STEVE HONG

stevehongforwork@gmail.com linkedin.com/in/steve-hong-1901 github.com/SteveHong1901

Education

University of Cambridge - MPhil Machine Learning & Machine Intelligence **Sep 2024 - Sep 2025**
- Candidate for the Machine Learning Track of the 11-month intensive programme that covers advanced topics in Deep Learning, Probabilistic Learning, Reinforcement Learning, Natural Language Processing and Computer Vision

University College London - BSc Statistics **Sep 2021 - Jun 2024**
- **Grades:** First-Class Grade in first and second year and ranked in the **top 10% of cohort**
- **Modules:** Machine Learning, MCMC methods, Bayesian Statistics, Financial Mathematics, Time Series Models, Data Structures
- Collaborated in 15+ technical individual and group projects that involves developing new libraries and data analysis

DLD College London - A Levels & GCSE **Sep 2018 - Sep 2021**
- **Grades:** Mathematics (A*), Further Mathematics (A*), Economics (A*), Physics (A), AEA Maths Awards (top 5% nationally)
- Awarded an academic excellence scholarship covering £80,000 of tuition fee

Technical Experience

Thesis - Non-Stationary & Multi-Task Gaussian Processes for Wind Turbine Monitoring **Sep 2023 – Jun 2024**
- Saving wind farms time and money by detecting early break-downs, with Prof. Petros Dellaportas and Miss Domna Ladopoulou
- Enhanced model RMSE by 20% using a Spectral Mixture Kernel, surpassing current wind farm benchmarks
- Contributed a GPyTorch extension for Non-Stationary Spectral Kernel leading to a further 5% reduction in RMSE and NLPD
- Publishing a research paper on the development and application of these methods to the Renewable Energy journal

J.P. Morgan - Data Science Summer Analyst **Jun 2023 – Aug 2023**
- Speed up by 10% J.P. Morgan's manual investment news analysis processes by developing a topic identification algorithm
- Fine-tuned a FinBERT model for customised financial named-entity recognition, gaining experience with Transformer architecture
- Utilised data preparation, regularisation, and optimisation techniques that resulted in a model in continued active development
- Return offer received for 2024 Summer Internship

Research Project - Bayesian Logistic Regression to Address High Multicollinearity **March 2024 – April 2024**
- Stabilised parameter estimates in logistic and cauchit regression models using Bayesian inference with MCMC
- Optimised candidate distribution selection through experimentation, emphasising heavy-tailed properties and preconditioning
- Diagnosed convergence and accuracy using Brier scores, effective sample sizes, and trace plot analyses

Research Project - Bayesian Inference in Heston's Model **Dec 2023 – Feb 2024**
- Investigated key results in the Heston's stochastic volatility model for option pricing and times series analysis
- Explored the literature on MCMC methods for approximating Bayesian inference of the parameters in Heston's model
- Produced code for MCMC parameter inference and achieved 5% improvement in RMSE compared to point estimation methods

Research Internship - UCL Clinical Operational Research Unit **May 2022 – Jul 2022**
- Conducted research on Self-Organising Maps to visualise high-dimensional clinical data – supervised by Dr. Ferran Espuny Pujol
- Researched different extensions in literature to rectify the incompatibility with categorical data and evaluated their effectiveness
- Developed the probabilistic extension (NCSOM) library to adapt to categorical data, in continued development by the CORU team

Extracurricular Activities

UCL Artificial Intelligence Society - Head of Events **Sep 2022 – Present**
- Collaborated with academic and industrial AI researchers to organise hackathons, workshops, and speaker events
- Conducted a fundraising campaign to host Climate-Hack in Jan 2023 with universities across the USA, Canada, and Australia

UCL Data Science Society - Head of Workshop **Jun 2022 – Present**
- Lead programming and workshops in Autumn 2022 and research projects in Spring 2023
- Responsible for delivering workshops on Matplotlib, Pandas and Scikit-Learn to an average audience size of 40-50 members

Skill Summary

Programming Languages: Python, R, PostgreSQL
Libraries/Tools: PyTorch, TensorFlow, Transformers, pandas, NumPy, Matplotlib, Git
Languages: English (Proficient), Vietnamese (Native)
Interests: Professional photography and Architecture