HAOHUI LU

lukhofai@gmail.com

EDUCATION

The University of Sydney, Australia *2020 – 2023*
Doctor of Philosophy
Thesis title: Chronic Disease Prediction using Graph Machine Learning

The University of Sydney, Australia *2020 – 2020*
Master of Data Science
Transferred to Doctor of Philosophy

The University of Sydney, Australia *2019 – 2019*
Graduate Certificate in Data Science

The University of Sydney, Australia *2012 – 2012*
Master of Project Management
Specialisation: Project Economics and Scheduling Management

The University of Sydney, Australia *2008 – 2011*
Bachelor of Commerce
Major: Operations Management and Decision Sciences (Business Analytics)

RESEARCH EXPERIENCE

Research Fellow May 2024 – Present
The University of Sydney *Sydney, Australia*

· Faculty Internal Funded Project on Child Undernutrition.
· Developed and fine-tuned large language models (LLMs) for analysing child undernutrition in Indonesia.
· Integrated LLMs into a web-based GUI, enhancing data interaction and visualisation.
· Compared LLM outcomes with traditional statistical models to validate effectiveness.
· Collaborated on a faculty-led project extending the Australia Awards Fellowship program.
· Coordinated with Indonesian policymakers and BAPPENAS, supporting strategic partnerships and workshops.

Research Assistant (in data analytics) October 2022 – May 2024
The University of Sydney *Sydney, Australia*

· Supported multidisciplinary experts in National Health and Medical Research Council (NHMRC) Ideas Grant project.
· Performed data preprocessing for Bilateral Investment Treaties (BITs) data, utilising the latest Natural Language Processing (NLP) techniques to analyse BITs similarities, genealogy, and network.
· Identified key countries' roles during specific periods using machine learning methods.
· Developed and maintained interactive dashboards.

Doctoral Researcher July 2020 – December 2023
The University of Sydney *Sydney, Australia*

- Chronic disease prediction using graph machine learning: Utilised administrative data and graph-based machine learning to find valuable insights from the patients, improve accuracy for disease risk prediction and identify high-risk patients.
- Used SQL to retrieve and manipulate data from administrative claims datasets and Python for machine learning and deep learning modelling.
- Developed dashboards using Tableau, PowerBI, Python package Plotly and Streamlit.

ACADEMIC EXPERIENCE

Casual Academic   July 2020 – Present
The University of Sydney   *Sydney, Australia*

- Tutoring and marking

  – PMGT5873 Project Economics and Procurement

  – PMGT6867 Quantitative Methods: Project Management

  – QBUS2310 Management Science

  – QBUS2810 Statistical Modelling for Business

  – QBUS6820 Business Risk Management

  – QBUS6840 Predictive Analytics

- Course design and Lecturer: PMGT5866 Quantitative Methods in Project Management (2023 Semester 2)

PROFESSIONAL EXPERIENCE

Peer Support Advisor/Senior Peer Support Advisor   June 2021 – June 2023
The University of Sydney   *Sydney, Australia*

- Helped students by addressing various inquiries, from understanding the support services offered by the university to exploring the social activities they could participate in.
- Gathered data analysis from various channels, derived insights from the information, and recommended activities for the student life team.

Personal Banker   July 2016 – April 2020
Australian and New Zealand Banking Group (ANZ Bank)   *Sydney, Australia*

- Provided a full range of professional sales advice to help customers achieve their financial needs and goals. Explained lending products' fees, interest rates, and current campaigns to customers, ensuring that the products met their needs.
- Held Tier 2 Personal Advice, Personal lending, and small business accreditation.

Data Analyst   August 2012 – July 2016
Dynamic Payment Pty Ltd   *Sydney, Australia*

- Identified fraudulent activity and took necessary actions for escalation. Analysed both unstructured and structured data to construct behavioural and predictive models, which doubled the number of business customers (merchants) during peak travelling periods.
- Performed data analytics, analysed sales, developed statistical models for sales forecasting and engaged in data visualisation.
- Developed and maintained interactive dashboards for daily transaction volume, amount, and new customer analysis, leading to a significant reduction in related data requests.

PEER-REVIEWED ACADEMIC PUBLICATIONS

- **Lu H.**, Uddin S. (2024). A parameterised model for link prediction using node centrality and similarity measure based on graph embedding. *Neurocomputing*. 127820.

- Uddin, S., **Lu, H**. (2024). Confirming the statistically significant superiority of tree-based machine learning algorithms over their counterparts for tabular data. *Plos one*, 19(3), e0298380.

- Uddin, S., **Lu, H**, Alschner, W., Patay, D., Frank, N., Gomes, F. S., Thow, A. M. (2024).(2024). An NLP-based novel approach for assessing national influence in clause dissemination across bilateral investment treaties. *Plos one*, 19(4), e0301541.

- Uddin, S., Yan, S., **Lu, H**. (2024). Machine learning and deep learning in project analytics: methods, applications and research trends. *Production Planning & Control*, 1-20.

- **Lu H.**, Uddin S. (2024). Unsupervised machine learning for disease prediction: a comparative performance analysis using multiple datasets. *Health and Technology*, 1-14.

