Duke Nguyen

dukeraphaelng.github.io
duke.nguyen.xyz@gmail.com
linkedin/duke-nguyen-syd
+61 410 705 820
Sydney NSW

SUMMARY

Research assistant, research graduate student in machine learning at the University of New South Wales, and former software engineer with strong background in backend development and database engineering. My research interests are Natural Language Processing (**NLP**), and Large Language Model (**LLM**). My Master's thesis was on formulating a unified framework for efficient transformers which utilise random features, supervised by **Dr. Aditya Joshi**.

SKILLS

Natural Languages: English, Japanese, Vietnamese
Formal Languages: Python, SQL, C, TypeScript, JavaScript, Crystal, Ruby, Latex
Libraries: LangChain, Pytorch, Sklearn, Pandas, Numpy, Matplotlib, React.js
Others: CI/CD, Docker, Cloud (AWS/ Azure), Statistics, Research, Documentation

PROJECTS

AI-Generated Text Detection

- Presented at ALTA 2023: **Duke Nguyen**, Khaing Myat Noe Naing, and Aditya Joshi. 2023. Stacking the Odds: Transformer-Based Ensemble for AI-Generated Text Detection. In *Proceedings of the 21st Annual Workshop of the Australasian Language Technology Association*, pages 173–178, Melbourne, Australia. Association for Computational Linguistics.
- Achieves a remarkable **95%** accuracy in detecting AI-generated text through an innovative ensemble of lightly fine-tuned Transformer models, boasting rapid training times without compromising performance. Serves as a robust solution to counter malicious use of text-generating AI, ensuring digital content integrity.

Fake News Detection

- Enhances current detection techniques, helping combating against misinformation, by integrating BERT and linguistic features (selected via ANOVA and Pearson correlation) with a novel similarity measure against topically aligned articles, as validated by Google News' PageRank. Grade: 90%

Authorship Identification

- Won the **Innovative Applications of AI Award** at Berrijam Jam April 2023 with *Team Spirit*.
- Tackled authorship identification on the Standardized Project Gutenberg Corpus, achieving significant results with a **95.12%** accuracy using a SDAE with SVM, surpassing the performance of finely-tuned BERT models.

EXPERIENCE

Research Assistant   2023 - 2024
UNSW, Sydney NSW

- Executes complex research tasks as part of a project funded by **Google's** exploreCSR grant. Tasks include data collection, analysis, reporting, and applying machine learning models, natural language processing (NLP) techniques, large language models to achieve research objectives.

Casual Academic 2024 - 2024
UNSW, Sydney NSW
- Led machine learning tutorials for 30 students across three groups, covering topics such as linear models, tree learning (kNN), ensemble learning, unsupervised learning, and neural networks; graded assignments for 60-80 students with detailed, constructive feedback to enhance understanding.

JavaScript Developer 2021 - 2022
QSStudio, Sydney NSW
- Contributed to the development, maintenance, and enhancement of the Servo-Track API, including authentication/authorization for an IoT cluster, serverless and Lambda application deployments, and CI/CD (Bitbucket, Docker).

Software Engineer 2020 - 2021
PlaceOS, Sydney NSW
- Engineered a cutting-edge e-commerce platform for IoT technologies and services (crystal, PostgreSQL, Docker, Azure AD B2C, Xero).
- Architected and managed a robust CI/CD pipeline utilizing Travis, GitLab, and GitHub Actions, paired with Docker for containerization. Additionally, maintained critical databases, ensuring data integrity and availability.

Software Engineer Intern 2020 - 2020
PlaceOS, Sydney NSW
- Collaborated on the design and development of a financial micro-service; leading the standardization of technical documentation format along IEEE standards.

EDUCATION

Master of Information Technology (Artificial Intelligence), Distinction 2022
University of New South Wales, Sydney NSW
- *Courses*: Neural Networks & Deep Learning, Machine Learning & Data Mining, Computer Vision, Applied AI, Big Data Management, Artificial Intelligence.

Diploma of Information and Technology, Distinction 2020
Academy of Information Technology (Coder Academy), Sydney NSW

Bachelor of Arts, Distinction 2019
University of Wollongong, Wollongong NSW
- Faculty Merit Scholarship & Undergraduate Excellence Scholarship (2016-2019)

WORKSHOPS

Speaker
- Research Seminar Association, Kernel Learning in Attention
- Crystal 1.0 Conference - Macros in Practice

Participant
- Collective and Robust Ubiquitous Intelligence (CRUISE) Lab
- Rainbow AI - Australia

AWARDS

IBM Quantum Certificate of Quantum Excellence in Quantum Machine Learning
IBM Quantum Advanced Award (Fall 2021, Africa 2021)

LIBRARIES

Data structure - Merkle Tree
Standards - JSON Web Key Set, Code Documentation Generator
Interface - etcd Client, PostgreSQL Client, HTTP Router, Microsoft Graph Client