Research Proposal for PhD Joint Training Program

Huchong Ling
Northeastern University
`xxxxxx@xxx.com`

2022-04-20

Contents

1   Introduction

The written text that we interact with on an everyday basis—news articles, emails, social media, books is the product of a profoundly social phenomenon with people at its core. With few exceptions, all of the text we see is written by people, and others constitute its audience. A vast amount of the content itself is centered on people: news (including classic NLP corpora such as the Wall Street Journal and the New York Times) details the roles of actors in current events, social media (including Twitter and Facebook) documents the actions and attitudes of friends, and books chronicle the stories of fictional characters and real people alike (Serban et al., 2016).

2   Persona Characterization and Variation1

2.1   Background and Motivation

The written text that we interact with on an everyday basis—news articles, emails, social media, books is the product of a profoundly social phenomenon with people at its core. With few exceptions, all of the text we see is written by people, and others constitute its audience. A vast amount of the content itself is centered on people: news (including classic NLP corpora such as the Wall Street Journal and the New York Times) details the roles of actors in current events, social media (including Twitter and Facebook) documents the actions and attitudes of friends, and books chronicle the stories of fictional characters and real people alike (Serban et al., 2016).

2.2   Research Purpose

The written text that we interact with on an everyday basis—news articles, emails, social media, books is the product of a profoundly social phenomenon with people at its core. With few exceptions, all of the text we see is written by people, and others constitute its audience. A vast amount of the content itself is centered on people: news (including classic NLP corpora such as the Wall Street Journal and the New York Times) details the roles of actors in current events, social media (including Twitter and Facebook) documents the actions and attitudes of friends, and books chronicle the stories of fictional characters and real people alike (Serban et al., 2016).

2.3   Methodology

The written text that we interact with on an everyday basis—news articles, emails, social media, books is the product of a profoundly social phenomenon with people at its core. With few exceptions, all of the text we see is written by people, and others constitute its audience. A vast amount of the content itself is centered on people: news (including classic NLP corpora such as the Wall Street Journal and the New York Times) details the roles of actors in current events, social media (including Twitter and Facebook) documents the actions and attitudes of friends, and books chronicle the stories of fictional characters and real people alike (Serban et al., 2016).

3   Persona Characterization and Variation2

3.1   Background and Motivation

The written text that we interact with on an everyday basis—news articles, emails, social media, books is the product of a profoundly social phenomenon with people at its core. With few exceptions, all of the

text we see is written by people, and others constitute its audience. A vast amount of the content itself is centered on people: news (including classic NLP corpora such as the Wall Street Journal and the New York Times) details the roles of actors in current events, social media (including Twitter and Facebook) documents the actions and attitudes of friends, and books chronicle the stories of fictional characters and real people alike (Serban et al., 2016).

3.2   Research Purpose

The written text that we interact with on an everyday basis—news articles, emails, social media, books is the product of a profoundly social phenomenon with people at its core. With few exceptions, all of the text we see is written by people, and others constitute its audience. A vast amount of the content itself is centered on people: news (including classic NLP corpora such as the Wall Street Journal and the New York Times) details the roles of actors in current events, social media (including Twitter and Facebook) documents the actions and attitudes of friends, and books chronicle the stories of fictional characters and real people alike (Serban et al., 2016).

3.3   Methodology

The written text that we interact with on an everyday basis—news articles, emails, social media, books is the product of a profoundly social phenomenon with people at its core. With few exceptions, all of the text we see is written by people, and others constitute its audience. A vast amount of the content itself is centered on people: news (including classic NLP corpora such as the Wall Street Journal and the New York Times) details the roles of actors in current events, social media (including Twitter and Facebook) documents the actions and attitudes of friends, and books chronicle the stories of fictional characters and real people alike (Serban et al., 2016).

4   Persona Characterization and Variation3

4.1   Background and Motivation

The written text that we interact with on an everyday basis—news articles, emails, social media, books is the product of a profoundly social phenomenon with people at its core. With few exceptions, all of the text we see is written by people, and others constitute its audience. A vast amount of the content itself is centered on people: news (including classic NLP corpora such as the Wall Street Journal and the New York Times) details the roles of actors in current events, social media (including Twitter and Facebook) documents the actions and attitudes of friends, and books chronicle the stories of fictional characters and real people alike (Serban et al., 2016).

