Phudish Prateepamornkul

+66894941342 | phudishp@gmail.com | phudish_p@berkeley.edu | linkedin.com/in/phudish-prateepamornkul

EDUCATION

University of California, Berkeley California, U.S.
Master of Arts in Statistics, GPA:3.923 *Aug. 2023 – Dec 2024*
- Relevant Modules: Probability and Statistics at an Advanced Level, Deep Unsupervised Learning, Natural Language Processing, Computer Vision, LLM Agents and Convex Optimization

University of Warwick Coventry, U.K.
BSc (Hons) Mathematics, First Class Honours (83.02%) *Sep. 2020 – Jul 2023*
- Overall Rank: Top 10 of cohort from 200+
- Relevant Modules: Linear Algebra, Multivariable Calculus, Mathematics of Machine learning, Matrix Analysis & Algorithms, Programming for Data Science, Neural Computing, Linear Statistical Modelling, Multivariate Statistics

EXPERIENCE

AI Scientist Jan 2025 – Present
SCB DataX *Bangkok, Thailand*
- Implemented an agentic AI chatbot with LangGraph and RAG to map deposit-related natural language queries to SQL, improving SCB finance inquiry processing and contributing to a 5M THB uplift
- Built a conversational AI chatbot that enables users to handle financial inquiries and carry out operations such as transferring money or paying credit-card bills through natural-language or voice interaction

Graduate Research Assistant Oct 2024 – Present
Berkeley AI Research (BAIR) *California, U.S.*
- Collaborating in the Speech Lab under the guidance of **Akshat Gupta** to develop techniques for editing knowledge in large language models (LLMs) and researching about continually learning AI-Scientist for AI-Research with small open-source models
- Published a **research paper** accepted to EMNLP 2025 Finding presenting a novel approach that scales knowledge editing to support up to 10,000 modifications while maintaining robust downstream performance

PUBLICATIONS

- **Lifelong Knowledge Editing Requires Better Regularization** EMNLP 2025 Findings
  Akshat Gupta*, **Phudish Prateepamornkul***, Maochuan Lu, Ahmed Alaa, Thomas Hartvigsen, Gopala Anumanchipalli

* Equal contribution.

PROJECTS

Learning Multimodal Representations in Token Space Mar 2024 – May 2024
- Extended M3AE architecture to text-audio-video triplets using tokenized representations from pretrained models (BERT, custom VQ-VAE, VideoGPT), training multimodal Transformers on Meta's MUGEN dataset that achieved lower reconstruction loss across all modalities compared to unimodal baselines at masking ratios from 50-95%
- Demonstrated that multimodal training in token space improves single-modality performance, with the model outperforming unimodal baselines even when evaluated unimodally, while reducing computational overhead by avoiding raw data processing

Mixture of Large Language Models Jan 2024 – May 2024
- Merged multiple 130M-parameter GPT-2 models using Linear, TIES, and DARE techniques, achieving 96% of GPT-2 XL's accuracy (65.3% vs 68%) while being 16x faster at inference (142s vs 2255s)
- Demonstrated the effectiveness of merged models achieving 58% accuracy on out-of-domain classification tasks a 6 percentage point improvement over individual fine-tuned models

TECHNICAL SKILLS

Languages: Java, Python, SQL, HTML/CSS, R
Frameworks: Flask, FastAPI, LangGraph, LangChain
Developer Tools: Git, Docker, Azure, VS Code
Libraries: NumPy, Pandas, SciPy, PySpark, Scikit-learn, PyTorch, HuggingFace, OpenCV