

Xueren Ge

•Tel: +1 470-662-2854 •Email: zar8jw@virginia.edu •Web: <https://xueren-ge.github.io/>

EDUCATION BACKGROUND

University of Virginia		<i>May. 2022 – Present</i>
Major: Computer Engineering	Degree: Doctor of Philosophy	GPA: 3.88/4.0
Relevant Courses: Benefits and risks of Large Language Models; Learning for Interactive Robotics		
Georgia Institute of Technology		<i>Aug. 2020 – May. 2022</i>
Major: Electrical and Computer Engineering	Degree: Master of Science	GPA: 3.91/4.0
Relevant Courses: Statistical Machine Learning; Natural Language Processing; Deep Learning; Convex Optimization;		
Chongqing University		<i>Sep. 2016 – Jun. 2020</i>
Major: Electrical Engineering and Automation	Degree: Bachelor of Engineering	GPA: 3.58/4.0
Relevant Courses: Computer Communication Networks; Intro to Database System; Advanced Programming Techniques;		
Scholarships: 2020 Georgia Tech Shenzhen Campus Level A “Merit-Based Scholarship” (5%);		
2018 Yangtze Power Scholarship (2/324);		
2016, 2018, 2019 Excellent Second Undergraduate Comprehensive Scholarship (6%)		
Awards:	2019 Chongqing University Excellent Undergraduate (10%);	
	2018 Chongqing University Excellent Student (5%);	
	2018 Chongqing University Science and Technology Innovation Advanced Individual (5%)	
Services:	Teaching Assistant for ECE Statistical Machine Learning 2022 Fall	
	Reviewer for ICRA 2024	
	Volunteer for EMNLP 2024	

PUBLICATIONS

Ge, Xueren, et al. "[DKEC: Domain Knowledge Enhanced Multi-Label Classification for Diagnosis Prediction](#)." *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*. 2024.

Weerasinghe, K., Janapati, S., Ge, X., Kim, S., Iyer, S., Stankovic, J.A., & Alemzadeh, H. (2024). [Real-Time Multimodal Cognitive Assistant for Emergency Medical Services](#). *2024 IEEE/ACM Ninth International Conference on Internet-of-Things Design and Implementation (IoTDI)*, 85-96.

Fang Xinxin, Wang Bingkai, Kong Hang, Ge Xueren, Yang Zhifang, Yu Juan, ... & Li Wenyan. (2023). [Human posture feature recognition method for neuropsychological comprehension test](#). *Journal of Chongqing University*, 46(4), 108-119.

Ma, Yuqing, and Xueren Ge. "[An Effective Method for Defect Detection of Copper Coated Iron Wire Based on Machine Vision](#)." *IOP Conference Series: Materials Science and Engineering*. Vol. 631. No. 2. IOP Publishing, 2019.

SKILLS

Programming Languages: Python, C/C++, Java, MySQL, HTML/XML, MATLAB

Machine Learning: PyTorch, Tensorflow, Wandb, Huggingface, Anaconda, Llama, GPT-4, Gemini, OpenCV

Others: Linux, Slurm, Docker, BeautifulSoup, Web APIs, OpenMP, OpenGL, Keil, Latex, Microsoft Suites

EXPERIENCES

University of Virginia	Research Assistant	Summer 2022 – Spring 2024
Intro.: Incorporate domain knowledge for multi-label classification in Electronic Health Records application		
➤ Developed an automated construction of heterogeneous knowledge graphs from multiple external websites (Wikipedia, MayoClinic and MedlinePlus) to alleviate imbalanced data distribution in EHRs		
➤ Designed a label-wise attention mechanism that incorporates heterogeneous knowledge graphs to train language models, to improve multi-label classification by 5% in micro f1 compared with SOTAs on MIMIC-III datasets		
➤ Used BeautifulSoup, Web API to scrape 3,000 medical web pages, extracted more than 5,000 medical entities by prompting LLMs (chain-of-thought) and normalized medical entities by UMLS API in Python		
SenseTime	Algorithm Developer Internship	Dec. 2020 – Jun. 2021

