XIAN LI

Email: xian.li@anu.edu.au | Phone: +61452080740 | Linkedin: in | Personal Website: 🌐

OBJECTIVE

Results-driven Data Scientist with a PhD in Statistics and a strong foundation in analytics, seeking a challenging quantitative role. Bringing expertise in data analysis, statistical modeling, machine learning, and a proven ability to derive actionable insights from complex datasets to deliver impactful solutions for clients.

WORK EXPERIENCE

Data Modeller/Data Analyst
Department of Health and Aged Care *Nov 2024 - present*
A department of the Australian Government responsible for health research, funding, promotion and regulation in Australia.
Health Workforce Division
Key responsibilities:

- Develop, enhance and maintain a range of mathematical models relating to the projection of the health workforce demand and supply.

- Apply in depth knowledge in forecasting, statistics, machine learning algorithms and quantitative techniques to produce actionable policy insights.

- Utilise complex analytic techniques and experimental methods independently to develop technological solutions to achieve business outcomes, using computational methods to discover data patterns and to develop classification, prediction and optimisation models.

- Work with a cross functional team of analysts, modellers, software developers and cloud engineers to develop a health workforce supply and demand modelling tool.

Data Analyst
Department of Employment and Workplace Relations *Nov 2023 - Oct 2024*
A department of the Australian Government to fill skill gaps to support Australia's economic growth.
Apprenticeships and Foundation Skills Division
Key responsibilities:

- Create visualizations, such as charts and graphs, to effectively communicate data findings using PowerBI

- Generate regular reports based on data analysis for stakeholders.

- Develop SQL queries to extract the required data. Ensure that the queries are efficient, accurate, and aligned with the specific criteria outlined in the ad hoc request.

- Modelling for policymakers involves developing analytical models to simulate and predict the effects of various policies, enabling informed decision-making through scenario analysis, sensitivity assessment, and effective communication of results using R.

- Efficiently manage and version control software projects using Azure DevOps and Git

Key achievements:

- Explore the impact of the Australian Apprenticeship Training Support Payment on the retention rate of the Apprenticeship Program 2023

- Change point detection for the Australian New Energy Apprenticeship Commencement 2023

- Factors that influence the trend of new energy commencements in 2023: A follow-up study

Research Support
Orima Research *Aug 2022 - July 2023*
Australia's leading provider of end-to-end research and data analytics services to the public and not-for-profit sectors.
High Performance Organisation Research Division
Key responsibilities:

- Analysing data and interpreting research results (using SPSS, R or Excel)

- Data processing and data file preparation, including automated data cleaning, data transformation, feature engineering, data integration and data validation, etc

- Developing and implementing an automated reporting system to generate charts and descriptive statistics using VBA, R, or Python

- Higher level statistical analysis including regression modeling where required

Key achievements:

- Develop automated data transformation in Australian Federal Police Census Reporting 2022

- Develop automated data integration in Department of Foreign Affairs and Trade Census Reporting 2022

- Implement developed methods in National Indigenous Australians Agency Census Reporting 2022

- Develop automated data validation in Department of the Prime Minister and Cabinet Census Reporting 2022

- Generate 10+ Gigabyte database using R/R Shiny parallel computing and distributed computing for Australian Capital Territory Public Servant Employee Survey Interactive Dashboard 2023

Evaluation Administrator
The Australian National University *Aug 2021 - July 2022*
A public research university and a member of the Group of Eight.
Planning & Performance Measurement Division
Key responsibilities:

- Text analytics and open text coding and working on survey design and analysis

- Conducting in-depth exploration of the data through various visualization techniques, including scatter plots, box plots, heatmaps, and time series plots.

- Building and training machine learning models to predict outcomes, classify data, or perform other tasks, using R or Python.

