

EDUCATION

---

**Georgia Institute of Technology**

Atlanta, USA

PhD in Computer Science

Aug. 2019 – Present

- Advisor: Diyi Yang, Mark Riedl
- GPA: 3.72/4.00

**Stanford University**

Atlanta, USA

Visiting Student Researcher

Sep. 2023 – Present

- Advisor: Diyi Yang

**Zhejiang University**

Hangzhou, China

B.Eng. in Computer Science and Technology

Sep. 2015 – Jul. 2019

- GPA: 3.86/4.00
- Minor in [Advanced Honor Class of Engineering Education, Chu Konchen Honors College](#) (40/6,000)

Research Interest

---

- Learning with structures for efficient NLP systems (data efficiency, training efficiency).
- Building social/linguistic/commonsense-aware NLP systems.
- Unlocking the ability to solve complex problems for large language models.

INTERNSHIP

---

**Tencent AI (US)**

Seattle, USA

Research Intern

May 2023 – Aug 2023

- Mentor: Jianshu Chen
- Research Topic: Solving Complex Problems via Skills-in-Context Prompting

**Amazon**

Santa Clara, USA

Applied Scientist Intern @ AWS

June 2022 – Jan 2023

- Mentor: Aston Zhang, Alex Smola
- Research Topic: Parameter Efficient Finetuning, Diffusion Language Model

**Allen Institute for AI**

Seattle, USA

Research Intern @ Mosaic Team

June 2021 – Sep. 2021

- Mentor: Keisuke Sakaguchi, Chandra Bhagavatula, and Yejin Choi
- Research Topic: Compositional Augmentation for Commonsense Question Answering

**Microsoft Research**

Seattle, USA

Researcher @ Microsoft Dynamics 365 AI

Dec. 2020 – May 2021

- Mentor: Dinghan Shen, and Weizhu Chen
- Research Topic: Data Augmentation, Efficient Finetuning

PUBLICATIONS

---

1. “Unlearn What You Want to Forget: Efficient Unlearning for LLMs”,  
**Jiaao Chen** and Diyi Yang  
*The 2023 Conference on Empirical Methods on Natural Language Processing, EMNLP 2023*
2. “A Cheaper and Better Diffusion Language Model with Soft-Masked Noise”,  
**Jiaao Chen** and Diyi Yang  
*The 2023 Conference on Empirical Methods on Natural Language Processing, EMNLP 2023*
3. “Is ChatGPT a General-Purpose Natural Language Processing Task Solver?”,  
Chengwei Qin, Aston Zhang, Zhuosheng Zhang, **Jiaao Chen**, Michihiro Yasunaga, and Diyi Yang  
*The 2023 Conference on Empirical Methods on Natural Language Processing, EMNLP 2023*