4.2   Research Purpose

The written text that we interact with on an everyday basis—news articles, emails, social media, books is the product of a profoundly social phenomenon with people at its core. With few exceptions, all of the text we see is written by people, and others constitute its audience. A vast amount of the content itself is centered on people: news (including classic NLP corpora such as the Wall Street Journal and the New York Times) details the roles of actors in current events, social media (including Twitter and Facebook) documents the actions and attitudes of friends, and books chronicle the stories of fictional characters and real people alike (Serban et al., 2016).

4.3 Methodology

The written text that we interact with on an everyday basis—news articles, emails, social media, books is the product of a profoundly social phenomenon with people at its core. With few exceptions, all of the text we see is written by people, and others constitute its audience. A vast amount of the content itself is centered on people: news (including classic NLP corpora such as the Wall Street Journal and the New York Times) details the roles of actors in current events, social media (including Twitter and Facebook) documents the actions and attitudes of friends, and books chronicle the stories of fictional characters and real people alike (Serban et al., 2016).

5 Persona Characterization and Variation4

5.1 Background and Motivation

The written text that we interact with on an everyday basis—news articles, emails, social media, books is the product of a profoundly social phenomenon with people at its core. With few exceptions, all of the text we see is written by people, and others constitute its audience. A vast amount of the content itself is centered on people: news (including classic NLP corpora such as the Wall Street Journal and the New York Times) details the roles of actors in current events, social media (including Twitter and Facebook) documents the actions and attitudes of friends, and books chronicle the stories of fictional characters and real people alike (Serban et al., 2016).

5.2 Research Purpose

The written text that we interact with on an everyday basis—news articles, emails, social media, books is the product of a profoundly social phenomenon with people at its core. With few exceptions, all of the text we see is written by people, and others constitute its audience. A vast amount of the content itself is centered on people: news (including classic NLP corpora such as the Wall Street Journal and the New York Times) details the roles of actors in current events, social media (including Twitter and Facebook) documents the actions and attitudes of friends, and books chronicle the stories of fictional characters and real people alike (Serban et al., 2016).

5.3 Methodology

The written text that we interact with on an everyday basis—news articles, emails, social media, books is the product of a profoundly social phenomenon with people at its core. With few exceptions, all of the text we see is written by people, and others constitute its audience. A vast amount of the content itself is centered on people: news (including classic NLP corpora such as the Wall Street Journal and the New York Times) details the roles of actors in current events, social media (including Twitter and Facebook) documents the actions and attitudes of friends, and books chronicle the stories of fictional characters and real people alike (Serban et al., 2016).

6 Persona Characterization and Variation5

6.1 Background and Motivation

The written text that we interact with on an everyday basis—news articles, emails, social media, books is the product of a profoundly social phenomenon with people at its core. With few exceptions, all of the

text we see is written by people, and others constitute its audience. A vast amount of the content itself is centered on people: news (including classic NLP corpora such as the Wall Street Journal and the New York Times) details the roles of actors in current events, social media (including Twitter and Facebook) documents the actions and attitudes of friends, and books chronicle the stories of fictional characters and real people alike (Serban et al., 2016).

6.2 Research Purpose

The written text that we interact with on an everyday basis—news articles, emails, social media, books is the product of a profoundly social phenomenon with people at its core. With few exceptions, all of the text we see is written by people, and others constitute its audience. A vast amount of the content itself is centered on people: news (including classic NLP corpora such as the Wall Street Journal and the New York Times) details the roles of actors in current events, social media (including Twitter and Facebook) documents the actions and attitudes of friends, and books chronicle the stories of fictional characters and real people alike (Serban et al., 2016).

