SUMYEONG AHN

428 S Shaw Ln # 3312, East Lansing, MI, United States, 48824
(+1)517-940-5197, *Firstname* (at) msu (dot) edu

WORK EXPERIENCE

Postdoc Researcher Sep. 2023 - Current
Michigan State University, United States (MSU)
Dept. of Computer Science and Engineering (CSE)
Host: Prof. Jiayu Zhou

EDUCATION

Ph. D. Sep. 2017 - Aug. 2023
Korea Advanced Institute of Science and Technology, S.Korea
Kim Jaechul Graduate School of Artificial Intelligence (KAIST AI)
Thesis: Mitigating Dataset Bias for Robust Deep Learning:
From Clean dataset to the Practical Noisy Dataset
Advisor: Prof. Seyouong Yun (KAIST AI), Prof. Yung Yi (KAIST EE)

Master of Science Mar. 2015 - Aug. 2017
Korea Advanced Institute of Science and Technology, S.Korea
Department of Electrical Engineering (KAIST EE)
Thesis: Multi-armed Bandit Problem with Intra- and Inter- Correlations
Advisor: Prof. Yung Yi (KAIST EE)

Bachelor of Science Mar. 2011 - Feb. 2015
Korea University, S.Korea
School of Electrical Engineering
Advisor: Taewoong Yun (KU EE)

RESEARCH INTERESTS

Current Interests

- Vision Language Models [CVPR'24, Under review (P4)]
- Machine Learning in Medical Domain [AMIA'24, Under review (P1)]
- Efficient Large Language Models [EMNLP'23]
- Noisy label [AAAI'23, ICLRW'23, Under review (P2)]
- Dataset Bias (a.k.a Spurious correlation) [ICLR'23, AAAI'23, Under review (P3)]

Previous Interests

- Class imbalance [ICLR'23 (Spotlight)]
- Unsupervised Domain Adaptation with GAN [ICPR'21]
- Stochastic Multi-armed Bandits with Additional Conditions [SIGMETRICS'18]
- AI for Computer Networking Systems [SIGMETRICS'18, MobHoc'21]

PUBLICATIONS (⋆ : EQUAL CONTRIBUTION & _ : EQUAL ADVISING)

Preprint or Preparing

[P4] **Sumyeong Ahn**, Siqi Liang, and Jiayu Zhou. Prompt Learning with Noisy Labels (Modified for anonimity). Under Review.

[P3] **Sumyeong Ahn** and Se-Young Yun. ORBIS: Open Dataset Can Rescue You from Dataset Bias Problem. Under Review.

[P2] **Sumyeong Ahn**, Sihyeon Kim, Jongwoo Ko, and Se-Young Yun. Fine-tuning Pre-trained Models for Robustness Under Noisy Labels. Under Review.

[P1] Xiaodan Zhang, Sandeep Vemulapalli, Nabasmita Talukdar, **Sumyeong Ahn**, Jiankun Wang, Han Meng, Sardar Mehtab Bin Murtaza, Aakash Ajay Dave, Dmitry Leshchiner, Dimitri F. Joseph, Martin Witteveen-Lane, Dave Chesla, Jiayu Zhou, and Bin Chen. Large Language Models in Medical Term Classification and Unexpected Misalignment Between Response and Reasoning, Under Review.

Conference

[C9] Jihwan Bang, **Sumyeong Ahn**, and Jae-Gil Lee. Active Prompt Learning in Vision Language Models, **CVPR 2024**. (To appear)

[C8] Xiodan Zhang, Nabasmita Talukdar, Sandeep Vemulapalli, **Sumyeong Ahn**, Jiankun Wang, Han Meng, Sardar Mehtab Bin Murtaza, Dmitry Leshchiner, Aakash Ajay Dave, Dimitri F. Joseph, Martin Witteveen-Lane, Dave Chesla, Jiayu Zhou, and Bin Chen. Comparison of Prompt Engineering and Fine-tuning Strategies in Large Language Models in Classification on Clinical Notes, **AMIA 2024 Informatics Summit**. (To appear)

[C7] Jongwoo Ko*, Seungjoon Park*, Yujin Kim, **Sumyeong Ahn**, Du-Seong Chang, Euijai Ahn, and Se-Young Yun. Nash: A Simple Unified Framework of Strutured Pruning for Accelerating Encoder-Decoder Language Models. **EMNLP 2023**(Findings).

[C6] **Sumyeong Ahn***, Seongyoon Kim*, and Se-Young Yun. Mitigating Dataset Bias by Using Per-sample Gradient. **ICLR 2023**.

[C5] **Sumyeong Ahn***, Jongwoo Ko*, and Se-Young Yun. CUDA: Curriculum of Data Augmentation for Long-tailed Recognition **ICLR 2023** (**Spotlight**).

[C4] **Sumyeong Ahn** and Se-Young Yun. Denoising After Entropy-based Debiasing a Robust Training Method for Dataset Bias with Noisy Labels. **AAAI 2023** (**Oral presentation**).

[C3] Sangwoo Moon, **Sumyeong Ahn**, Kyunghwan Son, Jinwoo Par, and Yung Yi. Neuro-DCF: Design of Wireless MAC via Multi-Agent Reinforcement Learning Approach. **MobiHoc 2021**.

[C2] Tran Hai H, **Sumyeong Ahn**, Taeyoung Lee, and Yung Yi. Enlarging Discriminative Power by Adding Extra Class in Unsupervised Domain Adaptation. **ICPR 2021**.

[C1] Donggyu Yun, **Sumyeong Ahn**, Alexandre Proutiere, Jinwoo Shin, and Yung Yi. Multi-Armed Bandit with Additional Observations. **SIGMETRICS 2018**.

