Xiao Shi

Madison, WI
Tel: +1 412 909 6275/ E-mail: xiao.shi@wisc.edu / Webpage: https://xiaoshi2610.github.io/

Education History

Ph.D Student | *Electrical and Computer Engineering*
University of Wisconsin-Madison, Madison, WI 01/2026-Present
Advisor: Prof. Ulugbek S. Kamilov

Master of Science | *Engineering Data Analytics & Statistics* 08/2024-12/2025
Washington University in St. Louis, St. Louis, MO
GPA: 3.84/4.00

Master of Science | *Management* 09/2022-06/2024
Carnegie Mellon University, Pittsburgh, PA
GPA: 3.78/4.00

B.Sc. in Economics | *Finance*
Central University of Finance and Economics, Beijing, China 09/2018-06/2022
GPA: 3.61/4.00

Awards

CMU Heinz College Dean List 2023, 2024
Carnegie Mellon University
Pittsburgh, PA

Publications

Preprint
EigenScore: OOD Detection using Covariance in Diffusion Models.
S Shoushtari, Y Wang, **X Shi**, MS Asif, US Kamilov
arXiv:2510.07206v1

Projects

Navigating Constraints to Provide Services in Allegheny County Jail 01/2024-05/2024
Advisor: Prof. Mariana Escallon Barrios
Python, Sklearn

- Utilized Python and Sklearn to set up a machine learning pipeline and optimized decision-making tools to identify individuals for service provision efficiently.

Large Language Model for the OpenWebText data 05/2024
Python, PyTorch

- A decoder-only (GPT structure) large language model designed for text summarization, question answering, sentiment analysis, and named entity recognition (NER) tasks. The model was pretrained on a subset of OpenWebText data (~65GB). We fine-tuned the model using additional datasets, including the CNN DailyMail Dataset and the Stanford Question Answering Dataset. Finally, we designed appropriate

prompts for the LLM to successfully complete all the aforementioned tasks.

Speech Recognition through a Transformer architecture network 04/2024
Python, PyTorch

- Developed an encoder-decoder and transformer-based CNN-LSTM network for transcribing spoken language into text. The network was pretrained on two sets of Mel Frequency Cepstral Coefficients (MFCCs) data to enhance its speech processing capabilities. The output from the pretrained network was then fed into a transformer encoder-decoder structure, effectively capturing deep relationships between words. This approach significantly improved the accuracy and efficiency of speech recognition.

VGGFace2 Dataset Face Classification and Verification 02/2024
Python, PyTorch

- A ConvNeXt-based CNN face classification network. The model was first pretrained on over 7000 samples to achieve high accuracy in the classification task. Then, fine-tuning was performed by removing the fully connected layer and implementing cosine similarity for the verification task. To classify or verify a face, the user simply needs to input it into the network and select the desired task. The network then generates results based on the user's selection.

Minority Language Captioning Contest 11/2022-12/2022
Python, PyTorch

- Developed an image captioning model using approximately 150,000 image-caption pairs from the Bloom Library. This model generates Hausa, Kyrgyz, and Thai descriptions for a given image, offering language support for companies looking to expand into these regions. Image features were extracted using Inspection V3, and an LSTM with Attention mechanism was employed to convert these image features into natural language descriptions.

Work Experience

Computational Imaging Group, WashU
Research Assistant 05/2023-09/2023
Saint Louis, MO

- Run experiments for an out-of-distribution detection project
- Research in Inverse Problem, Computational Imaging, Generative Model

PatSnap 05/2023-09/2023
Data Scientist ***Shanghai, China***

- Conducted data preprocessing, feature engineering, and set up an end-to-end machine learning pipeline for user persona and market prediction.
- Processed over 100,000 user data using Spark on AWS and maintained existing models.

Cinda Security 02/2022-06/2022
Data Analyst ***Beijing, China***

- Built linear regression and KNN models for bond interest rate pricing.
- Mined and preprocessed panel data, built time series models, and visualized results using Tableau.
- Established a VECM model to analyze financial data and predict stock price changes.

Related Courses

- **Graduate Level:**
  Dynamic System, Stochastic Process, Optimization, Bayes Learning, Time Series, Statistics Reasoning,

Machine Learning, Machine Learning with Big Dataset, Database Management, Econometrics, Deep Learning, Data Structure, Algorithm

- **Undergraduate Level:**
  Mathematic Analysis, Advanced Algebra, Probability, Statistics, Stochastic Process, Applied Time Series, Machine Learning & Pattern Recognition, Game Theory

Skills

Programming Language: Python, Java, R, MATLAB
Language: Mandarin (Native), English
Database: SQL