James **Dao**

Sydney, Australia

✉ james.kn.dao@gmail.com | ⌂ jmsdao.github.io | ◯ jmsdao | in jmsdao

Skills

Certifications	AWS Certified Cloud Practitioner a,
Programming	Python (numpy/pandas, PyTorch, matplotlib, scikit-learn), R (tidyverse, ggplot2, Bioconductor), Markdown/HTML/CSS
Tools/Services	Linux (WSL2, conda, git), Docker, Jupyter, AWS (EC2, S3, IAM, Lambda)

Work Experience

Longaevus Technologies *Remote*
BIOINFORMATICIAN (PART-TIME) *Mar 2022 - Present*

- Exploratory analysis of public datasets.
- Building web-based tools in R Shiny for analysts.

Tibra Capital *Sydney, Australia*
JUNIOR QUANT TRADER *Feb 2019 - Feb 2020*

- First junior trader of the cohort to be greenlit to trade solo.
- Managed market making operations of equity and index derivatives on the ASX and HKEX.
- Collaborated with quant researchers, software devs and key stakeholders to develop trading strategies, and improve efficiency of operations.

University of New South Wales *Sydney, Australia*
LAB DEMONSTRATOR *Jul 2016 - Jun 2018*

- Taught 1st to 3rd year laboratory sessions for electrical engineering students.
- Engaged in project management for student capstone projects.

Publications

2022

- Ilieva, M.; **Dao, J.**; Miller, H.E.; Madsen, J.H.; Bishop, A.J.R.; Kauppinen, S.; Uchida, S. Systematic Analysis of Long Non-Coding RNA Genes in Nonalcoholic Fatty Liver Disease. Non-Coding RNA 2022, 8, 56. https://doi.org/10.3390/ncrna8040056

Relevant Projects

Lipidomics Aging Clock (Ongoing)
IN COLLABORATION WITH THE BARSHOP INSTITUTE

- Tools used: Jupyter, Python (numpy/pandas, sklearn, statsmodels, matplotlib)
- Concepts covered: data wrangling/visualzation, PCA/dimensionality reduction, regularized linear models, tree-based learning, ensemble learning, cross-validation, feature importance analysis

Data Processing and RShiny Web App for LiverDB ◯
IN COLLABORATION WITH THE LAB OF CARDIOVASCULAR BIOINFORMATICS

- Tools used: Nextflow, Linux, R (shiny, edgeR, enrichR), Docker, HTML
- Concepts covered: parallelizing bioinformatics pipelines, DGE anaylsis, pathway enrichment analysis, interactive data visualization

Feedforward Neural Network on the MNIST Dataset Without Deep Learning Libraries
HTTPS://COLAB.RESEARCH.GOOGLE.COM/DRIVE/1FNRZCUNTFT53NKYN5XAXDXSZGDD8V__H

- Tools used: Google Colab, Python (numpy)
- Concepts covered: deep learning, multi-layer perceptron, backpropagation, stochastic gradient descent
- Trains at <1 sec per epoch on standard Colab CPU, and achieves approx. 93% accuracy on the MNIST test set

Education

University of New South Wales *Sydney, Australia*
BACHELOR OF ENGINEERING (ELECTRICAL) *Feb 2014 - Nov 2018*