NEATON ANG

(862) 400-8339 | neaton.ang@columbia.edu | https://www.linkedin.com/in/neatonang/ | NY, USA

EDUCATION

Columbia University, New York, NY   **Expected Dec 2026**
MS in Data Science (GPA: 3.73/4.00) - Agentic AI & Data Economy, NLP Topic Modelling Research

National University of Singapore (NUS), Singapore   **Aug 2019 – May 2023**
B. Hons Data Science & Analytics (CGPA 4.25/5.00)
- NUS Sports Scholar (Mid-Term)
- Nominee of Lee Hsien Loong Outstanding All-Round Achievement Award
- Semester Exchange at Technical University of Münich
- NUS Office of Admissions Campaign Ambassador (https://tinyurl.com/24nnmm9u)

PROFESSIONAL EXPERIENCE

DSO National Laboratories, Singapore   **Dec 2023 – Present**
Member of Technical Staff - Natural Language Processing Researcher
- Part of the team that successfully pre-trained multilingual 3B, 8B, 8x12B and 70B parameter LLMs from scratch
- Researched scalable data selection framework for LLM pre-training, decreasing noise and boosting benchmark scores in downstream tasks
- Demonstrated competency to work with large datasets (upwards of 1 Trillion tokens) via parallel computing methods and job schedulers
- Improved code throughput by optimising data processing pipelines via parallel computing methods (12x faster)
- DSO's social media feature: https://tinyurl.com/mw3jvn3r

Ethlas/Eleos Labs, Singapore   **May 2023 – Dec 2023**
Data Scientist - Generative AI
- Spearheaded company's Gen AI initiatives as the main Data Scientist (https://tinyurl.com/3wf6tcx5, https://tinyurl.com/4upyjhfj, https://tinyurl.com/yjj4ppc7, Speaker at Google Cloud event - https://www.youtube.com/watch?v=FIR670WYGfY)
- Fine-Tuning of Stable Diffusion image generation model for game asset creation, reducing design team's manual work by 65%
- Implemented memory storage for contextual dialogues powered by GPT, between gaming characters (Information Retrieval via BERT)
- Implemented a Machine Learning model to detect cryptocurrency transaction fraud with ~88% recall
- Implemented AI Personas into Non-playing characters to have open-ended conversations with users powered by GPT (RAG, LangChain)

DSO National Laboratories, Singapore   **Dec 2022 – May 2023**
Machine Learning (ML) Intern - NLP
- Created Fact Checking + Fake News detection pipeline via Large Language Models (GPT-3, BERT, BLOOM, FLAN-T5)
- Dealt with large unstructured data, perform zero-shot classification, few-shot classification, semi-supervised & unsupervised learning
- Prompt engineering with Large Language Models (GPT-3, BLOOM, FLAN-T5), improving current model by 5% accuracy
- Large Language Model fine-tuning with specific datasets for task-specific classification (PyTorch, TensorFlow)
- Improved fact checking and fake news detection via web-scraping for live updated evidence (Google, Bing, Wikipedia)

EXTRA CURRICULAR ACTIVITIES

National/Varsity Badminton, National University of Singapore   **Aug 2019 – May 2023**
- Represented Singapore in World University Games 2023 (https://tinyurl.com/mrx95rph)
- Represented NUS in Inter-Varsity Poly-ITE Games (2nd runner-up in 2019), Singapore University Games (Champions in 2019, 2022)

Badminton, Singapore Police Force, Singapore (National Service)   **Apr 2017 – Feb 2019**
- Hand-picked to represent Singapore Police Force in Inter-Forces Games, held in Hong Kong and Malaysia
- Achieved 1st in Men Doubles and 2nd in Men Singles in Intra-Singapore Police Force Badminton Tournament 2018

LEADERSHIP AND ACTIVITIES

National University of Singapore Badminton Varsity Team (Team Captain)   **May 2020 – June 2022**
- Managed the competitive badminton team and organised engagement activities for the broader NUS community to foster a sporting culture
- Launched NUS Recreational Badminton, currently with over 3,500 members

National University of Singapore Students' Sports Club Management Committee   **Jan 2020 – Jan 2021**
- Vice-Project Director of Marketing for the annual Inter-Faculty Games, the largest institution-wide project
- In charge of marketing efforts, such as contacting, negotiating and maintaining good relations with sponsors

Youth Executive Committee (Serangoon CC)   **Jan 2015 – June 2016**
- Joined the grassroots youth committee to organise events to provide food aid to the less privileged in Singapore and smartphone classes

TECHNICAL COMPETENCIES

Analytics/Programming: Python, R, Power BI, Tableau, SQL, MATLAB, Java, Slurm
Personal Projects: Natural Language Processing Projects/Visualisation - (https://bit.ly/3s4hpeh, https://bit.ly/3ggWjqD, https://bit.ly/3S3LxRK)
Hackathon: Shopee Code League – Trained an image classification model to identify and categorise product images for sellers on Shopee. Solved coding problems given by Shopee in Python and Java