

Tod Nestor

Data Scientist - ATO

✉ tod.m.nestor@gmail.com — [in linkedin.com/in/todnestor](https://www.linkedin.com/in/todnestor) — github.com/tmnestor

Professional Summary

Data scientist with a solid academic preparation and scientific research experience. Began scientific programming in the mid-80's, progressing to scientific applications software development by the late 80's. Expertise in machine learning, natural language processing, and statistical modeling with a keen interest in the training and application of Neural Networks.

Employment

Data Scientist (APS4) | Australian Taxation Office | November 2022 – Present

Document Understanding Group (2025)

- Vision-language model evaluation for automated document information extraction
- Comparative framework development for model performance assessment (Llama-3.2-Vision-11B, InternVL3-2B/8B)
- Technical liaison with data engineering for production model deployment
- Ground truth dataset preparation using a purpose designed, Microsoft Excel multi-annotator management system to facilitate mapping structured information from invoices, receipts, bank statements into 57 standardised fields

Individual Investor Portfolio Group (April 2024 – March 2025)

- Designed and implemented a “Free Text Categorisation” system to map in excess of 1 million Property Manager Claims from a SPAD Teradata Table into the ATO Rental Schedule categories
 - Modular architecture with contrastive learning (triplet-loss on all-mpnet-base-v2), Bayesian optimisation, and conformal prediction
 - Improved macro F1 from 72% (baseline) to 96% through advanced ML techniques
 - Deployed to KFP Staging with comprehensive GitLab Wiki documentation

Data Graduate Program (2023)

- Fraud detection using gradient boosted models (<1% fraud prevalence dataset)
- Contrastive learning for ANZSCO occupation classification
- Large Language Model ITAF submission and approval
- Shiny dashboard modularisation for improved code maintainability

Accounts Clerk | Pay&Advice | November 2020 – June 2021

Head of Mathematics | Beaconhills College | January 2007 – November 2020

Leading Teacher | Ouyen Secondary College | January 2005 – December 2006

Education

Master of Data Science (with Excellence) | UNSW | 2021 – 2022

Graduate Diploma Data Science | Monash University | 2019 – 2020

Doctor of Philosophy (Theoretical Geophysics) | Australian National University | 1992 – 1995

John Conrad Jaeger Scholar

Master of Science (Geophysics) | Monash University | 1989 – 1991

Australian Postgraduate Research Award

B.App.Sc (Mathematics) with Distinction | RMIT | 1985 – 1988

Technical Skills

Languages: Python, R, Teradata SQL, Fortran, C, Bash, AWK

Version Control: Git

Development Tools: RStudio, Visual Studio Code, VIM

Frameworks: Scikit-learn, Pandas, PyTorch, Transformers, Snorkel AI

Specialisations: Machine learning, NLP, contrastive learning, time series forecasting, explainable AI

Certifications

Machine Learning Specialisation | Coursera | 2023

3 courses, 91 hours: Supervised, Unsupervised and Reinforcement Learning

Python Programming | DataCamp | 2019

18 courses, 71 hours: Software Engineering, Packaging, Git, Unit Testing, OOP, Regular Expressions, Pandas