Dr. Siu Lun Chau

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

Current Position

Postdoc Researcher, CISPA Helmholtz Center for Information Security, Germany Sep.2023 - Feb.2025
- Advisor: Krikamol Muandet

Research Interests

My research aims to enhance machine learning models' ability to explain what they know (**explainability**) and be explicit about what they don't know (**uncertainty modelling**). I have also worked on:
- Kernel methods and Gaussian Processes
- Ranking and Preference Learning
- Cooperative game theory
- Causal Inference and Econometrics
- Bayesian Optimisation
- Graph Machine Learning

Education

DPhil in Statistical Science, University of Oxford, UK Oct.2018 - Aug.2023
- Thesis: *"Towards Trustworthy Machine Learning with Kernels"*
- Supervisor: Dino Sejdinovic, Mihai Cucuringu, Xiaowen Dong

MMATH in Mathematics and Statistics, University of Oxford, UK Oct.2014 - Jul.2018
- First Class Honors, ranked 2^{nd} in 4^{th} year and 1^{st} in 3^{rd} year.

Work Experience

Visiting Researcher, Australian Institute for Machine Learning, Australia Nov-Dec.2023
- Manager: Dino Sejdinovic

Postdoc Researcher, CISPA Helmholtz Center for Information Security, Germany Sep.2023-Feb.2025
- Advisor: Krikamol Muandet

Research Assistant, CISPA Helmholtz Center for Information Security, Germany Mar.2023-Aug.2023
- Completed my DPhil thesis while helping other members of the group conduct research.

Data Scientist, Ravio (HR Tech startup), UK Dec.2022-Feb.2023
- Developed a tree-based quantile regressor with monotonic constraints for compensation modelling
- Utilised pre-trained language models to align heterogeneous job titles across the industry

Applied Scientists Intern, Amazon, UK Jun.2022-Dec.2022
- Developed deep coherent probabilistic demand forecasts for logistic optimisation
- Produced research best practices and software development guidelines for the applied science team.

Visiting Researcher, Max Planck Institute of Intelligent Systems, Germany Oct.2021-Jun.2022
- Advisor: Krikamol Muandet
- Researched into relaxing assumptions in instrumental variable regression and examined non-parametric testing framework for regression discontinuity design.

Data Science Consultant, Catalyst AI, UK 2019-2021
- Built statistical models on crop yield data for an agricultural tech startup.
- Built demand forecasting models to predict pre-markdown sales for a fashion retail company.

Preprints

Masaki Adachi, Sebastian B. Orbell, Brady Planden, Natalia Ares, David A. Howey, Krikamol Maundet, Michael A. Osborne, **Siu Lun Chau** *"Looping in the human: Collaborative and Explainable Bayesian Optimisation"* arXiv preprint arXiv:2310.17273 (2023)..

Publications

CORE Conference rankings: $A^* = 7\%, A = 15\%, B = 28\%, C = 47\%$
Scimago Journal rankings: $Q1 = top25\%, Q2 = 25\% - 50\%, Q3 = 50\% - 75\%, Q4 = bottom25\%$

(A^*) Kiet Vo, Muneeb Aadil, **Siu Lun Chau**, Krikamol Muandet *"Causal Strategic Learning with Competitive Selection"* The 38th Annual AAAI Conference on Artificial Intelligence (AAAI) 2024.

(A^*) **Siu Lun Chau**, Krikamol Muandet*, Dino Sejdinovic*. (* denotes equal contributions) *"Explaining the Uncertain: Stochastic Shapley Values for Gaussian Process Models."* To appear in Advances in Neural Information Processing Systems 36 (NeurIPS) 2023. (Spotlight paper)

Simon Föll*, Alina Dubatovka*, Eugen Ernst†, **Siu Lun Chau**†, Martin Maritsch, Patrik Okanovic, Gudrun Thäter, Joachim M. Buhmann, Felix Wortmann, Krikamol Muandet. († denotes equal contributions) *"Gated domain units for multi-source domain generalization."* To appear in Transactions of Machine Learning Research (TMLR) 2023.

(A^*) **Siu Lun Chau**, Robert Hu, Javier Gonzalez, and Dino Sejdinovic. *"RKHS-SHAP: Shapley values for kernel*

methods." Advances in Neural Information Processing Systems 35 (NeurIPS) 2022.

(A^*) Robert Hu*, **Siu Lun Chau***, Jaime Ferrando Huertas, Dino Sejdinovic. (* denotes equal contributions) *"Explaining Preferences with Shapley Values."* Advances in Neural Information Processing Systems 35 (NeurIPS) 2022.

