

MICHIHARU YAMASHITA

Email: michiharu@psu.edu • Web: <https://mickeymst.github.io/>

Education

Pennsylvania State University

Ph.D. student in Information Science and Technology (Advisor: Prof. Dongwon Lee)

University Park, PA

August 2020 – Present

Thesis: Understanding and Modeling Company and Talent Behavior on Online Professional Networks

Tokyo Institute of Technology

M.Eng. in Computational Intelligence and Systems Science (Advisor: Prof. Kazuo Yano)

Tokyo, Japan

April 2015 – March 2017

Salutatorian (2nd place in the department)

Thesis: *Does Temperature Change Happiness?: An Empirical Analysis Using Wearable Sensors in Offices*

University of Tsukuba

B.S. in Management Science and Engineering (Advisor: Prof. Ushio Sumita)

Tsukuba, Japan

April 2011 – March 2015

Thesis: *Network Analysis and Visualization for Mobile Applications' Competitiveness*

Employment

PIKE Research Group at Penn State

Research Assistant (Advisor: Prof. Dongwon Lee)

University Park, PA

August 2020 – Present

- Research Topics: Computational Social Science, People Analytics, Machine Learning

Self-employment

Machine Learning Engineer

Tokyo, Japan

June 2019 – July 2020

- Understood and modeled user behavior with human resource data with a technology startup.
- Developed prediction model using wearable sensor data with a big tech company.
- Developed algorithms of advertising optimization for marketing with an educational startup.

Recruit Holdings

Research Engineer and Machine Learning Engineer

Tokyo, Japan

April 2017 – January 2019

- Analyzed cross-disciplinary service data of millions of user data.
- Understood and modeled user preference from GPS based location data.
- Developed machine learning based recommendation system

Megagon Labs (formerly Recruit Institute of Technology)

Research Engineer

Tokyo, Japan and Mountain View, CA

April 2017 – March 2018

- Developed an entity matching and entity extraction module and NLP related frameworks.
- Applied NLP modules into our companies.
- Promoted open-source components BigGorilla.

Hitachi Central Research Laboratory

Research Assistant (Advisor: Prof. Kazuo Yano)

Tokyo, Japan

April 2015 – March 2017

- Researched on computational social science, network science, and science of happiness.
- Developed a prediction model with wearable sensors.
- Exploratory data analysis to wearable sensors data and network data.

Ohma Inc.

Software Engineer (Advisor: Prof. Yutaka Matsuo)

Tokyo, Japan

March 2016 – March 2017

- Developed a network visualization search engine SPYSEE2 which had **1M+ visits per month**.
- Developed entity extraction, entity linking, and face recognition systems from unstructured and noisy web data.
- Crawled millions of web pages efficiently using programs on AWS.

Recruit Holdings

Data Science Intern

Tokyo, Japan

August 2015 – September 2015

- Developed a recommendation engine for an e-commerce service.

Fuller, Tokyo (Japan)

April 2014 – March 2015

Research Assistant (Advisor: Prof. Ushio Sumita)

- Developed an algorithm for competitive analysis and ranking of smartphone applications.

Taiyo Yuden Mobile Technology, Tokyo (Japan)

April 2014 – March 2015

Research Assistant (Advisor: Prof. Ushio Sumita)

- Developed a Java-based simulation system for a factory's logistics to improve productivity.
- Developed mathematical optimization models for product scheduling of a factory

Publications

- **Michiharu Yamashita**, Shota Katsumata, and Yusuke Fukasawa. "Discovery of User Preferences from Big Geospatial Data Using Topic Models." Proceedings of the 2018 IEEE International Conference on Big Data (The 3rd International Workshop on Application of Big Data for Computational Social Science). IEEE, 2018.
- **Michiharu Yamashita**, Hideki Awashima, and Hidekazu Oiwa. "A Comparison of Entity Matching Methods between English and Japanese Katakana." Proceedings of the 15th SIGMORPHON Workshop on Computational Research in Phonetics, Phonology, and Morphology at EMNLP, 2018.
- Kent Kawai, **Michiharu Yamashita**, and Yutaka Matsuo. "Face Recognition System Based on Convolutional Neural Network Robust to Occlusion and Low Quality Images." The 31st Annual Conference of the Japanese Society for Artificial Intelligence. JSAI, 2017.
- **Michiharu Yamashita**, Nobuo Sato, and Kazuo Yano. "Enhancing Collective Happiness by Controlling Room Temperature Using Big Data from Wearable Sensors." The 2016 IEICE General Conference. IEICE, 2016.

Honors and Awards

- Funds for Excellence in Graduate Recruitment Scholarships**, Penn State University August 2020 – July 2021
Prestigious and selected scholarships for incoming graduate students at Penn State
- Recruit Holdings Freshman Award First Prize 2018**, Recruit Holdings April 2018
1st in the whole of Recruit Holdings Group of over 500 freshmen
- Recruit Holdings R&D MVP Award 2017**, Recruit Holdings April 2018
1st in R&D department of Recruit Holdings of over 200 employees
- Full Repayment Exemption of Graduate Student Loan \$20,000 for Excellent Achievement**, JASSO August 2017
Top 10% of 26,987 Graduate Students using the Loan at all universities in Japan
- Salutatorian of Tokyo Institute of Technology**, Tokyo Institute of Technology March 2017
The 2nd representative in the department
- The Second Best Master Thesis Award**, Tokyo Institute of Technology February 2017
Master Thesis ranking 2nd in the department (2/62)
- Full Tuition Exemption for Students with Financial Difficulties but Outstanding Academic Performance at Tokyo Institute of Technology**, Tokyo Institute of Technology April 2015 – March 2017
- Recruit Holdings Internship Award First Prize 2016**, Recruit Holdings September 2015
1st of over 30 intern students
- Full Tuition Exemption for Students with Financial Difficulties but Outstanding Academic Capabilities at University of Tsukuba**, University of Tsukuba April 2011 – March 2015

Technical Strengths

Language: Japanese (Native), English (Fluent), Mandarin (Beginner)

Programming and Software: Python, Java, C/C++, Swift, R, SQL, Javascript, HTML, CSS, LaTeX

Framework and Tools: Tensorflow, Keras, Numpy/Scipy, scikit-learn, Flask, Elasticsearch, GCP, AWS