DONGKYU CHO

Machine Learning Researcher at LG AI Research — Ph.D. student at NYU

`https://umamicode.github.io` | sqvareinch@gmail.com | dongkyu.cho@lgresearch.ai

EDUCATION

New York University *New York City, New York*
Ph.D. in Computer Science *September 2024 –*

Seoul National University *Seoul, South Korea*
Master of Science in Data Science, advised by Prof. Sanghack Lee *March 2021 – August 2023*
- Research Assistant at Causality Lab
- Thesis: Learning to ignore- Single Source Domain Generalization via Oracle Regularization

Seoul National University *Seoul, South Korea*
Bachelor of Arts in Western History | Information Science and Culture (Interdisciplinary Major) *March 2014 – February 2021*
- Undergraduate Student Researcher on Computational Linguistics for Historical Analysis
- Graduated with Distinction (Cum Laude)

RESEARCH EXPERIENCE

LG AI Research *Seoul, South Korea*
Research Scientist Intern *July. 2023 –*
- *Research Area*: Foundation Models (Time-Series/Language), Loss Landscapes, Model Alignment
- *Projects*: Foundational Framework for Long Term Time-Series Forecasting, Causal Discovery in Time-Series Data

Seoul National University *Seoul, South Korea*
Graduate Research Assistant, advised by Prof. Sanghack Lee *2021 – 2023*
- *Research Area*: Causality, Causal Representation Learning, Out-of-Distribution Generalization
- *Projects*: Continual-learning framework for Financial Language Models (with NH Investments), Causal Language Models (with LG AI Research), Semi-Supervised Federated Learning Frameworks, Vision-Language Models

DaumSoft *Seoul, South Korea*
Data Analyst Intern *Jan. 2019 – Feb. 2019*
- *Research Area*: NLP-based Market Sentiment Analysis

SELECTED PUBLICATIONS   *: co-first author

2 A+ conference workshops — 1 preprints

Cho, D., Lee, S. (2023). Learning to ignore: Single Source Domain Generalization via Oracle Regularization, Accepted at *NeurIPS 2023, Causal Representation Learning Workshop*

Cho, D., Yang, J., Seo, J., Bae, S., Kang, D., Park, S., Choe, H., Ahn W., Lim, W. (2024). ShERPA: Leveraging Neuron Alignment for Knowledge-preserving Fine-tuning, Accepted at *ICLR 2024, Workshop on Mathematical and Empirical Understanding of Foundation Models (ME-FoMO)*

Cho, D., Hwang, I., Lee, S. (2024). Is Data Augmentation Reliable for Generalization?, *Working Paper*

TEACHING EXPERIENCE

Deep Learning and Machine Learning II, Teaching Assistant *Fall 2022*
- Assisted graduate course on statistical learning methods (e.g., Bayesian Frameworks).

Data Science Bootcamp, Teaching Assistant *Winter 2021*
- Assisted Undergraduate students with the introduction to machine learning.
- Lab sessions on machine learning libraries (e.g., Pytorch), Assignment design

PROJECTS

Stock Interrelation Research using Keyword and Supply Chain data *Seoul National University and NH Investment & Securities*
Collaborative Research *July 2021 – August 2022*
- *Research Area*: Natural Language Processing, Continual Learning, Financial Language Models
- A continual-learning based pipeline for a Financial Language Model
- Interactive Graph Database using the Bloomberg Supply Chain Data & Financial Corpus
- Launched as an app feature in the trading app of NH Investment & Securities

Towards Language Models Capable of Causal Reasoning *Seoul National University and LG AI research*
Collaborative Research *August 2022 – July 2023*
- *Research Area*: Natural Language Processing, Causality
- A Causal NLP Project, aborted during process to join LG AI Research.

Class Projects *Seoul National University*
- Class Project for *Special Topics in Data Science- Meta Learning*: "Semi-Supervised Federated Learning with Representations"
- Class Project for *Machine Learning for Visual Understanding*: "A combination approach for wikipedia Image/Caption Matching"

PROFESSIONAL EXPERIENCE

LG AI Research, Data Intelligence Lab *Seoul, South Korea*
Research Scientist Intern *July. 2023 –*
- Worked as a research intern at LG AI Research.

Seoul National University, Causality Lab *Seoul, South Korea*
***Graduate Research Assistant**, advised by Prof. Sanghack Lee* *2021 – 2023*
- Worked as a research assistant at Causality Lab, Seoul National University.

DaumSoft *Seoul, South Korea*
Data Analyst Intern *Jan. 2019 – Feb. 2019*
- Worked as a data analyst intern at DaumSoft (now known as VAIV company).

Republic of Korea Army *Icheon, South Korea*
Military Interpreter (English-Korean) *Mar. 2016 – Dec. 2017*
- Served as military interpreter at the Republic of Korea Army.

INVITED TALKS

- 2023 Student Researcher *Pizza* Seminar, Graduate School of Data Science, Seoul National University

SKILLS AND ACTIVITIES

- **Programming**: Python (PyTorch, TensorFlow, scikit-learn), R, SQL, MLOps Tools (e.g., Docker, WandB, GCP, AWS)
- **Languages**: English (Fluent), Korean (Native), French (Intermediate)
- **Test Scores**: TOEFL: 118 (2020.08), 116 (2022.08)