XIAOBO GUO

Xiaobo.Guo.GR@dartmouth.edu | (267) 353-5980 | github.com/xiaobo-guo

Education

Dartmouth College Hanover, NH, USA

PhD in Computer Science (computational social science) *Sep. 2019 – Now*

Peking University Beijing, China

LLB. in Law *Sep. 2015 – June.2019*

BS in Computer Science and Technology *Sep. 2015 – June.2019*

Work Experience

Amazon People Experience and Technology Central Science New York, NY, USA

Applied Scientist Intern *June. 2023 – Sep. 2023*

- Analyzed employment behavior, utilizing posts from the internal community as the primary data source.
- Analyzed employment-related opinions by examining both the comments and behavior patterns of community members.

Amazon People Experience and Technology Central Science New York, NY, USA

Applied Scientist Intern *June. 2022 – Sep. 2022*

- Developed a multi-tiered framework for summarizing employee feedback, employing clustering and summarization techniques for improved insights."
- Improved summarization accuracy by 28% by significantly reducing cumulative errors.

Research Projects

Analyzing the gap between people with different opinions on social media *Sep. 2023 - Now*

- Analyzed the features of posts referring to inner/outer community context.
- Figured out the similarity of different subreddits in terms of the religion topics.
- Proposed new metrics to evaluate the harmfulness of the text shown online.

Analyzing the Influential Users on Social Media *Sep. 2022 - Now*

- Figured out the features that influence the probability of being retweeted by sentence embeddings or perplexity
- Analyzed the opinion and topic similarity of the users and their followers with multiple metrics.
- Predicted the probability of following based on the current following graph and other meta information

Measuring Media Bias Through Masked Token Prediction *Aug. 2020 – June. 2022*

- Utilized lexical features to extract the stance of individual media outlets.
- Developed innovative metrics to evaluate the performance of models for media bias analysis.
- Enhanced the correlation rate between models and human evaluations by a r**elative 5%**, improving the **reliability and accuracy** of model outputs.

Emotion-based Modeling of Mental Disorders on Social Media *Aug. 2020 – Feb. 2021*

- Formulated mathematical methods to passively detect mental disorders, leveraging conversation data from Reddit.
- Utilized lexical features to predict user emotions in individual posts.
- Surpassed existing model performance by achieving a **2%** improvement in both **F1-score** and **accuracy**.

Identification of State-Sponsored Propaganda on Twitter *Aug. 2019 – June. 2020*

- Created a dataset for the state-sponsored propaganda identification task by leveraging the **Twitter API**.
- Utilized textual and visual features to tackle the task and achieve state-of-the-art performance.
- Analyzed the **distinguished features** among state-sponsored features of different countries.

Publication

1. **Xiaobo Guo** and Soroush Vosoughi "Length Does Matter: Summary Length can Bias Summarization Metrics". Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP), 2023
2. **Xiaobo Guo**, Weicheng Ma, and Soroush Vosoughi "Capturing Topic Framing via Masked Language Modeling". Findings of the Association for Computational Linguistics: EMNLP 202
3. **Xiaobo Guo,** Weicheng Ma, and Soroush Vosoughi "Measuring Media Bias via Masked Language Modeling". Proceedings of the International AAAI Conference on Web and Social Media (ICWSM), 2022
4. **Xiaobo Guo** and Soroush Vosoughi "A Large-Scale Longitudinal Multimodal Dataset of State-Backed Information Operations on Twitter". Proceedings of the International AAAI Conference on Web and Social Media (ICWSM), 2022
5. **Xiaobo Guo**, Yaojia Sun and Soroush Vosoughi "Emotion-based Modeling of Mental Disorders on Social Media". The 20th IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT), 2021
6. **Xiaobo Guo** and Soroush Vosoughi "Multi-modal Identification of State-Sponsored Propaganda on Social Media". International Conference on Pattern Recognition (ICPR), 2021

Language and Tools

1. Language: Python3, C++, SQL
2. Tools: Keras, NLTK, PyTorch, Scikit-learn, TensorFlow, Transformers, Numpy, Panda