4. “Mitigating Biases in Hate Speech Detection from A Causal Perspective”,  
Zhehao Zhang, **Jiaao Chen**, and Diyi Yang  
*The 2023 Conference on Empirical Methods on Natural Language Processing. EMNLP (Findings) 2023*
5. “Where Does Your News Come From? Predicting Information Pathways in Social Media”,  
Alexander K Taylor, Nuan Wen, Po-Nien Kung, **Jiaao Chen**, Violet Peng, and Wei Wang  
*Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval. SIGIR 2023*
6. “Controllable Conversation Generation with Conversation Structures via Diffusion Models”,  
**Jiaao Chen**, and Diyi Yang  
*The 61th Annual Meeting of the Association for Computational Linguistics (Findings). ACL 2023*
7. “Human-in-the-loop Abstractive Dialogue Summarization”,  
**Jiaao Chen**, Mohan Dodda, and Diyi Yang  
*The 61th Annual Meeting of the Association for Computational Linguistics (Findings). ACL 2023*
8. “Compositional Data Augmentation for Abstractive Conversation Summarization”,  
Siru Ouyang, **Jiaao Chen**, Jiawei Han, and Diyi Yang  
*The 61th Annual Meeting of the Association for Computational Linguistics. ACL 2023*
9. “Parameter-Efficient Fine-Tuning Design Spaces”,  
**Jiaao Chen**, Aston Zhang, Xingjian Shi, Mu Li, Alix Smola, and Diyi Yang  
*11th International Conference on Learning Representations. ICLR 2023*
10. “An Empirical Survey of Data Augmentation for Limited Data Learning in NLP”,  
**Jiaao Chen\***, Derek Tam\*, Colin Raffel, Mohit Bansal, and Diyi Yang  
*Transactions of the Association for Computational Linguistics. TACL*
11. “When FLUE Meets FLANG: Benchmarks and Large Pretrained Language Model for Financial Domain”,  
Raj Sanjay Shah, Kunal Chawla, Dheeraj Eidnani, Agam Shah, Wendi Du, Sudheer Chava, Natraj Raman, Charese Smiley, **Jiaao Chen**, and Diyi Yang  
*The 2022 Conference on Empirical Methods on Natural Language Processing. EMNLP 2022*
12. “E-VALUE: Understanding Dialect Disparity in NLU”,  
Caleb Ziem\*, **Jiaao Chen\***, Camille Harris, Jessica Anderson, and Diyi Yang  
*The 60th Annual Meeting of the Association for Computational Linguistics. ACL 2022*
13. “Focus on the Action: Learning to Highlight and Summarize Jointly for Email To-Do Items Summarization”,  
Kexun Zhang, **Jiaao Chen**, Diyi Yang,  
*The 60th Annual Meeting of the Association for Computational Linguistics. ACL 2022*
14. “Leveraging Expert Guided Adversarial Augmentation For Improving Generalization in Named Entity Recognition”,  
Aaron Reich, **Jiaao Chen**, Aastha Agrawal, Yanzhe Zhang, Diyi Yang,  
*The 60th Annual Meeting of the Association for Computational Linguistics (Findings). ACL 2022*
15. “Simple Conversational Data Augmentation for Semi-supervised Abstractive Dialogue Summarization”,  
**Jiaao Chen**, Diyi Yang,  
*The 2021 Conference on Empirical Methods on Natural Language Processing. EMNLP 2021*
16. “HiddenCut: Simple Data Augmentation for Natural Language Understanding with Better Generalization”,  
**Jiaao Chen**, Dinghan Shen, Weizhu Chen, and Diyi Yang,  
*The 59th Annual Meeting of the Association for Computational Linguistics. ACL 2021*
17. “Structure-Aware Abstractive Conversation Summarization via Discourse and Action Graphs”,  
**Jiaao Chen**, Diyi Yang,  
*2021 Annual Conference of the North American Chapter of the Association for Computational Linguistics. NAACL 2021*
18. “Continual Learning for Text Classification with Information Disentanglement Based Regularization”,  
Yufan Huang, Yanzhe Zhang, **Jiaao Chen**, Xuezhi Wang and Diyi Yang,  
*2021 Annual Conference of the North American Chapter of the Association for Computational Linguistics. NAACL 2021*
19. “Weakly-Supervised Hierarchical Models for Predicting Persuasive Strategies in Good-faith Textual Requests”,  
**Jiaao Chen**, Diyi Yang,  
*The Thirty-Fifth AAAI Conference on Artificial Intelligence. AAAI 2020*
20. “Multi-View Sequence-to-Sequence Models with Conversational Structure for Abstractive Dialogue Summarization”,  
**Jiaao Chen**, Diyi Yang,  
*The 2020 Conference on Empirical Methods on Natural Language Processing. EMNLP 2020*
21. “Local Additivity Based Data Augmentation for Semi-supervised NER”,  
**Jiaao Chen\***, Zhenghui Wang\*, Ran Tian, Zichao Yang and Diyi Yang,  
*The 2020 Conference on Empirical Methods on Natural Language Processing. EMNLP 2020*
22. “Examining the Ordering of Rhetorical Strategies in Persuasive Requests”,  
Omar Shaikh, **Jiaao Chen**, Jon Saad-Falcon, Polo Chau and Diyi Yang,

*The 2020 Conference on Empirical Methods on Natural Language Processing (Findings). EMNLP 2020*

23. “MixText: Linguistically-Informed Interpolation of Hidden Space for Semi-Supervised Text Classification”,  
**Jiaao Chen**, Zichao Yang, Diyi Yang,  
*The 58th annual meeting of the Association for Computational Linguistics (ACL). ACL 2020*
24. “Semi-supervised Models via Data Augmentation for Classifying Interactive Affective Responses”,  
**Jiaao Chen\***, Yuwei Wu\*, Diyi Yang,  
*AAAI Workshop on Affective Content Analysis, 2020*
25. “Let’s Make Your Request More Persuasive: Modeling Persuasive Strategies via Semi Supervised Neural Nets on Crowdfunding Platforms”,  
Diyi Yang\*, **Jiaao Chen\***, Zichao Yang, Dan Jurafsky, Eduard Hovy,  
*The 17th Annual Conference of the North American Chapter of the Association for Computational Linguistics. NAACL 2019.*
26. “Incorporating Structured Commonsense Knowledge in Story Completion”,  
**Jiaao Chen**, Jianshu Chen, Zhou Yu,  
*The Thirty-Fourth AAAI Conference on Artificial Intelligence. AAAI 2019*

