

EDUCATION

Utrecht University

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

– Working on data-centric efficient and robust NLP models

Utrecht, the Netherlands

07.2020–Current

East China Normal University

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

– Thesis: “Research on Detecting Gender Stereotypes in Texts”, long paper in EMNLP 2019

Shanghai, China

09.2017–07.2020

East China Normal University

B.Sc. in Psychology

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

Shanghai, China

09.2013–07.2017

WORKING EXPERIENCE

Amazon

Applied Scientist Intern

Worked on user-data-driven enrichment of knowledge graph based product catalog, mentored by Prof. Dr. Ziawasch Abedjan, long paper in ECML 2024.

Berlin, Germany

11.2022 - 05.2023

Sogou

Research Intern

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.

Hangzhou, China

07.2019 - 01.2020

SELECTED PUBLICATIONS

1. **Yupei Du**, Albert Gatt, and Dong Nguyen, “FTFT: efficient and robust Fine-Tuning by transFerring Training dynamics”, in *COLING 2025 and Data-centric Machine Learning at ICLR 2024*.
2. **Yupei Du**, Jacek Golebiowski, Philipp Schmidt, and Ziawasch Abedjan, “Guiding Catalogue Enrichment with User Queries”, in *ECML 2024*.
3. Goya van Boven, **Yupei Du**, and Dong Nguyen, “Transforming Dutch: Debiasing Dutch Coreference Resolution Systems for Non-binary Pronouns”, in *FAccT 2024*.
4. **Yupei Du** and Dong Nguyen, “Measuring the Instability of Fine-Tuning”, in *ACL 2023*.
5. **Yupei Du**, Qi Zheng, Yuanbin Wu, Man Lan, Yan Yang, and Meirong Ma, “Understanding Gender Bias in Knowledge Base Embeddings”, in *ACL 2022*.
6. **Yupei Du**, Qixiang Fang, and Dong Nguyen, “Assessing the Reliability of Word Embedding Gender Bias Measures”, in *EMNLP 2021*.
7. **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 and Lecturer** at Utrecht University 2024
Transformers: Applications in Language and Communication (INFOMTALC)
- **Teaching Assistant** at Utrecht University 2020
Methods in Artificial Intelligence Research (INFOMAIR)
- **Teaching Assistant** at Utrecht University 2021 - 2023
Applications of Machine Learning (INKUB3APML)
- **Teaching Assistant** at Utrecht University 2021 - 2023
Human-centered Machine Learning (INFOMHCML)

EXAMPLE SUPERVISION

- Goya van Boven, “Transforming Dutch: Debiasing Dutch Coreference Resolution Systems for Non-Binary Pronouns”, Master’s thesis, 2024. Long paper in FAccT 2024.
- Daan van der Weijden, “Dutch Data Augmentation For Low-Resource Tasks – can large language models be used for dutch data augmentation?”, Master’s thesis, 2023.
- Xinyu Zhang, “Study on Iterative Nullspace Projection Debiasing of Multiple Attributes of Word Embedding”, Master’s thesis, 2022.

SCHOLARSHIPS AND AWARDS

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