(+44) 07851241799
kathyka711@gmail.com

Ting-Wen (Kathy) Ko

Ting-Wen Ko
kerkathy

Research Interest

Causal reasoning, trustworthy AI, AI safety, model interpretability

Education

M.Sc., Computational Statistics and Machine Learning, University College London (in London, UK) **Sep 2024 — Sep 2025**

- Courses: Probabilistic and Unsupervised Learning, Supervised Learning, Bayesian Deep Learning, Statistical Model and Analysis, Open-Endedness and General Intelligence, Statistical NLP
- Supervised by Mengyue Yang and Jun Wang.

M.Sc., Computer Science, National Taiwan University (in Taipei, Taiwan) **Sep 2022 — Jun 2024**

- GPA: 4.17/4.3
- Supervised by Pu-Jen Cheng and Jyun-Yu Jiang.
- Thesis: Enhancing Retrieval Augmented Generation with Passage Combination.

B.Sc., Undergraduate Honors Program of Electrical Engineering and Computer Science **Sep 2018 - Aug 2022**

National Yang Ming Chiao Tung University (Previously National Chiao Tung University) (in Hsinchu, Taiwan)

- Overall GPA: 3.93/4.3

Research and Work Experience

University College London **Apr 2025 — Current**

Master's Thesis Research

- Exploring spurious correlation in LLMs from the perspective of OOD detection.

MediaTek Research UK **Jun 2024 — Aug 2024**

Deep Learning Intern

- Researched on the model representation when LLM reasons, layer optimization through adaptive pruning, and diffusion model techniques to improve transformer-based language models.

National Taiwan University **Sep 2023 — Jun 2024**

Master's Thesis Research (Link: 🔗) (Code:)

- Developed an end-to-end trainable dense passage retriever that optimizes passage combination selection for retrieval-augmented QA with LLM
- Implemented comprehensive baseline systems, including sparse/dense retrieval and reranking methods

Global Media Team, Yahoo! **Sep 2023 — Dec 2023**

Project Intern (Code:)

- Researched on multitask learning approaches for text readability assessment.
- Implemented unsupervised learning-to-rank to predict a comprehensive readability signal.

Global Media Team, Yahoo! **Jul 2023 — Aug 2023**

Research Engineering Intern

- Designed an internal LLM chatbot with vector database integration for Yahoo news articles.
- Applied retrieval-augmented generation methods for improved explainability and diverse search result by max marginal relevance search and asymmetric embeddings.

Selected Projects

Deconstructing LLM Faithfulness (Python) **Feb 2025 — Current**

- Revisited existing faithfulness tests from a causal framework.
- Investigate the application of RL techniques to train language models using faithfulness as a reward metric.
- Research areas: causality, RL fine-tuning, AI safety, trustworthy AI

BANDITPROMPT: Steering Creativity in Image Generation (Python / LLMs / Diffusion Models) **Feb 2025 — Current**

- Developed a novelty search-based approach for generating diverse sets of images using LLM prompt optimization.
- Implemented a Beta-Thompson multi-armed bandit algorithm to adaptively select specialized mutation strategies.
- Conducted comprehensive evaluations demonstrating superior diversity metrics compared to existing baselines.
- Research areas: Generative AI, evolutionary algorithms, multi-armed bandits, prompt optimization

Teaching Experience

System Programming Teaching Assistant (C) **Sep 2022 — Dec 2022**
National Taiwan University

- Designed and implemented a simulation of a context switch system utilizing non-local jumps and signals in a class assignment.
- Conducted TA sessions, providing guidance and support to 100+ students on academic coursework.

Service

Student Volunteer **Jul 2023**
The 46th International ACM SIGIR Conference

Awards & Honors

2024 Phi Tau Phi Scholastic Honor Society of the Republic of China Honorary Membership
2022 Ministry of Education (MOE) Overseas Exchange Student Financial Assistance Grant (**US $4.6K**)

Skills

ML/AI	Large Language Models, Retrieval-Augmented Generation, Language Modeling, Transformer Architectures, Causal Inference, Evolutionary Algorithms
Programming	Python, PyTorch, HuggingFace, C/C++, Java, SQL
Toolkits	Git, Linux, Pandas, NumPy, Scikit-learn, NLTK, Streamlit

Last updated April 23, 2025