Hongye Jin

Contact Information

EMAIL: jhy0410@tamu.edu HOMEPAGE: https://mooler0410.github.io/puguJin/

Research Interests

NLP (LLMs, data augmentation, safety), **Trustworthy Machine Learning** (Fairness& Out-of-Distribution Generalization & Security), **Data Mining** (Recommendation System)

Education

08/2020–now	Texas A&M University (TAMU) Ph.D student in Computer Science	Dept. of Computer Science & Engineering Advisor: Dr. Xia (Ben) Hu
09/2015–06/2020	Peking University (PKU) Bachelor of Science, Computer Science	Sch. of Electronics Engineering & Computer Science GPA: 3.6/4.0

Research Experience

09/2020–Now, TX
DATA Lab at Texas A&M University

Graduate Research Assistant, advised by Dr. Xia (Ben) Hu.
- Conduct research on trustworthy machine learning.
- Explore the properties of large language models to improve the LLMs training&inference, and enhance ML tasks by leveraging the capabilities of LLMs.

09/2022–12/2022, CA
Artificial Intelligence team at Visa Research

Research Intern, mentored by Dr. Huiyuan Chen and Dr. Hao Yang.
- Develop a new test-time-adaption framework to mitigate distribution shift problem caused by graph structures for Graph Neural Networks.

11/2020–02/2021, China
DAMO Academy, Alibaba

Research Assistant, mentored by Dingkun Long and Guangwei Xu
- Tackling the distant supervision challenge for NLP tasks. Propose to leverage BERT's language modeling ability to construct a denoiser for improving the quality of noisy text data.

09/2019–03/2020, Singapore
NExT++ Lab at National Univeristy of Singapore

Undergraduate Research Assistant, mentored by Dr. Xiang Wang and Dr. Tat-Seng Chua.
- Improve the performance and interpretability of collaborative filtering based recommendation models at the same time via an iterative disentangled representation learning strategy.

Publications(* co-first author)

1. Z. Jiang*, X. Han*, **H. Jin**, G. Wang, R. Chen, N. Zou, X. Hu, "Chasing Fairness under Distribution Shift: a Model Weight Perturbation Approach", NeurIPS2023
2. **H. Jin***, J. Yang*, R. Tang*, X. Han*, Q. Feng*, H. Jiang, B. Yin, X. Hu, "Harnessing the Power of LLMs in Practice: A Survey on ChatGPT and Beyond", Arxiv
3. **H. Jin***, X. Han*, Z. Jiang*, Z. Liu, N. Zou, Q. Wang, X. Hu, "Retiring ΔDP: New Distribution-Level Metrics for Demographic Parity", TMLR 2023
4. **H. Jin***, X. Han*, J Yang, Z Jiang, CY Chang, X Hu, "GrowLength: Accelerating LLMs Pretraining by Progressively Growing Training Length", Arxiv
5. **H. Jin**, F. Yang, C. Tilli, S. Mishra, X. Hu, "Transferring Fairness under Distribution Shift without Sensitive Information", Under Review

6. **H. Jin***, R. Tang*, C. Wigington, M. Du, R. Jain, X. Hu, "Exposing Model Theft: A Robust and Transferable Watermark for Thwarting Model Extraction Attacks", CIKM23(Short)
7. H. Chen, M. Das, V. Lai, Z. Jiang, **H. Jin**, X. Hu, M. Yeh, Y. Zheng, H. Yang, "Towards Mitigating Dimensional Collapse of Representations in Collaborative Filtering", Under review
8. X. Wang, **H. Jin**, A. Zhang, X. He, T. Xu, TS. Chua, "Disentangled graph collaborative filtering", SIGIR'20

Academic Activities

- Conference Reviewer: ICDM'22, WWW'23, KDD'23, NeurIPS'23, AAAI'24
- Journal Reviewer: ACM Transactions on Intelligent Systems and Technology
- Invited talk: Seminar in Visa Research