

2025–2026 学年短学期
课程综合实践 II: 数据要素交易基础

Lec 00: 课程简介

吴一航 yhwu_is@zju.edu.cn

浙江大学计算机科学与技术学院



声明

- 课程中使用的许多图片和资源都是从互联网、其他课程材料上收集的，仅用于教育目的，版权属于原始所有者。
- 相关课程
 - UChicago CMSC 35401: Topics in Machine Learning: The Interplay of Economics and Machine Learning (Winter 2024)
 - Harvard CS 236r (old CS 286r): Topics at the Interface between Computer Science and Economics: Value of Information and Data
 - Stanford CS 320: Value of Data and AI

数据是二十一世纪的新石油

Data is the new oil.

— British data scientist Clive Humby, 2006

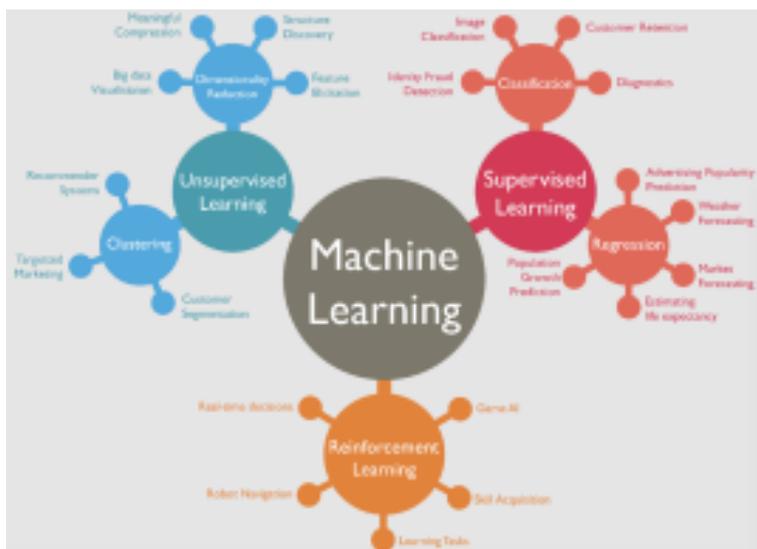


图 1 机器学习离不开数据



图 2 广告推荐与拍卖离不开数据

数据是二十一世纪的新石油



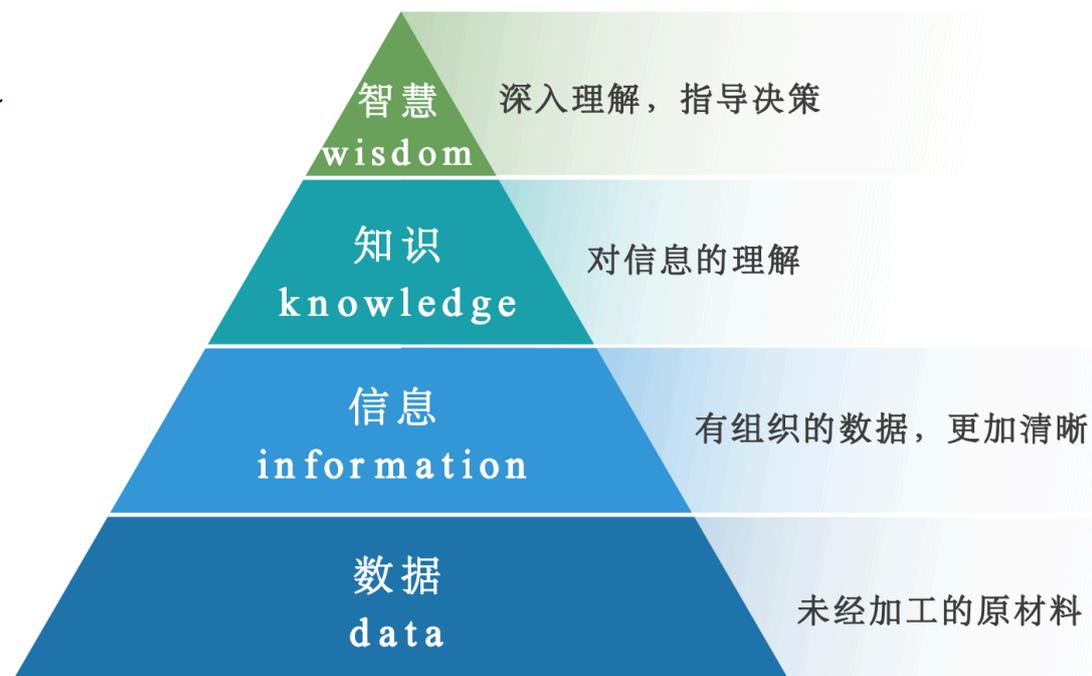
数据是二十一世纪的新石油



大数据时代的挑战

挑战一：数据的收集与整理 (Data Acquisition)