6.3 Methodology

The written text that we interact with on an everyday basis—news articles, emails, social media, books is the product of a profoundly social phenomenon with people at its core. With few exceptions, all of the text we see is written by people, and others constitute its audience. A vast amount of the content itself is centered on people: news (including classic NLP corpora such as the Wall Street Journal and the New York Times) details the roles of actors in current events, social media (including Twitter and Facebook) documents the actions and attitudes of friends, and books chronicle the stories of fictional characters and real people alike (Serban et al., 2016).

References

Siqi Bao, Huang He, Fan Wang, Hua Wu, Haifeng Wang, Wenquan Wu, Zhen Guo, Zhibin Liu, and Xinchao Xu. 2020. Plato-2: Towards building an open-domain chatbot via curriculum learning. *arXiv preprint arXiv:2006.16779*.

Marjan Ghazvininejad, Chris Brockett, Ming-Wei Chang, Bill Dolan, Jianfeng Gao, Wen-tau Yih, and Michel Galley. 2018. A knowledge-grounded neural conversation model. In *AAAI*.

Matthew Henderson, Blaise Thomson, and Steve Young. 2014. Word-based dialog state tracking with recurrent neural networks. In *Proceedings of the 15th Annual Meeting of the Special Interest Group on Discourse and Dialogue (SIGDIAL)*, pages 292–299.

Mike Lewis, Yinhan Liu, Naman Goyal, Marjan Ghazvininejad, Abdelrahman Mohamed, Omer Levy, Veselin Stoyanov, and Luke Zettlemoyer. 2020. Bart: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 7871–7880.

Yongkang Liu, Shi Feng, Daling Wang, Kaisong Song, Feiliang Ren, and Yifei Zhang. 2021. A graph reasoning network for multi-turn response selection via customized pre-training. In *AAAI*.

Liangchen Luo, Jingjing Xu, Junyang Lin, Qi Zeng, and Xu Sun. 2018. An auto-encoder matching model for learning utterance-level semantic dependency in dialogue generation. *arXiv preprint arXiv:1808.08795*.

Bishal Santra, Potnuru Anusha, and Pawan Goyal. 2020. Hierarchical transformer for task oriented dialog systems. *arXiv preprint arXiv:2011.08067*.

Iulian Serban, Alessandro Sordoni, Yoshua Bengio, Aaron Courville, and Joelle Pineau. 2016. Building end-to-end dialogue systems using generative hierarchical neural network models. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 30.

Lifeng Shang, Zhengdong Lu, and Hang Li. 2015. Neural responding machine for short-text conversation. *arXiv preprint arXiv:1503.02364*.

Haoyu Song, Yan Wang, Kaiyan Zhang, Wei-Nan Zhang, and Ting Liu. 2021. Bob: Bert over bert for training persona-based dialogue models from limited personalized data. *arXiv preprint arXiv:2106.06169*.

Ilya Sutskever, Oriol Vinyals, and Quoc V Le. 2014. Sequence to sequence learning with neural networks. *Advances in neural information processing systems*, 27.

Shiquan Yang, Rui Zhang, Sarah Erfani, and Jey Han Lau. 2022. An interpretable neuro-symbolic reasoning framework for task-oriented dialogue generation. *arXiv preprint arXiv:2203.05843*.

Haolan Zhan, Hainan Zhang, Hongshen Chen, Zhuoye Ding, Yongjun Bao, and Yanyan Lan. 2021. Augmenting knowledge-grounded conversations with sequential knowledge transition. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 5621–5630.

Yizhe Zhang, Siqi Sun, Michel Galley, Yen-Chun Chen, Chris Brockett, Xiang Gao, Jianfeng Gao, Jingjing Liu, and Bill Dolan. 2020. Dialogpt: Large-scale generative pre-training for conversational response generation. In *ACL (demo)*.

Wenya Zhu, Kaixiang Mo, Yu Zhang, Zhangbin Zhu, Xuezheng Peng, and Qiang Yang. 2017. Flexible end-to-end dialogue system for knowledge grounded conversation. *arXiv preprint arXiv:1709.04264*.

Signature Page

Signature of domestic supervisor

_____

Date(yy/mm/dd)

_____

Signature of hosting supervisor

_____

Date(yy/mm/dd)

_____