Charlie Rogers-Smith

Phone +447714626830

LinkedIn linkedin.com/in/charlierogers-smith

E-mail charlierogerssmith@gmail.com

Education

Oct 2019 - Oct 2020

University of Oxford

MSc Statistical Science

- **Distinction**, overall average = **77/100**, practical average = **88/100**.
- **Main areas**: adversarial robustness, learning theory, Bayesian ML, approximate inference, deep learning, Bayesian statistics, kernel methods, statistical inference.
- **Awards:** *Worcester College Prize* for academic excellence.

Sep 2016 - May 2019

University of St Andrews

BSc Mathematics

- **3.98 GPA** (converted), 1st class.
- **Main areas**: real analysis, functional analysis, measure and probability theory, Bayesian statistics.
- **Coding courses**: *'Advanced Symbolic Computation'*; *'Deep Learning'*; *'Machine learning (Datamining)'*: achieved 97% and ranked **1/60** students as a 2nd year in this master's-level course.
- **Awards:** runner up for the ***Best Senior Research Project in Maths, 2019***. The University of St Andrews *Deans' List for Academic Excellence 2016/17, 2017/18,* and *2018/19*.

Experience

Nov 2020 - Current

Oxford Applied and Theoretical ML Group

External Collaborator

- (About to submit:) Estimating the effectiveness of non-pharmaceutical interventions in the second wave of the COVID-19 pandemic, with *Jan Brauner* and *Soeren Mindermann*.
- (Upcoming:) Interpretability projects with *Jan Brauner* and *Soeren Mindermann*, (most probably) supervised by ***Yarin Gal***.

Jun 2020 - Sep 2020

University of Oxford: Dissertation

Understanding Adversarial Robustness

- Under ***Patrick Rebeschini***, I wrote a narrative literature review on adversarial robustness. It covers: the motivations for adversarial robustness, a holistic attack framework, a history of defence failures with corresponding recommendations for constructing and evaluating defence methods, a presentation of and comparison between promising defence methods, and an analysis of hypotheses for the origins of adversarial example. I intend to publish the dissertation. A manuscript is in preparation at bit.ly/36wyliD.

Nov 2019 -

Future of Humanity Institute, Research Fellow

Jan 2020 *Causal Identifiability of Agent Incentives*

- Supervised by ***Ryan Carey*** (Future of Humanity Institute, University of Oxford), as part of a collaboration with ***Tom Everitt*** (DeepMind). We asked the following question: for systems that suffer from perverse incentives, when is it possible to infer from data whether the incentive is being acted on? Our insight was to translate a given system into its structured causal model (SCM) form, and frame the question as determining the identifiability of the path-specific effect corresponding to the perverse incentive. If unidentifiable, it is impossible to tell whether the agent is acting on the incentive. We argue that such agents should not be deployed. My collaborators are continuing this work.

Jun 2019 - Sep 2019

Imperial College London & Aalto University: RA

Semi-Supervised Learning Via Deep GPs

- I initiated a collaboration between ***Marc Deisenroth*** and ***Samuel Kaski*** on methodology for deep Gaussian processes in the semi-supervised setting. I developed the first deep Gaussian process latent variable model that is scalable to a billion data points, alongside variational inference methodology for the first uncertain-input deep GP, which is also fully scalable.

Jun 2018 - Sep 2018

Aalto University: RA

ABC Via Population Monte Carlo And Classification

- Competitive, funded internship with ***Samuel Kaski***, working on approximate Bayesian computation (ABC). I proposed and implemented a population Monte Carlo approach to ABC using multi-class classification, which simultaneously solves the two core problems in ABC: sample inefficiency and the subjective choice of summary statistics. **Our method achieves the same accuracy as the state-of-the-art with an order of magnitude fewer simulations.** A preprint is available: arxiv.org/abs/1810.12233. We intend to submit to ICML 2021.

Funding & Awards

- (**$15,000**) grant from the **Open Philanthropy Project** for self-study and research at Oxford's *Applied and Theoretical Machine Learning Group,* 2021.
- (**$8,000**) grant from the ***Long Term Future Fund*** for post-master's career transition, 2020.
- (**$30,000**) funding from effective altruist donors for the *MSc Statistical Science, University of Oxford,* 2019 – 2020.
- (**$10,000**) funding from ***Effective Altruism Grants*** for research in Summer 2019.
- (**$4,000**) full scholarship for the ***Centre for Applied Rationality*** workshop, 2018
- Community Building Fellow at the ***Centre for Effective Altruism***, 2018.
- Brazilian Jiujitsu: **Gold @ *English National Championships***, **Silver @ *European Championships*** (2015).

Leadership

Founder and President: Effective Altruism: St Andrews

- Empowered students to pursue high-impact careers.
- Expanded the membership of the society to more than 2000, the **largest careers society**, and hosted world-renowned speakers such as *Peter Singer*.
- Managed a committee of 20+ people, running an average of 4 events per week.