Juyong Kim

juyongk@cs.cmu.edu | Homepage / Google Scholar / GitHub / LinkedIn | +1-412-616-5204

Professional Profile

Ph.D. candidate in the **Machine Learning Department** at **Carnegie Mellon University** with expertise in **natural language processing, multimodal learning, tabular ML, and computer vision**.

- **Research Focus**: Clinical NLP (multimodal timeline extraction, coding, normalization), Tabular ML (foundation models), Neural Architectures (generalization, compression), Computer Vision (3D motion)
- **Industrial Experience**: Applied research internships at Google Research, Abridge, Amazon
- **Publications**: ACL, EMNLP, ICML, CVPR, ECCV
- **Honors & Fellowship**: NLM Intramural Research Program (NIH)

Core Skills

- **AI/ML**: Multimodal Learning, Tabular ML, Generative Models, Scalable Training, Hyper-parameter Optimization, Parameter-Efficient Fine-Tuning (PEFT), Model Compression, Compositional Generalization, Transformers, GNNs
- **Natural Language Processing**: LLMs, Language Generation, Clinical Text Normalization, Clinical Coding (ICD)
- **Computer Vision**: Semantic Deep Networks, 3D Human Motion Reconstruction
- **Programming & Frameworks**: PyTorch, TensorFlow, Hugging Face, OpenCV; Python, C++, Java, MATLAB, SQL
- **Systems & Tools**: Docker, AWS, GCP, Git, Linux, Slurm, Flask, Spring, Android

Education

Machine Learning Department, Carnegie Mellon University Aug. 2018 – Present
- Ph.D. candidate in Machine Learning
- Advisor: Prof. Pradeep Ravikumar, Prof. Jeremy C. Weiss (NIH)

Vision & Learning Lab., Seoul National University Mar. 2016 – Feb. 2018
- M.S. in Computer Science and Engineering
- Advisor: Prof. Gunhee Kim, Prof. Sungju Hwang (UNIST)

Seoul National University Mar. 2008 – Feb. 2015
- B.S. in Electrical and Computer Engineering (Summa Cum Laude, Best Engineering Graduate Student Award)

Work & Research Experiences

Applied Scientist Intern May. 2024 – Aug. 2024
Amazon, Seattle, WA
- Developed a **multimodal multitask prompt tuning** framework with **automatic task grouping** for product attribute extraction, leveraging both visual and textual data.
- Evaluated on public and internal Amazon benchmarks, achieving competitive few-shot accuracy across 80+ tasks.

Student Researcher Nov. 2022 – May. 2024
National Library of Medicine, NIH, Bethesda, MD (Remote)
- Intramural Research Program, a federally funded to support clinical information processing research.
- Contributed to a project on **clinical event timeline extraction** from multimodal EHR data, developing annotation tools and **multimodal LLM** to improve temporal precision; Published at PAKDD 2024.

NLP Research Intern May. 2021 – Aug. 2021
Abridge, Pittsburgh, PA (Remote)
- Investigated **clinical LLMs** (T5) for converting doctor–patient conversations into structured SOAP note items.
- Designed and evaluated models using classical summarization metrics on proprietary datasets.

Research Intern May. 2020 – Aug. 2020
Google Research, Mountain View, CA (Remote)
- Developed a classification **benchmark for compositional generalization** from semantic parsing datasets.
- Demonstrated that **structure annotations** (parse trees, entity links) significantly improve Transformers' compositional generalization; Published at ACL 2021.

Software Engineer Sep. 2011 – Jul. 2014
Alternative Military Service, Seoul, South Korea
- Developed mobile web platforms and Windows CE applications.

Publications

International Conference

• **Juyong Kim**, C. Squires, P. Ravikumar, "**Knowledge-Enriched Machine Learning for Tabular Data**", in *International Conference on Neuro-symbolic Systems (NeuS), Oral Presentation*, 2025.

