

# Junhong Shen

5000 Forbes Ave, Pittsburgh, PA, 15213 | ☎ (+1)310-689-6721 | ✉ [junhongs@andrew.cmu.edu](mailto:junhongs@andrew.cmu.edu)

🏠 [Website](#) | 💻 [GitHub](#) | 🎓 [Scholar](#)

## Research Interests

Developing efficient, data-centric ML techniques for real-world ML applications

## Education

### Carnegie Mellon University

PH.D. IN MACHINE LEARNING, ADVISOR: [AMEET TALWALKAR](#) (QPA: 4.11/4.33)

Pittsburgh, PA

Sep. 2021 - Present

### University of California, Los Angeles

B.S. IN MATHEMATICS OF COMPUTATION (GPA: 4.0/4.0)

Los Angeles, CA

Sep. 2017 - June 2021

## Publications

### PREPRINT

#### CAT: Content-Adaptive Image Tokenization [\[paper\]](#)

Junhong Shen, Kushal Tirumala, Michihiro Yasunaga, Ishan Misra, Luke Zettlemoyer, Lili Yu\*, Chunting Zhou\*

#### ScribeAgent: Towards Specialized Web Agents Using Production-Scale Workflow Data [\[paper\]](#)

Junhong Shen, Atishay Jain, Zedian Xiao, Ishan Amlekar, Mouad Hadji, Aaron Podolny, Ameet Talwalkar

#### Specialized Foundation Models Struggle to Beat Supervised Baselines [\[paper\]](#)[\[code\]](#)

Zongzhe Xu, Ritvik Gupta, Wenduo Cheng, Alexander Shen, Junhong Shen, Ameet Talwalkar, Mikhail Khodak

#### L2G: Repurposing Language Models for Genomics Tasks [\[paper\]](#)

Wenduo Cheng, Junhong Shen, Mikhail Khodak, Jian Ma, Ameet Talwalkar

### ACCEPTED/PUBLISHED ARTICLES

#### UPS: Towards Foundation Models for PDE Solving via Cross-Modal Adaptation

TMLR 2024 & ICML 2024 AI4Science Workshop (Spotlight) [\[paper\]](#)[\[code\]](#)

Junhong Shen, Tanya Marwah, Ameet Talwalkar

#### Tag-LLM: Repurposing General-Purpose LLMs for Specialized Domains

ICML 2024 [\[paper\]](#)[\[code\]](#)

Junhong Shen, Neil Tenenholtz, James Brian Hall, David Alvarez-Melis, Nicolò Fusi

#### Cross-Modal Fine-Tuning: Align then Refine

ICML 2023 (Oral, top 158/6538) [\[paper\]](#)[\[code\]](#) [\[talk\]](#)[\[website\]](#)

Junhong Shen, Liam Li, Lucio M. Dery, Corey Staten, Mikhail Khodak, Graham Neubig, Ameet Talwalkar

#### Efficient Architecture Search for Diverse Tasks

NeurIPS 2022 [\[paper\]](#)[\[code\]](#)[\[blog\]](#)

Junhong Shen\*, Mikhail Khodak\*, Ameet Talwalkar

#### NAS-Bench-360: Benchmarking Neural Architecture Search on Diverse Tasks

NeurIPS 2022 Datasets and Benchmarks Track [\[paper\]](#)[\[website\]](#)[\[blog\]](#)

Renbo Tu\*, Nicholas Roberts\*, Mikhail Khodak, Junhong Shen, Frederic Sala, Ameet Talwalkar

#### AutoML Decathlon: Diverse Tasks, Modern Methods, and Efficiency at Scale

NeurIPS 2022 Competitions Track [\[paper\]](#)[\[website\]](#)

Nicholas Roberts, ... 24 authors ..., Junhong Shen, Evan Sparks

#### Iterative Teacher-Aware Learning

NeurIPS 2021 [\[paper\]](#)[\[code\]](#)

Luyao Yuan, Dongruo Zhou, Junhong Shen, Jingdong Gao, Jeffrey Chen, Quanquan Gu, Ying Nian Wu, Song-Chun Zhu

#### Theoretically Principled Deep RL Acceleration via Nearest Neighbor Function Approximation

AAAI 2021 [\[paper\]](#)[\[code\]](#)

Junhong Shen, Lin F. Yang

## Mathematical Reconstruction of Patient-Specific Vascular Networks Based on Clinical Images and Global Optimization

IEEE Access [[paper](#)] [[code](#)] [[talk](#)]

Junhong Shen, Abdul Hannan Faruqi, Yifan Jiang, Nima Maftoon

## Emergence of Pragmatics from Referential Game between Theory of Mind Agents

NeurIPS 2019 Emergent Communication Workshop [[paper](#)] [[code](#)]

Luyao Yuan, Zipeng Fu, Jingyue Shen, Lu Xu, Junhong Shen, Song-Chun Zhu

\* Equal Contribution

## Work Experience

### Fundamental AI Research (FAIR), Meta

Seattle, WA

RESEARCH INTERN (MENTOR: [CHUNTING ZHOU](#), [LILI YU](#), [LUKE ZETTLEMOYER](#))

May 2024 - Present

- Working on efficient image/video tokenization.

### AI Research Team, Scribe

Pittsburgh, PA

SENIOR MACHINE LEARNING RESEARCH ENGINEER

Feb. 2024 - May 2024

- Worked on LLM-based web agents.

