Zhaomin Xiao

Denton, TX 76205, USA
+1 940-703-9518 • zhaominxiao@my.unt.edu • linkedin.com/in/zhaominxiao/

EXPERTISE

○ Machine Learning ○ Natural Language Processing ○ Large Language Model ○ Deep Learning ○ Multimodal Learning

EDUCATION

University of North Texas Aug 2019 - Dec 2023 (Estimated)
Ph.D. Candidate in Computer Science

University of Pittsburgh May 2019
M.S. in Information Science

Shenzhen University Jun 2017
B.S. in Mathematics and Applied Mathematics

WORK EXPERIENCE

Machine Learning Engineer Intern Bellevue, WA
Meta (previously known as Facebook) *May. 2022 - Oct. 2022*

- Establish a data processing pipeline that standardizes features and prepare aggregated and anonymized data for the model trainer using the dataswarm pipeline. This pipeline ensures compliance with Apple's latest privacy guidelines.
- Build a privacy-preserving machine learning model for forecasting user ad clicks, leveraging aggregated and anonymized input data.
- Engage in collaboration with both the product and research teams to assess the model across various parameter and privacy configurations. Take the lead in weekly meetings aimed at crafting the initial prototype based on preliminary experimental results.

RESEARCH EXPERIENCE

University Enrollment Prediction Denton, TX
University of North Texas *Aug. 2022 - Present*

- Perform instruction-finetuning on large language models like LLaMA 2 and T5, employing diverse instructions and prompts.
- Utilize Parameter-Efficient Finetuning (PEFT) strategies, such as LoRA and Prompt-Tuning, to accelerate the finetuning procedure.

Corporate Event Prediction Denton, TX
University of North Texas *Aug. 2022 - Present*

- Scrape earnings call transcripts and audio recordings and construct a multimodal dataset for corporate event prediction.
- Develop multimodal neural networks to predict if certain types of corporate events, such as acquisitions, corporate bond issuances, and seasonal equity offerings, will occur within three months after the earnings call.

Spatial and Temporal Information Extraction Denton, TX
University of North Texas *Aug. 2019 - May. 2022*

- Scrape tweets and construct datasets for spatial and temporal information extraction. Create interactive web interfaces for collecting annotations using Amazon Mechanical Turk.
- Develop machine learning models, such as SVM, Logistic Regression, and Naive Bayes incorporating manually designed linguistic features, such as POS tags and named entities, to tackle the tasks. Utilize SVD and PCA for feature dimensionality reduction.
- Construct contextually-informed, multimodal neural networks for extracting spatial and temporal information from Twitter streams. Methods involving multimodal LLMs with zero-shot inference are also explored to address this challenge.

PUBLICATION

- Zhaomin Xiao, Eduardo Blanco, and Yan Huang. Context Helps Determine Spatial Knowledge from Tweets. In *Findings of the Association for Computational Linguistics: AACL 2023*, 2023. Accepted, To appear.
- Zhaomin Xiao, Yachen Cui, Zhelu Mai, Jiancheng Li, and Zhuoer Xu (2023). Corporate Event Prediction Using Earnings Call Transcripts. In *International Conference on Information Management and Big Data* (*SIMBIG*), 2023. Accepted, To appear.
- Zhaomin Xiao and Eduardo Blanco. Are People Located in the Places They Mention in Their Tweets? A Multimodal Approach. In *Proceedings of the 28th International Conference on Computational Linguistics* (*COLING*), 2022.

TECHNICAL SKILLS

- **Programming Languages:** Python, MySQL, Java, HTML, CSS, JavaScript, C
- **Tools:** PyTorch, Hugging Face, TensorFlow, Keras, Linux, Git, AWS, scikit-learn, Hive, Firebase, pandas, seaborn, aeneas
- **Pre-trained Models:** LLaMA 2, GPT 3, T5, BERT, VGG16, ResNet, RoBERTa, CLIP