JUNJIE WU

Email: junjie.wu@connect.ust.hk Website: https://wujunjie1998.github.io

EDUCATION

Hong Kong University of Science and Technology *September 2020 - July 2024 (Expected)*

Ph.D. in Artificial Intelligence, Information Hub (affiliated with the CSE Department)

Advised by **Prof. Dit-Yan Yeung**

Sun Yat-sen University, Guangzhou, China *September 2016 - June 2020*

(*Second-class, Third-class Scholarship of Sun Yat-sen University 2016-2017, 2017-2018*)

Bachelor in Statistics, School of Mathematics, Major GPA: 3.8/4.0

PUBLICATIONS

Jiajun Bao*, **Junjie Wu***, Yiming Zhang*, Eshwar Chandrasekharan and David Jurgens. "Conversations Gone Alright: Quantifying and Predicting Prosocial Outcomes in Online Conversations " (*Proceedings of the Web Conference (WebConf), 2021*) (*: Equal contribution. The order is alphabetical.)

Junjie Wu and Hao Zhou. "Augmenting Topic-Aware Knowledge-Grounded Conversations with Dynamic Built Knowledge Graphs" *(Proceedings of the second NAACL Workshop on Knowledge Extraction and Integration for Deep Learning Architectures (DeeLIO). 2021.)*)

FINISHED RESEARCH PROJECTS

CoAI Lab Tsinghua University, **Advisor: Prof. Minlie Huang** *Oct 2019 - Aug 2020*

Project: Tracking and controlling topic transition in document-grounded dialog system

- Employed a Transformer-based Encoder-Decoder framework to track topic transitions in the conversation history, along with a dynamic built KG module that automatically forms logical information between chatting topics during a conversation, which helps our system to learn from its conversational partner especially when they have different prior knowledge.
- Designed a topic predictor before the knowledge selector to determine whether to widen or deepen current topics, aimed to generate more informative responses and employ background knowledge effectively.

Blablablab University of Michigan, **Advisor: Prof. David Jurgens** *July 2019 - June 2020*

Project: Predicting prosocial (defined by many metrics like healthy, supportive, politeness) **outcomes in online conversations from a large-scale Reddit dataset**

- Constructed multiple text-based features from a large-scale Reddit dataset (over 100 million lines).
- Defined a set of prosocial metrics based on prior literature and coded them for our Reddit data corpus (Including employing a fine-tuned BERT model to produce several metrics. These models were wrapped to public python packages).
- Trained a combination of linear fixed/ mixed-effects models and BERT models (Including single fine-tuned BERT for each metric and a large BERT with a multi-metric based output) to forecast whether a conversation would have prosocial outcomes.
- Analyzed how different prosocial outcomes are impacted by our features, made suggestions for users and developers of social medias to build better communicating environments through a human studying.

NGN Lab Tsinghua University, **Advisor: Prof. Yongfeng Huang** *January 2019 - May 2019*

Project: English text emotion analysis and classification

- Preprocessed text data through tokenizing, filtering mispellings and special characters as well as text vectorization.
- Summarized the related research of emotion classification and implemented several baseline models including BiLSTM and attention mechanism to perform emotion classification in English text.
- Designed a novel multi-task learning framework to learn textual emotion features from emotion lexicon and corpus simultaneously, which improved the performance of emotion prediction on our dataset for nearly one percent (measured by F1-score).

NGN Lab Tsinghua University, **Advisor: Prof. Yongfeng Huang** *July 2018 - August 2018*

Project: Chinese text sentiment analysis

- Employed a bipartite graph-based sorting algorithm to extract pairs of opinion words and opinion targets from Chinese text corpus.

- Adopted a LSTM-based model to classify the sentiment polarity of each binary pair(an opinion word and an opinion target), then used these binary pairs to enrich existing Chinese sentiment lexicons.

TECHNICAL SKILLS AND OTHERS

Programming: Python, Pytorch, Matlab, R, Latex
TOEFL: 105 **GRE**: V155 Q170 AW4.0