Workshop

[W3] Jongwoo Ko*, Sumyeong Ahn*, and Se-Young Yun. Efficient Utilization of Pre-trained Model for Learning with Noisy Labels. **ICLR 2023 Trustworthy Workshop** (**Oral Presentation**).

[W2] **Sumyeong Ahn**$\star$, Jongwoo Ko*, and Se-Young Yun. CUDA: Curriculum of Data Augmentation for Long-tailed Recognition. **NeurIPS ML Safety Workshop and Distribution Shifts Workshop**.

[W1] **Sumyeong Ahn***, Seongyoon Kim*, and Se-Young Yun. Mitigating Dataset Bias by Using Per-sample Gradient. **NeurIPS ML Safety Workshop and Distribution Shifts Workshop**.

Journal

[J1] Donggyu Yun, **Sumyeong Ahn**, Alexandre Proutiere, Jinwoo Shin, and Yung Yi. Multi-Armed Bandit with Additional Observations. Proceedings of ACM on Measurement and analysis of Computing Systems **(POMACS)** , 2(1):1-22, 2018.

Patent

[Pat1] Nansol Seo, Jaemoon Lee, Jaeha Ahn, **Sumyeong Ahn**, Suho Shin, Yoonpyo Koo, and Yung Yi. Method for Controlling Multiple UAVs Based on Multi-Hop Wireless Mesh Networks, September 3rd 2019. KOR patent number: 10-2019874-0000.

MISC (Korean conference and journal)

[M1] Sangmin Bae, Taehyeon Kim, **Sumyeong Ahn**, Sangmook Kim, Jongwoo Ko, and Se-Young Yun. Client Sampling Algorithm in Federated Learning via Combinatorial Averaging and Multi-Armed Bandits. KIISE

[M2] Yoonpyo Koo, **Sumyeong Ahn**, Kyounghwan Son, Suho Shin, Jeonghun Yu, Jaesin Kim, and Yung Yi. An Implementation of Multi-Hop Voice Communication System Using Drones. KICS

PROJECT

- Student Leader

◦ Developing NAS (Neural Architecture Search) algorithm for detecting layer-wise communication network failures in multi-network-layer datasets. **KT (08/2021-08/2022)**
- Neural Architecture Search (NAS), Communication Network failure detection

◦ Implementing an Application Service for Open/Close Prediction of Commercial Stores using Power Consumption Data **KEPRI (04/2018-12/2019)**
- Predicting Open/Close Status DNN on Time-series Power Consumption Data

◦ Versatile Network System Architecture for Multi-dimensional Diversity **IITP (04/2016-08/2019)**
- Edge Networking, Fog Computing, FogOS, Matching between resources and requests

◦ Designing a Matching Algorithm for Efficient Resource Allocation in Distributed Edge Devices **KISA (06/2015-06/2016)**
- Matching algorithm, FogOS, Reseource-query matcher

◦ Modeling Multi-Hop Wireless Networks for Military Applications **ADD (09/2015-12/2016)**
- Flying Ad-Hoc Network (FANET), Ground Control Station (GCS), Voice Communication, UAVs

- Participant Student

◦ Video Transmission over Vehicular Networks **Bosch (01/2015-06/2015)**
- Vehicular Ad-Hoc Network (VANET), 802.11, WiFi, Video codec, real-time video communication

AWARDS

- KAIST EE Best Teaching Assistant Award (2019 Fall - Embedded System)

SKILLS

- Programming Languages & Frameworks (Selected)

- Machine learning tools: PyTorch, Tensorflow, Scikit-learn
- Database: MySQL, Oracle Database
- System Programming: Bash, C/C++
- Others: OpenCV, Numpy, Matlab

TEACHING

- Teaching Assistance (@ KAIST)

- EE209 Programming Structures for Electrical Engineering (2015 Spring, 2015 Fall, 2020 Spring)

- EE323 Computer Networks (2016 Spring)
- EE205 Data Structures and Algorithms for Electrical Engineering (2017 Fall)
- HSS190(J) Freshman Seminar 1 (2018 Spring)
- EE414 Embedded Systems (2016 Fall, 2019 Fall)
- AI505 Optimization for AI (2020 Fall)
- AI603 Machine Learning Theory (2021 Spring)

- Mentoring (F): Finish

- (Efficient LLM), Sangmin Bae and Joonki Kim, KAIST (2023-)
- (Active Learning) Jihwan Bang, KAIST (2023-)
- (Text-to-Speech) Sangmin Bae and Sione Ok, KAIST (2023-)
- (VLM Grounding) Sihyeon Kim and Boryeong Cho, KAIST (2023-)
- (Federated Learning) Seongyoon Kim, KAIST (2023-)
- (Federated Learning) Siqi Liang, MSU (2023-)
- (Medical LLM) Jiankun Wang, MSU (2023-)
- (Efficient LM), Jongwoo Ko and Seungjoon Kim, KAIST (-2023), **EMNLP 2023 (F)**

COURSES

- Machine Learning

- (EE837) Deep Learning for Computer Vision (Prof. Junmo Kim @ KAIST EE)
- (AI604) Deep Learning for Computer Vision (Prof. Jaegul Choo @ KAIST AI)
- (AI612) Machine Learning for Healthcare

- Networking

- (EE655) Economics in Communication Networks
- (EE827) Epidemics and Information Diffusion in Complex Networks
- (EE827) 3G/4G/5G Cellular Network Protocols
- (EE650) Optimization in Communication Networks
- (EE627) Performance Analysis of Communication Networks
- (EE827) Network Science
- (EE528) Engineering Random Processes

- Math

- (MAS241) Analysis 1
- (IE539) Convex Optimization

- System

- (EE516) Embedded Software
- (EE324) Network Programming