### Microsoft Research, New England

Cambridge, MA

RESEARCH INTERN (MENTOR: [DAVID ALVAREZ-MELIS](#), [NICOLÒ FUSI](#))

May 2023 - Aug. 2023

- Worked on adapting LLMs to specialized domains (e.g., low-resource languages, protein sequences, chemical formulas) via prompt tuning.

### Determined AI, Hewlett Packard Enterprise

Pittsburgh, PA

RESEARCH INTERN (MENTOR: [LIAM LI](#))

Jun. 2022 - Dec. 2022

- Worked on fine-tuning LLMs and vision transformers for scientific modalities via distribution alignment.

### Face ID Team, SenseTime

Beijing, China

PRODUCT MANAGER INTERN

Jun. 2018 - Sep. 2018

- Worked on 3D-structured-light Face ID; participated in 5 software version releases, bug fixing, and testing.

## Honors & Awards

2024 J.P. Morgan AI Ph.D. Fellowship, awardee (accepted)

2024 Bloomberg Data Science Ph.D. Fellowship, awardee (declined)

2021 UCLA Daus Prize, top-5 undergraduate students in mathematics

2017 - 21 UCLA Dean's Honors List, awardee

## Talks

Aug. 2024 [Repurposing LLMs for Long-Tail ML Applications](#), Research Summit, Ai4 2024

Nov. 2023 [Machine Learning for Diverse Tasks](#), Guest Lecture, ML with Large Datasets, CMU 10605

Nov. 2023 [Bridging LLMs and Long Tail ML Applications](#), Catalyst Reading Group, CMU

Mar. 2023 [Cross-Modal Fine-Tuning](#), AI4Science Talks

Dec. 2022 [DASH: How to Search Over Convolutions](#), The AutoML Podcast

Oct. 2022 [Tackling Diverse Tasks with Neural Architecture Search](#), DLML Journal Club, Mayo Clinic

## Professional Service

Co-organizer of [2024 CMU Agent Workshop](#)

Co-organizer of [AutoML Decathlon](#), NeurIPS 2022 Competition Track

CMU MSML Admissions Committee Fall 2022

CMU MLD Open House Committee Spring 2024/2025

Conference Reviewer NeurIPS (2022/2023/2024), ICLR (2024/2025), AAAI (2025), ICML (2024)

## Research Experience

---

### SAGE Lab

ADVISOR: [AMEET TALWALKAR](#) (CMU)

Pittsburgh, PA

June, 2021 - Present

- Ph.D. research on developing effective and efficient ML/AutoML tools for solving diverse tasks in practice.

### Lin Yang's Group

ADVISOR: [LIN F. YANG](#) (UCLA)

Los Angeles, CA

Jan. 2020 - June 2021

- Studied sample-efficient reinforcement learning; proposed an algorithm for estimating the value functions using nearest neighbor function approximator; provided theoretical justification on the sample complexity.

### Center for Vision, Cognition, Learning, and Autonomy (VCLA), UCLA

ADVISOR: [SONG-CHUN ZHU](#), [YING NIAN WU](#) (UCLA)

Los Angeles, CA

Jan. 2019 - June 2021

- Studied how theory of mind (ToM) can be integrated into various ML settings to improve algorithm efficiency.
- **Project 1: Multi-Agent Deep Reinforcement Learning with ToM**
  - Proposed an adaptive ToM algorithm in a referential game setting where the teacher and the student model each other's action likelihood while learning their own Q-functions; studied the emergent communication protocol between the agents.
- **Project 2: Efficient Learners in Iterative Machine Teaching**
  - Integrated ToM into machine teaching; improved teaching efficiency by having the learners model the teacher's training sample selection strategy with maximum likelihood estimation.
- **Project 3: Meta Machine Teaching**
  - Studied how meta-learning can be combined with machine teaching: the teacher monitors a group of students' learning processes, deduces their initial beliefs and learning models, and adapts its teaching scheme to each learner.

### Computational Metastasis Lab, Fields Institute for Research in Mathematical Sciences

ADVISOR: [NIMA MAFTOON](#) (UNIVERSITY OF WATERLOO)

Toronto, Canada

Jul. 2019 - Sep. 2019

- Developed a vascular network reconstruction framework that uses (i) the main vessel skeletons segmented from clinical images and (ii) global constructive optimization algorithm to generate patient-specific cerebral vascular models; validated the geometric (lengths, radii) and hemodynamic (pressure, shear stress) properties of the models via histogram analysis and blood flow simulation.


## Related Coursework

---

<b>CS</b>	Machine Learning, Artificial Intelligence, Information Theory, Computer Vision, Algorithms and Complexities, Advanced Programming, Operating Systems Principles, Deep Learning Systems, Computer Network Fundamentals
<b>Math</b>	Linear Algebra, Multivariable Calculus, Differential Equations, Discrete Mathematics, Probability Theory, Stochastic Processes, Mathematical Analysis, Complex Analysis, Optimization, Applied Numerical Methods, Game Theory
<b>Stats</b>	Statistical Modeling in Vision and Cognition, Computing and Inference in Vision and Cognition, Data Analysis and Regression

## Professional Skills

---

<b>Coding</b>	Proficient: Python, C, C++, Bash, R    Familiar: MATLAB, Java, Arduino
<b>Tools</b>	Git,  , PyTorch, Tensorflow, Scikit-learn, OpenCV, OpenAI Gym, Google Cloud Platform, Docker, SolidWorks