Min Gu Kwak

#329, ISyE Main Building, Georgia Institute of Technology,
755 Ferst Dr NW, Atlanta, GA 30318, USA
mkwak35@gatech.edu • min9kwak • https://min9kwak.github.io • +1-404-784-3064
CV compiled on 2024-07-19

EDUCATION

Korea University, Seoul, Republic of Korea

- Ph.D. in Industrial and Management Engineering — Mar 2016 – Feb 2022
  - Adviser: Prof. Seoung Bum Kim
  - Focus: Deep learning for handling out-of-distribution or limited number of data in various applications
  - Overall GPA: 4.42 / 4.50
  - Dissertation: Deep Representation Learning for Out-of-Distribution Data

Korea University, Seoul, Republic of Korea

- B.S. in Industrial and Management Engineering — Mar 2010 – Feb 2016
  - Overall GPA: 3.93 / 4.50

WORK EXPERIENCE

Georgia Institute of Technology, Atlanta, the United States of America

- Postdoctoral Fellow in H. Milton Stewart School of Industrial and Systems Engineering — Apr 2022 – Present
  - Supervisor: Prof. Jing Li
  - Focus: Artificial Intelligence for Medical Data Analysis and Integration for Clinical Usage

AWARDS & HONORS

- Research Funding, Next-Generation Science and Technology Leader NET, KOFST — 2024
  - Received funding for research activities aimed at improving the logical inference capabilities of large language models.
- Best Paper Award, Finalist, IISE 2024 Data Analytics & Information Systems — 2024
- Best Paper Award, Runners-Up, INFORMS 17th Workshop on Data Mining & Decision Analytics — 2022
- SAS Best Paper Award, Korea Business Intelligence Data Mining Conference — 2018
- National Science & Technology Scholarship, Korea Student Aid Foundation — 2014 – 2015
- Academic Excellence Scholarship, Korea University — 2010

SKILLS

- Programming Languages: Python (Advanced), R (Advanced), SQL (Proficient)
- Deep Learning Frameworks: Pytorch (Advanced), Keras (Advanced), Tensorflow (Proficient)
- Tools: Git (Proficient), Docker (Intermediate)

RESEARCH INTERESTS & ACTIVITIES

RESEARCH INTERESTS

- Artificial intelligence applications for medical data analysis
  - Neuroimaging and cone beam computed tomography image
  - Incomplete multi-modal learning with medical images
  - Conditional generative artificial intelligence with diffusion models
  - Interpretable graph neural networks for macromolecules
- Representation learning for out-of-distribution data
  - Anomaly detection, open-set classification, and their applications
  - Semi-supervised learning under class distribution mismatch scenario
- Learning with limited labeled data: semi-supervised and self-supervised learning
- Deep reinforcement learning
  - Robust and feedback-efficient preference-based reinforcement learning
  - Data-efficient reinforcement learning
- Large language model
  - Prompt engineering for logical inference
  - Automated data augmentation for self-correction
- Multi-channel signal data analysis

RESEARCH ACTIVITIES

- 7 published journal papers, 5 submitted journal papers, 2 conference proceedings, and 1 preprint
- 22 conference presentations

PUBLICATIONS

JOURNAL PAPERS

[1] Kim, J., Lee, Y. J., Kwak, M., Park, Y. J., & Kim, S. B. (2024). DynaSTI: Dynamics Modeling with Sequential Temporal Information for Reinforcement Learning in Atari. *Knowledge-Based Systems*, 112103. (citation: 0)

[2] Kwak, M.G., Su, Y., Chen, K., Weidman, D., Wu, T., Lure, F., Li, J., for the Alzheimer's Disease Neuroimaging Initiative. (2023). Self-Supervised Contrastive Learning to Predict the Progression of Alzheimer's Disease with 3D Amyloid-PET. *Bioengineering*, 10, 1141. (citation: 3)

[3] Kwak, M., Lee, Y. J., & Kim, S. B. (2023). SWaCo: Safe Wafer Bin Map Classification with Self-Supervised Contrastive Learning. *IEEE Transactions on Semiconductor Manufacturing*. (citation: 1)

[4] Lee, Y. J., Kim, J., Kwak, M., Park, Y. J., & Kim, S. B. (2022). STACoRe: Spatio-temporal and action-based contrastive representations for reinforcement learning in Atari. *Neural Networks*. (citation: 7)

