WANYU DU

wd5jq@virginia.edu; wyu-du.github.io; @WanyuDu

RESEARCH INTERESTS

Natural Language Processing, Large Language Models, Reinforcement Learning.

EDUCATION

University of Virginia (UVa) Aug. 2020 - May 2024 (Expected)
Ph.D. candidate in Computer Science
Advisor: Yangfeng Ji

University of Virginia (UVa) Aug. 2018 - May 2020
Master of Science in Computer Science
Advisor: Yangfeng Ji

Dongbei University of Finance and Economics (DUFE) Aug. 2014 - Jul. 2018
Bachelor of Management in E-Commerce

EXPERIENCE

Research Assistant - UVa ILP Lab Charlottesville, VA
Advised by Yangfeng Ji *Jan. 2019 - Present*

- Developed a novel reward function that approximates experts' preference to align large language models in generating faithful and informative responses.
- Proposed an uncertainty-based self-training algorithm for few-shot task-oriented dialogue generation.
- Developed an additive side network for controllable text generation.
- Designed a novel variational encoder-decoder model with Gaussian Process priors for text generation with high quality and diversity.
- Investigated efficient reinforcement learning algorithms for paraphrase generation.

Applied Scientist Intern - Amazon AWS AI New York, NY
Advised by Song Feng, James Gung, Yi Zhang, Saab Mansour *May 2023 - Aug. 2023*

- Developed a novel prompt-based controllable text generation method for large language models.
- Constructed large-scale simulation datasets that annotate task-oriented dialogues with dialogue flows.
- Performed parameter-efficient instruction-tuning with large language models on simulation datasets.

NLP Research Intern - Grammarly Engineering Vancouver, BC
Advised by Dongyeop Kang, Vipul Raheja *Jun. 2021 - Dec. 2021*

- 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.

NLP Research Intern - Tencent AI Lab Bellevue, WA
Advised by Liwei Wang, Yangfeng Ji *Jun. 2020 - Dec. 2020*

- Proposed a novel evaluation dimension, communicative intent flow, to evaluate 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.

PUBLICATIONS

[1] Blending Reward Functions via Few Expert Demonstrations for Faithful and Accurate Knowledge-Grounded Dialogue Generation
Wanyu Du, Yangfeng Ji.
arXiv preprint, 2023.

[2] Explaining Predictive Uncertainty by Looking Back at Model Explanations
Hanjie Chen, **Wanyu Du**, Yangfeng Ji.
AAAI Workshop on Uncertainty Reasoning and Quantification in Decision Making, 2023.

[3] Self-training with Two-phase Self-augmentation for Few-shot Dialogue Generation
Wanyu Du, Hanjie Chen, Yangfeng Ji.
In Findings of the Association for Computational Linguistics: EMNLP 2022, 2022.

[4] FlowEval: A Consensus-Based Dialogue Evaluation Framework Using Segment Act Flows
Jianqiao Zhao*, Yanyang Li*, **Wanyu Du***, Yangfeng Ji, Dong Yu, Michael R. Lyu, Liwei Wang.
In Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing (EMNLP 2022), 2022.

[5] Improving Iterative Text Revision by Learning Where to Edit from Other Revision Tasks
Zae Myung Kim, **Wanyu Du**, Vipul Raheja, Dhruv Kumar, Dongyeop Kang.
In Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing (EMNLP 2022), 2022.

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

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

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

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

[10] The GEM Benchmark: Natural Language Generation, its Evaluation and Metrics
Sebastian Gehrmann et al.
In Proceedings of the 1st Workshop on Natural Language Generation, Evaluation, and Metrics, 2021.

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

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

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

HONORS AND AWARDS

Best Paper Award at the First In2Writing Workshop	2022
UVa Computer Science Scholar Fellowship	2020
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

SERVICE

Program Committee	EMNLP, DialDoc Workshop, In2Writing Workshop	2022-2023
Reviewer	ACL, EMNLP, NAACL, CoNLL, ACL Rolling Review	2020-2023
Organizing Team Member	GEM Workshop at ACL 2021	2021

TEACHING

CS 4501 Reinforcement Learning, Teaching Assistant	2022
CS 6316 Machine Learning, Teaching Assistant	2022
CS 6501 Interpretable Machine Learning, Teaching Assistant	2022
CS 6501 Natural Language Processing, Teaching Assistant	2021

TECHNICAL SKILLS

Programming Language	Python, Java, C, SQL
Software & Tools	Pytorch, Tensorflow, MATLAB, Neo4j