- Uddin S., **Lu H.** (2024). Dataset meta-level and statistical features affect machine learning performance. *Scientific Reports*, 14 (1), 1670.

- Uddin S., Khan A., **Lu H.**, Zhou F., Karim S., Hajati F., Moni MA., (2024). Road networks and socio-demographic factors to explore COVID-19 infection during its different waves. *Scientific Reports*, 14 (1), 1551.

- **Lu H.**, Uddin S. (2023). Disease Prediction Using Graph Machine Learning Based on Electronic Health Data: A Review of Approaches and Trends. *Healthcare*. 11(7), 1031.

- Uddin, S., Wang, S., Khan, A., & **Lu, H.** (2023). Comorbidity progression patterns of major chronic diseases: The impact of age, gender and time-window. *Chronic Illness*, 19(2), 304-313.

- Uddin, S., Ong, S., **Lu, H.**, & Matous, P. (2023). Integrating machine learning and network analytics to model project cost, time and quality performance. *Production Planning & Control*, 1-15.

- Uddin, S., Khan, A., **Lu, H.**(2023). Impact of COVID-19 on Journal Impact Factor. *Journal of Informetrics*, 17(4), 101458.

- **Lu H.**, Uddin, S., Hajati, F., Moni, M. A., & Khushi, M. (2022). A patient network-based machine learning model for disease prediction: The case of type 2 diabetes mellitus. *Applied Intelligence*, 52(3), 2411-2422.

- **Lu H.**, Uddin, S. (2022). Embedding-based link predictions to explore latent comorbidity of chronic diseases. *Health Information Science and Systems*, 11(1), 2.

- **Lu H.**, Uddin, S. (2022). A disease network-based recommender system framework for predictive risk modelling of chronic diseases and their comorbidities. *Applied Intelligence*, 52(9), 10330-10340.

- **Lu, H.**, Uddin, S. (2022). Explainable Stacking-Based Model for Predicting Hospital Readmission for Diabetic Patients. *Information*, 13(9), 436.

- Uddin, S., Haque, I., **Lu, H.**, Moni, M. A., & Gide, E. (2022). Comparative performance analysis of K-nearest neighbour (KNN) algorithm and its different variants for disease prediction. *Scientific Reports*, 12(1), 1-11.

- Uddin, S., Ong, S., & **Lu, H.** (2022). Machine learning in project analytics: a data-driven framework and case study. *Scientific Reports*, 12(1), 15252.

- Uddin, S., Khan, A., **Lu, H.**, Zhou, F., & Karim, S. (2022). Suburban Road Networks to Explore COVID-19 Vulnerability and Severity. *International Journal of Environmental Research and Public Health*, 19(4), 2039.

- Uddin, S., **Lu, H.**, Khan, A., Karim, S., & Zhou, F. (2022). Comparing the Impact of Road Networks on COVID-19 Severity between Delta and Omicron Variants: A Study Based on Greater Sydney (Australia) Suburbs. *International Journal of Environmental Research and Public Health*, 19(11), 6551.

- Uddin, S., Wang, S., **Lu, H.**, Khan, A., Hajati, F., & Khushi, M. (2022). Comorbidity and multimorbidity prediction of major chronic diseases using machine learning and network analytics. *Expert Systems with Applications*, 205, 117761.

- Wang, S., **Lu, H.**, Khan, A., Hajati, F., Khushi, M., & Uddin, S. (2022). A machine learning software tool for multiclass classification. *Software Impacts*, 13, 100383.

- **Lu H.**, Uddin, S. (2021). A weighted patient network-based framework for predicting chronic diseases using graph neural networks. *Scientific reports*, 11(1), 22607.

CONFERENCES AND PAPER PRESENTATIONS

- **Lu, H.**, Uddin, S., Hajati, F., Khushi, M., & Moni, M. A. (2022). Predictive risk modelling in mental health issues using machine learning on graphs. In *Australasian Computer Science Week 2022* (pp. 168-175).

- **Lu, H.**, Uddin, S.(2022). Predicting readmission risk for diabetic patients: Make artificial intelligence work in real life with interpretable machine learning. In *Digital Health Week 2022*

- **Lu, H.**, Uddin, S.(2022). A network approach to explore COVID-19 vulnerability and severity In *INSNA Sunbelt 2022*.

- **Lu, H.**, Uddin, S.(2023). KNN-Based Patient Network and Ensemble Machine Learning for Disease Prediction In *Health Information Science Society 2023*.

HONOURS AND AWARDS

- Postgraduate Research Support Scheme 2022 - The University of Sydney

- Paulette Isabel Jones Career Award 2023 - The University of Sydney

PROFESSIONAL SERVICE

- Session chair: Sunbelt 2022. Session 93-1. Migration and social networks in the COVID-19 context.

- Paper reviewer: Advanced Engineering Informatics, BMC Health Services Research, BMC Medical Informatics and Decision Marking, Complexity, Computers in Biology and Medicine, IScience, Machine Learning with Applications, Scientific reports, Journal of Computational Social Science, Journal of Medical Systems, Health Data Science.

CERTIFICATES

- SAS Certified Specialist: Base Programming Using SAS 9.4

- SAS Certified Advanced Programmer for SAS 9