

中国海洋大学本科生课程大纲

课程名称	机器学习 Machine Learning	课程代码	075302101220
课程属性	专业知识	课时/学分	64/3
课程性质	选修	实践学时	32
责任教师	冯彦	课外学时	64(32×2)

课程属性：公共基础/通识教育/学科基础/专业知识/工作技能，**课程性质：**必修、选修

一、课程介绍

1. 课程描述：

机器学习是研究计算机如何模拟或实现人类的学习行为，以获得新知识和新技能，重新组织已有的知识结构，不断完善自身系统性能的学科。目前，机器学习已经成为智能数据分析和处理的重要手段，广泛地应用在网络数据分析处理、计算机视觉、自然语言处理、生物信息学等领域。本课程作为机器学习的入门课程，详细阐述机器学习的基本理论、模型和算法，并结合上机实践，进行实际问题的数据分析和处理。

2. 设计思路：

本课程针对数学类专业的本科生开设，讲授部分本着从基础入手、注重解决问题的原则而展开，介绍经典的机器学习理论、模型及方法，其中有监督学习包括回归模型、决策树、神经网络、支持向量机，贝叶斯分类器等，无监督学习包括聚类方法等。

上机实践部分针对各种机器学习方法，选取典型的实际问题、并有效利用在线平台的课程资源进行实验，使学生从数据的展示、算法的实现到数据处理结果分析等方面有直观形象的认识，更好地理解 and 掌握机器学习理论、模型及其在实际问题中的应用。

3. 课程与其他课程的关系：

先修课程：数学分析，高等代数，概率论，数理统计，数学实验基础

并行课程：数据挖掘

后置课程：软件工程，并行计算基础

二、课程目标

使学生掌握机器学习的基本理论、模型及方法，理解有监督、无监督学习方法在实际问题中的应用，并能针对实际问题采用合适的模型、方法利用计算机进行求解，对数据处理的结果给出合理的分析，提供合理的决策依据，进一步提高数学理论分析和综合应用的能力，为今后从事机器学习、数据处理等领域的工作和研究打下良好的基础。

三、学习要求

要完成所有的课程任务，学生必须：

- (1) 按时上课，认真听讲，积极参与课堂讨论；
- (2) 多思考多实践，对学习中的难点、重点问题主动和老师、同学交流；
- (3) 认真完成上机作业，遵守学术诚信规范。

四、进度安排

序号	主题	计划课时	主要内容概述	实验实践内容
1	绪论	2+0	机器学习的发展和应用 基本概念和术语 模型的评估方法	
2	线性模型——回归	4+4	线性回归模型 过拟合与欠拟合 收缩方法 多项式回归	作业 1: 线性回归实验
3	线性模型——分类	2+2	Logistic 回归模型 线性判别分析	作业 2: Logistic 回归实验
4	决策树	4+4	决策树基本流程 划分选择 剪枝处理	决策树实验练习
5	神经网络	4+4	感知机与多层网络 误差逆传播算法 全局最小与局部极小	作业 3: 神经网络实验
6	支持向量机	4+4	间隔与支持向量 对偶问题 核函数	作业 4: 支持向量机实验
7	贝叶斯分类器	2+0	贝叶斯决策论 朴素贝叶斯分类器	
8	集成学习	4+4	Boosting Bagging 与随机森林	Boosting 实验练习
9	聚类	4+4	聚类的性能度量 距离计算	作业 5: K 均值聚类实验

			原型聚类	
10	项目设计及答辩	2+6	项目设计答辩	项目设计实践

五、参考教材与主要参考书

1. 参考教材（不强制购买）

周志华.《机器学习》. 清华大学出版社, 2016

2. 主要参考书

[1] 李航.《统计学习方法》. 清华大学出版社, 2012

[2] Mitchell T.M.《机器学习》. 机械工业出版社, 2008

[3] Trevor Hastie, Robert Tibshirani, Jerome Friedman.《统计学习基础——数据挖掘、推理与预测》. 译者: 范明, 柴玉梅, 咎红英. 电子工业出版社, 2005

[4] Christopher Bishop.《Pattern Recognition and Machine Learning》. Springer, 2006

六、成绩评定

(一) 考核方式 C: A. 闭卷考试 B. 开卷考试 C. 论文 D. 考查 E. 其他

(二) 成绩综合评分体系:

成绩综合评分体系	比例%
1. 平时表现、课下作业	20
2. 上机作业	30
3. 项目设计论文	50
总计	100

附：评分标准及说明

1) 平时表现、课下作业的评分标准:

评分标准	得分
------	----

1. 按时上课，保证全勤；学习态度认真主动，对学习中的难点和热点问题积极和老师、同学交流讨论；认真完成课下作业。	80-100 分
2. 基本能够保证出勤；学习主动性一般，对学习中的难点和热点问题基本能够和老师、同学交流讨论；基本完成课下作业。	30-79 分
3. 课堂出勤不足三分之一；学习态度不认真，主动性较差，对学习中的难点和热点问题不和老师、同学交流讨论；未完成课下作业。	0-29 分

2) 上机作业的评分标准:

评分标准	得分
1. 按时上交作业，严格按照要求实现全部或绝大部分功能，求解思路清晰，程序运行结果正确，代码符合规范，可读性强。	80-100 分
2. 按时上交作业，能够按照要求实现基本功能，求解思路较为清晰，程序运行结果存在部分错误，代码基本符合规范，可读性一般。	50-79 分
3. 按时上交作业，未能按照要求实现基本功能，程序无法运行，代码不符合规范，可读性差。	20-49 分
4. 未按时上交作业；或不遵守学术诚信规范，作业抄袭。	0 分

3) 项目设计论文的评分说明:

评价项目	主要评价内容	满分
书面报告质量	结构是否严谨，逻辑性强； 文字表达是否准确、流畅，条理清楚，重点突出； 格式、图、表是否规范； 参考文献的引用和书写是否规范；	30 分
项目功能完成情况	是否完成项目设计要求中指定的各项基本功能； 是否对项目设计中的难点和热点问题给出详细的分析和解决步骤； 是否对项目设计中存在的问题、项目的扩展和改进做必要的说明； 是否具有一定的学术水平或实际价值；	30 分
程序代码质量	代码组织是否良好，逻辑性强； 是否有必要的文档说明，确保程序的可读性； 程序代码是否符合规范； 引用他人的代码是否有明确的说明；	20 分

答辩表现	项目汇报是否思路清晰，重点突出，详略得当； 回答问题是否准确，有一定的专业性； 仪表、精神风貌是否良好； 报告时间是否符合要求；	20分
------	---	-----

七、学术诚信

学习成果不能造假，如考试作弊、盗取他人学习成果、一份报告用于不同的课程等，均属造假行为。他人的想法、说法和意见如不注明出处按盗用论处。本课程如有发现上述不良行为，将按学校有关规定取消本课程的学习成绩。

八、大纲审核

教学院长：

院学术委员会签章：