

浙江大学研究生课程教学大纲

课程编号	3413175	开课院系	海洋学院		
中文课程名称	数据挖掘技术			授课语言	中文
英文课程名称	Data Mining				
课程性质	专业学位课	课程类别	博士生课	课程体系	通用课程
任课教师姓名	马东方	工号	0012767	职称	讲师（高校）
学历		E-mail	mdf2004@zj. ed. cn	联系电话	15657123809
辅讲教师1姓名		工号		职称	
学历		E-mail		联系电话	
教学学时	32	实验学时	0	实践学时	0
其他学时	0	总学时	32	自学学时	0
学分数	2	考核方式	课程论文	开课学期	冬
课程内容中文简介	数据挖掘是一门新兴的交叉型学科，是在信息技术方向速度兴起的一门新技术，在很多重要领域都发挥着积极作用。本课程主要介绍数据挖掘的基本概念、内涵、方法和技术，具体包括：多源异构大数据的基本特征、数据的预处理方法、分类预测方法、关联挖掘方法、聚类分析方法、时间序列数据挖掘方法等内容，重点介绍各类数据挖掘方法的基本原理、使用条件、使用过程及优缺点。在强调课程基本知识的基础上，本课程也将讲授数据挖掘方面的前沿探索和当前所面临的挑战，并结合学生的研究方向有针对性的设置课程设计题目。通过本门课程的学习，学生应能全面了解数据挖掘的专门知识，具备解决实际问题的能力，可以胜任海洋大数据或其他大数据的知识挖掘与提取工作。				
课程内容英文简介	Data mining has become a new interdisciplinary subject with rise of information technology now, and it is playing an important role in many fields. This course mainly introduces the basic concepts, methods and techniques of data mining, including the basic features of multi-source heterogeneous data, data preprocessing method, classification forecasting method, association mining method, clustering analysis method, time series data mining method and so on. We will lay special stress on analyzing the basic principle, the use condition, the use guideline and the advantages and disadvantages of various data mining methods. On the basis of emphasizing the basic knowledge, this course will also discuss the frontier exploration and the challenge of data mining, and design individualized topic for each student considering their research directions. By studying this course, students should be able to fully understand the special knowledge of data mining, and have the ability to solve practical problems. They should be competent to mining the marine or other big data and extract valuable information.				
预备知识要求	数据挖掘技术需要学生具有扎实的数据结构、概率论与数理统计、高等数学等方面的基础知识。人工智能等相关课程对本课程学习有很大帮助，但并非必须。				
教学目标	本课程面向海洋信息科学与工程、海洋电子工程或者其他信息类专业研究生所开设的一门专业选修课。针对“数据挖掘技术”课程内容实践性、应用性强的特点，以“理论指导实践、实践强化理论”为指导，通过讲授+练习+讨论的形式开展多层面的学习，使学生融会贯通，深入了解和认知问题。课程目标为：通过本门课程的学习旨在让学生了解数据挖掘的总体概貌，掌握数据挖掘的基本内涵、原理、算法及其他相关技术，培养学生熟练应用数据挖掘技术来解决实际问题的能力，准确把握数据挖掘的前沿动态、追踪学术热点，为今后从事的科学研究工作或高级应用开发任务打下坚实基础。本课程要求学生广泛查阅相关资料，及时完成课堂作业，认真总结并思考课程难点，创新性的完成报告设计。				
参考文献					
	书名	著者	出版社	出版年份	

参考书目	数据挖掘教程		Richard J. Roiger, Michael W. Geatz	清华大学出版社	2003
	数据挖掘概念与技术		Jiawei Han、Micheline Kamber	机械工业出版社	2007
教学日历	周次	教学内容（包括课堂讲授、实验、讨论、考试等）			
	1	数据挖掘基础知识：大数据的基本特征、数据挖掘内涵及功能、数据挖掘系统的应用、数据挖掘的热点问题；数据仓库与联机事务处理：数据集市、数据仓库的基本内涵			
	2	数据仓库与联机事务处理：数据仓库实现方法、数据仓库系统结构、OLAP的典型操作等；数据预处理：数据清理、数据集成与变换			
	3	数据预处理：数据规约、离散化与概念分层；挖掘频繁模式、关联和相关：基本概念及路线图、频繁项集挖掘			
	4	挖掘频繁模式、关联和相关：关联规则挖掘、关联挖掘与相关分析；分类与预测：基础知识、基于决策树的分类、贝叶斯分类			
	5	分类与预测：神经网络分类、线性与多变量回归、非线性预测、其他回归方法等；翻转课堂			
	6	聚类分析：聚类分析中的数据类型、聚类方法分类、聚类方法			
	7	聚类分析：层次方法、密度聚类方法、基于约束的聚类分析			
	8	课程总结，讲授典型案例，结合个人研究方向的拟定课程设计大作业题目，独立完成任务。（讨论+课程论文）			
申请理由	培养方案调整				
涉及培养方案调整情况 （在所涉培养类型下打“√”）	学科/专业学位类别（领域）名称及代码	年级	硕士	博士	直博生
学科/专业学位类别（领域）意见	负责人签名：_____年____月____日				
院系意见	主管院长（系主任）签名（盖院系章）：_____年____月____日				