

# Yuchen Li

**Email:** yuchenl4@cs.cmu.edu

**LinkedIn:** <https://www.linkedin.com/in/yuchenli01/>

**Website:** <https://www.cs.cmu.edu/~yuchenl4/>

**Twitter:** [https://twitter.com/\\_Yuchen\\_Li\\_](https://twitter.com/_Yuchen_Li_)

## EDUCATION

---

### Carnegie Mellon University

August 2020 - May 2025 (expected)

- Ph.D. student in Machine Learning, advised by Prof. Andrej Risteski Pittsburgh, PA
- Research interests: machine learning, natural language processing, and data mining GPA: 4.24/4.33
- I currently work on (1) improving mathematical understanding of how neural network models process language, and (2) developing principled approaches to self-supervised learning

### University of Illinois at Urbana-Champaign

August 2015 - May 2019

- B.S. Statistics and Computer Science; second major in Mathematics; minor in Electrical Engineering Urbana-Champaign, IL
- Research mentors: Prof. Jiawei Han, Prof. AJ Hildebrand, Prof. Pramod Viswanath GPA: 4.0/4.0
- Summa Cum Laude; Highest Distinction in both majors
- Bronze Tablet Scholar: top 3% in the graduating class
- John R. Pasta Outstanding Undergraduate Award 2018: 2 out of about 400 computer science undergraduate junior students

## WORK

---

### Research Intern, Google LLC

May 2023 - August 2023

- Working on improving language model decoding / sampling Remote or New York City, United States

### Research Intern, ByteDance Ltd. (parent company of TikTok)

May 2022 - August 2022

- Working on a research project about machine translation at the ByteDance AI Lab Remote, United States
- Investigating the effect of language similarity in machine learning models for multi-lingual translation

### Machine Learning Engineer, Quora, Inc.

August 2019 - August 2020

- Developed neural network models for predicting ads clickthrough rate Mountain View, CA
- Led team-wide initiative of improving system robustness and reducing on-call burden in the Ads Ranking team
- Developed features for modeling the ads and the users
- Improved backend for cost-efficient low-latency online serving of machine learning models
- Launched online A/B testing to compare different model variants based on a set of key business metrics

### Software Engineer Intern, Facebook, Inc.

May 2018 - August 2018

- Designed and implemented a platform for investigating machine learning feature importance Menlo Park, CA
- Proposed and developed feature perturbation and statistical analysis methods

## PUBLICATIONS

---

(\* indicates equal contribution or alphabetical order)

1. **Yuchen Li**, Yuanzhi Li, Andrej Risteski. *How Do Transformers Learn Topic Structure: Towards a Mechanistic Understanding*. International Conference on Machine Learning (ICML) 2023.
2. Ashwini Pokle \*, Jinjin Tian \*, **Yuchen Li** \*, Andrej Risteski. *Contrasting the landscape of contrastive and non-contrastive learning*. Conference on Artificial Intelligence and Statistics (AISTATS) 2022.
3. **Yuchen Li**, Andrej Risteski. *The Limitations of Limited Context for Constituency Parsing*. Association for Computational Linguistics (ACL) 2021.

4. Xinwei He \*, A.J. Hildebrand \*, **Yuchen Li** \*, Yunyi Zhang \*. *Complexity of Leading Digit Sequences*. Journal of Discrete Mathematics & Theoretical Computer Science, vol. 22 no. 1, Automata, Logic and Semantics 2020.
5. Yu Shi \*, Jiaming Shen \*, **Yuchen Li**, Naijing Zhang, Xinwei He, Zhengzhi Lou, Qi Zhu, Matthew Walker, Myunghwan Kim and Jiawei Han. *Discovering Hypernymy in Text-Rich Heterogeneous Information Network by Exploiting Context Granularity*. Conference on Information and Knowledge Management (CIKM) 2019.
6. Hongyu Gong, **Yuchen Li**, Suma Bhat and Pramod Viswanath. *Context-Sensitive Malicious Spelling Error Correction*. The Web Conference (WWW) 2019.
7. **Yuchen Li** \*, Zhengzhi Lou \*, Yu Shi, and Jiawei Han. *Temporal Motifs in Heterogeneous Information Networks*. In Proceedings of the 14<sup>th</sup> International Workshop on Mining and Learning with Graphs (MLG) 2018.

## TALKS

---

### ***How Do Transformers Learn Topic Structure: Towards a Mechanistic Understanding***

- University of Illinois at Urbana-Champaign (UIUC)'s Machine Learning Reading Group Virtual, April 2023

### ***The Limitations of Limited Context for Constituency Parsing***

- Association for Computational Linguistics (ACL) Conference Virtual, August 2021
- NEC Laboratories Europe Virtual, July 2021
- Approximately Correct Machine Intelligence (ACMI) Lab, CMU Virtual, June 2021

## TEACHING

---

### ***Teaching Assistant, Advanced Introduction to Machine Learning***

August 2022 - December 2022

- Taught one recitation session on optimization Pittsburgh, PA
- Held weekly office hours; created and graded homework problems
- Course webpage: <https://www.cs.cmu.edu/~nihars/teaching/10715-Fa22/index.html>
- Instructor: Prof. Nihar Shah

### ***Teaching Assistant, Probabilistic Graphical Models***

January 2022 - May 2022

- Taught two recitation sessions (about algorithmic reduction and causality, respectively) Pittsburgh, PA
- Mentored nine teams of student projects (topics include natural language processing, information retrieval, 3D computer vision, and causality)
- Held weekly office hours; created and graded homework and quiz problems
- Course webpage: <https://andrejristeski.github.io/10708-22/>
- Instructors: Prof. Andrej Risteski and Prof. Hoda Heidari

### ***Volunteer Teaching Assistant, Data Science Bootcamp***

May 2019

- Contributed to creating a Python tutorial focusing on NumPy and Matplotlib libraries. Urbana, IL
- Helped in the class section for about 20 math graduate students and answered their programming questions
- Program webpage: <https://ravat1.github.io/2019DSB-website/>
- Instructor: Prof. Uma Ravat

### ***Member, Eta Kappa Nu (HKN) Honors Society Alpha Chapter at UIUC***

August 2016 - May 2019

- Contributed to its educational website Weber's Wiki Urbana, IL
- Developed and taught two Analog Signal Processing midterm review sessions, each with about 40 students attending

## SERVICE

---

### ***Reviewer, Academic Research Conferences***

- International Conference on Learning Representations (ICLR) 2023
- Artificial Intelligence and Statistics (AISTATS) 2022
- Neural Information Processing Systems (NeurIPS) 2021

### ***Member, CMU Machine Learning Admitted PhD Student Open House Committee***

2022

- Planned and organized information sessions and social events for admitted PhD students Pittsburgh, PA

### ***Mentor, CMU Computer Science Graduate Application Support Program***

November 2020

- Connected with potential applicants, particularly those from underrepresented groups Pittsburgh, PA
- Offered feedbacks on their application materials and answered their questions related to graduate school application

## SKILLS

---

- Proficient in *Python, PyTorch, TensorFlow, C, SQL, MATLAB, PHP*;    intermediate in *C++, R, Haskell*;    basic in *Java, JavaScript*