Szilvia **Ujváry**

Machine Learning PhD student · Mathematics & Machine Learning Graduate

sru23@cam.ac.uk | szilviaujvary | szilviaujvary

Personal Statement

I am a second-year Ph.D. student in machine learning at the University of Cambridge, with a strong mathematical/statistical background and coding skills. My research focuses on **explaining emergent abilities of LLMs**, such as in-context learning and out-of-distribution generalisation, as well as related foundational questions in (algorithmic) information theory. I am very keen to chat about potential collaborations in these areas. My publications received **spotlight awards** at ICML and NeurIPS.

Education

University of Cambridge — *Cambridge, UK*

PhD in Machine Learning — *Sept 2023 - Aug 2026 (expected)*

- Supervisor: Dr Ferenc Huszár
- Research interests: LLMs, deep learning theory, Bayesian inference, algorithmic information theory
- Teaching Assistant in "Deep Neural Networks" (year 3 module)

University of Cambridge — *Cambridge, UK*

MPhil Machine Learning and Machine Intelligence — *Sept 2022 - Aug 2023*

- Result: **Distinction, rank 2/44**
- Modules include: Computer Vision, Deep Learning, Probabilistic Machine Learning, Information Theory, Reinforcement Learning
- Disstertation topic: Probabilistic Machine Learning, Statistical Learning Theory, **best dissertation in cohort**

Imperial College London — *London, UK*

BSc Mathematics — *Sept 2019 - July 2022*

- Result: First class honours (87%), **top 5-10%, Dean's List** in all years
- Modules include: Python, Data Science, Statistics, Stochastic Simulation, Time Series, R, Network Science, Probability, Applied Probability
- 3rd year Academic Representative, Undergraduate Teaching Assistant

Awards

2024 **Spotlight Paper Award**, NeurIPS 2024
2024 **Spotlight Paper Award**, ICML 2024
2024 **ICML Travel Scholarship**, QuantCo
2023 **Distinction, rank 2/44**, University of Cambridge, *for exceptional performance in my MPhil*
2022 **Full Scholarship**, Emmanuel College, University of Cambridge, *for my MPhil*
2021 **G-Research Prize**, Imperial College London, *for academic excellence*
2020-2022 **Dean's List Years 1, 2, 3**, Imperial College London, *top 10% of cohort in all three years*
2020-2022 **Marjorie McDermott Scholarship**, Imperial College London, *for academic excellence and good citizenship*
2020 **Individual Research Project Prize**, Imperial College London, *for my project on Covid modelling*

Tech Stack

Python, PyTorch, JAX, Tensorflow, NumPy, Matplotlib, Pandas, R, MATLAB, git, GitHub, Linux, LaTeX

Publications

Meszaros, A., Reizinger P., **Ujváry, S.**, Bethge M., Huszár F. (2024) Rule Extrapolation in Language Modeling: A Study of Compositional Generalisation on OOD Prompts. **Spotlight**, *Neural Information Processing Systems (NeurIPS), 2024.* `https://arxiv.org/abs/2409.13728`

Reizinger P.*, **Ujváry, S.***, Meszaros, A.*, Kerekes, A.*, Bethge M., Huszár F. (2024) Position Paper: Understanding LLMs Requires More than Statistical Generalization. **Spotlight**, *International Conference on Machine Learning (ICML), 2024.* `https://arxiv.org/abs/2405.01964`

Ujváry, S., Flamich G., Fortuin V., Hernández-Lobato JM. (2023) Estimating optimal PAC-Bayes bounds with Hamiltonian Monte Carlo. In *1st Mathematics of Modern Machine Learning Workshop, NeurIPS 2023.* `https://arxiv.org/abs/2310.20053`

Ujváry, S.*, Telek, Z.*, Kerekes A.*, Mészáros A.* and Huszár F. (2022) Rethinking Sharpness-Aware Minimization as Variational Inference. In *14th International OPT Workshop on Optimization for Machine Learning, NeurIPS 2022.* `https://arxiv.org/abs/2210.10452`

Experience

Spring into Quant Finance Attendee

G-Research — *April 2024*

Deep Tech Fellow

LakeStar Advisors LLP — *April 2024-June 2024 (part-time)*

- Conducted a Deep Dive assessing novel technologies, market/industry and company landscape in the topic of "Next-generation deep learning architectures". Provided technical evaluation and investment advice on one of LakeStar's investment applicants.

Research Assistant: Flatness and Generalization

University of Cambridge *July 2022 - October 2022*

- Established a link between the Sharpness-Aware Minimization (SAM) and Mean-Field Variational Inference (MFVI) algorithms, and developed VariationalSAM, a method that combines the advantages of both.
- Computed implicit regularisation equations for the above algorithms, and identified biases towards flatness.
- Supervisor: Dr Ferenc Huszár.

Research Assistant: Canonical Correlation Analysis for Feature Selection

Imperial College London *July 2021 - Sept 2021*

- Worked on an optimisation problem of selecting subsets of features for Canonical Correlation Analysis with optimal p—value.
- Conducted an evaluation study comparing several evolutionary algorithms on synthetic data.
- Supervisor: Dr Marina Evangelou

Research Assistant: Clustering of Single Cell RNA-seq Data

Imperial College London *July 2020 - Aug 2020*

- Learnt about clustering and dimensionality reduction algorithms, and applied them on single cell RNA-seq data.
- Supervisor: Dr Vahid Shahrezaei

Further Relevant Research Projects

- **Bayesian Transcendence (current project):** Autoregressive Language Models trained on mixture distributions implicitly approximate Bayesian inference. However, Bayesian inference fails in out-of-distribution scenarios (model misspecification), but Language Models perform better. To understand why, we study this phenomenon on different toy datasets controlling the degree of model misspecification.
- **An Algorithmic Information Theory framework for Exchangeability and Causal Inference (current project):** This fundamental topic focuses on extending links between statistical and Algorithmic Information Theory, in the topics of exchangeability and causal discovery. The main practical relevance is that approximating methods in AIT allows us to trade off data for compute: for instance, causal conclusions can be made using single samples, but infinite compute, as opposed to statistics, which requires infinite data. We have lots of ideas in this domain, including a) normative models of OOD behaviour, b) algorithmic de Finetti and in-context learning, and c) Transformer complexity.
- **Tighter PAC-Bayesian Bounds with Hamiltonian Monte Carlo (MPhil thesis, 2023):** This project focused on improving statistical bounds on the generalisation error of neural networks. I developed a method that computes data-independent PAC-Bayesian bounds using approximate (Hamiltonian Monte Carlo) samples from neural network posteriors, and achieved tighter generalisation guaratees on MNIST.
- **InfoWGAN and MINEGAN (2023):** Analyzed information regularization methods for learning disentangled representations with GANs. Extended the InfoGAN method to Wasserstein GANs, and proposed an alternative method for information regularization, based on the Mututal Information Neural Estimator (MINE).
- **Covid modelling (2020):** Modeled the spread of Covid-19 by extending the SIR model of epidemics, and stochastic network models. Constructed forecasts based on early Covid data, and modeled the effects of cyclic lockdown strategies in London. Developed a Python dashboard to visualize my forecasts.