Yibo HU

Email: yibo.hu@gatech.edu

Website: https://snowood1.github.io/ | Google Scholar

RESEARCH INTERESTS

Data mining; artificial intelligence; natural language processing, including natural language generation, information extraction, and interdisciplinary applications to social science and cyber security; uncertainty quantification.

APPOINTMENT

Postdoctoral Fellow, Georgia Institute of Technology, July 2023 - present

EDUCATION

Ph.D. in Computer Science, The University of Texas at Dallas, expected in June 2023

M.S. in Computer Science, The University of Texas at Dallas, Dec. 2017

M.S. in Electronics & Communication Engineering, Xiamen University, July 2014

B.E. in Communication Engineering, Xiamen University, July 2011

PUBLICATIONS

Sadaf Md Halim, Saquib Irtiza, **Yibo Hu**, Latifur Khan, and Bhavani Thuraisingham. WokeGPT: Improving Counterspeech Generation Against Online Hate Speech by Intelligently Augmenting Datasets Using a Novel Metric. In IEEE International Joint Conference on Neural Networks (IJCNN), 2023. [To appear]

Erick Skorupa Parolin, **Yibo Hu**, Latifur Khan, Javier Osorio, Patrick T. Brandt, and Vito D'Orazio. Confli-T5: An AutoPrompt Pipeline for Conflict Related Text Augmentation. In IEEE International Conference on Big Data (Big Data), 2022. [link]

Yibo Hu, Yu Lin, Erick Skorupa Parolin, Latifur Khan, and Kevin Hamlen. Controllable Fake Document Infilling for Cyber Deception. In Findings of the Association for Computational Linguistics: (EMNLP), 2022. [link]

Khandakar Ashrafi Akbar, Sadaf Md Halim, **Yibo Hu**, Anoop Singhal, Latifur Khan, and Bhavani Thuraisingham. Knowledge Mining in Cybersecurity: From Attack to Defense. In IFIP Annual Conference on Data and Applications Security and Privacy (DBSec), 2022. [link]

Zhuoyi Wang, **Yibo Hu**, Latifur Khan, Kevin Hamlen, and Bhavani Thuraisingham. CAPT: Contrastive Pre-Training based Semi-Supervised Open-Set Learning. In IEEE International Conference on Multimedia Information Processing and Retrieval (MIPR), 2022. [link]

Erick Skorupa Parolin, MohammadSaleh Hosseini, **Yibo Hu**, Latifur Khan, Patrick T. Brandt, Javier Osorio, and Vito D'Orazio. Multi-CoPED: A Multilingual Multi-Task Approach for Coding Political Event Data on Conflict and Mediation Domain. In Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society (AIES), 2022. [link]

Yibo Hu, MohammadSaleh Hosseini, Erick Skorupa Parolin, Javier Osorio, Latifur Khan, Patrick Brandt, and Vito D'Orazio. ConfliBERT: A Pre-trained Language Model for Political Conflict and Violence. In Conference of the North American Chapter of the Association for Computational Linguistics (NAACL), 2022. [link]

Erick Skorupa Parolin, **Yibo Hu**, Latifur Khan, Javier Osorio, Patrick T. Brandt, and Vito D'Orazio. CoMe-KE: A New Transformers Based Approach for Knowledge Extraction in Conflict and Mediation Domain. In IEEE International Conference on Big Data (Big Data), 2021. [link]

Yibo Hu, and Latifur Khan. Uncertainty-aware reliable text classification. In Proceedings of the ACM SIGKDD Conference on Knowledge Discovery & Data Mining (KDD), 2021. [link]

Yibo Hu, Yuzhe Ou, Xujiang Zhao, Jin-Hee Cho, and Feng Chen. Multidimensional uncertainty-aware evidential neural networks. In Proceedings of the AAAI Conference on Artificial Intelligence (AAAI), 2021. [link]

TEACHING EXPERIENCE

Teaching Assistant, UT Dallas

CS6375 Machine Learning (Graduate), Summer 2018, Spring 2019, Spring 2020

CS6364 Artificial Intelligence (Graduate), Spring 2018, Fall 2020

CS6350 Big Data Management and Analytics (Graduate), Spring 2020

CS4395 Human Language Technologies (Undergraduate), Fall 2019

CS4347 Database Systems (Undergraduate), Fall 2019

CS4375 Introduction to Machine Learning (Undergraduate), Fall 2018

CS6320 Natural Language Processing (Graduate), Fall 2018

CS3345 Data Structures and Introduction to Algorithmic Analysis (Undergraduate), Spring 2018

MENTORING EXPERIENCE

Graduate Research Mentors, UT Dallas

Mentored three graduates on developing multilingual language models for political science, resulting in two master's thesis and an academic paper, Summer 2022 - Spring 2023

Undergraduate Research Mentors, UT Dallas

Clark Summer Research Program, Summer 2023

ACM Student Research Project: "Adversarial Attacks on Hate Speech Detection", Spring 2023

Data Science Summer Camps, Summer 2021

SERVICE

Program Committee or Reviewers

The Annual Meeting of the Association for Computational Linguistics (ACL) 2023

International Joint Conferences on Artificial Intelligence (IJCAI) 2023

SIGKDD Conference on Knowledge Discovery and Data Mining (KDD) 2020, 2021, 2023

The International Conference on Database Systems for Advanced Applications (DASFAA) 2023

The Conference on Empirical Methods in Natural Language Processing (EMNLP) 2022

SIAM International Conference on Data Mining (SDM) 2022

The 24th International Conference on Artificial Intelligence and Statistics (AISTATS) 2021

ACM Transactions on Internet Technology (TOIT) 2021

TALKS

"Recent Advances in Big Data Natural Language Processing", guest talk in the class Big Data Management and Analytics, UT Dallas, Feb. 2023

"Transfer Learning and Uncertainty Quantification in Natural Language Processing for Political Science and Cyber Security", guest talk in the class Recent Advances in Computing, UT Dallas, Nov. 2022

"Fake Document Generation for Cyber Deception", invited talk with CVS Health team, UT Dallas, Nov. 2022

"Improving Reliability in Deep Learning Models' Prediction", Yahoo Targeting Science group, June 2022

ACADEMIC AND INDUSTRY INTERNSHIPS

Research Intern, Targeting Science Group at Yahoo, May 2022- Aug. 2022
Proposed a novel text generation model for contextual advertising.

Research Assistant, Big Data Analytics Lab at UT Dallas, Sept. 2020- present
Conducted interdisciplinary research on AI, social science, and cybersecurity.

Lead Research Assistant, Power Optimization Center at Vistra Energy, TX, Feb. 2020-Aug. 2020
Developed forecasting models for Vistra's Moss Landing Energy Storage.