Ansong Ni

ansong.ni@yale.edu • http://niansong1996.github.io • +1 (412) 641-9684 • New Haven, CT 06511

I teach large language models to write code and reason about program behavior like human programmers. Specifically, I curated training and testing data, designed learning algorithms, developed training and evaluation pipelines, and analyzed model outputs. During my Ph.D., I divided my time between Yale and top industrial research labs, including Google DeepMind, Meta AI, and MSR.

EDUCATION

Yale University - Computer Science Department, New Haven, CT — Sep 2020 - Jul 2024 (Expected)
Ph.D. student in Computer Science | GPA: 4.0/4.0

Carnegie Mellon University - School of Computer Science, Pittsburgh, PA — Aug 2018 - Dec 2019
M.S. in Computer Science | GPA: 4.11/4.33

Nanjing University -Software Institute, Nanjing, China — Aug 2014 - Jun 2018
B.Eng. in Software Engineering | GPA: 4.41/5.0

PROFESSIONAL EXPERIENCE

Google DeepMind - Learning for Code Team, Mountain View, CA — Summer 2023
Research Intern | Hosts: Pengcheng Yin, Charles Sutton

- Fine-tuning LLMs (PaLM 2-L) to reason about code execution using execution traces. (Arxiv TBA)

Meta AI (FAIR) - FAIR Accel/Labs, Menlo Park, CA — Summer 2022
Research Intern | Mentors: Victoria Lin, Sida Wang

- Training T5 models to verify LLM-generated code with execution results. (Published in ICML'23)

Microsoft Research - Deep Learning Group, Redmond, WA — Jun 2021 - Apr 2022
Research Intern | Managers: Alex Polozov, Chris Meek, Chenglong Wang, Jeevana Priya Inala

- Self-improving LMs with partially-correct programs identified w/ exec. traces. (Published in ICLR'23)

Allen Institute for AI - AllenNLP Team, Seattle, WA — Summer 2020
Research Intern | Managers: Pradeep Dasigi, Matt Gardner

- A new formulation to jointly train retrievers and QA models for multi-doc QA. (Published in EMNLP'21)

Microsoft Research Asia - Software Analytics Group, Beijing, China — Jun 2017 – Dec 2017
Research Intern | Manager: Shi Han

- Automatic mining of insights from multi-dimensional data in Excel. (Product released [link])

ACADEMIC RESEARCH

Yale University - LILY Lab, New Haven, CT — Sep 2020 - Current
Research Assistant | Advisors: Dragomir Radev, Arman Cohan

- Semantic parsing and code generation.
- Long document and dialogue summarization.

Carnegie Mellon University - Institute for Software Research, Pittsburgh, PA — Spring 2020
Research Assistant | Advisor: Claire Le Goues, Ruben Martins

- Combining NLP and program synthesis for automatic API migration (e.g., Tensorflow ↔ PyTorch).

Carnegie Mellon University - Language Technology Institute, Pittsburgh, PA — Spring 2019
Research Associate | Advisor: Graham Neubig

- Active learning for weakly-supervised semantic parsing.

Nanjing University - LAMDA Group, Nanjing, China — 2017-2018
Research Assistant | Advisor: Ming Li

- Learning to predict failures in continuous integration.

TECHNICAL SKILLS

- Programming Languages (ranked by proficiency): Python, Shell, CUDA, C/C++, Java, SQL
- Frameworks and Tools: PyTorch, Vim, Linux, Git, wandb, Pytorch-Lightning, DeepSpeed

- Featured Repositories:
  - NLP4Code (LM training and evaluation for code generation) [Link]
  - SummerTime (Text summarization toolkit for non-experts, 250+ stars) [Link]
  - LEVER (Experiment code for the LEVER paper) [Link]

SELECTED PUBLICATIONS

[See Google Scholar for a full list] *(*: denotes equal contribution)*

NExT: Teaching Large Language Models to Reason about Code Execution
A. Ni, M. Allamanis, A. Cohan, Y. Deng, K. Shi, C. Sutton, P. Yin *Preprint'24*

Quantifying Contamination in Evaluating Code Generation Capabilities of Language Models
M. Riddell, A. Ni, A. Cohan *Preprint'24*

L2CEval: Evaluating Language-to-Code Generation Capabilities of Large Language Models
A. Ni, P. Yin, Y. Zhao, M. Riddell, T. Feng, R. Shen, S. Yin, Y. Liu, S. Yavuz, C. Xiong, S. Joty, Y. Zhou, D. Radev, A. Cohan *(To Appear) TACL'24*

LEVER: Learning to Verify Language-to-Code Generation with Execution
A. Ni, S. Iyer, D. Radev, V. Stoyanov, W-T. Yih, S. I. Wang, V. X. Lin** *ICML'23*

Learning Math Reasoning from Self-Sampled Correct and Partially-Correct Solutions
A. Ni, J. P. Inala, C. Wang, O. Polozov, C. Meek, D. Radev, J. Gao ICLR'23

Explicit Knowledge Transfer for Weakly-Supervised Code Generation
Z. Azerbayev, A. Ni, H. Schoelkopf, D. Radev DL4C @ ICLR'23

Mitigating False-Negative Contexts in Multi-document QA with Retrieval Marginalization
A. Ni, M. Gardner, and P. Dasigi EMNLP'21

SOAR: A Synthesis Approach for Data Science API Refactoring
A. Ni, D. Ramos*, A. Yang, I. Lynce, V. Manquinho, R. Martins, and C. Le Goues* ICSE'21

Merging Weak and Active Supervision for Semantic Parsing
A. Ni, P. Yin, and G. Neubig AAAI'20

INVITED TALKS

Enhancing Language Models for Code Generation using Execution
- Invited Talk @ MIT CSAIL | Microsoft PROSE | UT Austin | HKUST 2023 - 2024

Foundation Models for Code and Math
- Guest Lecture @ Yale CPSC 488/588 "AI Foundation Models" Dec 2023

ACADEMIC ACTIVITIES

- Program Committee / Reviewer
  - *ML Conferences*: NeurIPS (2022-), ICML (2023-), ICLR (2024-)
  - *NLP Conferences*: COLM (2024-), ACL ARR (2021-), EMNLP (2022-)
  - *Workshops*: DL4C, SUKI, IntEx-SemPar
- Teaching Assistant:
  - Yale CPSC 477/577 Natural Language Processing
  - Yale CPSC 482/582 Topics in Applied Machine Learning
- Mentored undergraduates (projects) at Yale:
  - Martin Riddell (code generation) – Currently at Yale
  - Ziming Mao (long-input summarization) → CS PhD @ UC Berkeley
  - Zhangir Azerbayev (code generation) → CS PhD @ Princeton
  - Hailey Schoelkopf (code generation) → Research Scientist @ Eleuther AI

HONORS & AWARDS

- AWS Cloud Credits for Research Program ($10,000) Yale University, Apr 2023
- University Nominee for Google Fellowship Yale University, Oct 2022
- University Nominee for Microsoft Fellowship Yale University, July 2022
- Outstanding Graduate Nanjing University, Jun 2018