[5] Kwak, M., & Kim, S. B. (2021). Unsupervised Abnormal Sensor Signal Detection With Channelwise Reconstruction Errors. *IEEE Access*, 9, 39995-40007. (citation: 11)

[6] Lee, J., Do, H., Kwak, M., Kahng, H., & Kim, S. B. (2021). Hierarchical segment-channel attention network for explainable multichannel signal classification. *Information Sciences*, 567, 312-331. (citation: 5)

[7] Lee, S., Kwak, M., Tsui, K. L., & Kim, S. B. (2019). Process monitoring using variational autoencoder for high-dimensional nonlinear processes. *Engineering Applications of Artificial Intelligence*, 83, 13-27. (citation: 138)

JOURNAL SUBMITTED

[1] Kwak, M. G., Mao, L., Zheng, Z., Su, Y., Lure, F., & Li, J. (2024). A Mutual Knowledge Distillation Framework for Alzheimer's Disease Diagnosis Using Incomplete Multi-Modal Images. *IEEE Transactions on Automation Science and Engineering*, under revision.

[2] Lee, Y., Kwak, M. G., Chen, R. Q., Yan, H., Mupparapu, M., Lure, F., Setzer, F. C., & Li, J. (2024). Oral-Anatomical Knowledge-Informed Semi-Supervised Learning for 3D Dental CBCT Segmentation and Lesion Detection. *IEEE Transactions on Automation Science and Engineering*, under revision.

[3] Heo, J., Lee, Y. J., Kim, J., Kwak, M., Park, Y. J., & Kim, S. B. (2024). Mixing Corrupted Preferences for Robust and Feedback-Efficient Preference-Based Reinforcement Learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, submitted.

[4] Lee, Y. J., Kim, J., Park, Y. J., Kwak, M., & Kim, S. B. (2024). Masked and Inverse Dynamics Modeling for Data-Efficient Reinforcement Learning. *IEEE Transactions on Neural Networks and Learning Systems*, under revision.

[5] Kwak, M., Kahng, H., & Kim, S. B. (2023). Safe Semi-Supervised Contrastive Learning Using In-Distribution Data as Positive Examples. *IEEE Transactions on Knowledge and Data Engineering*, under revision.

CONFERENCE PROCEEDINGS

[1] Jung, J. R., Seo, H. D., Hwang, K. R., Do, H. R., Kwak, M., & Kim, S. B. (2020, January). Optimization of Health Indices for Power Assets in Substation Using Machine Learning Method. *CIGRE GS 2020*.

[2] Kwak, M., & Kim, S. B. (2019, September). Channel-Wise Reconstruction-Based Anomaly Detection Framework for Multi-channel Sensor Data. In *Proceedings of SAI Intelligent Systems Conference* (pp. 1222-1233). Springer, Cham.

PREPRINTS

[1] Kwak, M. G., Su, Y., Chen, K., Weidman, D., Wu, T., Lure, F., & Li, J. (2023). A Mutual Knowledge Distillation-Empowered AI Framework for Early Detection of Alzheimer's Disease Using Incomplete Multi-Modal Images. *medRxiv*, 2023-08. (citation: 0)

PRESENTATIONS **INTERNATIONAL CONFERENCES**

[1] Lee, Y., Kwak, M. G., Chen, R. Q., Yan, H., Mupparapu, M., Lure, F., Setzer, F. C., Li, J., Oral-Anatomical Knowledge-driven Semi-supervised Semantic Segmentation for Dental CBCT Image, 2024 IISE Annual Conference, Montreal, Canada (May 18-21)

[2] Kwak, M., Su, Y., Chen, K., Weidman, D., Wu, T., Lure, F., Li, J., A Mutual Knowledge Distillation-Empowered AI Framework for Early Detection of Alzheimer's Disease Using Incomplete Multi-Modal Images, INFORMS 18th Workshop on Data Mining & Decision Analytics, PHX, AZ, US, 2023 (Oct. 14)

[3] Kim, J. H., Lee, Y. J., Kwak, M. G., Park, Y. J., Kim, S. B., Reinforcement Learning with Non-Contrastive Learning to Enhance Sample Efficiency in Atari, 2023 INFORMS Annual Meeting, PHX, AZ, US, 2023 (Oct. 15-18)

