《大数据分布式应用》教学大纲

课程编号：2121473B

课程类型：□通识教育必修课 □通识教育选修课

□学科基础课 □专业核心课

☑专业提升课 □专业拓展课

总 学 时：48 讲课学时：32 实验（上机）学时：16

学　　分：3

考试类型：□考试 ☑考查

适用对象：信息管理与信息系统专业

□是 ☑否 适合作为其他专业学生的个性化选修课

先修课程：《面向对象程序设计》 《计算机网络技术与应用》

一、教学目标

本课程是信息管理与信息系统专业（商务智能）的专业提升课，是大数据技术理论与应用的基础课程；通过本课程的学习学生将掌握大数据分布式计算的基本理论和应用方法实，了解并行计算、分布式计算、云计算和大数据计算等大数据技术及其应用；探索如何应用大数据技术挖掘数据中所蕴含的潜在价值。

目标1：掌握并行计算的基本理论和关键技术；

目标2：掌握分布式计算的基本理论和关键技术；

目标3：掌握云计算的基本理论和关键技术；

目标4：掌握大数据计算的基本理论和关键技术；

目标5：践行课程思政教学。

二、教学内容及其与毕业要求的对应关系

（一）教学内容

本课程主要教学内容是在学生掌握基本的面向对象程序设计和计算机网络技术的基础上，学习大数据分布式的相关理论和技术。具体内容包括并行计算、分布式计算、云计算和大数据计算及其应用等。其中细讲和精讲的内容包括并行计算关键技术及应用、分布式计算关键技术及应用、大数据计算关键技术及应用，粗讲内容包括大数据概述和云计算。并行计算是分布式应用的基础，分布式计算是分布式应用的核心，大数据计算是大数据时代分布式应用的核心，故这三部分内容为本课程的重点。

（二）教学方法和手段

根据教学目标，结合教学内容，拟采用的教学方法有：理论讲解、案例分析、计算机实验、课后作业等；通过课堂提问和课后预留作业的方式对所学知识进行温习和巩固；通过实例化的编程实验增强学生对所学知识的掌握和理解。

（三）实践教学环节要求

根据教学进度和要求布置相应的课后作业，结合课题内容设计上机实验环节。课后作业需要学生结合课堂内容查阅文献来完成，旨在扩宽学生的知识面，强化学生对课堂内容的理解；上机实验在教师启发指导下由学生独立完成，旨在强化学生对理论知识的理解。

（四）学习要求

为有效学习本课程，要求学生首先具备面向对象程序设计和计算机网络技术 等方面的基本知识。应该熟读课程大纲，提纲挈领地掌握大数据概述、并行计算、分布式计算、云计算和大数据计算的相关理论、关键技术和应用，随后按照大纲熟读教材，并通过课后思考和上机实践进行多角度和多层次的反复学习。

（五）与毕业要求的关系

如何有效地使学生掌握大数据分布式应用中关键技术的基本理论、方法和原理，能够结合实际的领域数据进行分布式应用，并根据分布式应用的结果为管理决策服务，发挥大数据时代数据在商务智能运营与决策中的作用。

（六）教学中应注意的问题

由于面向对象程序设计和计算机网络与应用是该课程的先修基础，如果学生先修基础课没有学好，学习大数据分布式相关理论和技术时，则出现半知不解的情况。因此，教学中需要根据学生掌握先修课程基础情况，由易到难循序渐进、结合实际案例进行由浅入深的教学。

三、各教学环节学时分配

以表格方式表现各章节的学时分配，表格如下：

**教学课时分配**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 序号 | 章节内容 | 讲课 | 实验 | 其他 | 合计 |
| 1 | 第一章大数据概述 | 2 | 1 | 0 | 3 |
| **2** | **第二章并行计算与应用** | **8** | **4** | **0** | **12** |
| **3** | **第三章分布式计算与应用** | **6** | **3** | **0** | **9** |
| **4** | **第四章云计算与应用** | **4** | **2** | **0** | **6** |
| **5** | **第五章大数据计算与应用** | **12** | **6** | **0** | **18** |
| **合计** |  | **32** | **16** | **0** | **48** |

