Wenbo Zhang

PhD Student · Informatics

E343, Westgate Building, Pennsylvania State University (University Park), State College, PA, 16802

+1 814-852-9399 | wenbozhangjs@gmail.com | https://wenbozhangjs.github.io/

Research Interests

Natural Language Processing, Multi-lingual Language Model, Code-mixed Text Processing, AI for Social Impact, Deep learning, Data Mining, Speech Processing.

Education

Pennsylvania State University (PSU) *Pennsylvania, USA*

Doctor of Philosophy in Informatics (GPA: 4.0/4.0) *Aug. 2021 - Present*

- Advisor: Dr. Amulya Yadav

University of Southern California (USC) *California, USA*

Master of Science in Electrical Engineering *Aug. 2016 - May. 2018*

- Advisor: Dr. Cauligi Raghavendra

University of Electronic Science and Technology of China (UESTC) *Sichuan, China*

Bachelor of Engineering in Renewable Energy Materials and Devices *Sept. 2011 – Jul. 2015*

Publications

Under Review

[1] Hangzhi Guo, Xinchang Xiong, **Wenbo Zhang**, Amulya Yadav. ReLax: Efficient and Scalable Recourse Explanation Benchmarking using JAX. Journal of Open Source Software (JOSS), 2024. (**Currently Under Review)**

Peer Reviewed Publications

[1] **Wenbo, Zhang**, Hangzhi Guo, Prerna Ranganathan, Jay Patel, Sathyanath Rajasekharan, Nidhi Danayak, Manan Gupta, Amulya Yadav. A Continual Pre-training Approach to Tele-Triaging Pregnant Women in Kenya. In Proceedings of the 37th AAAI Conference on Artificial Intelligence, 2023. (**System has been full-time deployed by Jacaranda Health**)

[2] Hangzhi Guo, Xinchang Xiong, **Wenbo Zhang**, Amulya Yadav. Efficient and Scalable Recourse Explanation Benchmark using JAX. In the NeurIPS 2023 Workshop on XAI in Action: Past, Present, and Future Applications, 2023.

[3] **Wenbo Zhang**, Hangzhi Guo, Prerna Ranganathan, Jay Patel, Sathyanath Rajasekharan, Nidhi Danayak, Manan Gupta, Amulya Yadav. TRIM-AI: Harnessing Language Models for Providing Timely Maternal & Neonatal Care in Low-Resource Countries. In the AAAI Workshop on AI for Social Good (AI4SG), 2023.

Archival

[1] **Wenbo Zhang**, Hangzhi Guo, Ian Kivlichan, Vinodkumar Prabhakaran, Davis Yadav and Amulya Yadav. A Taxonomy of Rater Disagreement: Surveying Challenges & Opportunities from the Perspective of Annotating Online Toxicity. arXiv preprint arXiv:2311.04345

In Prep

[1] **Wenbo Zhang**, Amulya Yadav. Benchmarking the Ability of LLMs across Wide Variety of Code-Mixed Tasks

[2] **Wenbo Zhang**, Amulya Yadav. Code-Mixed LLMs: Harnessing LLMs Annotation Abilities for Reinforcement Learning from Human Feedback

Research Experience

Adapt LLMs to the scenarios of code-mixed contents understanding through RLHF *Pennsylvania, USA*

Advisor: Dr. Amulya Yadav *November. 2023 - Present*

- This work attempts to answer the question: Do large language models understand code-mixed contents well? We focuses on analyzing the LLMs' ability to understand and deal with code-mixed text between prompt-based models and small fine-tuned based models specifically designed for code-mixed contents processing. We further enhance the existing LLMs better to understand code-mixed contents through reinforcement learning from human feedback (RLHF).

Assist case assessment of online sexual abuse and exploitation among children *Pennsylvania, USA*

Advisor: Dr. Amulya Yadav *Feb. 2024 - Present*

- Children in the Philippines are vulnerable to online sexual abuse and bullying. This research aims to tackle the following research question in partnership with World Hope International-Philippines (WHI-PH): Can we build an AI-driven decision support system which can automate the preliminary triage and case assessments of Online Sexual Abuse and Exploitation of Children survivors?

Machine learning for phenotypic pattern identification of adolescents with drug usage *Pennsylvania, USA*

Advisor: Dr. Amulya Yadav *Feb. 2023 - Jul. 2023*

- This research focuses on identifying potential students (in high school) who may use alcohol, cigarette or marijuana in the future. We analyze potential patterns which may lead such behaviors through the machine learning perspective.