[4] Kwak, M., Su, Y., Chen, K., Weidman, D., Wu, T., Lure, F., Li, J., Self-Supervised Contrastive Learning to Predict Alzheimer's Disease Progression with 3D Amyloid-PET, INFORMS 17th Workshop on Data Mining & Decision Analytics, IND, US, 2022 (Oct. 15)

[5] Mok, C., Kwak, M., Kim, S.B., Self-Supervised Learning for Anomaly Detection on Multivariate Sensor Data, 2022 INFORMS Annual Meeting, IND, US, 2022 (Oct. 16-19)

[6] Kwak, M., Kim, S.B., Aggregating In-Distribution Data into Positive Examples for Safe-Semi Supervised Contrastive Learning, 2021 INFORMS Annual Meeting, Virtual Conference, 2021 (Oct. 24-27)

[7] Kwak, M., Kim, S.B., Safe Semi-Supervised Learning with Self-Supervised Approach, 2021 ICIEA, Virtual Conference, 2021 (Apr. 23-26)

[8] Kwak, M., Kim, S.B., Explainable Failure Prediction for Multi-channel Sensor Data, 2019 INFORMS Annual Meeting, Seattle, WA, US, 2019 (Oct. 20-23)

[9] Kwak, M., Kim, S.B., Convolutional Autoencoder-Based Multichannel Signal Monitoring Method, 2018 INFORMS International Conference, Taipei, Taiwan, 2018 (Jun. 17-20)

DOMESTIC CONFERENCES (SOUTH KOREA)

[1] Kwak, M., Kim, S.B., Safe Semi-Supervised Contrastive Learning for Out-of-Distribution Data, 2021 Korean Institute of Industrial Engineers, Virtual Conference, 2021 (Nov. 12)

[2] Shin, J.Y., Kim, K.H., Lee, S.J., Kim, S.B., Kwak, M., Lee, M., Lee, J., Artificial Intelligence System for Safety Helmet Detection, 2021 Korean Institute of Industrial Engineers, Jeju-do, Korea, 2021 (Jun. 2-5)

[3] Kwak, M., Kang, S.H., Kim, S.B., Critical Test Item Selection in Mobile Manufacturing Process, 2021 Korean Institute of Industrial Engineers, Jeju-do, Korea, 2021 (Jun. 2-5)

[4] Kwak, M., Do, H., Kahng, H., Lee, J., Kim, S.B., Anomaly Detection and Diagnosis of Engine Valve with Attention Mechanism, 2019 Korean Institute of Industrial Engineers, Seoul, Korea, 2019 (Nov. 8)

[5] Min, D., Kwak, M., Kim, S.B., Waste Classification System Based on Convolutional Neural Network, 2019 Korean Institute of Industrial Engineers, Seoul, Korea, 2019 (Nov. 8)

[6] Lee, M., Kwak, M., Lee, J., Jo, Y., Mok, C., Kim, S.B., Early Prediction of 112 Reports Based on Machine Learning Algorithms, 2019 Korean Institute of Industrial Engineers, Seoul, Korea, 2019 (Nov. 8)

[7] Lee, J., Do, H., Kahng, H., Kwak, M., Kim, S.B., Hierarchical Feedforward Attention Network for Explainable Multi-sensor Signal Data Analysis, 2019 Korean Institute of Industrial Engineers, Seoul, Korea, 2019 (Nov. 8)

[8] Kwak, M., Kim, S.B., Monitoring and Diagnosis for Multi-Channel Sensor Data, Korea Business Intelligence Data Mining Society, Seoul, Korea, 2018 (Nov. 30)

[9] Kwak, M., Do, H., Lee M., Chae S., Kim, S., Improving the Criminal Classification System in Response to New Crime Methods, Industrial Engineering & Management Science Conference, Seoul, Korea, 2018 (Nov. 9)

[10] Kwak, M., Park, Y., Kim, S.B., Anomaly Detection for Equipment Status Monitoring with Multichannel Sensor Data, Industrial Engineering & Management Science Conference, Gyeongju, Korea, 2018 (Apr. 5-6)

