Hiroki Tanabe

Data Scientist / PhD in Computer Science / YouTuber

Experience

2022 –Present
Oct

Data Scientist, *Yahoo Japan Corporation*, Osaka, Japan

Responsible for the development of CVR (Conversion Rate) prediction models for the display ads, which is used for the automated bidding. Experienced data aggregation using Hive on a Hadoop cluster, development of machine learning models using Python (XGBoost, LightGBM), construction of DevOps and MLOps pipelines with Argo Workflows, Airflow, and Screwdriver.cd, and A/B testing in the production environment. All processes are performed on the AWS-like (such as S3) or GCP-like (such as Vertex AI) private cloud platform.

2017 –Present
Feb

YouTuber

Manage two YouTube channels, the Gaming channel with 16,100 subscribers / 7,063,000 total views and the Music Cover channel with 4,500 subscribers / 1,170,000 total views. A speedrun video has more than 1 million views. Also have a virtual YouTuber channel that is automated by a *generative AI*, basically written in Python and uses the OpenAI API and the Azure Kubernetes Service.

2020 – 2022
Apr   Sept

Research Fellowship for Young Scientists (DC1), *Japan Society for the Promotion of Science*, Tokyo, Japan

Awarded to excellent PhD students with a selection ration of 20%, this fellowship allows the fellows to focus on a freely chosen research topic based on their innovative ideas.

Research subject: Development, theoretical analysis, and implementation of multiobjective optimization algorithms, particularly first-order methods, i.e., methods that use only the first-order derivatives of the objective functions. They are useful in machine learning, statistics, signal and image processing.

2020 – 2020
Sept   Sept

Software Engineering Internship, *VOYAGE Group Inc.*, Tokyo, Japan

Developed a coding test tool in a four-person engineering team to improve efficiency in HR. Communicated with the HR staff and asked about their recruiting problems. Designed the prototype with Miro. Implemented authentication by Google account, uploading source codes, and grading each submission, using Vue.js and Firebase.

2020 – 2020
Aug   Sept

Data Scientist Internship, *Yahoo Japan Corporation*, Osaka, Japan

Estimated heterogeneous treatment effect on earnings of Yahoo! Japan Auction of a recommendation system by using Machine Learning with Python on Jupyter Notebook. Analyzed seven days of preprocessed traffic data of Yahoo! Japan Auction, including the click data of the recommendation banner, in Pandas DataFrame. Implemented X-Learner, a kind of algorithm called "meta-learner," which estimates CATE (Conditional Average Treatment Effect) using any machine learning estimator. Used LightGBM and linear regression as the estimator of X-Learner. Prevented leakage through correlation analysis. Verified if the recommendation system impacts the earnings via Student's t-test and Cohen's d effect size. Note that such estamates of the treatment effect is useful when it is difficult to conduct A/B testing.

2020 – 2020
Aug   Aug

Software Engineering Internship, *Cookpad Japan Inc.*, Tokyo, Japan

Built a new Ruby on Rails web application from scratch. It enables users to record and rate recipes they have used. Using the Lean Startup Methodology, developed the MVP (Minimum Viable Product) that satisfies the value hypothesis and utilized a feedback loop through user testing. Presented a prototype of the product by using Figma. Implemented functions of recording recipes with title, memo, image, and rating, editing and destroying them, showing them in a grid view, and sorting them by rating and date, using Ruby on Rails, Docker, Boostrap.

2019 – 2020
Feb   Mar

Data Scientist Internship, *HACARUS Inc.*, Kyoto, Japan

Designed and developed some machine learning models for the manufacturing and agricultural industry, mainly through analyzing a small number of image datasets, typically using Python. In particular, developed models for surface defect detection through various techniques, including dictionary learning and fused lasso and hyperspectral image segmentation via online dictionary learning and typical classification algorithms (e.g., support vector machine, random forest). Surveyed scientific papers and implemented algorithms from them, using some machine learning libraries (e.g., scikit-learn), image processing libraries (e.g., open-cv, scikit-image), data visualization libraries (e.g., matplotlib, seaborn), and notebooks environments (e.g., Jupyter).

