

大数据与机器学习

[关闭](#)

01510243,,大数据与机器学习,Big Data and Machine Intelligence

总学分: 3,总学时: 96,周学时:48/48/64

开课单位: 基础工业训练中心,课程类别: 全校性选修课,课程特色: 文化素质课 实践课 通识选修课,课程面向: 本科生

考核方式: 考查,教学方式: 实践

课程团队: 李超 郭敏 章屹松 马晓东 高英 周晋 杨建新 李双寿

课程内容简介:机器学习是指智能机器和智能的软件依赖算法对观测到的数据进行分析 (reason) 来做出有用的预测 (prediction) 和决策 (decision)。(机器学习)系统依赖机器学习、人工智能,组合了计算,数据、建模和算法。课程学习机器学习系统的基本原理和操作,包括:机器学习基本概念,深度学习基本概念,大模型基础原理与前沿。课程进行相应动手实践,包括机器学习scikit-learn实践、深度学习框架TensorFlow2实践。课程学习的前沿探索模块包括基于大模型LLM的机器学习,包括:现代深度学习架构,Transformer,LLM前沿研究进展,如高效率推断等。课程选题具有一定挑战性与前沿研究性。

Course Description: Intelligent machines and intelligent software rely on algorithms that can reason about observed data to make predictions or decisions that are useful. Such systems rely on machine learning and artificial intelligence, combining computation, data, models, and algorithms. This course targets to learn the fundamentals of machine intelligence systems, including: Basic concepts of machine learning, Basic concepts of deep learning, The basic principles and frontiers of LLMs. This course also provide hands on practice, including scikit-learn practice, TensorFlow2 Practice and course project. The cutting-edge exploration module of the course study includes machine intelligence based on the Large Language Model (LLM), including: modern deep learning architectures, the Transformer, and cutting-edge research advances in LLM, such as high-efficiency inference. The topics selected for the course are challenging and cutting-edge research.

先修要求:大二及以上,有高数、概率论和线性代数基础,以及有一门编程语言基础。

适用院系专业:全校

教学目标:对实践技能而言,学习Python, Scikit-learn, TensorFlow2和SQL的基本知识。对未来发展而言,了解数据管理与机器智能的方法论与实践状况,为未来发展,做好产业认识、知识准备和必要的实践技能储备

预期学习成效:基于Python语言的学习,学习机器学习Scikit-learn工具,学习TensorFlow2/Keras2机器学习库,进行数据处理,数据分析,训练机器学习模型,学习并部署应用模型。

参考书:参考书: 奥雷利安·杰龙(Aurélien Géron), 机器学习实战: 基于Scikit-Learn、Keras和TensorFlow(原书第2版), 机械工业出版社, 2020-10. Aurélien Géron, Hands-on Machine Learning with Scikit-Learn, Keras and TensorFlow, O'Reilly book, 2019.