

# Jaeseung Heo

Graduate School of Artificial Intelligence  
POSTECH (Pohang University of Science and Technology)

✉ jsheo12304@postech.ac.kr  
🌐 jason9464.github.io

## EDUCATION

---

### POSTECH

*Integrated M.S./Ph.D. Program, Graduate School of Artificial Intelligence*  
*Advisor: Prof. Dongwoo Kim*

*Pohang, South Korea*  
Mar, 2021 - Present  
Expected Graduation: Feb, 2027

### Sungkyunkwan University

*B.S. in Computer Science and Engineering*  
*B.S. in Systems Management Engineering (Double Major)*

*Suwon, South Korea*  
Mar, 2016 - Aug, 2020

## RESEARCH INTERESTS

---

My research aims to identify which training data causes specific model behaviors, with a longer-term goal of extending this to problematic behaviors relevant to safety and alignment. I use training data attribution (TDA), particularly influence functions, to characterize how training examples drive model behavior, and translate these signals into data-centric interventions such as augmentation, label smoothing, and selection. Methodologically, I develop influence functions that capture dependence between training examples, both explicit (as in graph neural networks) and implicit (arising from joint loss minimization). Going forward, I aim to connect TDA with mechanistic interpretability, and ultimately to trace behaviors such as subliminal learning back to their origins in training data.

## PUBLICATIONS

---

### Conference Publications

- **Posterior Label Smoothing for Node Classification** *AAAI '26*  
Jaeseung Heo, MoonJeong Park, Dongwoo Kim Oral Pres.
- **Influence Functions for Edge Edits in Non-Convex Graph Neural Networks** *NeurIPS '25*  
Jaeseung Heo, Kyeongheung Yun, Seokwon Yoon, MoonJeong Park, Jungseul Ok, Dongwoo Kim
- **EPIC: Graph Augmentation with Edit Path Interpolation via Learnable Cost** *IJCAI '24*  
Jaeseung Heo\*, Seungbeom Lee\*, Sungsoo Ahn, Dongwoo Kim
- **Mitigating Oversmoothing through Reverse Process of GNNs for Heterophilic Graphs** *ICML '24*  
MoonJeong Park, Jaeseung Heo, Dongwoo Kim

### Preprints

- **Interaction-Aware Influence Functions for Group Attribution** *arXiv '26*  
Jaeseung Heo, Kyeongheung Yun, Youngbin Choi, Sehyun Hwang, Jungseul Ok, Dongwoo Kim
- **Transductive Generalization and Its Application to Graph Node Classification** *arXiv '26*  
MoonJeong Park\*, Seungbeom Lee\*, Kyungmin Kim, Jaeseung Heo, Seunghyuk Cho, Shouheng Li, Sangdon Park, Dongwoo Kim
- **The Oversmoothing Fallacy: A Misguided Narrative in GNN Research** *arXiv '25*  
MoonJeong Park, Sunghyun Choi, Jaeseung Heo, Eunhyeok Park, Dongwoo Kim

\*Equal contribution.

## HONORS AND AWARDS

---

- **BK21 Outstanding Research Paper Award, POSTECH** *2025, 2026*
- **Dean's List, Sungkyunkwan University** *2018, 2019, 2020*

## SERVICE AND TEACHING

---

- **Reviewer**, ICML '22, NeurIPS '26, ICML Workshop on Mechanistic Interpretability '26
- **Teaching Assistant**, CSED700H: Convex Optimization, POSTECH *Fall 2022*
- **Lecturer**, Youth AI & Big Data Academy, AI Expert Training Course, POSCO *2021, 2022*

## REFERENCES

---

### **Dongwoo Kim**

Associate Professor  
POSTECH  
dongwoo.kim@postech.ac.kr

### **Jungseul Ok**

Associate Professor  
POSTECH  
jungseul.ok@postech.ac.kr

### **MoonJeong Park**

Ph.D. Candidate  
POSTECH  
mjeongp@postech.ac.kr