Tod Nestor

Data Scientist - ATO

tod.m.nestor@gmail.com — linkedin.com/in/todnestor — github.com/tmnestor

Professional Summary

Data scientist with extensive mathematical foundation (PhD Theoretical Geophysics). I have been a programming hobbyist since the mid-80's, progressing to scientific applications software development in the late 80's. Expertise in machine learning, natural language processing, and statistical modeling with proven ability to deliver production-ready solutions for complex business problems.

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
- Technical liaison with data engineering for production model deployment

Individual Investor Group (2024)

- Machine learning solutions for taxpayer compliance verification
- Model development and evaluation for classification tasks

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 modularization 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
Specializations: 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