Other minor engagements included: Added some new algorithms (e.g., alternating direction methods of multipliers for the generalized lasso, matching pursuit for sparse coding) to the open-source library "spm-image," a scikit-learn compatible library of sparse modelling and compressive sensing. Implemented batch-OMP, an improved version of the orthogonal matching pursuit algorithm, to contribute to the research project from NEDO (the New Energy and Development Organization), the Japanese largest governmental R&D organization. Developed front-end of an in-house web application that generates an interactive visualization of image datasets' quality, using flask, pandas, and bokeh.

DUMMY ADDRESS

□ ****-****-*****   •   ✉ *tanabe.hiroki.45n@kyoto-u.jp*   •   in *hiroki-tanabe*   •   🐦 *zalgo3*
○ *zalgo3*

Education [*]

2019 – 2022 Oct Sept	**PhD, Computer Science**, *Kyoto University*, Kyoto, Japan
2018 – 2019 Apr Sept	**MSc, Computer Science**, *Kyoto University*, Kyoto, Japan, Valedictorian, early degree completion
2014 – 2018 Apr Mar	**BSc, Computer Science**, *Kyoto University*, Kyoto, Japan, Valedictorian

Skills Summary

Technical

Industry Knowledge	Optimization, Machine Learning, Data Engineering, DevOps, MLOps, Generative AI	Tools & Technologies	Python, TypeScript, Next.js, Vim, LaTeX, Apache Hive, Apache Hadoop, GitHub Actions, Git, Docker, Kubernetes, OpenAI API, LangChain

Language

English	Professional working proficiency	Japanese	Native or bilingual proficiency

Publications

Papers

- Hiroki Tanabe, Ellen H. Fukuda, and Nobuo Yamashita, An accelerated proximal gradient method for multiobjective optimization. *To appear in Computational Optimization and Applications*, 2023. [pdf]
- Hiroki Tanabe, Ellen H. Fukuda, and Nobuo Yamashita, Convergence rates analysis of a multiobjective proximal gradient method. *Optimization Letters*, 17, pp. 333–350, 2023. [doi, pdf]
- Hiroki Tanabe, Ellen H. Fukuda, and Nobuo Yamashita, Proximal gradient methods for multiobjective optimization and their applications, *Computational Optimization and Applications*, 72(2), pp. 339–361, 2019. [doi, pdf]
- Hiroki Tanabe, Ellen H. Fukuda, and Nobuo Yamashita, New merit functions for multiobjective optimization and their properties. *Submitted*, 2023. [pdf]
- Hiroki Tanabe, Ellen H. Fukuda, and Nobuo Yamashita, A globally convergent fast iterative shrinkage-thresholding algorithm with a new momentum factor for single and multi-objective convex optimization. *Submitted*, 2022. [pdf]

Books

- Takashi Someda, Naoki Kitora, Ippei Usami, Ryuji Masui, and Hiroki Tanabe, An Introduction to Sparse Modeling for IT Engineers (in Japanese), Shoeisha, 2021.

Invited Talks

- Takeshi Koshizuka, Kohei Harada, Yoshiaki Inoue, Ayumi Igarashi, Hiroki Tanabe, and Kiyohito Nagano, Open Roundtable: The Present and Future of OR Research (in Japanese), Operations Research Society of Japan, 2022.

Awards

Aug 2020	**Student Thesis Award 2020 [given to 6 students]**, *Operations Research Society of Japan*
July 2020	**Full Exemption from Repayment for Graduate School Students with Particularly Outstanding Achievements for Category 1 Loans [1,456 out of 21,538]**, *Japan Student Services Organization*
Feb 2020	**Grand Prize in Data Analytics Competition 2019 [1 out of 14 teams]**, *Japan Institute of Marketing Science*
Feb 2020	**Excellent Research Award at Kyoto University's 14th ICT Innovation [8 out of 54]**, *Kyoto University*
Sep 2018	**A finalist for the Young Author's Award in SICE Annual Conference 2018 [4 out of 11]**, *Society of Instrument and Control Engineers*

References

Available upon request.

---

[*]In Japan, Doctorate Degree = Doctor of Philosophy (PhD) = Doctor of Science (D.Sc.), Master's Degree = Master of Science (MSc), Bachelor's Degree = Bachelor of Science (BSc), and Informatics = Information Science = Computer Science are often used interchangeably. For this reason, PhD, MSc, BSc, and Computer Science are used here as universal translations to reduce confusion.