• S. Shin, **Juyong Kim**, E. Halilaj, M. J. Black, "**WHAM: Reconstructing World-grounded Humans with Accurate 3D Motion**", in *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2024.

• **Juyong Kim***, G. Frattallone-Llado*, C. Cheng, D. Salazar, S. Edakalavan, J. C. Weiss, "**Using Multimodal Data to Improve Precision of Inpatient Event Timelines**", in *Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD)*, 2024.

• W. Zhang, Z. Wang, **Juyong Kim**, C. Cheng, T. Oommen, P. Ravikumar, J. C. Weiss, "**Individual Fairness under Uncertainty**", in *European Conference on Artificial Intelligence (ECAI)*, 2023.

• **Juyong Kim**, A. Sharma, S. Shanbhogue, P. Ravikumar, and J. C. Weiss, "**AnEMIC: A Framework for Benchmarking ICD Coding Models**", in *Conference on Empirical Methods in Natural Language Processing (EMNLP, System Demonstrations)*, 2022.

• **Juyong Kim**, J. C. Weiss, P. Ravikumar, "**Context-Sensitive Spelling Correction of Clinical Text via Conditional Independence**", in *Conference on Health, Inference, and Learning (CHIL)*, 2022.

• **Juyong Kim**, P. Ravikumar, J. Ainslie, S. Ontañón, "**Improving Compositional Generalization in Classification Tasks via Structure Annotations**", in *Proceedings of the Association for Computational Linguistics (ACL)*, 2021 (Short Paper).

• **Juyong Kim**, L. Gong, J. Khim, J. C. Weiss, P. Ravikumar, "**Improved Clinical Abbreviation Expansion via Non-Sense-Based Approaches**", in *Machine Learning for Health (ML4H) NeurIPS Workshop*, 2020.

• **Juyong Kim**, Y. Park, G. Kim, S. Hwang, "***SplitNet*: Learning to Semantically Split Deep Networks for Param-eter Reduction and Model Parallelization**", in *International Conference on Machine Learning (ICML)*, 2017.

• W. Goo, **Juyong Kim**, G. Kim, S. Hwang, "**Taxonomy-Regularized Semantic Deep Convolutional Neural Networks**", in *European Conference on Computer Vision (ECCV)*, 2016.

• J. Kim, **Juyong Kim**, S. You, Y. Oh, and S. Oh, "**Actionable Topological Mapping for Navigation Using Nearby Objects",** in *Proc. of the IEEE International Conference on Automation Science and Engineering (CASE),* 2012.

Honors & Awards

- **NLM Intramural Research Program Fellowship, NIH** (2022-2024)
- **ILJU Overseas Ph.D. Scholarship** (2018-2023)
- **NVIDIA Deep Learning Contest**, 2nd place (2016, Korea)
- **Hyundai Motor Chung Mong-Goo Scholarship** (2016-2018)
- **National Science and Engineering Scholarship** (2008-2015)

Service & Teaching Experience

- **Peer Reviewer**: ACL-IJCNLP (2021), CHIL (2022), ML4H (2022–2023); Journals: ACL Rolling Review (2021–2024)
- **Teaching Assistant**: Advanced Deep Learning (CMU, Spring 2022), Advanced Introduction to Machine Learning (CMU, Fall 2019), Computer Vision (SNU, Spring 2016)

Additional

Selected Coursework

- **Machine Learning**: (CMU) Advanced ML, Deep Reinforcement Learning, ML with Graphs, Theoretical and Empirical Foundations of Modern ML, (SNU) Probabilistic Graphical Models
- **Computer Vision**: (CMU) Learning-based Image Synthesis
- **Statistics & Optimization**: (CMU) Convex Optimization, Foundations of Causal Inference, (SNU) Estimation Theory
- **Robotics & Control**: (SNU) Fundamentals of Control Engineering, Advanced Control Techniques

Others

- **Competitive Programming**: Codeforces Master (2100+ rating)

(Last update: 09/11/2025)