Yeqiao Fu

Tel: +86 138 1118 3009; +852 5640 9355 | Email: fyq5166@gmail.com

EDUCATION

University of Hong Kong (HKU)	**Hong Kong, China**
BEng in Computer Science, double major in Finance	Sept. 2022- Jun. 2026 (Expected)

- CGPA: 3.70/4.30
- **Relevant coursework:** Artificial Intelligence, Linux Based Data Analysis, R Based Data Visualization and Analysis, Computer Programming in Python, Computer Programming in C++, Object-oriented Programming and Java, Data Structures and Algorithms, Calculus and Ordinary Differential Equations, Linear Algebra, Probability and Statistics

RESEARCH EXPERIENCE

Machine Learning and the Application within Cyber Incident Response	**OSRP Cambridge**
Independent research, supervised by Professor Kieren Lovell at Cambridge University	Jun. 2023 - Sep. 2023

- Text Preprocessing Excellence: Developed robust text preprocessing workflows using the NLTK library, including tokenization, POS tagging, lemmatization, and stop-word filtering, while optimizing downstream text vectorization for improved model performance.
- Advanced Text Analysis: Implemented state-of-the-art GloVe and Word2Vec models to create comprehensive text representations, capturing intricate semantic relationships and contextual nuances in textual data.
- Scam Evaluation Framework: Designed a sophisticated six-category scam evaluation system based on the STP-II model, enabling detailed analysis and categorization of potential threats.
- Predictive Modeling: Utilized Sklearn's multi-output regression and iterative model tuning to forecast scam threat levels with enhanced accuracy, delivering reliable predictions for real-world applications.
- Achievements: Achieved a 90% accuracy rate in predicting features of less abstract scam attributes, showcasing the model's high reliability and precision.

Spider2-V: How Far Are Multi-modal Agents From Automating Data Science and Engineering Workflows?	**XLANG Lab HKU**
Research assistant, supervised by Assistant Professor Tao Yu at HKU	Mar. 2024 - Jun. 2024

- Benchmark Development: Designed and implemented a benchmark to evaluate multi-modal agents' ability to automate data science workflows, featuring 494 real-world tasks across 20 enterprise-level applications.
- Performance Optimization: Improved baseline success rates from 14.0% to 35.0% through grounding algorithm optimization and prompt tuning, demonstrating significant advancements in automating complex workflows.
- Model Training: Trained and evaluated vision-language models in Linux-based environments, focusing on tasks requiring both code generation and GUI operations.
- Algorithm Development: Developed grounding algorithms to enhance performance on multi-step data science tasks, addressing challenges in executing fine-grained, knowledge-intensive GUI interactions.
- Prompt Engineering: Applied advanced prompt tuning techniques to refine model responses, improving accuracy and context awareness in professional tool interactions.
- Publication: Co-authored Spider2-V: How Far Are Multimodal Agents From Automating Data Science and Engineering Workflows, featured as a spotlight presentation at the NeurIPS D&B track in September 2024.

Interactive Website Modeling and Trajectory Generation	**XLANG Lab HKU**
Research assistant, supervised by Assistant Professor Tao Yu at HKU	Jul. 2024 - Present

- World Modeling: Constructed interactive world models for websites to enable AI agents to explore and record interaction outcomes, capturing structured intra- and inter-page data with ally-trees and URL maps.
- Pipeline Development: Designed a comprehensive pipeline for trajectory generation, including exploration, task execution, and evaluation.
- Efficiency Enhancements: Reduced redundancy in web interaction via component classification (mechanical & semantic algorithms) and URL navigation reduction.
- Trajectory Collection: Achieved 100% grounding and replay success using ally-tree-based localization, collecting high-level descriptions and step-by-step trajectories to guide task generation.
- Evaluation Metrics: Designed an AI evaluator to assess trajectories based on Completeness, Complexity, Conciseness, Concreteness, and Diversity, achieving alignment with human evaluation (68–74%) using 4omini.
- Scalable System Design: Developed modular and extensible project structures with automated pipelines from configuration to evaluation.
- Future Directions: Exploring vision-based approaches for web interaction and evaluating world model methods on established benchmarks to refine agent performance.

SKILLSET&INTERESTS

Languages: Mandarin (Native), Cantonese (Intermediate), English (fluent, IELTS: 8)
Technical Skills: Python (Intermediate), R (Intermediate), C(Basic), C++ (Intermediate), Java(Basic); Database: MySQL (Intermediate); Operating System: Windows, Linux
Interests: gym, reading, racing, diving, eating