Luwei Wang

Edinburgh | luwei.wang@ed.ac.uk | Web: demi-wlw.github.io | github.com/Demi-wlw
demi.lw.wang@gmail.com | linkedin.com/in/luwei-demi-wang

Profile

I am a PhD candidate in Biomedical AI at the School of Informatics, University of Edinburgh, focusing on developing interpretable and robust AI methods for personalised medicine and clinical decision support. My research lies at the intersection of machine learning and biomedicine, where I collaborate closely with domain experts to advance AI models by integrating theoretical insights into more flexible, operationalizable and practical solutions for real-world healthcare applications.

Education

University of Edinburgh, PhD in Biomedical AI — Sept 2023 – August 2027

- **Coursework:** Natural Language Processing, Reinforcement Learning, Next Generation Genomics

University of Edinburgh, MSc in Statistics with Data Science — Sept 2021 – August 2022

- **Grade:** Distinction
- **Dissertation:** Understanding and Clustering Trajectories of Multiple Long-Term Conditions
- **Coursework:** Machine Learning and Pattern Recognition, Methods for Causal Inference, Bayesian Data Analysis, Text Technologies for Data Science

Hong Kong Baptist University, BSc (Honours) in Mathematics and Statistics, Minor in Computer Science — Sept 2017 – August 2021

- **Grade:** First Class
- **Final Year Project:** Mathematical Trading Model Based on Improved MACD for Stock Market
- **Coursework:** Computer Vision and Pattern Recognition, Data Mining, Design and Analysis of Algorithms, Markov Chain and Queueing Theory, Multivariate Statistical Methods

Work Experience

Research Assistant, Data Science Unit, School of Informatics, University of Edinburgh — Oct 2022 – April 2023

- Contributed to an Alan Turing Institute project on supervised clustering of critically ill COVID-19 patients to explore heterogeneity of treatment effects.
- Applied and compared supervised clustering methods—including Supervised K-means, Mahalanobis Metric Learning, Supervised Hierarchical Clustering, and a VAE-based pairwise method—to identify subphenotypes using protein biomarker data.
- Investigated subgroups likely to benefit from convalescent plasma therapy, using clustering results to inform patient selection and support precision treatment strategies.
- Delivered a comprehensive report evaluating the effectiveness of supervised clustering for uncovering treatment-response heterogeneity.

Statistician Intern, Census and Statistics Department, Hong Kong — June 2019 – July 2019

- Analyzed the current and forecast expenditure of annual training programs in Hong Kong to determine its capital contribution, taking into account the floating exchange rate between Hong Kong dollars and Japanese Yen, as well as the inflation rate.

Publications

Nonparametric Bayesian Multi-facet Clustering for Longitudinal Data — Upcoming 2025
Conference on Uncertainty in Artificial Intelligence (UAI)

Luwei Wang, Kieran Richards, Sohan Seth

https://openreview.net/forum?id=o7xZLAheJ5

Projects

Detecting the Abnormal Behaviors from Online Examination Videos 2021

- Investigated two common forms of abnormal behavior during online exams and curated a representative dataset for training, validation, and testing.
- Built a suspicious behavior detection system using convolutional neural networks (CNNs) and YOLO for object detection, focusing on anomalous facial and hand movements.
- Applied a support vector machine (SVM) classifier to detect abnormal videos, achieving an average classification accuracy of 86%.

Clustering Trajectories of Multiple Long-term Conditions [Website]

- Focused on clustering patient disease trajectories to identify phenotypes and predict adverse outcomes (e.g., mortality) using GP data from ∼150,000 UK Biobank participants.
- Assessed three established clustering techniques: Non-negative Matrix Factorization, Mixture Hidden Markov Models, and Dynamic Time Warping.
- Developed a novel method that integrates K-means with Recurrent Neural Networks (RNNs) to extract temporal features and yield more distinct, disease-dominant clusters.
- Compared all four methods based on model complexity and clustering performance, consistently identifying three representative trajectory clusters.
- Contributed to improved understanding of multimorbidity progression through the discovery of common disease development patterns.

Academic Paper Search Engine [GitHub]

- Collaborated on the development of a search engine for academic papers supporting Boolean, proximity, and semantic query search capabilities.
- Led the design and implementation of the document embedding pipeline, utilizing BERT to encode papers and queries into semantic feature vectors.
- Achieved superior performance metrics, outperforming arXiv's search engine in relevance and accuracy of returned results.

Skills

Programming: Python, R, MATLAB, JAVA, SQL

Tools: PyTorch, Stan, LaTeX