#### TEACHING EXPERIENCE

---

<b>School of Computer Science, Georgia Institute of Technology</b>	Atlanta, USA
Teaching Assistant at NLP: CS-4650/7650 with Professor Diyi Yang	Jan. 2020 – May. 2020
<b>Department of Computer Science, Zhejiang University</b>	Hangzhou, China
Teaching Assistant at Computer Organization with Professor Qingsong Shi	Mar. 2018 – Jun. 2018

#### ADVISING

---

##### Master Students

- Aaron Reich (Master, 2021 Spring – 2022 Fall)
- Aastha Agrawa (Master, 2021 Spring – 2022 Fall)
- Yufan Huang (Master, 2020 Fall - 2021 Spring)
- Kunal Chawla (Master, 2020 Fall – 2021 Fall)
- Nikhil Gupta (Master, 2020 Fall – 2021 Spring)

##### Undergraduate Students

- Zhehao Zhang (Undergraduate, 2022 Summer –)
- Siru Ouyang (Undergraduate, 2022 Summer – 2023 Spring)
- Mohan Dodda (Undergraduate, 2022 Summer – 2023 Spring)
- Wenna Qin (Undergraduate, 2022 Fall – 2023 Spring)
- Kexun Zhang (Undergraduate, 2021 Spring – 2022 Spring)
- Dheeraj Eidnani (Undergraduate, 2020 Fall – 2022 Spring)
- Yanzhe Zhang (Undergraduate, 2020 Fall - 2021 Spring)
- Omar Shaikh (Undergraduate, 2019 Fall – 2020 Spring)
- Jon Saad-Falcon (Undergraduate, 2019 Fall - 2020 Spring)

#### SERVICE EXPERIENCE

---

##### Conference/Journal Reviewing Activities

- ACL 2021, 2022, 2023
- EMNLP 2020, 2021, 2022, 2023
- NAACL 2021
- IJHCI 2023
- CHI 2023
- UIST 2022
- WWW 2021
- AAAI 2020, 2021
- ICLR 2022

##### Conference Volunteer

- ACL 2020

##### Georgia Tech NLP Seminar

- Organizer (2019 – 2021)

#### RECENT PRESENTATION & INVITED TALKS

---

- “An Empirical Survey of Data Augmentation for Limited Data Learning in NLP”. Conference Talk at ACL 2023.
- “Human-in-the-loop Abstractive Dialogue Summarization”. Invited Talk at Adobe 2023.
- “Simple Conversational Data Augmentation for Semi-supervised Abstractive Dialogue Summarization”. Conference Talk at EMNLP 2021.
- “HiddenCut: Simple Data Augmentation for Natural Language Understanding with Better Generalization”. Conference Talk at ACL 2021.
- “Structure-Aware Abstractive Conversation Summarization via Discourse and Action Graphs”. Conference Talk at NAACL 2021.
- “Conversation Summarization with Multiple Dialogue Views”. Invited Talk at Google NLP, Nov. 13<sup>th</sup>, 2020.
- “Multi-View Sequence-to-Sequence Models with Conversational Structure for Abstractive Dialogue Summarization”. Conference Talk at EMNLP 2020.
- “Local Additivity Based Data Augmentation for Semi-supervised NER”. Conference Talk at EMNLP 2020.
- “MixText: Linguistically-Informed Interpolation of Hidden Space for Semi-Supervised Text Classification”. Conference Talk at ACL 2020.
- “Incorporating Structured Commonsense Knowledge in Story Completion”. Conference Talk at AAAI 2019.

#### SELECTED AWARDS AND HONORS

---

- |   |      |
|---|------|
| • Outstanding Undergraduate Thesis, Zhejiang University.                          | 2019 |
| • Honorable Mention at MCM/ICM Contest.   | 2017 |
| • Academic Physics Contest at Zhejiang University, fourth place.                  | 2017 |
| • ACEE Honor Class at Chu Konchen Honors College, Zhejiang University (40/6,000). | 2016 |
| • Scholarship for Academic Achievement at Zhejiang University.                    | 2016 |
| • Scholarship for All-around Achievement at Zhejiang University.                  | 2016 |
| • ACM School Trial at Zhejiang University, third prize.                           | 2016 |

#### ADDITIONAL INFORMATION

---

##### **Technology and Language Skills:**

- Programming Languages: C/C++, Python, Verilog, Matlab, Labview, SQL, C#, Lex/Yacc.
- Development Skills: Pytorch, Tensorflow, OpenCV, QT.

##### **Interests:**

- Piano (Level 10 Certificate), Basketball.