The Oxford Economics Summer School:

Foundations of Machine Learning

8 September – 12 September 2025

instructor	Maximilian Kasy
email	teachingmaxkasy@gmail.com
class time	14:00-16:00
practice sessions	16:30-17:30
webpage	https://maxkasy.github.io/home/ML_Oxford_summerschool_2025/
location	Manor Road Building

1 Overview and Objectives

The goal of this course is to provide you with a quick but solid introduction to the **theoretical foundations of machine learning**. This will allow you to

1. become critical consumers of machine learning research, including an understanding when new methods might or might not be useful for empirical work in economics,

2. develop your own research agenda around importing ideas from machine learning into economic and econometric theory, and

3. speak to the machine learning literature, contributing ideas from economics.

We first introduce probably approximately correct **learning theory** for classification and prediction. We then consider regularization and data-driven choice of tuning parameters. We will discuss how to perform supervised learning tasks using **Python** and the scikit-learn package. We will discuss the canonical normal means model. In this model, we will motivate **shrinkage estimators** in different ways, and will prove the famous result that shrinkage estimators can uniformly dominate conventional estimators.

We will next introduce deep **neural nets**, a highly successful method for supervised learning. In this context, we will also consider numerical methods used for training neural nets, such as stochastic gradient descent. We finish this part of class by discussing **conformal inference**, a simple general technique to calculate valid confidence intervals for the predictions of any supervised learning algorithm.

The next part of class will cover different frameworks for **online and adaptive learning**. We will start with the adversarial online learning setting, where no probabilistic assumptions about data generation are made at all. We will next consider multi-armed bandits, and review some theoretical results providing performance guarantees (regret bounds) for algorithms used for learning in bandit settings.

The class will conclude with a discussion of **ethics** and the **social impact** of artificial intelligence. We will, in particular, review debates surrounding fairness and discrimination, as well as differential privacy.

Practice sessions The summer school will be complemented by practice sessions. In these sessions you will implement some of the methods discussed in lectures in Python. You will in particular conduct simulation exercises, to verify some of our theoretical results numerically.

2 Course outline and readings

Day 1 **Supervised learning, Python**

Shalev-Shwartz, S. and Ben-David, S. (2014). *Understanding machine learning: From theory to algorithms.* Cambridge University Press, chapters 2 to 6.

VanderPlas, J. (2022). *Python Data Science Handbook.* O'Reilly, chapter 5.

Day 2 **Shrinkage in the normal means model**

Stigler, S. M. (1990). The 1988 Neyman memorial lecture: a Galtonian perspective on shrinkage estimators. *Statistical Science*, pages 147–155.

Morris, C. N. (1983). Parametric empirical Bayes inference: Theory and applications. *Journal of the American Statistical Association*, 78(381):pp. 47–55.

Stein, C. M. (1981). Estimation of the mean of a multivariate normal distribution. *The Annals of Statistics*, 9(6):1135–1151.

Abadie, A. and Kasy, M. (2019). Choosing among regularized estimators in empirical economics - the risk of machine learning. *Review of Economics and Statistics*, 101(5).

Fessler, P. and Kasy, M. (2019). How to Use Economic Theory to Improve Estimators: Shrinking Toward Theoretical Restrictions. *The Review of Economics and Statistics*, 101(4):681–698.

Day 3 **Deep neural nets and conformal inference**

Goodfellow, I., Bengio, Y., and Courville, A. (2016). *Deep learning.* MIT Press, chapters 6-8.

Bartlett, P. L., Montanari, A., and Rakhlin, A. (2021). Deep learning: a statistical viewpoint. *arXiv preprint arXiv:2103.09177*

Angelopoulos, A. N. and Bates, S. (2021). A gentle introduction to conformal prediction and distribution-free uncertainty quantification. *arXiv preprint arXiv:2107.07511*

Day 4 **Online learning and multi-armed bandits**

Cesa-Bianchi, N. and Lugosi, G. (2006). *Prediction, learning, and games.* Cambridge University Press, chapter 2.

Bubeck, S. and Cesa-Bianchi, N. (2012). Regret Analysis of Stochastic and Nonstochastic Multi-armed Bandit Problems. *Foundations and Trends® in Machine Learning*, 5(1):1–122.

Russo, D. J., Roy, B. V., Kazerouni, A., Osband, I., and Wen, Z. (2018). A Tutorial on Thompson Sampling. *Foundations and Trends® in Machine Learning*, 11(1):1–96.

Wager, S. and Xu, K. (2021). Diffusion asymptotics for sequential experiments. *arXiv preprint arXiv:2101.09855.*

Kasy, M. and Sautmann, A. (2021). Adaptive treatment assignment in experiments for policy choice. *Econometrica*, 89(1):113–132.

Caria, S., Gordon, G., Kasy, M., Osman, S., Quinn, S., and Teytelboym, A. (2024). Job search assistance for refugees in Jordan: An adaptive field experiment. *Journal of the European Economic Association*, 22(2).

Cesa-Bianchi, N., Colomboni, R., and Kasy, M. (2025). Adaptive maximization of social welfare. *Econometrica*, 93(3).

Day 5 **Ethics and machine learning**

Kasy, M. (2025). *The means of prediction: How AI Really Works (and Who Benefits).* University Of Chicago Press.

Kearns, M. and Roth, A. (2019). *The Ethical Algorithm: The Science of Socially Aware Algorithm Design.* Oxford University Press.

Dwork, C. and Roth, A. (2014). The algorithmic foundations of differential privacy. *Foundations and Trends® in Theoretical Computer Science*, 9(3–4):211–407.

Kasy, M. and Abebe, R. (2021). Fairness, equality, and power in algorithmic decision making. *ACM Conference on Fairness, Accountability, and Transparency.*

Kasy, M. (2024). Algorithmic bias and racial inequality: A critical review. *Oxford Review of Economic Policy.*