- Preparing reports for committees based on data interpretation using Latex

Key achievements:

- Use an ensemble of logistics regression and random forest to make accurate predictions and improve the ANU application procedure. International/Domestic Offer Conversion Modeling 2005-2016

- Use linear regression to detect any fraud applications or identify areas where there are issues, bottlenecks, or challenges that need to be addressed and present the report to the Academic Quality Assurance Committee. Cohort Performance Analysis: an exploration of undergraduate and postgraduate progress rates and attrition

- Use logistic regression to detect any fraud applications or identify areas where there are issues, bottlenecks, or challenges that need to be addressed and present the report to the Academic Quality Assurance Committee. Cohort Performance Analysis: An Exploration of the 2005-2016 Undergraduate Completion

- All reports that have been audited by an external auditor

EDUCATION

Doctor of Philosophy, Major in Statistics *Feb 2020 - Sept 2024*
The Australian National University, Canberra
College of Business & Economics
Key achievements:

- The Statistical Society of Australia Canberra rego grant prize winners, 2023
  https://statsoc.org.au/News-and-media-releases

- The Statistical Society of Australia Dennis Trewin Prize, 2021
  https://www.statsoc.org.au/Canberra-branch-Dennis-Trewin-prize

- Australian Government Research Training Program Fee-Offset Scholarship, 2020

- Australian Government Research Training Program Scholarship, 2020

Bachelor of Statistics (Honours) with First Class, Major in Statistics *Feb 2019 - Nov 2019*
The Australian National University, Canberra   GPA: 7.00/7.00
College of Business & Economics
Key achievements:

- The College of Business and Economics Research School of Finance, Actuarial Studies and Statistics Honours Scholarship, 2019

- Thesis: Circle Detection in Social Networks

Bachelor of Science, Major in Mathematics and Statistics *July 2015 - July 2018*
The Australian National University, Canberra   GPA: 6.78 /7.00
College of Science
Key achievements:

- Burton and Garran Hall Distinguished Scholar Award, 2017

- Joint Colleges of Science Prize, Scholarship or Award, 2016

- Deputy Dean (Education) Commendation, 2015

KEY SKILLS

- **Massive and Complex Data Solutions**
  Highly proficient skills in large and complex datasets through my multiple experiences including handling 10+ Gigabyte survey datasets and administration datasets using parallel computing or cloud computing for advanced modelling.

- **Automated Data Pipeline (R)**
  Developed excellent skills in automated data pipeline including missing data imputation, data validation, data processing and transformation, multidimensional data analysis and modelling, model simulation assessment and generation of reports.

- **Advanced Statistical Theories, Statistical Modelling and Machine Learning Alogrithom**
  Highly developed knowledge and skills in the deployment and development of complex statistical and machine learning modelling to address real-world data challenges, including but not limted to regression analysis, machine learning algorithms, clustering analysis, time series analysis, large-scale data, data imputation, bootstrap resampling, network analysis, robust statistics and causal inference, etc.

- **Data Visualization (R package ggplot or PowerBI)**
  Developed proficient skills in data visualization to deliver technical and complex statistical inference results clearly and concisely to stakeholder with various backgrounds.

CONFERENCE & PUBLICATION

- Conference
  - Contributed Papers at Australian Statistical Conference (ASC 2023), Wollongong, Australia
    Big data methods
    https://virtual.oxfordabstracts.com/#/event/public/4133/program
  - Contributed Papers at Joint Statistical Meetings (JSM 2023), Toronto, Canada
    Theory and Methods for Variable Selection
    https://ww2.aievolution.com/JSMAnnual
  - Contributed Papers at Early Career and Student Statisticians Miniconference (ECSS 2022), Perth, Australia
    Subbagging Variable Selection for Massive Data
    https://ecssmini2022.netlify.app/programme
  - Invited speaker at 5th International Conference on Econometrics and Statistics (EcoSta 2022), Ryukoku University, Kyoto, Japan
    Change-point Detection and Variable Selection for Large-scale Data
    http://www.cmstatistics.org/EcoSta2022/fullprogramme.php
  - Invited speaker at SSA Canberra October Branch Meeting (2021), Canberra, Australia
    Subbagging Variable Selection
    https://www.statsoc.org.au/event-4528615

- Publication
  - On the Subbagging Estimation for Massive Data
    https://arxiv.org/abs/2103.00631 (Submitted to Journal of Econometrics)
  - Unaware Fairness: Hierarchical Random Forest for Protected Classes
    https://arxiv.org/abs/2106.15767
  - Robust Distributed Computing
    Working project

CERTIFICATIONS & PROFESSIONAL MEMBERSHIPS

- Certified Practising Accountant, Foundation Program, 2018

- The Statistical Society of Australia (SSA), Student Member, 2022 - current

REFERENCES

Available upon request