SHAOJIE JIANG

Machine Learning Scientist

Homepage: https://shaojiejiang.github.io/

Email: shaojie.jiang1@gmail.com

Phone: +31 6 82928861

I am a Machine Learning Scientist at Huawei R&D Center Amsterdam. My work involves chatbots, information retrieval and conversational QA systems. In parallel, I am completing my PhD at the University of Amsterdam. I have interned at Replika.ai and Amazon.com.

EDUCATION

2017 - present	University of Amsterdam, The Netherlands PhD Candidate Research area: *Open-domain Dialogue Systems* Supervisors: Prof. dr. Maarten de Rijke & Dr. Svitlana Vakulenko Thesis plan: Expected to finish by June 2023
2014 - 2017	Northwest A&F University, Yangling, China Master of Engineering Thesis: *Research on Feature Representation and Optimization Methods in Structured Object Tracking* Supervisor: Prof. Jifeng Ning
2010 - 2014	Northwest A&F University, Yangling, China Bachelor of Engineering Thesis: *Implementation of Single Image Haze Removal Using Dark Channel Prior* Advisors: Dr. Yaojun Geng and Prof. Jifeng Ning

WORKING EXPERIENCE

2022.07 - present	Huawei R&D Center, Amsterdam, the Netherlands Machine Learning Scientist Area: Chatbots and Conversational QA Systems for Petal Search Responsibility: • Lead of ChatEval challenge project at DSTC 11 • Contact person of the Conversational AI project • Contact person of DReaMS Lab work package 3
2021.02 - 2021.07	Replika, San Francisco, USA Research Intern Project: Improving Generative Dialogue Engagingness

Team: AI Research Team
Mentor: Artem Rodichev
Outcome: A chatbot model with improved engagingness, and a follow-up research paper accepted at CHIIR 2023 (see below)

2020.07 - 2020.09   Amazon.com, Inc., Amsterdam, The Netherlands
Applied Science Intern
Project: Customer Review Summarisation
Team: Subjective NLP
Mentor: Alex Klementiev
Outcome: An automatic aspect extraction system for customer reviews

PUBLICATIONS

Shaojie Jiang, Ruqing Zhang, Junfei Guo, Evangelos Kanoulas, and Maarten de Rijke. *"Reproducing WebGPT Using Open-source Materials"* To be submitted to SIGIR 2023 Reproducibility Track.

Shaojie Jiang, Ruqing Zhang, Svitlana Vakulenko, and Maarten de Rijke. *"A Simple Contrastive Learning Objective for Alleviating Neural Text Degeneration."* Submitted to SIGIR 2023 Full Papers Track.

Shaojie Jiang, Svitlana Vakulenko, and Maarten de Rijke. *"Weakly Supervised Turn-level Engagingness Evaluator for Dialogues."* To appear in CHIIR 2023 as a full paper. Source code link below (Lit-Seq).

Shaojie Jiang, Ruqing Zhang, Svitlana Vakulenko, and Maarten de Rijke. *"A Simple Contrastive Learning Objective for Alleviating Neural Text Degeneration."* arXiv preprint arXiv:2205.02517. 2022. Source code link below (Lit-Seq).

Shaojie Jiang, Thomas Wolf, Christof Monz, and Maarten de Rijke. *"TLDR: Token Loss Dynamic Reweighting for Reducing Repetitive Utterance Generation."* arXiv preprint arXiv:2003.11963. 2020. Source code link below (TLDR).

Shaojie Jiang, Pengjie Ren, Christof Monz, and Maarten de Rijke. *"Improving Neural Response Diversity with Frequency-Aware Cross-Entropy Loss."* In The World Wide Web Conference, pp. 2879-2885. ACM, 2019. Source code link below (FACE).

Shaojie Jiang, and Maarten de Rijke. *"Why are Sequence-to-Sequence Models So Dull? Understanding the Low-Diversity Problem of Chatbots."* In Proceedings of the 2018 EMNLP Workshop SCAI: The 2nd

International Workshop on Search-Oriented Conversational AI, pp. 81-86. 2018.

Shaojie Jiang, Jifeng Ning, Cheng Cai, and Yunsong Li. *"Robust Struck tracker via color Haar-like feature and selective updating."* Signal, Image and Video Processing 11, no. 6 (2017): 1073-1080.

Jifeng Ning, Jimei Yang, **Shaojie Jiang**, Lei Zhang, and Ming-Hsuan Yang. *"Object tracking via dual linear structured SVM and explicit feature map."* In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 4266-4274. 2016.

OPEN SOURCE CONTRIBUTIONS

Lit-Seq project	https://github.com/ShaojieJiang/lit-seq
TLDR project	https://github.com/ShaojieJiang/tldr
FACE project	https://github.com/ShaojieJiang/FACE

AWARDS

2019	Invited by NLP Summit at Google Zurich
2018	Student Travel Grant for the 2nd SCAI Workshop
2017	Faculty Outstanding Master Thesis Award
2013	National Endeavor Scholarship
2012	National Endeavor Scholarship

TEACHING AND SUPERVISION

2021.02 - 2020.03	Teaching Assistant for Master AI course Information Retrieval 1
2019.11 - 2020.06	Master AI Thesis Supervision: Traian Vidraşcu, *"Link Prediction for Software Graphs using Inbound and Outbound Neighborhoods"*
2019.11 - 2020.06	Master AI Thesis Supervision: Kai Liang, *"Neural Hierarchical Text Classification with Transfer Learning Approaches"*
2019.04 - 2019.07	Master Data Science Thesis Supervision: Savvina Daniil, *"Training Support Vector Machines Using Multi-Party Computation"*
2019.06 - 2019.06	Teaching Assistant for Master AI Project-AI course on "Repetition Reduction for Chatbots"
2018.10 - 2018.12	Teaching Assistant for Master AI course Information Retrieval 2 on "Hierarchical Neural Response Generation"
2018.01 - 2018.06	Master AI Thesis Supervision: Ricardo Fabián Guevara Meléndez, *"Task-Oriented Dialog Agents Using Memory-Networks and Ensemble Learning"*

COMMUNITY SERVICE

2022	Reviewer of TOIS
2021	Reviewer for JSCT and TALLIP
2020.02-2021.02	Member of PhD Council at the Informatics Institute of the University of Amsterdam
2020	Reviewer for TOIS, ACL, EMNLP and EACL
2019.09-2020.11	Representative of Informatics Institute on PhD Council of the Faculty of Science of the University of Amsterdam
2019.03-2019.06	Subreviewer for SIGIR and CIKM

SKILLS Python, PyTorch, Git, Bash, LaTeX, Photography

REFERENCES

Prof. dr. Maarten de Rijke   Distinguished University Professor AI & IR at University of Amsterdam, Director Innovation Center for Artificial Intelligence. Email: m.deRijke@uva.nl

Dr. Jun Luo   Machine Translation team lead at Huawei R&D Center Amsterdam. Email: junluo@huawei.com

January 29, 2023