Duke Nguyen

unsw/duke-nguyen Google Scholar duke.nguyen@unsw.edu.au +61 410 705 820 linkedin/duke-nguyen-syd Sydney NSW

SUMMARY

PhD student in natural language processing (NLP) at the University of New South Wales, researching on NLP for English dialects.

RESEARCH

Chan, Fai Leui, **Duke Nguyen**, and Aditya Joshi. 2014. "Is Hate Lost in Translation?": Evaluation of Multilingual LGBTQIA+ Hate Speech Detection. arXiv preprint arXiv:2410.11230.

Duke Nguyen, Aditya Joshi, and Flora Salim. 2024. Spectraformer: A Unified Random Feature Framework for Transformer. arXiv preprint arXiv:2405.15310v3

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, Melbourne, Australia. Association for Computational Linguistics.

EXPERIENCE

Casual Academic

2024 - Now

UNSW, Sydney NSW

Invigilated exams; led tutorials and graded assignments for the following courses

- COMP9444, covering neural network and its applications (e.g., CNN, LSTM, transformer, deep reinforcement learning, auto-encoders, adversarial learning).
- COMP9414, covering autonomous agents, optimisation, knowledge representation, reinforcement learning.
- COMP9417, covering linear models, tree learning (kNN), ensemble learning, unsupervised learning, and neural networks along with rigorous mathematical foundations.

Research Assistant

2023 - 2024

UNSW, Sydney NSW

• Executed 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, NLP techniques, large language models.

JavaScript Developer

2021 - 2022

QSStudio, Sydney NSW

• Contributed to the development of the ServoTrack 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. Architected and managed a robust CI/CD pipeline utilizing Travis, Git-Lab, and GitHub Actions, and Docker; maintained critical databases.

Software Engineer Intern

2020 - 2020

PlaceOS, Sydney NSW

• Collaborated on the development of a financial micro-service; standardized technical documentation using IEEE standards.

EDUCATION

Doctor of Philosophy (Computer Science and Engineering) University of New South Wales, Sydney NSW

2028

• University Postgraduate Award (2024-2028)

Master of Information Technology (Artificial Intelligence), Excellence Award 2024 University of New South Wales, Sydney NSW

- Thesis: Spectraformer, a unified framework for fast Transformers (via RF), helps discover novel SOTA fast transformer models.
- Research Project: 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.
- Research Project: 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.

Diploma of Information and Technology, Distinction Academy of Information Technology, Sydney NSW

2020

Bachelor of Arts, Distinction

2019

University of Wollongong, Wollongong NSW

• Faculty Merit Scholarship & Undergraduate Excellence Scholarship (2016-2019)

WORKSHOPS

Speaker

• Crystal 1.0 Conference - Macros in Practice

Membership

- Collaborative Human-Centric AI Systems (CRUISE) Lab
- HCC ML NLP Reading Group (UNSW)

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

SKILLS

Natural Languages: English, Japanese, Vietnamese

Formal Languages: Python, SQL, C, TypeScript, JavaScript, Crystal, Ruby, Latex

Libraries: Pytorch, Sklearn, Pandas, Numpy, Matplotlib, React.js

Others: CI/CD, Docker, Cloud (AWS/Azure), Statistics, Research, Documentation