

# Yijia Shao

Ph.D. Student

Computer Science Department  
Stanford University

Email: [shaoyj@stanford.edu](mailto:shaoyj@stanford.edu)

Homepage: <https://shaoyijia.github.io/>

## Education

2023–now     Ph.D., Computer Science, Stanford University

2019–2023     B.S., Data Science, Peking University

Thesis Title: Continual Learning in Natural Language Processing (*Top 10 Bachelor Thesis in School of EECS, Peking University, 2023*)

Supervisors: Prof. [Bing Liu](#) and Prof. [Dongyan Zhao](#)

## Research Interests

Natural Language Processing, Machine Learning

## Awards & Honors

- School of Engineering Fellowship, Stanford, 2023
- SenseTime Scholarship, 2022
- National Scholarship by the Ministry of Education of the People's Republic of China, 2020, 2022
- May 4th Scholarship, Peking University, 2021
- First prize in 12th Chinese Mathematics Competition Final, 2020
- First Class Scholarship for Freshmen (rank 2<sup>nd</sup> in National College Entrance Examination), 2019

## Experience

- **Plus Lab, University of California, Los Angeles** *July 2022 - Oct 2022*  
- Visiting Summer Intern, advised by Prof. Nanyun (Violet) Peng
- **Data, Knowledge, and Intelligence, Microsoft Research Asia** *Jan 2022 - July 2022*  
- Research Intern, advised by Dr. Mengyu Zhou
- **TensorFlow Lite, Google** *July 2021 - Sep 2021*  
- Engineering Intern, hosted by Tiezhen Wang and Chao Mei

# Publications

## Peer-Reviewed Conference

- [8] *ACCENT: An Automatic Event Commonsense Evaluation Metric for Open-Domain Dialogue Systems*  
Sarik Ghazarian\*, **Yijia Shao\***, Rujun Han, Aram Galstyan, Nanyun Peng  
The 61st Annual Meeting of the Association for Computational Linguistics. ACL 2023
- [7] *Class-Incremental Learning based on Label Generation*  
**Yijia Shao**, Yiduo Guo, Dongyan Zhao, Bing Liu  
The 61st Annual Meeting of the Association for Computational Linguistics. ACL 2023 short
- [6] *AnaMeta: A Table Understanding Dataset of Field Metadata Knowledge Shared by Multi-dimensional Data Analysis Tasks*  
Xinyi He, Mengyu Zhou, Mingjie Zhou, Jialiang Xu, Xiao Lv, Tianle Li, **Yijia Shao**, Shi Han, Zejian Yuan, Dongmei Zhang  
The 61st Annual Meeting of the Association for Computational Linguistics. ACL 2023 (Findings)
- [5] *Continual Pre-training of Language Models*  
Zixuan Ke\*, **Yijia Shao\***, Haowei Lin\*, Tatsuya Konishi, Gyuhak Kim, Bing Liu  
The Eleventh International Conference on Learning Representations. ICLR 2023
- [4] *FormLM: Recommending Creation Ideas for Online Forms by Modelling Semantic and Structural Information*  
**Yijia Shao**, Mengyu Zhou, Yifan Zhong, Tao Wu, Hongwei Han, Gideon Huang, Dongmei Zhang  
The 2022 Conference on Empirical Methods in Natural Language Processing. EMNLP 2022
- [3] *Adapting a Language Model While Preserving its General Knowledge*  
Zixuan Ke, **Yijia Shao**, Haowei Lin, Hu Xu, Lei Shu, Bing Liu  
The 2022 Conference on Empirical Methods in Natural Language Processing. EMNLP 2022
- [2] *Continual Training of Language Models for Few-Shot Learning*  
Zixuan Ke, Haowei Lin, **Yijia Shao**, Hu Xu, Lei Shu, Bing Liu  
The 2022 Conference on Empirical Methods in Natural Language Processing. EMNLP 2022 short
- [1] *CMG: A Class-Mixed Generation Approach to Out-of-Distribution Detection*  
Mengyu Wang\*, **Yijia Shao\***, Haowei Lin, Wenpeng Hu, Bing Liu  
The European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases. ECML-PKDD 2022

## Archive Preprints

- [1] *LUNA: Language Understanding with Number Augmentations on Transformers via Number Plugins and Pre-training*  
Hongwei Han, Jialiang Xu, Mengyu Zhou, **Yijia Shao**, Shi Han, Dongmei Zhang  
ArXiv Preprint (arXiv:2212.02691)

## Teaching

- Teaching Assistant, Introductory Topics in Data Science  
*Peking University*

*Fall 2022*

## Professional Service

### Reviewer

- ACL 2023.
- EMNLP 2023.
- ACL Rolling Review 2023.
- NeurIPS 2023.

## Invited Talks

- "Thoughts on NLP Research in the (Post-)LLM Era"  
Guest talk at SMS seminar, Peking University, Apr 2023

Last update: September 23, 2023