- 石油在使用前也需要提炼加工
- 数据的聚合有利于打破数据孤岛
- 处理不善，容易形成“数据沼泽”
- 从数据到智慧的流程如图



大数据时代的挑战

挑战二：如何保障数据的安全与合规（Data Security and Compliance）

- 数据泄漏风险大、成本高
- 数据滥用需要规范



大数据时代的挑战

挑战三：如何将数据流动到最需要它的人手中？

- 数据市场（Data Market）
- 数据市场的基本架构？数据市场有哪些基本要求与设计目标？

挑战四：如何实施数据定价（Data Pricing）

- 数据定价与传统定价的区别？有什么新的要求？
- 如何设计适合于数据的定价机制？

挑战五：如何衡量数据的价值（Data Valuation）

- 合作博弈中的概念：Shapley 值
- 如何解决 Shapley 值的计算困难性？
- Data Shapley 如何应用于提升机器学习模型准确率？

课程安排

时间	主题	备注
6.25	课程简介，数据要素市场概述	
6.26	数据合规与安全	
6.27	非合作博弈论基础（一）	
6.28	非合作博弈论基础（二）	
6.29	合作博弈与数据估值	
6.30	多臂老虎机算法基础与应用	
7.1		休息
7.2	拍卖与机制设计基础	
7.3	最优机制，拍卖与数据定价	
7.4	数据版本化定价	
7.5	贝叶斯劝说	
7.6	信息定价	特邀讲座
7.7	隐私，外部性与数据定价	课堂练习

课程目标

- 了解数据要素市场这一新兴交叉学科的前沿知识
 - 涉及微观经济学、博弈论与机制设计、数据库系统、机器学习、理论计算机、数据安全.....
 - 仅有的博弈论相关课程，少有的理论课程，“无用之用”
- 掌握数据要素市场这一新兴交叉学科各方向的研究基础
 - Data acquisition, data security, data pricing, data valuation.....
 - 判断自己是否对相关方向感兴趣或适合于相关方向研究
 - (我相信劝退效果是远大于劝进效果的)
 - (但有时候知道自己不适合什么也很重要，不是吗:))
- 培养基本的文献阅读能力，综合运用目前所掌握的思维、编程能力
 - 课程综合实践的意义所在

课程基本信息

- 上课时间：6.25-7.7（除 7.1 外），每天 9:30-12:00
 - 下午课程自行完成作业/实验等，本教室仍然可以使用
- 上课地点：紫金港北 3-211
- 教师：刘金飞 (jinfeiliu@zju.edu.cn)
- 助教：吴一航 (yhwu_is@zju.edu.cn)
- 课程教材：《数据要素市场：以数据交易为核心》（暂未出版，预计 2025 年末出版）
- 课程网站
 - 学在浙大平台：<https://courses.zju.edu.cn>，用于提交作业实验等
 - https://yhwu-is.github.io/Teach/ec/data_market/2025/，发布课件、作业、实验指导、大作业指南等，往年资料也可查看

参考书籍

本课程涉及领域较新，且为交叉领域，暂无系统性全面覆盖的书籍（除正在施工中的自编教材）。

如果对计算经济学感兴趣，以下是可供参考的教材：

综合性教材

- *Economics and Computation*, David C. Parkes Sven Seuken
 - 可能暂未出版

微观经济学

- 《微观经济学：现代观点》 [美] 哈尔·范里安
 - 入门教材
- 《微观经济理论》 [美] 安德鲁·马斯-克莱尔，迈克尔·D·温顿，杰里·R·格林
 - 简称 MWG，微观经济学圣经级教材，当然难度也较大

拍卖理论

- 《拍卖理论》 [美] 维佳·克里斯纳

博弈论

- 入门级别
 - 《博弈论基础》 [美] 吉本斯
 - 《博弈入门》 [美] 马丁·J·奥斯本
- 拓展级别：
 - 《博弈论》 [以] 迈克尔·马希勒，埃隆·索兰，什穆埃尔·扎米尔
 - 《博弈论：矛盾冲突分析》 [美] 罗杰·B·迈尔森
 - 《博弈论》 [美] 弗登伯格，[法] 让·梯若尔

算法博弈论

- 《斯坦福算法博弈论二十讲》 [美] 蒂姆·拉夫加登
 - 尝试在科普和专业之间寻找平衡，但效果一般
- *Algorithmic Game Theory*, Noam Nisan, Tim Roughgarden, Eva Tardos, Vijay V. Vazirani
 - 专著类型，非常不适合入门

评分标准

模块	理论	实践
课堂参与	课堂练习 (10%)	课堂互动 (5%)
平时练习	作业 ($2 * 10 = 20\%$)	实验 ($2 * 10 = 20\%$)
综合评价	期末练习 (20%)	大作业 (25%)

课堂参与

- 课堂练习：随机课堂 quiz，考察当堂课所学内容，难度较低
- 课堂互动：主要是博弈论游戏，参与即可获得分数

平时练习

- 作业：2 次理论作业，每次 10 分，题量和难度适中
- 实验：2 次代码实验，每次 10 分，代码填空，无需写报告，难度较低

综合评价

- 期末练习：开卷考，允许联网搜索（但不允许使用生成式人工智能），难度适中，和上课、作业内容强相关
- 大作业：允许 1-3 人组队，分为实践和理论两类，具体参考大作业要求

评分标准

其他要求:

- 平时练习允许讨论，但不允许抄袭
 - 如果认定抄袭，则当次作业所有参与抄袭的同学都记 0 分；
- 期末练习不允许讨论，独立作答；
 - 如果认定作弊，则根据《浙江大学学生违纪处理办法》处理；
- 关于生成式人工智能：
 - 平时作业、实验等可以使用，但需要对生成的内容负责，建议独立思考为主，生成式人工智能仅作为辅助工具；
 - 期末练习不允许使用，因为目前大模型能力较强，即使完全没有学习也可能可以使用大模型获得很多分数，而期末练习希望考察同学课程期间自己的学习成果。

上述分数可能会根据实际情况进行调整，总而言之请选课的同学不要太担心成绩的问题，而是将重点放在理解课程内容上，相信这样可以同时收获满意的课程体验和结果。