

Market Research and Analysis

Syllabus and References

- Zhenyu Zhao
- Nankai Institute of International Economics
- Feishu: 2120253538
- Email: zzynankai@outlook.com
- Website: xishanyu2.github.io

Course Contents

Lecture 1: Big Data Analysis in AI Era

Lecture 2: Introduction to Machine Learning

Lecture 3: Natural Language Process and Text Analysis

Lecture 4: Large Language Model Annotation

Lecture 5: Double Machine Learning for Causal Inference

Textbooks

- [1]周志华著.机器学习.清华大学出版社.2016
- [2]李航著.统计学习方法.第2版.清华大学出版社.2019
- [3]陈强编著.机器学习及Python应用.高等教育出版社.2021
- [4]郭峰等编著.机器学习与社会科学应用.上海财经大学出版社.2024
- [5]刘征驰编著.机器学习与经济大数据分析:基于Python实现.北京大学出版社.2024
- [6]姚加权, 洪永淼编著.Python经济大数据分析.高等教育出版社.2024
- [7]林建浩, 张一帆, 陈良源编著.经济金融文本分析.高等教育出版社.2025
- [8] Hastie T, Tibshirani R, Friedman J. The elements of statistical learning[J]. 2009.
- [9] James G. An introduction to statistical learning with applications in Python[J]. 2023.
- [10] [Visualize-ML \(Iris Series 鸢尾花书\) · GitHub](#)
- [11] [Machine Learning for Economics](#)

Readings for Lecture 1

- [1]洪永淼,汪寿阳.大数据如何改变经济学研究范式? [J].管理世界,2021,37(10):40-55+72+56.
- [2]王永进.数字时代的经济学研究[J].新经济,2023,(12):5-19.
- [3]Varian, H. R., "Big Data: New Tricks for Econometrics", Journal of Economic Perspectives, 2014, 28(2), 3-28.
- [4]Schwabish J A. An economist's guide to visualizing data[J]. Journal of Economic Perspectives, 2014, 28(1): 209-234.
- [5]Korinek A. Generative AI for Economic Research: Use Cases and Implications for Economists[J]. 2023.

Readings for Lecture 2

TBU

Readings for Lecture 3

综述:

- [1]沈艳,陈赞,黄卓.文本大数据分析在经济学和金融学中的应用:一个文献综述[J].经济学(季刊),2019,18(04):1153-1186.
- [2]姚加权,张锬澎,罗平.金融学文本大数据挖掘方法与研究进展[J].经济学动态,2020,(04):143-158.

数字化转型:

- [3]-[6]吴非等, 2021; 袁淳等, 2021; 张永坤等, 2021; 赵宸宇等, 2021
- [7]左祥太.企业数字化转型的“言行不一”与股价风险[J].中国经济学,2025,(01):81-105+245-247.

政策不确定性:

- [8]-[9]Baker et al., 2016; 张峰等, 2019; [Trade Policy Uncertainty](#)

LDA:

- [10] Bellstam G, Bhagat S, Cookson J A. A text-based analysis of corporate innovation[J]. Management Science, 2021, 67(7): 4004-4031.

Readings for Lecture 4

- [1]金星晔,左从江,方明月,等. 企业数字化转型的测度难题: 基于大语言模型的新方法与新发现[J].经济研究,2024,59(03):34-53.
- [2]陆瑶,施函青,周欣怡. 中国企业数字技术风险暴露对企业价值的影响——来自大语言模型的文本分析证据[J].经济研究,2025,60(02):73-89.
- [3]Juhász R, Lane N J, Oehlsen E, et al. Measuring industrial policy: A text-based approach[R]. National Bureau of Economic Research, 2025.
- [4]Fang H, Li M, Lu G. Decoding China's industrial policies[R]. National Bureau of Economic Research, 2025.
- [5]He L Y, Wang L. Can artificial intelligence curb greenwashing? Firm-level evidence based on large language model[J]. Energy Economics, 2025: 108954.

Readings for Lecture 5

- [1] Chernozhukov V, Chetverikov D, Demirer M, et al. Double/debiased machine learning for treatment and structural parameters[J]. 2018.
- [2]张涛,李均超. 网络基础设施、包容性绿色增长与地区差距——基于双重机器学习的因果推断[J].数量经济技术经济研究,2023,40(04):113-135.
- [3]胡诗云,江弘毅,解海天.双重机器学习的理论与应用——从“黑箱”到“工具箱”的实践指南[J].数量经济技术经济研究,2026,43(03):154-179.

Other Resources

Videos:

- [Hung-yi Lee - YouTube](#)
- [金融科技工具箱——面向经管金融同学的Python、爬虫、机器学习课_哔哩哔哩_bilibili](#)
- [金融人工智能“十讲”、“十练”、“十赛”](#)
- [机器学习因果推断之机器学习_哔哩哔哩_bilibili](#)
- [拿铁一定要加冰的个人空间-哔哩哔哩视频](#)

Official Accounts:

- [连享会](#)、[AI and Economics](#)、[机器学习与数字经济实验室](#)、[经金数智实验室](#)、[社科学术汇](#)、[数量经济学.....](#)

All questions are welcomed!