WANYU DU

Email: wd5jq@virginia.edu; Homepage: `wyu-du.github.io`

RESEARCH INTERESTS

My research interest is natural language generation. In particular, I am interested in conversation modeling, few-shot learning, interactive and iterative text generation, open-domain dialogue generation & evaluation, and reinforcement learning for dialogue systems.

EDUCATION

University of Virginia (UVa), US — August 2020 - Present
PhD student in Computer Science, Advisor: Yangfeng Ji

University of Virginia (UVa), US — August 2018 - May 2020
Master of Science in Computer Science

Dongbei University of Finance and Economics (DUFE), China — August 2014 - July 2018
Bachelor of Management in E-Commerce (Top 1%)

PUBLICATIONS

1. Jianqiao Zhao*, Yanyang Li*, **Wanyu Du***, Yangfeng Ji, Dong Yu, Michael R. Lyu, Liwei Wang. FlowEval: A Consensus-Based Dialogue Evaluation Framework Using Segment Act Flows, *preprint*, 2022.

2. **Wanyu Du**, Vipul Raheja, Dhruv Kumar, Zae Myung Kim, Melissa Lopez, Dongyeop Kang. Understanding Iterative Revision from Human-Written Text, *In Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (ACL 2022)*, 2022.

3. **Wanyu Du***, Zae Myung Kim*, Vipul Raheja, Dhruv Kumar, Dongyeop Kang. Read, Revise, Repeat: A System Demonstration for Human-in-the-loop Iterative Text Revision, *In Proceedings of the First Workshop on Intelligent and Interactive Writing Assistants (In2Writing 2022)* - **Best Paper Award**, 2022.

4. **Wanyu Du**, Jianqiao Zhao, Liwei Wang and Yangfeng Ji. Diverse Text Generation via Variational Encoder-Decoder Models with Gaussian Process Priors, *ACL 2022 6th Workshop on Structured Prediction for NLP (SPNLP 2022)*, 2022.

5. **Wanyu Du** and Yangfeng Ji. SideControl: Controlled Open-domain Dialogue Generation via Additive Side Networks, *In Findings of the Association for Computational Linguistics: EMNLP 2021*, 2021.

6. Sebastian Gehrmann et al., The GEM Benchmark: Natural Language Generation, its Evaluation and Metrics, *In Proceedings of the 1st Workshop on Natural Language Generation, Evaluation, and Metrics (GEM 2021)*, 2021.

7. Stephanie Schoch, **Wanyu Du**, Yangfeng Ji. Contextualizing Variation in Text Style Transfer Datasets, *In Proceedings of the 14th International Conference on Natural Language Generation (INLG 2021)*, 2021.

8. **Wanyu Du** and Yangfeng Ji. An Empirical Comparison on Imitation Learning and Reinforcement Learning for Paraphrase Generation, *In Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing (EMNLP 2019)*, 2019.

9. Xinyu Wang, Tsan-Ming Choi, Xiaohang Yue, Mengji Zhang, and **Wanyu Du**. An Effective Optimization Algorithm for Application Mapping in Network-on-Chip Designs, *IEEE Transactions on Industrial Electronics*, 2019.

RESEARCH EXPERIENCE

Iterative Text Revision June 2021 - August 2021
NLP Research Intern *Grammarly Engineering*

· Proposed a novel approach to model human iterative text revision behaviors.
· Constructed new data resources for iterative text revision tasks.
· Designed human-in-the-loop iterative text revision systems. [**ACL 2022 & In2Writing Workshop**]

Controlled Open-domain Dialogue Generation January 2021 - May 2020
Research Assistant *UVa ILP Lab*

· Proposed a novel approach to control the generation of Transformer-based pretrained language models under the low-data setting.
· Designed novel control attributes losses to incorporate useful control signals into large pretrained language models.
· Conducted empirical experiments to show that our approach has better controllability, higher generation quality and better sample-efficiency than existing gradient-based and weighted-decoding methods. [**EMNLP 2021 (Findings)**]

Communicative Intent Flow for Multi-turn Dialogue Evaluation June 2020 - Dec 2020
NLP Research Intern *Tencent AI Lab*

· Proposed a novel evaluation dimension, communicative intent flow, to evaluation the coherence between interactive conversations in open-domain dialogues.
· Designed a communicative intent annotation schema and conducted crowd-sourcing human annotations on two benchmark open-domain dialogue datasets.
· Conducted empirical experiments to analyze the correlation between communicative intent flow and the overall quality of multi-turn conversations, and showed that the communicative intent flow is an essential evaluative dimension to multi-turn conversations.

Gaussian Process Priors for Diverse Text Generation June 2019 - May 2020
Research Assistant *UVa ILP Lab*

· Proposed a novel approach to enrich contextual representation learning of encoder-decoder models by introducing a stochastic function to map encoder hidden states into random context variables.
· Applied Gaussian process to explicitly model the dependency among random context variables.
· Conducted empirical experiments on downstream text generation tasks (e.g. paraphrase generation, text style transfer and dialog generation), and demonstrated that the proposed method outperforms baseline models in both quality and diversity.

Efficient Learning Algorithms for Paraphrase Generation January 2019 - May 2019
Research Assistant *UVa ILP Lab*

· Developed a unified framework for different learning algorithms (e.g. REINFORCE, DAgger, etc.) in a sequence-to-sequence model.
· Proposed variant learning algorithms based on reinforcement learning, and conducted empirical comparison on two benchmark datasets to see which algorithms would best alleviate exposure bias.
· Figured out the most effective learning algorithm for paraphrase generation, which outperformed the state-of-art model with 12.21 BLEU score. [**EMNLP 2019**]

HONORS

- **Best Paper Award at the First In2Writing Workshop**, 2022.
- **UVa Academic Excellence Fellowship**, 2018.

- **DUFE Outstanding Graduates**, 2018.
- **DUFE First Place Scholarship**, 2014-2018.
- **National Data Driven Innovation Research Competition: National 2nd Prize**, 2018.
- **Mitsubishi UFJ Trust Scholarship**, 2017.
- **Citi Financial Innovation Application Competition: National 3rd Prize**, 2016.