# Siu Lun Chau

## Postdoctoral Researcher @ CISPA

✉ siu-lun.chau@cispa.de | 📞 +44 7415 137484 | ⌂ chau999 | 📍 Saarbrücken, Germany

## About me

Hi, I am a postdoctoral researcher at the CISPA Helmholtz Center for Information Security researching at interface of economic theory and machine learning, with a focus on promoting trustworthy AI models. Previously at Oxford, Max Planck Institute, and Amazon.

## Education

**DPhil in Statistical Science** — *Oxford, United Kingdom*
St.Peter's College, University of Oxford — *Oct 2018 - Apr 2023*

- THESIS: *"Towards Trustworthy Machine Learning with Kernels"*
- SUPERVISORS: Prof. Dino Sejdionvic, Prof. Mihai Cucuringu, and Prof. Xiaowen Dong.
- FUNDING: ESPRC and MRC studentship for DPhil in Statistics and Machine Learning

**MMATH Mathematics and Statistics (First Class Honor)** — *Oxford, United Kingdom*
Lady Margaret Hall, University of Oxford — *Sep 2014 - Jul 2018*

- RANK: 2nd in 4th year and 1st in 3rd year.
- MASTER'S SUPERVISOR: Prof. Mihaela Van Der Shaar
- MASTER'S THESIS: *"Modelling Diseases Trajectories with Infinite Mixture of Gaussian Processes"* (Distinction)
- UNDERGRADUATE SUPERVISOR: Prof. Francois Caron
- UNDERGRADUATE ESSAY: *"Adaptive and Gradient Bossting"* (Distinction)

## Work Experience

**CISPA Helmholtz Center for Information Security** — *Saarbrucken, Germany*
Postdoctoral Researcher — *Mar 2023 - Present*

- SUPERVISOR: Dr. Krikamol Muandet
- PROJECT: *"Towards Trustworthy AI through Synergy between Machine Learning and Economics"*

**Amazon** — *London, United Kingdom*
Applied Scientists Intern — *Jun 2022 - Dec 2022*

- PROJECT: Coherent Multi-granularity Demand Forecasting for the Transportation Service Outbound Network
  - Developed Deep Coherent Probabilistic Forecasts on the Amazon EU transportation network for logistic optimisation. Solutions are delivered into production-ready AWS infrastructure.
- Produced research best practice and software developement guidelines for the Applied Science Team.

**Max Planck Institute of Intelligent System** — *Tübingen, Germany*
Visiting researcher — *Oct 2021 - Jun 2022*

- SUPERVISOR: Dr. Krikamol Maundet
- PROJECT: Interface between Machine Learning and Economics
  - Researched into relaxing restrictive assumptions in Instrumental Variable Regression and examined non-parametric hypothesis testing framework for Regression Discontinuity Design.

**Cambridge Spark** — *Cambridge, United Kingdom*
Machine Learning Content Developer — *Aug 2017 - Oct 2021*

- RESPONSIBILITIES: Designed and delivered ML courses to up-skill students and companies. Topics covered include: ML fundamentals, graphs, model explainability using LIME and SHAP, time series forecasting, and Gaussian processes.

**Oxford Strategy Group Digital** — *Oxford, United Kingdom*
Cofounder & Managing Director — *Apr 2017 - Apr 2019*

- RESPONSIBILITIES: Cofounded and managed Oxford first student-led machine learning consulting group.

## Research Interests

The goal of my research is to develop trustworthy algorithms for practitioners using tools from machine learning, statistics, and economics. In particular, I am interested in the following:

- Econometrics and Experimental design
- Kernel methods and Gaussian Processes
- Ranking and Preference learning

- Causal Inference
- Explainability and Uncertainty modelling
- Graph Machine Learning

## Publications

### CONFERENCE PROCEEDINGS

7. **Siu Lun Chau**, Robert Hu, Javier Gonzalez, Dino Sejdinovic, *"RKHS-SHAP: Shapley Values for Kernel Methods"*, Conference on Neural Information Processing Systems (NeurIPS 2022)

6. **Siu Lun Chau\***, Robert Hu\*, Jaime Ferrando Huertas, Dino Sejdinovic, *"Explaining Preference with Shapley Values"*, Conference on Neural Information Processing Systems (NeurIPS 2022)

5. Robert Hu, **Siu Lun Chau**, Dino Sejdinovic, Joan Alexis Glaunès, *"Giga-scale Kernel Matrix-Vector Multiplication on GPU"*, Conference on Neural Information Processing Systems (NeurIPS 2022)

4. **Siu Lun Chau**, Mihai Cucuringu, Dino Sejdinovic, *"Spectral Ranking with Covariates"*, European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML PKDD 2022)

3. **Siu Lun Chau**, Javier Gonzalez, Dino Sejdinovic, *"Learning Inconsistent Preference with Gaussian Processes"*, International Conference on Artificial Intelligence and Statistics (AISTATS 2022)

2. **Siu Lun Chau\***, Jean Francois Ton\*, Yee Whye Teh, Javier Gonzalez, Dino Sejdinovic, *"BayesIMP: Uncertainty Quantification for Causal Data Fusion"*, Conference on Neural Information Processing Systems (NeurIPS 2021)

1. **Siu Lun Chau\***, Shahine Bouabid\*, Dino Sejdinovic, *"Deconditional Downscaling with Gaussian Processes"*, Conference on Neural Information Processing Systems (NeurIPS 2021)