四、教学内容

**第1章 大数据概述**

教学内容包括： 大数据产生的技术、数据产生方式、大数据的概念、大数据的特性、大数据技术概述和大数据应用；

教学重点：本章教学的重点就是让学生了解大数据的起源、特性和应用。

课程思政切入点：以国家大数据中心为案例，介绍大数据在国内的应用情况以及大数据在精准扶贫攻坚中作用，让学生真正体会到大数据的价值，激发学生的学习兴趣，激发学生致力于国家大数据建设的热情。

课程的考核要求：了解大数据的起源和应用领域；理解大数据的概念和大数据技术；掌握大数据的特性；

复习思考题：（1）大数据的概念和特性；（2）大数据技术；（3）大数据的应用领域。

**第2章 并行计算与应用**

教学内容包括：并行计算概念、并行计算分类及应用、并行计算应用模型和方法、基于多进程的并行应用、基于多线程的并行应用、基于消息机制的并行应用等；

课程思政切入点：以国家并行计算工程技术研究中心为案例，使学生了解国家并行计算中心的发展例程，以及国内并行计算行业科学家，激发学生向前辈学习，激发学生学习前沿技术、报效祖国。

课程的考核要求：（1）并行计算分类及应用、基于多进程的并行应用、基于多线程的并行应用和基于消息机制的并行应用。

复习思考题：（1）并行计算的分类；（2）并行计算应用模型和方法；（3）并行应用中安全机制。

**第3章 分布式计算与应用**

教学内容包括：分布式计算概念、分布式计算分类及应用、分布式计算应用模型和方法、基于socket的分布式应用、基于webservice的分布式应用。

课程思政切入点：以中国分布式计算总站为案例，介绍分布式计算的意义和价值，激发学生立足于自身知识，积极投身教育公益事业，进行技术科普。

课程的考核要求：（1）分布式计算分类及应用、基于socket的分布式应用、基于webservice的分布式应用。

复习思考题：（1）分布式计算的分类；（2）基于socket的分布式应用流程；（3）基于webservice的分布式应用。

**第4章 云计算与应用**

教学内容包括：云计算概念、云计算分类及应用、云计算关键技术、云计算应用案例分析等；

课程思政切入点：以中国科学院云计算中心和华为云计算中心为案例，使学生了解国内云计算的技术水平，激发学生投身高精尖核心技术的研发中，为接近卡脖子技术而努力。

课程的考核要求：（1）云计算分类及应用和云计算关键技术。

复习思考题：（1）云计算的分类；（2）云计算的关键技术。

**第5章 大数据计算与应用**

教学内容包括：大数据计算概念、大数据信息采集技术、大数据信息存储技术、大数据分析技术、大数据挖掘技术和大数据可视化及其应用等；

课程思政切入点：以百度无人驾驶、google无人驾驶和特斯拉无人驾驶系统为案例，通过市场、技术和应用前景多个维度的对比，让学生深刻体会到高精尖技术的重要性，激发学生学习热情和兴趣。

课程的考核要求：（1）大数据信息分布式共享存储体系的组成与实现、大数据自动聚类与实现、基于机器学习的大数据挖掘技术及应用；高维数据的可视化及其应用。

复习思考题：（1）大数据存储模式；（2）大数据分析关键技术；（3）机器学习在不同模态数据中的应用。

五、考核方式、成绩评定

本课程的考核分为平时考核及期末考查两种形式。本课程平时成绩占40%，期末考查成绩占60%。

平时考核采用撰写上机实验、实验报告、课堂案例讨论等方式。期末考核采用案例论文分析的方式，案例内容将尽可能包括《大数据分布式应用》所讲授的重点内容，使学生通过案例将理论知识和实践结合起来，为将来学生从事相关大数据行业的工作奠定基础。

六、主要参考书及其他内容

[1] Hermann Kopetz著，吴际，龙翔，尚利红译.分布式实时系统原理与设计.北京：机械工业出版社,2019年1月

[2] 林伟伟,刘波.分布式计算、云计算与大数据.北京：机械工业出版社.2015年12月

执笔人：曹海青 教研室主任：胡磊　　　　系教学主任审核签名：