Keyang Xuan

keyangds.github.io linkedin.com/in/keyang-xuan/ keyangx3@illinois.edu

Education

University of Illinois, Urbana-Champaign Aug 2023 - Dec 2024 (Expected)
Master in Computer Science *GPA: 4.0/4.0*

University of Minnesota, Twin Cities Sep 2019 - May 2023
Bachelor of Science in Computer Science, Data Science ***(Dual Degree)*** *GPA: 3.92/4.0*

Relevant Coursework: Algorithm and Data Structure, Discrete Mathematics, Theory of Probability, Applied Statistics, Operating System, Adv. Machine Learning, Optimization for ML, Mathematical Modeling, DB Systems, Deep Learning in NLP, Computer Vision, Adv. NLP, Distributed System, Web Programming, Text Information System, Data Mining Principle, ML System

Skills

Languages: C/C++, Java, Python, JavaScript, HTML/CSS, MATLAB, R, Node.js, LaTeX
Tools: Git/GitHub, Unix Shell, VS Code, IntelliJ IDEA, Atom, R Studio, Postman

Publication

Bai,W.,**Xuan,K.**,Huang,P.,Wu,Q.,Wen,J.,Wu,J.,Lu,K. **APILOT: Navigating Large Language Models to Generate Secure Code by Sidestepping Outdated API Pitfalls**. *Under Review*, 2024.

Xuan,K.,Yi,L.,Yang,F.,Wu,R.,Fung,Y.,Ji,H. **LEMMA: Towards LVLM Enhanced Multimodal Misinformation Detection with External Knowledge Augmentation**. *KnowledgeLM @ ACL*, 2024.

Ravirathinam,P.,Ghosh,R.,Wang,K.,**Xuan,K**.,Khandelwal,A.,Dugan,H., Hanson,P. and Kumar,V. **Spatiotemporal Classification with limited labels using Constrained Clustering for large datasets**. *SDM*, 2023.

Projects

LVLM-Based Multimodal Misinformation Detection | *GPT4V, Duckduckgo* Sep 2023 – Feb 2024

- Led the research in conducting empirical study of evaluating Large Vision Language Model performance on Multimodal misinformation detection.
- Developed a novel LVLM-based solution for multimodal misinformation detection by integrating external knowledge from search engine with LVLM reasoning capability, which demonstrates a better performance on two representative datasets in the field.

Self-Created Face ID Verification System | *PyTorch, Numpy, Matplotlib, Github* Jan 2023 – May 2023

- Led a team of four in the conceptualization and exploration of a computer vision-related project
- Developed a cutting-edge deep learning algorithm (Siamese NN + Auto-encoder) and implemented it using PyTorch, simulating the functionality of Face ID

Optimization Method with Deep Learning | *PyTorch, Numpy, Matplotlib, Gitub* Jan 2023 – May 2023

- Implemented Anderson acceleration method as a new deep learning optimizer in PyTorch version
- Compared with typical optimization methods (Adam, Momentum, SGD) and develop global strategy to improve Anderson acceleration's performance
- Wrote survey paper about reviewing typical deep learning optimization algorithms

Cross-lingual Irony Detection | *Python, Transformers, Huggingface, Gitub* Sep 2023 – Dec 2023

- Explored potential outperforming on irony detection by using multilingual model by comparing the performance between several pretrained model (XLM-Roberta, GPT2, Bert).
- Implemented soft-prompting technique to increase the performance of multilingual model on irony detection.

Spatiotemporal Classification on ReaLSAT | *PyTorch, Matplotlib, Geopandas, Github, QGIS* Jun 2022 – Jun 2023

- Implemented Deep Learning Model (CNN, LSTM, ResNet) to classify global water body on ReaLSAT dataset.
- Explored the feature of either spatial information or temporal information on each type of waterbody based on model performance

Experience

National Renewable Energy Laboratory | *Research Intern* Jan 2024 – Present
Investigate grid edge resource flexibility prediction with social impact integrated.

University of Minnesota Security Lab | *Research Assistant* Jun 2023 – Present
Designing filter based moderation on Copilot to improve robustness of code recommendation
Crawling the commit information from Github for issue collection

University of Minnesota Data Mining Lab | *Research Assistant* Jun 2022 – Jun 2023
Conducted Spatiotemporal classification research and published paper on top data mining conference
Applied constrained loss and Spatiotemporal modalities to help improve global waterbody classification

University of Minnesota Neuroscience Department | *Research Assistant* Apr 2022 – Dec 2022
Conducted research related to brain and behavior relationships
Designed motion tracking algorithm based on Fourier Transform to track mice motion status

CITIC Future | *Quantitative Analyst Intern* Jun 2021 – Aug 2021
Implemented Double Moving Average Model and rigorously assessing its suitability across diverse commodity futures markets

Shanghai Metro Data Tech | *Data Analyst Intern* May 2021 – Jun 2021
Crafted SQL queries to precisely extract essential data in alignment with BI product design requirements

University of Minnesota Math Department | *Teaching Assistant* Sep 2020 – May 2021
Facilitated structured discussion and review sessions, guided students through worksheet completion, and contributed to the grading process for weekly assignments and exams