TRIM-AI: Harnessing language models for providing timely maternal health care *Pennsylvania, USA*

Advisor: Dr. Amulya Yadav *Sept. 2021 - Jun. 2022*

- Maternal health situation are servere in Kenya, Africa. This work focuses on developing an NLP framework, using multi-lingual pretraining and continual pretraining, to predict the user's medical situation (emergency level) based on code-mixed SMS messages.
- This framework has been deployed inside the PROMPTS (digital health system developed by Jacaranda Health). According to the feedback from Jacaranda Health, this framework reduces the monthly system management cost by 22.8% and PROMPTS helpdesk's workload by ∼12%.

Awards, Fellowships, & Grants

2023 **AAAI-23 student scholarship**, AAAI Conference on Artificial Intelligence (AAAI)

PSU Student Travel Award, College of Information sciences and technology, Pennsylvania State University

2022 **AI Societal Impact Award**, Center for Artificial Intelligence Foundations and Engineered Systems (CAFÉ) at Pennsylvania State University

2014 **3rd Class of National People's Scholarship (top 15%)**, University of Electronic Science and Technology of China

2013 **3rd Class of National People's Scholarship (top 15%)**, University of Electronic Science and Technology of China

2012 **3rd Class of National People's Scholarship (top 15%)**, University of Electronic Science and Technology of China

Teaching Experience

Fall 2023 **DS 442 Artificial Intelligence**, Teaching Assistant at Pennsylvania State University
Spring 2022 **DS 442 Artificial Intelligence**, Teaching Assistant at Pennsylvania State University
Spring 2024 **DS 442 Artificial Intelligence**, Teaching Assistant at Pennsylvania State University

Industrial Experience

Machine Learning Engineer *Beijing, China*

AI Lab, Kingsoft Co., Ltd. *Jan. 2019 - Jul. 2021*

- Applied recent advanced NLP techniques to develop information extraction modules inside knowledge graph.
- Employed NLP seq2seq models and speech processing techniques to construct the text-to-speech (TTS) system.

Direction 1: knowledge graph

Open domain knowledge graph construction *Beijing, China*

Project Participant *Jan. 2021 - Jul. 2021*

- Designed modules (name entity recognition and relation extraction) for Chinese knowledge graph construction.
- The knowledge graph has been deployed inside the Kingsoft electronic notebook website.

Direction 2: speech processing (especially TTS, generative AI for speech synthesis)

English multi-speaker text-to-speech (TTS) system for novel website *Beijing, China*

Project Participant *Jul. 2020 - Dec. 2020*

- Developed a system which generated speech with someone's tone through few minutes' voice recordings.
- Created a prototype for audiobook reading on English novel translation website to support multiple human voices.

End-to-end framework for Chinese polyphone pronunciation prediction *Beijing, China*

Project Leader *Apr. 2020 - Jul. 2020*

- Built end-to-end framework for pronunciation prediction of Chinese polyphone with multi-phonemic values.
- Improved the pronunciation correctness of Chinese speech synthesis system.

NLP based Chinese text prosody prediction *Beijing, China*

Project Leader *Jan. 2020 - Mar. 2020*

- Modeled the prosody (short pronunciation break among Chinese words) prediction as the sequence tagging problem.
- Improved the naturalness and quality of the synthesized speech generated from Chinese speech synthesis system.

End-to-end Chinese text-to-speech (TTS) system *Beijing, China*

Project Leader *Jan. 2019 - Apr. 2020*

- Implemented whole pipeline of end-to-end Chinese TTS system, including the text processing module (which extracts semantic information from input sentences), acoustic model (which predicts acoustic features based on the semantic information), and the vocoder model (which transforms acoustic features into speech signals).
- Applied on the Kingsoft policy question answer (QA) system.

Past Internship and Research Visit

Intern *Suzhou, China*

Comprehend Information Technology Co., Ltd. *Jun. 2017 - Aug. 2017*

- NLP-based data mining on the traffic data (electronic checkpoints data) accessed from Suzhou City Brain.
- Predicted and partitioned whole city into different function regions (education area, central business area etc.).

Skills

DevOps	Google Cloud Platform, Alibaba cloud, Docker
Back-end	Django
Programming	Python, R, C, LaTeX, Shell
Framework	Tensorflow, Pytorch