Xueren Ge

•Tel: +1 470-662-2854 •Email: zar8jw@virginia.edu •Web: <https://xueren-ge.github.io/>

Intro.: Designed a contactless algorithm to measure heart rate and body temperature in SenseThunder Air product

- Extracted Region of Interest from thermal images by landmark detection and homography based on OPENCV, then used FFT, smoothing and bandpass filters to extract weak breath rate, heart rate signals from continuous thermal images
- Built an end-to-end algorithm pipeline (Python), including data preprocessing, implementation, evaluation, error analysis

PROJECTS

LLMs for Diagnosis Prediction

Aug. 2024 – Dec.2024

Intro: Explored prompting and finetuning strategies for diagnosis prediction using Electronic Health Records (EHRs).

- Developed a “chain-of-diagnosis” prompt template, enabling effective finetuning of Llama-3.1-8B with LORA for enhanced diagnostic accuracy.
- Conducted comprehensive comparisons between our finetuned model and state-of-the-art Medical LLMs using chain-of-thought prompting to evaluate improvements in diagnosis prediction.

N-Version Programming on LLMs

Aug. 2023 – Dec.2023

Intro: Improve multiple-choice question answering accuracy by merging multiple LLMs

- Designed a majority voter and weighted majority voter to combine responses from GPT-3.5, PALM-1 and Llama-2 and Improved multiple-choice question answering by 5% in accuracy

BERT Visualization and Interpretation

Aug. 2022 – Dec. 2022

Intro: Analysis of whether BERT create reasonable embeddings in each layer

- Visualized anisotropy problem by randomly sampled two words and calculate cosine similarity in 12 layers
- Conducted biased analysis by removing sentences related to “female” concepts in SNLI training datasets
- Discussed the redundancy of BERT model by freezing each layer for classification problems

Phishing Websites detection based on Machine Learning

Jan. 2021 – May. 2021

Intro.: Designed phishing websites detection algorithms and a phishing web inquiry application

- Extracted and encoded URL features (domain, IP address, DNS record, redirection, protocol and etc.) by Python
- Trained SVM, RF, GDBT and etc. models separately, then built ensemble model to boost overall performances
- Developed a web [application](#) based on ensemble model interface and PhishTank API

Intelligent Diagnosis of Alzheimer's disease Based on AI Image Technology

Jul. 2019 – Nov. 2019

Intro.: Established an intelligent diagnosis system to assess the severity of Alzheimer's disease for patients automatically.

- Developed a human eye state recognition GUI based on eye aspect ratio and Dlib toolkit by Python
- Proposed a video analysis method of human posture estimation based on image morphology processing and Fast R-CNN
- Finetuned CNN with Resnet-50 as backbone by PyTorch to classify hand-drawn geometric figures

Research on Bus Station Waiting People Statistics Based on Wireless Interactive System

Jun. 2017 – Jul.2018

Intro.: Developed a wireless interactive terminal to get around the problem of difficulty in dispatching campus sightseeing buses in Chongqing University and random number of waiting people.

- Implemented wireless communication, LCD display and other logical C programming based on STM32 core
- Designed and soldered circuit board with STM32 as the core and ZigBee wireless transmitter, matrix keyboard module, power module and LCD display module as the peripheral circuit
- Modeled campus bus dispatching as TSP problem and derived optimal solution by ant colony algorithm on MATLAB

ACADEMIC COMPETITIONS

2018 Mathematical Contest in Modeling (MCM)

Meritorious Winner

Feb. 2018

- Data Cleaning: used the grey predictive modeling method to supplement the missing data, exclude outliers
- Used NARX autoregressive neural network to predict energy development and discussed the model Robustness
- Converted energy development strategy into nonlinear optimization problem and solved with Genetic Algorithm