### JOURNAL ARTICLES

1. Xingyue Pu, **Siu Lun Chau**, Xiaowen Dong, Dino Sejdinovic, *"Kernel-based Graph Learning from Smooth Signals: A Functional viewpoint"*, IEEE Transactions on Signal and Information Processing over Networks (IEEE 2020)

## Invited Talks

| | |
|---|---|
| CISPA Helmholtz Center for Information Security | *Saarbrücken, Germany* |
| Title: Explaining Kernel methods and preference models with RKHS-SHAP | *Feb 2023* |
| Oxford Strategy Group Digital | *Oxford, United Kingdom* |
| Title: Introduction to Explainable ML | *Feb 2023* |
| ECML PKDD 2022 | *Grenoble, France* |
| Title: Spectral Ranking with Covariates | *Sep 2022* |
| ELISE Theory Workshop on ML Fundamentals at Eurecom | *Antibes, France* |
| Title: Explainability for Kernel methods | *Sep 2022* |
| S-DCE Alan Turing Institute seminar | *London, United Kingdom* |
| Title: Deconditional Gaussian Processes | *Jun 2022* |
| UCL Gatsby Unit | *London, United Kingdom* |
| Title: Explaining Kernel methods with RKHS-SHAP | *Apr 2022* |

| UCL Statistical Machine Learning Group | *London, United Kingdom* |
|---|---|
| Title: Deconditional downscaling with Gaussian Processes | *Feb 2022* |

| Imperial & Oxford StatML seminar | *London, United Kingdom* |
|---|---|
| Title: Shapley values for Model Explanations | *Feb 2022* |

| Warwick ML Group | *Warwick, United Kingdom* |
|---|---|
| Title: Uncertainty Quantification for Causal Data Fusion | *Jun 2021* |

## Supervision Experience

### Masters Students

| Oscar Yung (University of Oxford) | *Oxford, United Kingdom* |
|---|---|
| THESIS: *"Two Sample Testing for Regression Discontinuity Design"* | *Feb 2022 - May 2022* |

| Samuel Weinman (University of Oxford) | *Oxford, United Kingdom* |
|---|---|
| THESIS: *"Analysis of Price-Volume Interplay in Financial Markets via Machine Learning"* | *May 2020 - Aug 2022* |

### Undergraduate Students

| Mohammad Mehdi Mojarradi, Jihong Lee (Williams College) | *Oxford, United Kingdom* |
|---|---|
| Williams-Exeter Exchange Program at Oxford University | *Mar 2021 - Nov 2021* |

| William Conyers, Daniel Park (Williams College) | *Oxford, United Kingdom* |
|---|---|
| Williams-Exeter Exchange Program at Oxford University | *Jan 2020 - Mar 2020* |

## Teaching Experience

### University of Oxford

TUTOR

| | |
|---|---|
| A12 Simulation and Statistical Programming | 2020 |
| SB1.1 Applied Statistics | 2020 |
| SB1.2 Computational Statistics | 2020 |
| SB2.2 Statistical Machine Learning | 2021 |

TEACHING ASSISTANT

| | |
|---|---|
| SB2.1 Foundations of Statistical Inference | 2019 |

## Awards

| Sep 2018 | ESPRC and MRC studentship for DPhil in Statistics and Machine Learning |
|---|---|
| Sep 2017 | Department Prize for FHS Mathematics and Statistics Part B (Top of the year) |

## Consulting Projects

| Ravio (HR Tech Startup) | *London, United Kingdom* |
|---|---|
| PROJECT: MONOTONIC QUANTILE REGRESSION FOR SALARY BENCHMARKING | *Dec 2022 - Feb 2022* |

- Developed a tree-based quantile regressor with monotonic constraints to incorporate business logic and avoid quantile crossing.

| Ravio (HR Tech Startup) | *London, United Kingdom* |
|---|---|
| PROJECT: JOB TITLE ALIGNMENT USING NLP MODELS | *Apr 2022 - May 2022* |

- Utilised pre-trained languague models such as RoBERTa and GPT3 to create word embeddings to compare and align job titles across companies to standardise compensations.

gini (Fin tech startup)                                                 *Hong Kong*
PROJECT: EXPLAINABLE FORECASTING SPREADSHEET PLUG-IN            *Dec 2020 - Jan 2021*

- Developed a Gaussian Processes based explainable time series model using SAGE for giniPredict, a forecasting spreadsheet plug-in built for non technical decision-makers.

Greenvale (Agricultural tech startup)                          *Cambridge, United Kingdom*
PROJECT: STATISTICAL ANALYSIS ON CROP YIELD DATA                   *Aug 2019 - Jan 2020*

- Conducted statistical analysis on crop yield data to examine seasonal effect on tuber growth across varieties.
- Developed a short-term forecasting model using Gaussian Processes for canopy development based on groundcover observations.

Bonmarché (Fashion retail), UK                                  *London, United Kingdom*
PROJECT: MARKDOWN PRICE OPTIMISATION                              *Mar 2019 - May 2019*

- Developed a demand forecasting model to predict pre-markdown sales and solved for the optimal discount and markdown price to reach the user-defined target sell-through.

## Advisory Position

Hop3 Rewards                                                   *California, United States*
Data Science Advisor                                              *Jan 2023 - Present*

- Advise the data science development in hop3, a web 3.0-powered rewards app that reinvents the way people find fun things to do and interact with brands.

Juvenate Consulting                                                     *Hong Kong*
Advisor                                                          *Jan 2019 - Jan 2021*

- Advised in the early-stage developement in Juvenate, a student-led consultancy striving for greater social impact while helping young people acquire essential skills in the workplace.