Alexander **Soen**

alexandersoen@gmail.com | alexandersoen.github.io | alexandersoen | Alexander Soen

Summary

My research focus is on utilising boosting algorithms, information geometric tools, and the theory of loss functions with a focus on fairness and privacy in machine learning. Recently, I have been exploring topics involving generalization bounds and theory involving classification with rejection and importance weighting. Previously I have worked on topics including formal methods / theorem provers, visualisation in academic influence, knowledge graphs, universal approximation theorems, and point process models.

Education

Doctor of Philosophy in Computer Science — *Anticipated Graduation 2025 July*
The Australian National University; Canberra, Australia — *2021 – Pres.*

Bachelor of Advanced Computing (R&D) (First Class Honours, University Medal) — *GPA: 7.0/7.0*
The Australian National University; Canberra, Australia — *2016 – 2019*

Secondary School — *ATAR: 98.75*
Radford College; Canberra, Australia — *2014 – 2015*

Publications

Published

[1] Zhu H, **Soen A**, Cheung YK, Xie L, "Online Learning in Betting Markets: Profit versus Prediction" — *ICML 2024*
[2] Wang EX, **et al.**, "3D NLTE Lithium abundances for late-type stars in GALAH DR3" — *MNRAS 2024/3*
[3] **Soen A**, Husain H, Nock R, "Fair Densities via Boosting the Sufficient Statistics of Exponential Families" — *ICML 2023*
[4] **Soen A**, Alabdulmohsin I, Koyejo S, Mansour Y, Moorosi N, Nock R, Sun K, Xie L, "Fair Wrapping for Black-box Predictions" — *NeurIPS 2022*
[5] Rizoiu MA, **Soen A**, Li S, Calderon P, Dong L, Menon AK, Xie L, "Interval-censored Hawkes processes" — *JMLR 2022*
[6] **Soen A**, Sun K, "On the Variance of the Fisher Information for Deep Learning" — *NeurIPS 2021*
[7] **Soen A**, Mathews A, Grixti-Cheng D, Xie L, "UNIPoint: Universally Approximating Point Processes Intensities" — *AAAI 2021*
[8] Shin M, **Soen A**, Readshaw BT, Blackburn SM, Whitelaw M, Xie L, "Influence flowers of academic entities" — *IEEE VAST 2019*

Preprints (arXiv)

[9] **Soen A**, Nielsen F, "pyBregMan: A Python library for Bregman Manifolds" — *2408.04175*
[10] **Soen A**, Husain H, Schulz P, Nguyen V, "Rejection via Learning Density Ratios" — *2405.18686*
[11] **Soen A**, Sun K, "Tradeoffs of Diagonal Fisher Information Matrix Estimators" — *2402.04163*
[12] Nock R, Amid E, Nielsen F, **Soen A**, Warmuth MK, "Tempered Calculus for ML: Application to Hyperbolic Model Embedding" — *2402.04163*
[13] Li S, Walder C, **Soen A**, Xie L, Liu M, "Sampled transformer for point sets" — *2302.14346*
[14] Calderon P, **Soen A**, Rizoiu MA, "Linking Across Data Granularity: Fitting Multivariate Hawkes Processes to Partially Interval-Censored Data" — *2111.02062*

Work Experience and Projects

Student Trainee — *Remote + Tokyo, Japan*
RIKEN — *2023 – Pres.*

- 6 month internship at the RIKEN Imperfect Information Learning Team.
- Machine learning research on generalized exponential families, importance weighting, and PAC-Bayesian generalization bounds.

PhD Student — *Canberra, Australia*
Australian National University — *2021 – Pres.*

- In collaboration with the interdisciplinary Humanising Machine Intelligence group at the Australian National University.
- Developing novel algorithms using tools from theoretical machine learning and information geometry, with applications in algorithmic fairness.

