

Changho Shin

cshin23@wisc.edu
<https://ch-shin.github.io>

1210 W Dayton St, Madison, WI 53706

INTERESTS *Large Language Models, Multimodal Models, Weak Supervision, Optimal Transport*

EDUCATION **University of Wisconsin-Madison** Sep. 2020 –

- Ph.D. Student in School of Computer, Data & Information Sciences
- Advisor: Professor Frederic Sala
- Pursuing additional MS degree in Math

Seoul National University Mar. 2015 – Feb. 2017

- M.S. in Department of Transdisciplinary Studies
- Thesis: Application of Traditional ML and DNN Techniques on Energy Disaggregation with 10 Hz AMI Data
- Advisor: Professor Wonjong Rhee, IEEE Fellow

Seoul National University Mar. 2011 – Feb. 2015

- B.A. in Psychology
- B.S. in Computer Science and Engineering
- Graduated with honors (Cum Laude)

PUBLICATIONS

- [1] Joon Suk Huh*, **Changho Shin***, Elina Choi, “Pool-Search-Demonstrate: Improving Data-wrangling LLMs via better in-context examples”, *NeurIPS 2023 Workshop: Table Representation Learning (TRL)* 2023.
- [2] Dyah Adila*, **Changho Shin***, Linrong Cai, Frederic Sala, “Zero-Shot Robustification of Zero-Shot Models With Auxiliary Foundation Models”, *NeurIPS 2023 Workshop: Robustness of Few-shot and Zero-shot Learning in Large Foundation Models (R0-FoMo)* 2023 (**Best paper award honorable mention**).
- [3] **Changho Shin***, Tzu-heng Huang*, Sui Jiet Tay, Dyah Adila, Frederic Sala, “Multimodal Data Curation via Object Detection and Filter Ensembles”, ICCV 2023 Datacomp Workshop (Rank #1 in DataComp competition filtering track (small)).
- [4] **Changho Shin**, Sonia Crompt, Dyah Adila, Frederic Sala, “Mitigating Source Bias for Fairer Weak Supervision”, *NeurIPS* 2023.
- [5] **Changho Shin**, Alice Schoenauer-Sebag, “Can we get smarter than majority vote? Efficient use of individual rater’s labels for content moderation”, *NeurIPS 2022 Workshop: Efficient Natural Language and Speech Processing (ENLSP)* 2022.
- [6] **Changho Shin**, Winfred Li, Harit Vishwakarma, Nicholas Roberts, Frederic Sala “Universalizing Weak Supervision”, *International Conference on Learning Representations (ICLR)* 2022.
- [7] **Changho Shin**, Eunjung Lee, Jeongyun Han, Jaeryun Yim, Hyoseop Lee, Wonjong Rhee, “The ENERTALK Dataset, 15 Hz Electricity Consumption Data from 22 Houses in Korea”, *Nature Scientific Data*, 2019 (Impact Factor = 5.929).
- [8] **Changho Shin**, Seungeun Rho, Hyoseop Lee, Wonjong Rhee, “Data Requirements for Applying Machine Learning to Energy Disaggregation”, *Energies*, May 2019 (Impact Factor = 2.707).
- [9] **Changho Shin**, Sunghwan Joo, Jaeryun Yim, Hyoseop Lee, Taesup Moon, Wonjong Rhee, “Subtask Gated Networks for Non-Intrusive Load Monitoring”, *AAAI Conference on Artificial Intelligence* 2019.

JOB EXPERIENCE	Twitter , San Francisco, USA <i>ML Engineer Intern</i> • Mentor: Alice Schoenauer Sebag • Manager: Milind Ganjoo Jun. 2022 – Aug. 2022
	Encored Technologies , Seoul, Korea <i>Data Scientist</i> • Advisor: Dr. Hyoseop Lee Jan. 2018 – Jul. 2020
	Korea Institute for Defense Analyses , Seoul, Korea <i>Researcher</i> Jan. 2017 – Dec. 2017
TEACHING EXPERIENCE	University of Wisconsin-Madison • Teaching assistant for CS 839 (Foundation Models) Fall 2023 • Teaching assistant for CS 300 (Programming II) Fall 2022, Spring 2023 • Teaching assistant for CS 760 (Machine Learning) Fall 2021, Spring 2022 • Teaching assistant for CS 320 (Data Programming II) Spring 2021 • Teaching assistant for CS 220 (Data Programming I) Fall 2020
HONORS	Best Paper Award Honorable Mention 2023 NeurIPS R0-FoMo Workshop
	CS Departmental Scholarship 2020 University of Wisconsin-Madison
	1st Creative National Defense Conference - 2nd Place 2016 Ministry of National Defense • Topic: Cooperative unmanned aircraft system with reinforcement learning
	Merit-based Scholarship 2015 Seoul National University
Graduate Coursework	<ul style="list-style-type: none"> • M2680.001300 Machine Learning for Information Studies @ SNU • M2680.001400 Social Computing @ SNU • 493.613 Mathematics for Intelligent Systems (Numerical Linear Algebra) @ SNU • 493.701 Learning and Applications of Deep Neural Networks @ SNU • M0000.005400 Convex Optimization @ SNU • M0000.005400 Neural Networks @ SNU • CS537 Introduction to Operating Systems @ UW • CS639.004 Introduction to Computational Learning Theory @ UW • CS726 Nonlinear Optimization 1 • CS744 Big Data Systems @ UW • CS761 Mathematical Foundations of Machine Learning @ UW • CS784 Foundations of Data Management @ UW • CS787 Advanced Algorithms @ UW • CS839 Probability and Learning in High Dimension @ UW • CS880 Advanced Topics in Learning Theory @ UW • Math521 Analysis I @ UW • Math522 Analysis II @ UW • Math551 Elementary Topology @ UW • Math629 Introduction to Measure and Integration @ UW • Math621 Analysis III (Analysis on Manifolds) @ UW • Math721 A First Course in Real Analysis @ UW • Math733 Theory of Probability I @ UW • Math761 Differentiable Manifolds @ UW • Math833 Modern Discrete Probability @ UW • Math888 Randomized Linear Algebra @ UW • Stat992 Optimal Transport and Applications to Machine Learning @ UW

**TECHNICAL
SKILLS**

Machine Learning / Deep Learning / Data Science

PyTorch, TensorFlow, Keras, scikit-learn, NumPy, Pandas, SciPy

DBMS

MySQL, MongoDB, PySpark

Research & Development Tools

Jupyter, PyCharm, Docker, GitHub, CircleCI, Shell, Amazon Web Services

Programming Languages

Python, R, MATLAB, Java, Go, C, L^AT_EX