Taeyoung **Yun**

Black-box Optimization · Generative Models · AI Safety

291 Daehark-ro, Yuseong-gu, Daejeon 34141

99yty@kaist.ac.kr | dbsxodud-11.github.io | dbsxodud-11 | Taeyoung Yun

Personal Profile

My research interest lies in solving complex and high-dimensional black-box optimization problems through the lens of conditional generative modeling. I'm interested in Diffusion Models, Generative Flow Networks (GFlowNets), and their applications to real-world tasks, e.g, biological sequence design, material discovery, and mechanical design.

Recently, I have been particularly interested in fine-tuning models with GFlowNets to sample from posterior distributions. I conduct research on fine-tuning LLM with GFlowNets to generate diverse and effective attacker prompts that can be used for robust red-teaming. I also conduct research on fine-tuning diffusion models with GFlowNets to solve offline model-based optimization problems.

Education

KAIST (Korea Advanced Institute of Science and Technology) *Daejeon, Korea*

Ph.D in Industrial and Systems Engineering *March 2024 - Current*

- Supervised by Jinkyoo Park

KAIST *Daejeon, Korea*

M.S in Graduate School of AI *September 2022 - February 2024*

- Supervised by Jinkyoo Park

KAIST *Daejeon, Korea*

B.S in Industrial and Systems Engineering & Computer Science (Double Major) *March 2018 - August 2022*

Internships

Hong Kong University of Science and Technology (HKUST) *Remote*

Visiting Intern *September 2024 - Current*

- Internship at HKUST hosted by Ling Pan.
- Fine-tune LLM to generate diverse and effective attacker prompts for recently developed LLMs
- Fine-tune LLM to generate model-preferred prompts for generating high-quality images from text-to-image diffusion models

Kakao Recommendation Team *Seoul, Korea*

Research Intern *March 2021 - August 2021*

- Develop contextual bandit algorithms for a personal recommendation.
- Analyze the gap between simulation and real-world deployment.

Industrial Projects

Incentive Design for Managing Taxi Fleet *Daejeon, Korea*

Collaborate with ETRI *March 2023 - March 2024*

- Develop an RL-based incentive design algorithm for rebalancing taxi fleet to resolve taxi imbalance problem.

Traffic Light Optimization *Seoul, Korea*

Collaborate with KT *March 2022 - March 2023*

- Develop a Bayesian optimization algorithm for managing multiple traffic lights in the real world to reduce congestion.

Publications

Guided Trajectory Generation with Diffusion Models for Offline Model-based Optimization *Arxiv*

Taeyoung Yun, Sujin Yun, Jaewoo Lee and Jinkyoo Park *2024*

- paper
- code

An Offline Meta Black-box Optimization Framework for Adaptive Design of Urban Traffic Light Management Systems *KDD*

Taeyoung Yun*, Kanghoon Lee*, Sujin Yun, Ilmyung Kim, Won-Woo Jung, Min-Cheol Kwon, Kyujin Choi, Yoohyeon Lee, and Jinkyoo Park (*: Equal Contribution) *2024*

- paper
- code

Learning to Scale Logits for Temperature-conditional GFlowNets *ICML*

Minsu Kim*, Juhwan Ko*, **Taeyoung Yun***, Dinghuai Zhang, Ling Pan, Woochang Kim, Jinkyoo Park, and Yoshua Bengio (*: Equal Contribution) *2024*

- paper
- code

GTA: Generative Trajetory Augmentation with Guidance for Offline Reinforcement Learning *ICLR GenAI4DM Workshop (Spotlight)*

Jaewoo Lee*, Sujin Yun*, **Taeyoung Yun**, and Jinkyoo Park (*: Equal Contribution) *2024*

- paper
- code

Local Search GFlowNets *ICLR (Spotlight)*

Minsu Kim, **Taeyoung Yun**, Emmanuel Bengio, Dinghuai Zhang, Yoshua Bengio, Sungsoo Ahn, and Jinkyoo Park *2024*

- paper
- code

Honors & Awards

Dean's List (Top 2%)

KAIST *2021*

NH Big Data Competition (Excellence Award)

DACON *2021*

Teaching Experiences

Teaching Assistant

IE437: Data-Driven Decision Making and Control *2023, 2024*

Teaching Assistant

MAS480: Introduction to Scientific Machine Learning *2022*

Academic Services

NIPS Reviewer *2024*