

Chiyu Wei

Phone: (213)-725-1539 | Email: chiyuwei@usc.edu

EDUCATION

University of Southern California

Master of Science, Computer Science

Los Angeles, California

Expected: December 2023

Wuhan University of Technology

Bachelor of Engineering, Communications Engineering

Wuhan, China

July 2021

PUBLICATIONS

La Gatta, V.*, Wei, C.*, Luceri, L., Pierri, F., Ferrara, E. (2023). Retrieving False Claims on Twitter during the Russia-Ukraine Conflict. In Companion Proceedings of the ACM Web Conference 2023. doi: 10.1145/3543873.3587571

* denotes equal contribution.

RESEARCH

Multimodal Partisan Sentiment Framing in Inflation News Coverage

Research Assistant, USC Viterbi School of Engineering

Los Angeles, California

Advised by Dr. Mohammad Soleymani and Dr. Meiqing Zhang

May 2023 - Present

- Embarked on a comprehensive examination of partisan biases in inflation news coverage across multimodal media channels. This research entailed a deep dive into how various media outlets, with different political leanings, chose to select and present information about inflation.
- Executed a detailed multimodal content analysis on a variety of sources, including newspaper editorials, cable television transcripts, and video clips from January 2020 to December 2022.
- Initial findings revealed significant differences in partisan emotion framing between television news and newspapers. The presence of stronger partisan cues leads to greater emotionality, and extremity, revealing the pivotal role emotion framing plays in shaping partisan narratives.

Classification of Political Extremists and Moderate Users on Twitter

Research Assistant, USC Information Sciences Institute

Los Angeles, California

Advised by Dr. Emilio Ferrara and Dr. Luca Luceri

March 2023 - Present

- Pioneered a novel methodology, False-Supervised Learning, to accurately classify Twitter users based on their political leanings, distinguishing between extremist and moderate users.
- Through this innovative technique, extremist or moderate labels are attributed to users according to their ideological scores. A distinct BERT classification model is trained and fine-tuned for each proportion split using the labeled data.
- The model demonstrating superior precision in distinguishing between moderate and extremist ideologies is chosen as the optimal solution. This model will help define an ideological

threshold, effectively categorizing users into the respective ideological categories.

Retrieving False Claims on Twitter during the Russia-Ukraine Conflict

Research Assistant, USC Information Sciences Institute

Los Angeles, California

Advised by Dr. Emilio Ferrara and Dr. Luca Luceri

June 2022 - March 2023

- Collected and manually annotated 83 false claims circulated on Twitter during the initial weeks of the Russia-Ukraine Conflict. This task involved handling and categorizing 5,872 original tweets into four distinct classes: related, not related, support, and refute.
- Designed and implemented a transformer-based automated pipeline capable of detecting and retrieving tweets that discussed the identified false claims. This model achieved an F1-score of 80.57%. Further, the tweet retrieval model obtained a Top-3 accuracy of 96.35%.
- Successfully demonstrated the practicality and effectiveness of the developed approach in retrieving false claims. The model showed consistent performance in real-world conditions.

PROJECTS

Self-Introduction Generator Using NLP Techniques

Team course project

August 2022 - December 2022

- Leveraged a BERT-based question-answering model to extract essential introductory information and integrated the data into a template, which achieved a BERTScore of 0.817.

Multimodal Features Extraction and Sentiments Classification

Solo course project

January 2022 - May 2022

- Applied BERT, OpenFace, and OpenSmile to extract text, visual, and audio features of 2,199 video clips. Designed LSTM sentiments classifiers and late fusion strategy for each modality.

Kaggle competition: H&M Personalized Fashion Recommendations

Team course project

January 2022 - May 2022

- Developed product recommendation models by harnessing two years' worth of real-world transaction data coupled with item images, all aimed at predicting the top 12 products for each customer. Our approach secured a position within the 9% of all competitors.

SKILLS

- **Programming Languages:** Python, Java, C++, C, and Matlab.
- **Machine Learning Libraries:** Extensive experience with PyTorch, Keras, scikit-learn, Sentence-Transformers, matplotlib, and seaborn.
- **Operating Systems:** Experience in working with Linux-based systems.
- **Networking and Computing:** Experience with SSH and High-Performance Computing.