Data Science for Economics

Course Description:

This course introduces methods of data science that are useful for economists. The course will cover machine learning and artificial intelligence techniques with a focus on their applications in economics. Topics include foundational AI/ML concepts, predictive modeling, natural language processing, data retrieval, web scraping, and the ethical implications of AI in economic contexts. Students will learn to apply these tools to economic data, exploring issues such as causal inference, forecasting, and economic policy analysis.

Weekly breakdown (timing orientative, will surely deviate):

Week 1: Introduction to AI and ML in Economics

- Overview of AI and ML concepts within economics
- Prediction vs. causal inference, with applications in policy evaluation
- Understanding overfitting and how to evaluate economic predictions
- Suggested Readings: Athey & Imbens (2017), Varian (2014), Korinek (2023)

Week 2: Prediction Models

- Random forest, neural networks, and ensemble methods, matrix factorization techniques
- Suggested Readings: Mullainathan & Spiess (2017)

Week 3: Data Acquisition in Economic Research

- Techniques for OCR and web scraping in economic research
- Suggested Readings: Dell et al (2024)

Week 4: Text and Image Analysis for Economic Insights

- Working with images and text data in economic contexts
- Introduction to topic modeling (LDA) and text-based economic indicators
- Suggested Readings: Bandiera et al (2020), Voth and Yanagizawa-Drott (2024)

Week 5: Text Embeddings

- Text as a prediction problem, embeddings
- Suggested Readings: Ash and Hansen (2023)

Week 6: Sentiment Analysis and Economic Indicators

- Sentiment analysis in economic forecasting, speech recognition
- **Suggested Readings:** Baker, Bloom and Davis (2016)

Week 7: Transformer Models

- Introduction to transformer models
- Applications of BERT, GPT
- Suggested Readings: Vaswani et al (2017)

Week 8: Ethics and Bias in Economic AI Applications

- Exploring ethical concerns of AI in economic decisions, fairness, and transparency
- Examining potential biases in economic ML models and their impact on policy
- Suggested Readings: Korinek (2019)

Week 9: Interpretability and uncertainty

- Introduction to interpretability methods, including SHAP and LIME
- Conformal predictions

Week 10: Fine-Tuning Large Language Models (LLMs) for Economic Applications

- Introduction to the concept of fine-tuning large language models (LLMs) such as GPT and BERT for specialized tasks in economics.
- Suggested Readings: Bana (2022)

Bibliography

Here's the list in alphabetical order:

- Ash, E., & Hansen, S. (2023). Text algorithms in economics. Annual Review of Economics, 15(1), 659-688.
- Athey, S., & Imbens, G. W. (2019). Machine Learning Methods That Economists Should Know About. Annual Review of Economics, 11, 685-725.
- Baker, S. R., Bloom, N., & Davis, S. J. (2016). Measuring economic policy uncertainty. The Quarterly Journal of Economics, 131(4), 1593-1636.
- Bana, S. H. (2022). Work2vec: using language models to understand wage premia. Unpublished Manuscript.
- Bandiera, O., Prat, A., Hansen, S., & Sadun, R. (2020). CEO behavior and firm performance. Journal of Political Economy, 128(4), 1325-1369.
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet Allocation. Journal of Machine Learning Research, 3(Jan), 993-1022.
- Dell, M. (2024). Deep learning for economists (No. w32768). National Bureau of Economic Research.
- Dell, M., Carlson, J., Bryan, T., Silcock, E., Arora, A., Shen, Z., ... & Heldring, L. (2024). American stories: A large-scale structured text dataset of historical US newspapers. Advances in Neural Information Processing Systems, 36.
- Kleinberg, J., Ludwig, J., Mullainathan, S., & Obermeyer, Z. (2015). Prediction policy problems. American Economic Review, 105(5), 491-495.
- Kleinberg, J., Lakkaraju, H., Leskovec, J., Ludwig, J., & Mullainathan, S. (2018). Human Decisions and Machine Predictions. The Quarterly Journal of Economics, 133(1), 237-293.
- Korinek, A. (2019). Integrating ethical values and economic value to steer progress in artificial intelligence (No. w26130). National Bureau of Economic Research.

- Korinek, A. (2023). Generative AI for economic research: Use cases and implications for economists. Journal of Economic Literature, 61(4), 1281-1317.
- Michalopoulos, S., & Rauh, C. (2024). *Movies* (No. w32220). National Bureau of Economic Research.
- Mueller, H., & Rauh, C. (2018). Reading between the lines: Prediction of political violence using newspaper text. American Political Science Review, 112(2), 358-375.
- Mullainathan, S., & Spiess, J. (2017). Machine Learning: An Applied Econometric Approach. Journal of Economic Perspectives, 31(2), 87-106.
- Vaswani, A. et al. (2017). Attention is all you need. Advances in Neural Information Processing Systems.
- Varian, H. R. (2014). Big Data: New Tricks for Econometrics. Journal of Economic Perspectives, 28(2), 3-28.
- Voth, H. J., & Yanagizawa-Drott, D. (2024). Image(s) (No. 19219). CEPR Discussion Papers.