Hadoop 集群构建，能进行相应的部署及配置；
3. 熟悉主流应用服务器的架构体系以及各种中间件技术。

## 五、毕业标准

学生按专业人才培养方案要求修完规定的课程，考核合格，达到毕业最低总学分和《国家学生体质健康标准》相关要求，获得本专业要求的证书，准予毕业，颁发毕业证书。

### （一）学分要求

本专业按学年学分制安排课程，学生最低要求修满总学分 124 学分。

必修课要求修满 94 学分，占总学分的 75.81%。

其中：基本素质课要求修满 21 学分，占总学分的 16.94%；

职业能力课要求修满 73 学分，占总学分的 58.87%。

选修课要求修满 30 学分，占总学分的 24.19%。

其中：基本素质课要求修满 20 学分，占总学分的 16.13%；

职业能力课要求修满 10 学分，占总学分的 8.06%。

## （二）证书要求

### 1. 获得以下英语证书之一

- (1) 广东省英语教学指导委员会颁发的高级职业英语证书；
- (2) 全国大学英语四、六级考试委员会颁发的全国大学生英语四级或六级考试证书；

### 2. 获得以下专业或职业资格认证证书之一

- (1) 广东省嵌入式软件公共技术中心组织的嵌入式 Linux 软件开发工程师职业资格认证证书；
- (2) 国家劳动和社会保障部职业技能鉴定中心和广东省 LINUX 公共服务技术支持中心的“全国计算机高新技术 Linux 系统管理模块管理员级”职业资格认证证书；
- (3) 通过国家人事部“计算机技术与软件专业技术资格（水平）考试”中计算机程序员初级资格考试，获取相应的“程序员”技术资格证书；

## 六、职业能力核心课程

### 1. Hadoop 大数据存储与运算

课程能力目标：了解 Hadoop 的架构、原理、Hadoop 集群配置及安装（JDK、SSH），熟悉 Hadoop IDE 开发环境配置（Eclipse 配置）和 Hadoop Java API 编程实例、具备 Hadoop 编程开发能力。

课程主要内容：Hadoop 简介、架构、原理、Hadoop 集群配置及安装（JDK、SSH）、Hadoop IDE 开发环境配置（Eclipse 配置）、Hadoop Java API 编程实例、Hadoop Java API 编程实例、Hadoop 命令（hadoop fs , hadoop job, Hadoop jar）、Hadoop 基础编程（Mapper、Reducer、Driver）、Hadoop 高级编程（FileInputFormat、Combiner、Partitioner、FileOutputFormat）、Hadoop 案例—基于 KNN 的鸢尾花类别预测（算法实现及 Hadoop 编程）、Hadoop 实例—基于 KMeans 的客户价值分析（算法实现及 Hadoop 编程）。

### 2. Hbase 大数据快速读写

课程能力目标：熟悉 HBase 集群安装配置、掌握 HBase 架构与数据模型、实现基于 HBase 的冠字号查询系统。

课程主要内容：HBase 简介、HBase 集群安装配置（Zookeeper 简介、Zookeeper 安装配置、配置文件解读）、HBase 架构与数据模型（组件原理功能、Rowkey 设计、Schema 设计）、HBase 命令行（hbase shell）、HBase 开发环境配置、Java API 操作 HBase 删/建表、HBase 增删改查、HBase 与 Hadoop Mapreduce 交互（HBase to HDFS, HDFS to HBase, HBase to HBase）、基于 HBase 的冠字号查询系统（案例背景、架构、数据处理、案例实现）。

### 3. 大数据查询与处理

课程能力目标：熟悉 Pig 架构与原理、Pig Latin 运行环境、Pig 数据类型，Hive 与 Pig 区别，掌握 Pig 安装、配置，熟悉 Pig Latin 数据加载/输出、数据转换、数据存储。

课程主要内容：Pig 简介、架构与原理、Pig Latin 运行环境，Pig 数据类型、Hive 与 Pig 区别、Pig 安装、配置、Pig 运行模式、Pig 基本指令，Pig Latin 数据加载/输出、数据转换、数据存储，Pig 内置函数、自定义函数（UDF），使用 Hcatalog、Hive 与 Pig 交互，Pig 案例——SOGO 案例。

#### 4. Spark 大数据快速运算

课程能力目标：掌握 Spark 安装配置，了解其原理与架构，掌握 Spark 常用编程技术，实现基于 Spark ALS 的电影推荐系统。

课程主要内容：Spark 简介（Spark 简介、应用场景）、Spark 安装配置（如何兼容 Hadoop、HBase 集群）、Spark 原理与架构（组件功能、RDD 原理）、Spark 编程（Scala 及编程简介、Spark Transformation/Action 编程）、Spark IntelliJ IDEA 开发环境配置、基于 Spark ALS 电影推荐系统（案例背景、系统架构、Spark Shell 实现、系统业务逻辑实现）。

#### 5. Oozie 大数据工作流

课程能力目标：了解 Oozie 工作原理，熟悉 Oozie 环境配置及页面监控、Oozie Workflow 配置。

课程主要内容：Oozie 工作原理、Oozie 环境配置及页面监控、Oozie Workflow 配置（Hadoop MR 工作流、Hive 工作流、Pig 工作流、Spark 工作流）、定时任务配置。