(A^*) Robert Hu, **Siu Lun Chau**, Dino Sejdinovic, and Joan Glaunès. *"Giga-scale Kernel Matrix-Vector Multiplication on GPU."* Advances in Neural Information Processing Systems 35 (NeurIPS) 2022

(A) **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

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

(A^*) **Siu Lun Chau***, Jean Francois Ton*, Yee Whye Teh, Javier Gonzalez, Dino Sejdinovic (* denotes equal contributions) *"BayesIMP: Uncertainty Quantification for Causal Data Fusion"*, Advances in Neural Information Processing Systems 34 (NeurIPS) 2021

(A^*) **Siu Lun Chau***, Shahine Bouabid*, Dino Sejdinovic (* denotes equal contributions) *"Deconditional Downscaling with Gaussian Processes"*, Advances in Neural Information Processing Systems 34 (NeurIPS) 2021

($Q1$) 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

Australian Data Science Network Conference 2023 — 2023
- *"Stochastic Shapley Values for Gaussian Processes"*

Data 61 Melbourne — 2023
- *"Collaborative and Explainable Bayesian Optimisation"*

School of Computing, Australian National University — 2023
- *"Stochastic Shapley Values for Gaussian Processes"*

School of Computing and Information Systems, The University of Melbourne — 2023
- *"Collaborative and Explainable Bayesian Optimisation"*

Australian Institute for Machine Learning (AIML) — 2023
- *"Explaining the uncertain: Stochastic Shapley values for Gaussian process models"*

Department of Management, Technology, and Economics at ETH Zurich — 2023
- *"Explaining the uncertain: Stochastic Shapley values for Gaussian process models"*

ETH AI Center — 2023
- *"Explaining the uncertain: Stochastic Shapley values for Gaussian process models"*

Oxford-Man Institute — 2023
- *"Explaining the uncertain: Stochastic Shapley values for Gaussian process models"*

CISPA Helmholtz Center for Information Security — 2023
- *"Explaining Kernel methods and preference models with RKHS-SHAP"*

Oxford Strategy Group Digital — 2023
- *"Introduction to Explainable ML"*

ECML PKDD — 2022
- *"Spectral Ranking with Covariates"*

ELISE Theory Workshop on ML Fundamentals at Eurecom — 2022
- *"Explainability for Kernel methods"*

S-DCE Alan Turing Institute Seminar — 2022
- *"Deconditional Gaussian process"*

UCL Gatsby Unit — 2022
- *"Explaining kernel methods with RKHS-SHAP"*

UCL Statistical Machine Learning Group — 2022
- *"Deconditional downscaling with Gaussian processes"*

Imperial & Oxford StatML Seminar — 2022
- *"Shapley values for model explanations"*

Warwick ML Group — 2021

- *"Uncertainty quantification for causal data fusion"*

Research Funding & Awards

Helmholtz Association (Postdoc funding)	2023
ESPRC and MRC Studentship for DPhil in Statistics and Machine Learning	2018-2023
St.Peter's College, University of Oxford Travel Award	2018
Department Prize for FHS Mathematics and Statistics Part B (Top of the year)	2017

Supervision Experience

PhD Students

Kiet Vo (CISPA) 2023

Anurag Singh (CISPA) 2023

Master's Student

Oscar Yung (University of Oxford) 2022
- Thesis: *"Two Sample Testing for Regression Discontinuity Design"*

Samuel Weinman (University of Oxford) 2020
- Thesis: *"Analysis of Price-Volume Interplay in Financial Markets via Machine Learning"*

Undergraduate Students

Mohammad Mojarradi, Jihong Lee, William Conyers, Daniel Park (William's College) 2020-2021
- Williams-Exeter Exchange Program at Oxford University

Teaching Experience

University of Oxford

Tutor, SB2.2 Statistical Machine Learning	2021
Tutor, SB1.2 Computational Statistics	2020
Tutor, SB1.1 Applied Statistics	2020
Tutor, A12 Simulation and Statistical Programming	2020
Teaching Assistant, SB2.1 Foundations of Statistical Inference	2019

Reviewer Service

Mathematics: Mathematical Methods for Decision Making and Optimization	2023
Transactions on Knowledge and Data Engineering	2023
AAAI	2023
NeurIPS	2021, 2022, 2023
ICML	2022
AISTATS	2022
ECML PKDD	2022
IEEE	2021