ECON4061 Machine Learning in Economics

Module Outline

Instructor: Marit Hinnosaar

Autumn 2025

Lectures: 10 2-hour sessions, at Monica Partridge D04; Thu 9-11am
Computer classes: 8 1-hour sessions in October & November, at Hallward LG101, Tue 12-1pm
Office hour sign-up:
`https://outlook.office365.com/owa/calendar/SchoolofEconomicsOfficeHours@UniofNottm.onmicrosoft.com/bookings/`

Module aims and objectives

The module introduces students to Machine Learning and the analysis of large datasets in economics. Vast amounts of unstructured data have become available, especially because many production and consumption activities leave digital footprints. Machine Learning algorithms provide powerful tools for analyzing such data. The module focuses on using Machine Learning to generate new data in economics. In particular, it demonstrates how to create structured data from unstructured sources such as text or images. This enables economists to provide better answers to long-standing questions and to address new ones that could not be answered before. The module teaches coding in Python and provides basic knowledge and hands-on practice in implementing Machine Learning algorithms, including Large Language Models. The module gives an introduction to those methods, applies these to real-world datasets, and discusses applications and current academic research that illustrate their use in economics.

Assessment

The assessment consists of group coursework and in-person closed book written 2-hour exam. The exam gives 75% of the mark in the module, group coursework gives 25%. Specifically, group coursework presentation gives 15%, and group coursework essay gives 10% of the final mark. The exam will take place during the Autumn examination period in January. Group courseworks are presented in the last three weeks of the module during lecture times. The deadline for the group coursework essay is Friday, December 12, 2025 at 15:00.

Tentative schedule of lectures and corresponding references

1. Introduction and overview of the module
2. Basic methods in Natural Language Processing: dictionaries, dictionary-based sentiment analysis, cosine similarity, applications in economics
   - Matthew Gentzkow, Bryan Kelly, and Matt Taddy. Text as Data. *Journal of Economic Literature*, 57(3):535–574, September 2019. https://doi.org/10.1257/jel.20181020
   - Matthew Gentzkow and Jesse M. Shapiro. What Drives Media Slant? Evidence From U.S. Daily Newspapers. *Econometrica*, 78(1):35–71, 2010. https://doi.org/10.3982/ECTA7195
   - Tim Loughran and Bill Mcdonald. When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks. *The Journal of Finance*, 66(1):35–65, 2011. https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1540-6261.2010.01625.x

- Gerard Hoberg and Gordon Phillips. Text-Based Network Industries and Endogenous Product Differentiation. *Journal of Political Economy*, 124(5):1423–1465, 2016. https://doi.org/10.1086/688176
- David Deming and Lisa B. Kahn. Skill Requirements across Firms and Labor Markets: Evidence from Job Postings for Professionals. *Journal of Labor Economics*, 36(S1):S337–S369, 2018. https://doi.org/10.1086/694106

3. Neural networks, word embeddings, word2vec, applications in economics
   - Elliott Ash, Daniel L. Chen, and Arianna Ornaghi. Gender Attitudes in the Judiciary: Evidence from US Circuit Courts. *American Economic Journal: Applied Economics*, 16(1):314–350, 2024. 10.1257/app.20210435
   - Elliott Ash and Stephen Hansen. Text Algorithms in Economics. *Annual Review of Economics*, 15(1):659–688, 2023. https://doi.org/10.1146/annurev-economics-082222-074352
4. Recurrent Neural Networks, Long Short-Term Memory Networks, Sequence to sequence (seq2seq) models, applications in economics
   - Yuriy Gorodnichenko, Tho Pham, and Oleksandr Talavera. The Voice of Monetary Policy. *American Economic Review*, 113(2):548–584, 2023. https://www.aeaweb.org/articles?id=10.1257/aer.20220129
5. Attention, Transformers, Large Language Models and other Foundation Models, applications in economics
   - Marit Hinnosaar and Toomas Hinnosaar. Influencer Cartels. SSRN Scholarly Paper 3786617, May 2024
6. Applications of Foundation Models: open-ended survey data
   - Ingar Haaland, Christopher Roth, Stefanie Stantcheva, and Johannes Wohlfart. Understanding Economic Behavior Using Open-Ended Survey Data. *Journal of Economic Literature*, 2025
   - Sebastian Link, Andreas Peichl, Christopher Roth, and Johannes Wohlfart. Attention to the Macroeconomy. SSRN Scholarly Paper 4697814, 2024
7. Large Language Models: fine-tuning
   - Edward J. Hu, Yelong Shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, and Weizhu Chen. LoRA: Low-Rank Adaptation of Large Language Models. arXiv 2106.09685, June 2021
   - Josh Lerner, Junxi Liu, Jacob Moscona, and David Y. Yang. Appropriate Entrepreneurship? The Rise of China and the Developing World. NBER Working Paper 32193, NBER, March 2024
   - Luca Braghieri, Sarah Eichmeyer, Ro'ee Levy, Markus Mobius, Jacob Steinhardt, and Ruiqi Zhong. Article-level Slant and Polarization of News Consumption on Social Media. CEPR DP19807, December 2024
8. Group coursework presentations
9. Group coursework presentations
10. Group coursework presentations

Tentative schedule of computer classes

1. Introduction to Python (assigining values, logical operators and conditionals, lists and indexing, loops, defining functions), Jupyter Notebooks, Pandas, importing and exporting data, taking a look at the data, data aggregation in Pandas, summary statistics, data visualization
2. Regular expressions, dictionaries, sentiment analysis
3. Cosine similarity
4. Word embeddings, Word2vec
5. Large Language Models for text summarization and information extraction

6. Transformer-based models to work with text and images
7. Model fine-tuning
8. Review