#### 6. 顶岗实习

课程能力目标：经过顶岗实习，进一步巩固学生所学的数据库应用技术、数据挖掘编程语言、Hadoop 大数据存储与运算、Hbase 大数据快速读写、大数据查询与处理、Spark 大数据快速运算、Oozie 大数据工作流等学科知识，提高学生的实际工作能力和技能，为今后的工作打下基础。

课程主要内容：结合所学的数据库应用技术、数据挖掘编程语言、Hadoop 大数据存储与运算、Hbase 大数据快速读写、大数据查询与处理、Spark 大数据快速运算、Oozie 大数据工作流等学科知识进行顶岗实习，实习课程的主要内容包括：常用数据挖掘算法的实际应用，数据平台的运营管理和安全维护，保障数据的存档、保密工作；搭建基于 Hadoop/Spark/Shark 的大数据平台，大数据采集、清洗、整合等。

### 七、实践教学条件

#### （一）校内实训室（中心、基地）

##### 1. 软件技术工程中心 1

规模：可供 105 名学生（员）实训教学（技能培训、资格考证），理论实训一体化教学。

功能：软件技术专业教学、软件产品（系统）开发及软件工程实施环境与条件。

主要设备：

序号	设备名称	基本配置	单位	基本数量	适用范围：专业课程教学，（职业资格培训、认证）
1	浪潮服务器集群	CPU 2.8GHZ 内存 1G 73G*2 RAID	台	4	网络运行支撑环境（操作系统 WIN/LINUX 平台、数据库平台、中间件平台）、软件设计、开发与测试平台
2	TCL PC 机	内存 512M	台	105	专业课程教学、软件系统设计、开发与测试

##### 2. 软件技术工程中心 2

规模：可供 128 名学生（员）实训教学（技能培训、资格考证）、理论实训一体化教学。

功能：软件技术专业教学、软件开发和工程实施环境与条件

主要设备:

序号	设备名称	基本配置	单位	基本数量	适用范围: 专业课程教学, (职业资格培训、认证)
1	联想万全 R525 服务器集群	双核系统 CPU 3.0G 16M/RAID 系统/2GRAM	台	4	网络运行支撑环境 (操作系统 WIN/LINUX 平台、数据库平台、中间件平台)、软件设计、开发与测试平台
2	联想品牌台式计算机	P4 3.0/1G/80G/15	台	128	专业课程教学、软件系统设计、开发与测试
3	手机	Android、Symbian、Windows Phone	台	20	移动互联网应用开发

## (二) 校外实训基地

1. 广州泰迪智能科技有限公司

规模: 30-80 名学生

功能: 顶岗实习、项目合作、咨询服务、师资培训

## 八、课程设置与教学安排 (附表 1)

## 九、各类课程学时学分比例表 (附表 2)

## 十、部分专业课程教材推荐

大数据工程和数据挖掘作为新兴产业, 如何实现教学与企业需求相吻合具有重要意义, 广州泰迪智能科技有限公司作为深耕大数据及数据挖掘行业多年的企业, 在大数据及数据挖掘方面拥有丰富的实际项目经验和独到的行业见解。同时, 经过与各高校多年的深入合作, 广州泰迪也更清楚高校育人与企业用人如何有效对接, 故对相关专业课程的教材进行推荐, 教材的绝大部分内容也是源于企业实际项目, 更具有实践意义。

推荐的每本教材皆有配套的 PPT、源代码及原始数据, 教师教学能更加轻松顺畅, 学生也更能体会到实际企业项目的过程, 提升教学质量。

部分专业课程教材推荐表

课程编号	课程名称	学时数	推荐教材	章节
01050056Z	Python 语言基础	72	Python 与数据挖掘 / 张良均等著. 一北京: 机械工业出版社	第 1 章~第 5 章
01050056Z	R 语言基础	64	R 语言与数据挖掘 / 张良均等著. 一北京: 机械工业出版社	第 1 章~第 5 章
01050056Z	数据挖掘基础算法 (Python 或 R 语言)	54	R 语言与数据挖掘 / 张良均等著. 一北京: 机械工业出版社 Python 与数据挖掘 / 张良均等著. 一北	第 6 章~第 10 章

			京：机械工业出版社	
01050383Z	Hadoop 大数据存储与运算	54	Hadoop 与大数据挖掘 / 张良均等著. — 北京：机械工业出版社	第 1 章~第 2 章
01050383Z	Hbase 大数据快速读写	54	Hadoop 与大数据挖掘 / 张良均等著. — 北京：机械工业出版社	第 3 章
01050385Z	大数据查询与处理	54	Hadoop 与大数据挖掘 / 张良均等著. — 北京：机械工业出版社	第 4 章~第 5 章
01050386Z	Spark 大数据快速运算	54	Hadoop 与大数据挖掘 / 张良均等著. — 北京：机械工业出版社	第 6 章
01050387Z	Oozie 大数据 workflow	54	Hadoop 与大数据挖掘 / 张良均等著. — 北京：机械工业出版社	第 7 章~第 8 章