[11] Sung, Y., Do, H., Kwak, M., Kim, S.B., Machine Learning-Based Heatlth Index for Facility Failure Prediction and Diagnosis, Industrial Engineering & Management Science Conference, Gyeongju, Korea, 2018 (Apr. 5-6)

[12] Kwak, M., Kim, S.B., Ensemble Pruning with Optimization Problem Framework, Industrial Engineering & Management Science Conference, Yeosu, Korea, 2017 (Apr. 27-28)

[13] Kwak, M., Kim, S.B., Data-Driven Forecasting Method for Intermittent Demand, Industrial Engineering & Management Science Conference, Seoul, Korea, 2016 (Nov. 19)

RESEARCH PROJECTS (GEORGIA TECH)

Image-based Models of Tumor-Immune Dynamics in Glioblastoma

- NIH — Oct 2023 – Present
  - Developed conditional diffusion & transformer models for translating tumorous brain images into healthy brain images for reducing patient heterogeneity.
  - Trained the model on benchmark datasets and applied it to internal datasets.
  - Utilized generated normal and original images to construct an epidermal growth factor receptor (EGFR) classification model.
  - Developed a brain image preprocessing tool that transforms MRIs, region of interests (ROIs), and biopsy locations into BraTS atlas. (Link)

AIDen: An AI-Empowered Detection and Diagnosis System for Jaw Lesions Using CBCT

- NIH — Aug 2022 – Present
  - Integrated domain knowledge into a deep semantic segmentation model for precise lesion detection in 3D CBCT images.
  - Applied knowledge of lesion occurrence near tooth roots to regularize and guide the segmentation model.
  - Significantly improved the detection and segmentation performance of small lesions.

Multi-Modality Image Data Fusion and Machine Learning Approaches for Personalized Diagnostics and Prognostics of MCI due to AD

- NIH — Apr 2022 – Present
  - Developed a self-supervised contrastive model for classifying MCI converters and non-converters using 3D amyloid-PET images, incorporating label information during the pre-training step.
  - Created a mutual knowledge distillation model for handling incomplete multi-modal 3D image data (MRI and amyloid-PET) in MCI conversion classification.
  - Designed a novel teacher model that enhances knowledge distillation by focusing on modality-common representation.

RESEARCH PROJECTS (KOREA UNIVERSITY)

Developing Non-Invasive Lipid Measurement Algorithm Based on 2D Array Sensor

- Samsung Advanced Institute of Technology — May 2020 – Apr 2021
  - Developed a method to predict blood lipid concentrations using optical sensor data, eliminating the need for blood draws.
  - Designed a data preprocessing framework to predict lipid levels from sensor data.
  - Implemented a hybrid approach combining autoencoders and machine learning algorithms.

Congestion-Aware Control of Overhead Hoist Vehicles in Semiconductor Fabrication Logistics

- Samsung Electronics — May 2020 – Dec 2020
  - Developed an adaptive agent to control transportation vehicles in semiconductor FABs, aiming to minimize traffic congestion.
  - Applied imitation learning and data augmentation techniques within a deep reinforcement learning framework.

Text Mining and Trend Analysis on Venture Companies and Startups

- Korea Institute of Startup and Entrepreneurship Development — Mar 2020 – Jul 2020
  - Conducted web crawling to collect news articles on venture companies and startups and performed text mining analysis to identify key trends and keywords by year.
  - Used community detection algorithms to group keywords and built a pipeline for hierarchical trend analysis.
  - Performed sentiment analysis to evaluate positive and negative impacts of government policies, and included findings in government agency reports.

Conversational Platform R&D: Machine Reading Comprehension with Large Language Models

- Hanwha ICT — May 2019 – Dec 2019
  - Developed text question answering methods for both Korean news articles and in-house regulation documents to ensure compliance.

- Trained large language models (e.g., BERT) and distilled them into smaller models for deployment (e.g., DistilBERT).

Durability Monitoring System for Road Simulator

- Hyundai Motors and DS-eTrade — Apr 2019 – Dec 2019
  - Developed an algorithm to detect abnormal states and problematic parts of vehicles during road simulator operations.
  - Implemented a hierarchical feedforward attention network to detect abnormal states and explain the causes.

