Wenqian Ye

✆ (+1) 571-519-9957
✉ wy2029@nyu.edu
🖥 My Webpage
⚫ Github   in Linkedin

Research Interests

My main research interest is to develop improved representation learning methods that have solid interpretability based on the foundation of machine learning or probabilistic explanation. I am eager to persistently improve my levels of understanding in diverse areas of study including but not limited to 1). Bayesian Machine Learning, 2). Self/Semi-Supervised Learning, 3). Fairness/Robustness in Machine Learning, 4). AI for Social Good.

Education

2023 – Now   **PhD, Computer Science**, *School of Engineering and Applied Science*, University of Virginia.

2020 – 2022   **MS, Computer Science**, *Courant Institute of Mathematical Sciences, New York University*.

2017 – 2020   **BS, Mathematics**, University of Illinois Urbana-Champaign, Champaign, IL, Minor in Computer Science and Electrical Engineering, High Distinction.

Publications

Preprints

2021   **Wenqian Ye, Fei Xu, Yaojia Huang, Cassie Huang, Ji A**, Adversarial Examples Generation for Reducing Implicit Gender Bias in Pre-trained Models, ArXiv preprint arXiv:2110.01094, 2021. CSCI-GA. 2590 NLP Course Project.

Published

2023   **Xu Cao*, Wenqian Ye*, Elena Sizikova, Xue Bai, Megan Coffee, Hongwu Zeng, Jianguo Cao** , ViTASD: Robust ViT Baselines for Autism Spectrum Disorder Facial Detection, In *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*.

2023   **Wenqian Ye*, Yunsheng Ma*, Xu Cao** , Uncertainty Estimation in Deterministic Vision Transformer , In *AAAI Workshop on Uncertainty Reasoning and Quantification in Decision Making (UDM-AAAI)*.

2021   **Guoxuan Li, Songmao Zhang, Jiayi Wei, Wenqian Ye** , Combining FCA-Map with Representation Learning for Aligning Large Biomedical Ontologies , In *International Semantic Web Conference, Workshop on Ontology Matching (ISWC)*.

Research Experience

NYU Center of Data Science

Aug,2022 – Present   ***Masked Domain Adaptation (MDA): Lipschitz Regularized Masked Self-supervised Learning for Chest X-ray Domain Adaptation***.

Introduced Masked Domain Adaptation (MDA), a novel and robust self-supervised domain adaptation procedure with Lipschitz Regularized Self-Attention (LRSA) module for domain adaptation in chest X-ray image analysis. Investigated the connection between MDA and inverse problems, and analyze the ability of the LRSA module and masked self-supervised learning strategy to learn the true data distribution. Empirically evaluated MDA on two Tuberculosis classification benchmarks (Montgomery and Shenzhen datasets), and demonstrate state-of-the-art classification results compared to existing self-supervised domain adaptation methods.

Advisor :   **Dr. Elena Sizikova**, *Assistant Professor*, Center of Data Science, NYU (*Personal Web-page*)

IBM-Illinois Center for Cognitive Computing Systems Research (C3SR)

Jan 2019 – May 2020 **Co-founder**, *LiveSensus*.
Built a machine learning model and open-sourced dataset consisting of 30 hours of audio labeled with MOS scores for quality estimation during Vo-IP. Designed and developed both simulators to re-create quality degradation in videos and audios for dataset and survey launched on AWS and LiveSensus website. Collaborated with four other founders, Professor Sanjay Patel and a leading live streaming company, five founders selected from 40 students under Alchemy Foundry at UIUC Coordinated Science Laboratory(CSL).

Advisor : **Prof. Sanjay J. Patel**, *Professor of Electrical and Computer Engineering and Computer Science*, University of Illinois at Urbana-Champaign (*Personal Web-page*)

Chinese Academy of Science

Jun 2021 – Aug 2021 ***Combining FCA-Map with Representation Learning for Aligning Large Biomedical Ontologies***.
Developed FCA-Map to utilize the Formal Concept Analysis (FCA) formalism for aligning ontologies in an incremental way. Combined FCA-Map with the representation learning technique Siamese BERT so as to take advantage of the semantic representation in numerical latent space. Evaluated our method on the OAEI 2020 LargeBio small version tasks. Our method obtains the highest recall and F-measure for FMA-NCI (92.3% and 93.9%) and FMA-SNOMED (83.1% and 87.4%)

Advisor : **Prof. Songmao Zhang**, *Professor*, Academy of Mathematics and Systems Science, Chinese Academy of Science

Work Experience

Present **Research Fellow**, *AI Laboratory*, Shenzhen Children's Hospital.
Focus on AI for Social Good, Pediatrics AI.

2022 – 2023 **Software Engineer**, *Cirrus Logic*, Austin, TX.
Conduct Embedded software validation and testing for audio and haptics application – unit test design, automation, analysis, and report. Work on internal and customer-facing UI design and implementation System-level testing for components including device driver, firmware, and UI Software test automation. Implement DSP algorithm prototype in Python/Matlab and fixed-point firmware in C/C++.

Fellowships & Awards

2023 AAAI Student Scholarship Grant
2023 UVA Computer Science Scholar
2016 VEX U Excellence Award

Teaching Experience

Fall 2021 **CSCI-GA. 2590: Natural Language Processing**, New York University.
Graded the written assignments, exams and final projects. Set up the autograder for code assignments. Held the TA section during weekly office hours. Answered questions and provide guidelines for students on the CampusWire forum.

Services

Program Committee **Workshops**.
UDM-KDD (SIGKDD 2023)

Reviewer **Journals**.
IEEE Internet of Things Journal

Conferences.
ICASSP 2023; AAAI 2023; ICML 2022

Workshops.
IEEE Workshop on Machine Learning for Signal Processing (MLSP); VTTA (NeurIPS 2022)

Membership	**Member**. Institute of Electrical and Electronics Engineers (IEEE); Association for Computing Machinery (ACM); IEEE Signal Processing Society (SPS); ACL Year-Round Mentorship Program

Technical Skills

Languages	Python, C/C++, R, MATLAB, Golang, SystemVerilog, LaTeX
Packages	PyTorch/TensorFlow, PyTorch Lightening, Huggingface, Scikit-learn
Others	AWS, CUDA, MySQL, Git, Jenkins