Keyang Xuan

Minneapolis, MN 55455| (651) 246 4533| xuan0008@umn.edu | LinkedIn: https://www.linkedin.com/in/keyang-xuan/

EDUCATION

University of Minnesota – Twin Cities (College of Science and Engineering) **Expected Graduation: June 2023**

- B.S., Data Science & B.S., Computer Science, **CGPA 3.91**
- Algorithm and Data Structure, Discrete Mathematics, Operating System, Advanced Machine Learning, Optimization for Machine Learning, Applied Statistics, Mathematical Modeling, Practice of DB Systems, Deep Learning in NLP, Computer Vision

Research Interest

My research interest broadly spans Deep Learning, Environmental AI and Natural Language Processing. My current interests primarily include finding more reasonable training setting in deep learning, using prompting technique to improve reasoning of language AI, explainable AI and deep learning application in environment.

Publication

Ravirathinam, P., Ghosh, R., Wang, K., **Xuan, K.,** Khandelwal, A., Dugan, H., Hanson, P., & Kumar, V. (2022). Spatiotemporal Classification with limited labels using Constrained Clustering for large datasets. Submitted to SDM 2023. [Link]

Computer Skills

Programming Language: Java, C++, JavaScript, Python, MATLAB, C, OCaml, Assembly, Julia
IDE: VSCode, Spyder, IntelliJ, Pycharm
OS: Linux, Windows, MacOS

WORK EXPERIENCES

Research Assistant **Department of Computer Science & Engineering, University of Minnesota, Twin Cities** **June 2022 – Present**

- Working in Prof. Vipin Kumar's lab to assist in data mining research on global waterbody dataset (ReaLSAT)
- Implementing deep learning model (CNN, ResNet, LSTMCNN) based on either spatial information or temporal information to classify typical types of water body.
- Our research paper "Spatiotemporal Classification with limited labels using Constrained Clustering for large datasets" right now is under review in 23rd SIAM International Conference on Data Mining (SDM23) https://arxiv.org/abs/2210.07522

Research Assistant. **Neuroscience Department, University of Minnesota, Twin Cities** **April 2022 – Present**

- Working in Prof. Timothy Ebner's lab to assist in research about exploring further relations between brain and behavior.
- Implementing DeepLabCut and design motion tracking algorithm based on Fourier Transform to track mice motion status.

MURAJ Reviewer **Minnesota Undergraduate Research & Academic Journal.** **October 2021 – March 2022**

- Peer reviewer for checking computer science field undergraduate research work's qualification in comprehensive scope.
- Anonymously provide revision suggestions to the author to help them strengthen their papers' logic, level of reader friendly.

Quantitative Analyst, Intern **CITIC Futures, Research Department** **June 2021 – August 2021**

- Build multiple regression in R to assist with Ferrous Metal Future Price Analysis, optimizing model based on R^2 and MSE performance.
- Implementing Double Moving Average Model and testing model's applicability on various types of commodity futures.

Data Analyst, Intern **Shanghai Metro Data Tech** **May 2021 – June 2021**

- Using company self-designed application to assist BI product design.
- Write SQL to extract necessary data based on BI products design needs.

Teaching Assistant **University of Minnesota, Math Department** **September 2020 – May 2021**

- Courses: 'Precalculus I', 'Precalculus II'.
- Supervised discussion and review sessions; Instructed students to complete worksheets; Assisted weekly assignment and exams grading.

SELECTED PROJECTS

Optimization Method with Deep Learning **September 2022 – Present**

- Implement Anderson acceleration method as a new deep learning optimizer in Pytorch version.
- Comparing it with typical optimization methods (Adam, Momentum, SGD) and figure out global strategy to improve Anderson acceleration's accuracy.

Tools: Python (PyTorch, Numpy, Matplotlib)

Cross-lingual Transfer Learning on Irony Detection **September 2022 – Present**

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

Tools: Python (Transformers, Numpy, PyTorch, Matplotlib)

Spatial-temporal Classification on ReaLSAT **June 2022 – Present**

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

Tools: Python (PyTorch, NumPy, Matplotlib, Sklearn.linear_model, Geopandas), QGIS

Simplified Commend-Line Shell **February 2022 – March 2022**

- Create a simplified shell which includes basic function including tokenization, command running, redirection, signal management and basic execution of program pipelines

Tools: C

California Housing Price Analysis **November 2021 – January 2021**

- Individual work on California Housing Price Analysis aims to select the most significant factors from the database.

- Optimized the penalty term's weight in shrinkage method based on plotting K-fold method cross-validation result.
- Comparing the performance of shrinkage methods with the simple linear regression model by comparing MSE and R^2 value.

Tools: Python (NumPy, Matplotlib, Sklearn.linear_model, Statsmodels)

Double MA Strategy Optimization **July 2021 – August 2021**

- Optimize time-length parameter based on practicing model in various commodity futures' price data in past decades.
- The work figures out the best combination of long average and short average depending on different futures' feature.

Tools: Python (NumPy, Matplotlib)

AWARDS

- Dean List of 2019, 2020, 2021
- Complete Bachelor Degree with High Distinction