Detecting and Categorizing Failure Patterns of EGR Valve

- Hyundai Motors and DS-eTrade — Apr 2019 – Dec 2019
  - Predicted failures in EGR valves of diesel cars and analyzed sensors causing these failures.
  - Implemented a hierarchical feedforward attention network to detect failures and identify critical sensors and time steps.
  - Utilized sensor-level attention scores to cluster failure patterns.

Classification of Signal Patterns for Abnormal Cause Analysis of Semiconductor Logistics Systems

- Samsung Electronics — Mar 2019 – Nov 2019
  - Developed a framework consisting of anomaly detection, anomaly pattern clustering, and classification of logistics indices.
  - Discovered meaningful anomaly patterns by clustering channelwise reconstruction errors.
  - Employed an open-set model capable of classifying known classes and detecting unseen classes not present in training data.

AI-Based Smart Construction to Reduce Costs by 20%

- Hyundai Heavy Industries and Youngshine D&C — Sep 2016 – Dec 2020
  - Conducted a study to predict construction equipment failures by analyzing sensor data.
  - Employed incremental PCA to adapt to changing data distributions over time, creating a lightweight model that can be easily embedded in equipment control systems.

Deep Learning-Based Reliability Diagnosis Process Improvement

- Samsung Electronics — May 2018 – Apr 2019
  - Conducted a study on machine learning methods for early diagnosis and prediction of wafer quality in sub-10nm logic technology.
  - Applied domain knowledge-based and machine learning-based methods for missing data imputation.

ICT-Based Crime Risk Prediction and Response Platform Development for Early Awareness of Risk Situations

- Electronics and Telecommunications Research Institute — Apr 2018 – Dec 2018
  - Developed a method combining criminal data with various public data to predict near-future crime reports.
  - Considered a combination of recurrent neural networks and ensemble algorithms as the prediction model.
  - Developed a software tool for visualizing crime report statuses.

Advancing Health Index Module for 154kV Substation Facilities

- Hyosung Heavy Industries — Sep 2018 – Dec 2018
  - Conducted a study to improve the accuracy of a pre-developed health index module.
  - Applied the advanced module to various types of facilities to demonstrate its generality.

Optimal Decision-Making System for Maintenance and Health Index Module for 154kV Substation Facilities

- Hyosung Heavy Industries — Dec 2017 – Mar 2018
  - Developed a system for optimal maintenance decisions for multiple facilities within a limited budget.
  - Applied dynamic programming and integer programming to optimize the decision-making process.
  - Designed a decision-making pipeline with gradient boosting machines to predict health indices and calculate feature importance for inspection items.
  - Implemented the pipeline into a monitoring system for real customers.

Forecasting Demand for Construction Equipment Parts Using Big Data Analysis

- Hyundai Heavy Industries — Oct 2016 – Jul 2017
  - Predicted monthly demand for construction equipment parts and designed an objective dealer evaluation indicator by comparing predicted values to actual sales.
  - Developed software for visual comparisons of actual sales, predicted demands, and evaluation indicators by region, dealer, and part.

SERVICE AND LEADERSHIP

- Reviewer, INFORMS Journal of Computing 2024
- Reviewer, IEEE Transactions on Automation Science and Engineering 2023
- Sergeant, Korean Augmentation to the United States Army Sep 2011 – Jun 2013

MEMBERSHIP

- Member, Institute for Operations Research and the Management Sciences (INFORMS)
- Member, Korean-American Scientists and Engineers Association (KSEA)

REFERENCES

Jing Li

- Professor, H. Milton Stewart School of Industrial & Systems Engineering, Georgia Institute of Technology
- Ph.D., Industrial and Operations Engineering, University of Michigan
- Telephone: +1-404-894-6515
- E-mail: jing.li@isye.gatech.edu

Seoung Bum Kim

- Professor, Department of Industrial and Management Engineering, Korea University
- Ph.D., Industrial and Systems Engineering, Georgia Institute of Technology
- Telephone: +82-2-3290-3379
- E-mail: sbkim1@korea.ac.kr

Hyungrok Do

- Assistant Professor, Division of Biostatistics, Department of Population Health, NYU Grossman School of Medicine
- Ph.D., Industrial and Management Engineering, Korea University
- Telephone: +1-646-501-3409
- E-mail: doh03@nyu.edu