PyBregMan - Co-creator — *Remote + Tokyo, Japan*
Australian National University / RIKEN — *2024 - Pres.*

- An open source Python library for geometric computing on BREGman MANifolds with applications. Available on GitHub and PyPi.
- Tutorial "Data Representations on the Bregman Manifold" accepted at ICML'24 GRaM workshop with Google Colab.
- Website at: `https://franknielsen.github.io/pyBregMan/index.html`. Reference documentation for software in preparation [9].

Applied Scientist Intern
Canberra, Australia

Amazon *2023*

- 6 month internship at the Amazon Australian Machine Learning team.
- Machine learning research on causal inference, uncertainty quantification, and learning with rejection. Paper in review [9].
- Working on business projects in the retail product space.
- Analysing data, building models, and using Amazon's tool-kits (AWS, Python, Tensorflow).

Interval-Censored Point Processes – Research Assistant
Sydney, Australia

University of Technology Sydney *2020*

- Worked in a Facebook funded project which involves the collaboration of computer scientists and social scientists to study hate speech.
- Built and deployed various web-crawlers from scratch in Python using numerous APIs.
- Developed new algorithms to fit interval-censored data to Hawkes Process; which resulted in publications [5, 14].

Knowledge Graphs – Research Assistant
Canberra, Australia

Australian National University *2020*

- Collaborated with departments of the Australian Government to integrate different data sources for analysis.
- Created a software pipeline to create knowledge graphs using various technologies (RDF, SPARQL, external APIs).

Point Processes and Neural Networks – Summer Research + Research Student
Canberra, Australia

Australian National University + Australian Signals Directorate *2018 – 2019*

- Collaborated with the Australian Signals Directorate in linking different types of Hawkes process models.
- Proposed a novel architecture for incorporating universal approximation of neural networks for Hawkes process models.
- The work resulted in publication [7].

Visualisation of Academic Influence – Research Assistant
Canberra, Australia

Australian National University *2017 – 2019*

- Maintained and developed the InfluenceMap website (`https://influencemap.cmlab.dev`): a visualisation tool for examining citation and publication based influence patterns in research.
- Worked with Microsoft Academic API to gather the data used for visualisation.
- Presented and demoed the project at the 2018 ACM Multimedia Conference business meeting in Seoul, South Korea.
- The insights and tools developed resulted in publication [8].

Theorem Provers – Summer Research
Canberra, Australia

Australian National University *2016*

- Investigated translating formal semantics defined in HOL4 to executable code in CakeML.
- Presented a talk with a poster at the Fifth Data61 Software Systems Summer School.

Other Experience

Teaching Assistant
Canberra, Australia

Australian National University (Various Courses) *2017, 2020 – 2024*

- Taught courses ranging in topics from machine learning (primarily), data management, to logic with various conveners.
- Helped design and release course material, including, original assignments and lecture plans.
- Taking a head tutor role in 2022 and part of 2023 for a machine learning course of 250+ students, which includes overseeing course design and day-to-day logistics. I have advised in the material and topics taught in the course; and have been strongly involved in developing and creating all course content including examination material.

Honors & Awards

- NeurIPS Scholar Award (Registration + Accommodation Cover) *2022*
- Australian Government Research Training Program *2021*
- Australian National University: University Medal [Top 2 First Class Honours Graduates] *2019*
- Ian Ross Honours Scholarship [High-performing Honours Student] ($5000) *2019*
- Honours Scholarship with the Australian Signal Directorate ($8000) *2019*
- Summer Scholarship with the Australian National University ($5000) $\times 3$ *2016 – 2018*

Coding Proficiency

Programming	Python (Adv.), R (Inter.), C (Inter.), Julia (Basic), Coq (Basic), ML (Basic), Haskell (Basic), Rust (Basic)
Machine Learning	PyTorch (Adv.), scikit-learn (Adv.), Tensorflow (Inter.)
Other	AWS (Inter.), Bash (Adv.), LaTeX (Adv.), Git (Inter.)