

EDUCATION

Utrecht University

Utrecht, the Netherlands

Ph.D. in Computer Science, Advisor: Dr. Dong Nguyen

07.2020–Current

- Working on the robustness of NLP models, including their training instability, societal biases, and generalization to unseen test data.

East China Normal University

Shanghai, China

M.Sc. in Computer Science, Advisor: Dr. Yuanbin Wu

09.2017–07.2020

- Thesis: “Research on Detecting Gender Stereotypes in Texts”, leads to a long paper in EMNLP 2019.

East China Normal University

Shanghai, China

B.Sc. in Psychology

09.2013–07.2017

- Thesis: “Research on the Attention of Negative Emotional Stimuli”

WORKING EXPERIENCE

Amazon

Berlin, Germany; Amsterdam, the Netherlands

Applied Scientist Intern

11.2022 - Current

Working on catalog enrichment.

Sogou

Hangzhou, China

Research Intern

07.2019 - 01.2020

Developed a VAE-based question generation system to improve the generation diversity. Used crowd-sourcing to annotate a Chinese question generation dataset consisting of 20k+ samples.

SELECTED PUBLICATIONS

1. **Yupei Du** and Dong Nguyen, “Measuring the Instability of Fine-Tuning”, *arXiv pre-print*.
2. **Yupei Du**, Qi Zheng, Yuanbin Wu, Man Lan, Yan Yang, and Meirong Ma, “Understanding Gender Bias in Knowledge Base Embeddings”, in *ACL 2022*.
3. **Yupei Du**, Qixiang Fang, and Dong Nguyen, “Assessing the Reliability of Word Embedding Gender Bias Measures”, in *EMNLP 2021*.
4. **Yupei Du**, Yuanbin Wu, and Man Lan, “Exploring Human Gender Stereotypes with Word Association Test”, in *EMNLP 2019*.

ACADEMIC SERVICE

- **Program Committee** 2021 - Current
ACL, EMNLP, NAACL, EACL, ACL Rolling Review
- **Outstanding Reviewer**
ACL 2022

TEACHING

- **Teaching Assistant** at Utrecht University Fall 2020 - 2021
Methods in Artificial Intelligence Research (INFOMAIR)
- **Teaching Assistant** at Utrecht University Spring 2021 - 2022
Applications of Machine Learning (INKUB3APML)
- **Teaching Assistant** at Utrecht University Spring 2021 - 2022
Human-centered Machine Learning (INFOMHCML)

SUPERVISION

- Joris Albers, “Comparing Contextualised Embeddings for Predicting the (Graded) Effect of Context in Word Similarity”, Bachelor’s thesis, 2021.
- Alexander van Leeuwen, “Detecting Toxic Text Spans using SHAP”, Bachelor’s thesis, 2022.
- Xinyu Zhang, “Study on Iterative Nullspace Projection Debiasing of Multiple Attributes of Word Embedding”, Master’s thesis, 2022.

SCHOLARSHIPS AND AWARDS

- ECNU Zhihui Award 2019
Top 5% of the master’s students in East China Normal University.
- Utrecht University Teaching Award 2022
In recognition of distinguished teaching of Applications of Machine Learning.