

Personal Statement

Proven leader in driving data science and digital transformation portfolios. Strong product mindset and stakeholder management skills, delivering impactful, value-driven solutions. Specialist in NLP with extensive R&D and consulting experience across various industries.

Work Experience

Atlassian

Principal Machine Learning Engineer

July 2024 – Now

Sydney, Australia

- > Adopt GenAI at scale with quality

Judo Bank

GenAI and Data Science Lead

Dec 2023 – May 2024

Sydney, Australia

- > Led and shaped GenAI/LLM strategy, prioritising use cases, building capability roadmaps, and defining ROI metrics
- > Owned and drove a GenAI POC which significantly enhances internal banker's work efficiency, validated with trial user groups

Endeavour Group

Data Science Chapter Lead

Apr 2022 – Dec 2023

Sydney, Australia

- > Set goals, priorities and roadmaps for the digital transformation data science team (as a line manager of 12 permanent data scientists from Junior to Principal levels), and provide support for tribe agile delivery
- > Drive and execute initiatives that empower data-driven business and optimisation, e.g. build a generic forecast engine and tailor it for different departments from logistics, merchandise to store rostering
- > Enable team to build data science portfolios for efficient and effective MLOps, e.g. Feature Store, Machine Learning CookieCutter

Accenture

Tech Arch Delivery Manager and Delivery Lead

Sep 2018 – Mar 2022

Sydney, Australia

- > Led the new platform design and data migration for a client's business critical legacy system
- > Oversaw tech architecture design and implementation for a client's 5G telco system roll-out
- > Set up and improved process, culture and capability for a diverse team as a delivery lead
- > Coordinated business units and tech teams to unblock barriers and ensure shared goals
- > Brought next best actions and improvement plans for managements in quarterly roadmap plannings

Kaplan

Senior Data Scientist/NLP Specialist

May 2017 – Aug 2018

Sydney, Australia

- > Optimised a text-based financial compliance system with more than 100x speed up with efficient algorithms
- > Drove a serverless design in cloud which significantly boosted the system throughput and scalability
- > Applied active learning and other pragmatic methods to improve compliance system accuracy
- > Co-developed an in-house data anonymisation system for PII Redactions

Sydney Startup

Data Scientist

Mar 2016 – Mar 2017

Sydney, Australia

- > Designed and implemented an automatic user profile summarisation system. The work features data crawling, cleaning, reconciliation and natural language generation from various social media sources. The deliverables facilitate online recruiters for more effective and efficient talent acquisitions

IBM Research

Research Scientist

Apr 2014 – Mar 2016

Melbourne, Australia

- › Led the design and implementation for a number of social media analytics with ML and NLP, including (a) a text-based [geotagger](#), (b) [Social Media Analytics in Real Time](#) and (c) a biomedical [named entity recogniser](#). The deliverables have been utilised by clients from fashion, sports and insurance

Melbourne University/NICTA

PhD and Research Assistant

Jul 2010 – Mar 2014

Melbourne, Australia

- › Applied NLP/ML methods to improve the utility of social media for various analytics, including "[Early detection and mitigation of disastrous events with broadband-enabled social interaction technologies](#)"

Microsoft Research Asia

Fulltime Research Intern

Jan 2007 – Aug 2007, Sep 2009 – Jun 2010

Beijing, China

- › Co-developed an English writing assistant which has been productised as [Bing Search Dictionary](#) in China. My work includes bilingual text mining, short sentence categorisation, incremental bilingual text crawling. This product had won [2010 Wall Street Journal Asian Innovation Award](#)

Professional Skills

Cloud Best practices for multi-tier scalable and reliable solutions, high availability and disaster recovery design, serverless/containerization architecture, complex program delivery, data migration and integration, enterprise hybrid solutions

Frontend HTML, CSS, Javascript (jQuery and React)

Programming Python, R, Bash, Java, C#, C/C++

Modelling Jupyter, pandas, scikit-learn, spaCy, word2vec, PyTorch, transformers, AWS SageMaker, GCP Vertex AI

Backend MySQL, Postgres, Redis/Memcached, Kafka, MongoDB

Container/DevOps Docker, Kubernetes/OpenShift, git/bitbucket, JIRA/Trello, Jenkins/Bamboo

Education

Melbourne University

PhD in Computer Science (Natural Language Processing)

2010 – 2014

Melbourne, Australia

Dalian University of Technology

Bachelor and Master in Software Engineering

2003 – 2009

Dalian, China

Extra Information

My research focuses on [Improving the Utility of Social Media with Natural Language Processing](#). In particular, I studied Twitter text normalisation and geolocation prediction in depth. A tweet length summary of this work is *U wanto undrstnd TWEETS lik dis? Make them readable like this; and if there are too many to read, select those from a region.*

Professional achievements

- › Certified [AWS Solution Architect Associate, Professional](#) and [Google Cloud Digital Leader](#)
- › Published a few top-tier Natural Language Processing (NLP) conference and journal papers (e.g. ACL, EMNLP, COLING and JAIR). My H10-index is 15 and the most cited paper is over 500 (by 2020). See [Google Scholar](#).
- › Co-organised two international workshops on social media analytics ([W-NUT 2015](#) and [W-NUT 2016](#)).
- › Held two issued patents in the United States (US9569551B2, US9697245B1).
- › Held one software copyright in China (2015SR225534).