Masahiro NEGISHI

negishi-masahiro1110@g.ecc.u-tokyo.ac.jp

Education

The University of Tokyo *Tokyo, Japan*

Master of Information Science and Technology *Apr. 2023 - (Mar. 2025)*

- GPA: 4.00/4.00 (tentative)
- Courses: Advanced Data Analysis, Information-Theoretic Learning Theory, Natural Language Processing, etc.

The Vienna University of Technology *Vienna, Austria*

Research exchange student *Mar. 2024 - Jul. 2024*

- Grade: 1 (excellent), the highest rating on a 5-point scale
- Course: Project in Computer Science

The University of Tokyo *Tokyo, Japan*

Bachelor of Science in Information Science *Apr. 2019 - Mar. 2023*

- GPA: 3.85/4.00 (the top of 39 students in the department)
- Courses: Statistical Machine Learning, Statistics and Optimization, Discrete Mathematics, etc.

Research Experience

The University of Tokyo *Tokyo, Japan*

Bachelor's and master's theses research, under Prof. Masashi Sugiyama *Oct. 2022 - (Mar. 2025)*

- (Bachelor): Thesis: "Pairwise-constraint classification in weakly supervised machine learning: Risk-consistent approach and classifier-consistent approach"
- (Master): Thesis: "Weakly Supervised Disentanglement from Distance Supervision" (tentative)

King Abdullah University of Science and Technology *Thuwal, Saudi Arabia*

Visiting student researcher, under Prof. Di Wang *Sep. 2024 - (Feb. 2025)*

- Developing a symmetry-aware generative model for sampling physically and chemically valid crystals
- Gaining experience in evaluating generated crystals with first-principles calculations

Vienna University of Technology *Vienna, Austria*

Research exchange student, under Prof. Thomas Gärtner *Feb. 2024 - Jul. 2024*

- Investigated metric properties of the distance between embeddings of graph neural networks
- Identified fragments of molecules that graph neural networks consider important for prediction

OMRON SINIC X Corporation *Tokyo, Japan*

Research internship *Nov. 2023 - Feb. 2024, Aug. 2024*

- Developed a coefficient estimation algorithm for symbolic regression for scientific discovery
- Achieved a success rate of 58% on complex physics dataset, which is 33% higher than existing methods

Matsuo Institute, Inc *Tokyo, Japan*

Research internship *Oct. 2022 - Aug. 2023*

- Verified the scaling law of generative models for autonomous driving and robotics
- Acquired experience in distributed training of up to a billion parameter models using deepspeed

Awards & Scholarships

Scholarship for studying abroad *Tokyo, Japan*

Japan Student Services Organization *Feb. 2024 - Jul. 2024*

- Received $540 USD per month for 6 months, plus airfare

Honorable mention *Tokyo, Japan*

Faculty of Information Science, School of Science *Mar. 2023*

- Ranked 1st in the department (38 students) based on the overall evaluation of my thesis and coursework

7th place in global hackathon on privacy-preserving machine learning *Virtual*

United Nations Privacy Enhancing Technologies Lab *Nov. 2022*

- Ranked 7th among student teams, and 11th among all teams

Publications

Domestic(peer-reviewed)

Masahiro NEGISHI, Makoto SATO, Ryosuke UNNO, Koudai TABATA, Taiju WATANABE, Junnosuke KAMOHARA, Taiga KUME, Ryo OKADA, Yusuke IWASAWA, and Yutaka MATSUO. Scaling Laws of Dataset Size for VideoGPT. Proceedings of the Annual Conference of JSAI, 2023, Volume JSAI2023, 37th (2023), Pages 2G6OS21f05.

Koudai TABATA, Junnosuke KAMOHARA, Ryosuke UNNO, Makoto SATO, Taiju WATANABE, Taiga KUME, **Masahiro NEGISHI**, Ryo OKADA, Yusuke IWASAWA, and Yutaka MATSUO. Construction and Validation of Action-Conditioned VideoGPT. Proceedings of the Annual Conference of JSAI, 2023, Volume JSAI2023, 37th (2023), Pages 1G4OS21a02.

Makoto SATO, Ryosuke UNNO, **Masahiro NEGISHI**, Koudai TABATA, Taiju WATANABE, Junnosuke KAMOHARA, Taiga KUME, Ryo OKADA, Yusuke IWASAWA, and Yutaka MATSUO. Scaling Laws of Model Size for World Models. Proceedings of the Annual Conference of JSAI, 2023, Volume JSAI2023, 37th (2023), Pages 2G5OS21e02.

International(peer-reviewed)

Masahiro NEGISHI, Ryo IGARASHI, Yoshitaka USHIKU, Yoshitomo MATSUBARA, Naoya CHIBA. Two-Stage Coefficient Estimation in Symbolic Regression for Scientific Discovery. Machine Learning and the Physical Sciences @ NeurIPS 2024 workshop

International(under-review)

Masahiro NEGISHI, Pascal WELKE, Thomas GÄRTNER. WILTing Trees: Interpreting the Distance Between MPNN Embeddings. https://openreview.net/forum?id=9pBnp90o2D (Notice of acceptance will be given on January 22, 2025.)

Oral Presentation

The 37th Annual Conference of the Japanese Society for Artificial Intelligence, Kumamoto, Japan, 2023. “Scaling Laws of Dataset Size for VideoGPT”, 20 minutes

Skills

Machine Learning	Graph Neural Networks, Graph Kernels, explainable AI, Generative Models, Symbolic Regression, Weakly supervised learning, Disentanglement
Programming	Python(PyTorch, PyTorch Geometric, pymatgen, deepspeed, etc), C++, LaTeX, Wandb, Docker, GitHub, HTML, JavaScript
Language	Japanese(Native), English(TOEFLiBT 104, C1)

Extracurricular Activity

Educational workshop on responsible AI for peace and security *Malmö, Sweden*

United Nations and Stockholm International Peace Research Institute *Nov. 2023*

- Learned about what AI risks are and how to mitigate them with responsible research and innovation
- Discussed with selected Master’s and Ph.D. students from around the world