

Aki Ariga

www.linkedin.com/in/aki-ariga/ • chezou@gmail.com

GitHub: github.com/chezou • Website: chezo.uno

Customer-Centered “Machine Learning in Production” Engineer

- More than 10 years of experience in the application of machine learning research, implementation, and product development at start-ups and big enterprises, with experience in recommendation systems and production deployment.
 - Strongly motivated to deliver customer values with cross-team collaboration between product manager, UX designer, customer-facing engineers, and sales representatives, from analyzing customer requirements to execution of the development. Led multiple projects to release as an engineering lead.
 - An experienced backend engineer designed backend systems as a micro-service while considering machine learning data and model lifecycles.
-

TECHNICAL PROFICIENCIES

Platforms: macOS, Linux, Windows

Languages: Python, Ruby, Kotlin, JavaScript, R, C/C++, Java

Applications/Frameworks: scikit-learn, LightGBM, Docker, Ruby on Rails, Presto, Terraform, Apache Spark, TensorFlow

PROFESSIONAL EXPERIENCE

Treasure Data, Vancouver, Canada

April 2021 – Present

Staff Software Engineer

Relocated from Tokyo to Vancouver. See the details in the Tokyo section.

Treasure Data, Tokyo, Japan

August 2018 – April 2021

Staff Software Engineer

Technical Scope: Ruby on Rails, Python, Machine Learning, Data pipeline, Kotlin, Terraform

Led multiple projects as a Directly Responsible Individual (DRI) to release new features for a marketing solution on an enterprise big data platform to leverage data and machine learning, mainly developing in the backend API and data pipeline.

Key Accomplishments:

- Led the development of an ML application for marketers, from writing a report on the technical direction to the product leadership team, designing the end-to-end machine learning system, supporting the product manager and UX designer in understanding customers' expectations, and implementing the backend component. This project was released three weeks ahead of schedule and met the challenging schedule of providing a live demo at [CDP World](#).
- Led a geo-diverse and cross-team project for ID unification as a DRI from technical design, implementation API server with Kotlin, and collaboration with UX designers for prototyping and supporting product managers to solidify product requirements with a better understanding of technical feasibility.
- Implemented a container-based Python execution system for machine learning and [implemented out-of-the-box machine learning solutions](#) with Python machine learning frameworks and PySpark.
- Implemented a backend system for customer marketing profile visualization with Ruby on Rails and a workflow engine.
- Developed a paid training hands-on course for Python and machine learning to enable customers' Python knowledge of our big data platform.
- Consults a machine learning project for a marketing campaign in the automotive industry and succeeds in improving revenue.
- Authored an [MLOps tutorial article](#) in a technical magazine published by Lambda Note in 2019.
- Presenter at [The 22nd Information-Based Induction Science Workshop](#) held in Nagoya 2019.

LegalOn Technologies, Tokyo, Japan

January 2020 – March 2021

Technical Consultant, Parttime

Technical Scope: MLOps, Machine Learning Infrastructure

Aki Ariga

Page 2 of 4

As a second job, consulted designing Kubernetes-based next-generation machine learning platform architecture and MLOps capability to achieve more frequent experiments, model serving, and model management.

Key Accomplishments:

- Advised to architect Kubernetes-based ML platform, which resulted in having [a presentation of its architecture](#) at an academic workshop.

Cloudera, Tokyo, Japan

March 2016 – June 2018

Sales Engineer

Technical Scope: Python, PySpark, R

Engaged clients to enhance data-driven use cases employing machine learning techniques. Developed prototypes of machine learning demo apps contributing OSS.

Key Accomplishments:

- Documented machine learning deployment patterns and authored the book [“Machine Learning at Work.”](#) Published by O’Reilly, Japan in 2017.
- Implemented six data science and machine learning demo use case programs contributing to open sources, such as [sparklyr](#) and [fastFM](#), which enhanced product capabilities.
- Subject matter expert delivering five presentations on machine learning project strategy and deployment patterns.
- Created four detailed technical entries on Cloudera’s blog to promote the company’s products about machine learning.
- Recipient of an award from the Cloudera APAC region for writing multiple documents on machine learning and using Python and PySpark to educate colleagues.
- Presenter at the [Strata Data Conference](#) held in Singapore in 2017.

Cookpad Inc.

July 2013 - March 2016

Software Engineer

Technical Scope: Ruby on Rails, Apache Solr

Designed and implemented the firm’s machine learning and natural language processing-based recipe recommendation service delivered through a Ruby on Rails-based web application.

Key Accomplishments:

- Designed and implemented a recipe recommendation service using machine learning and natural language processing on Ruby on Rails. The service became the highest-paid user acquisition feature for a half-year.
- Developed a word2vec model based on recipe corpus, which is published for academic use, and the innovation resulted in an interview with a leading Japanese technical magazine.
- Organized a machine learning hackathon event resulting in the hiring of three graduate students.

Toshiba Corporation

April 2008 - July 2013

Research Engineer

Technical Scope: Apache Nutch, Apache Hadoop, Apache Solr, NLP, Machine Learning

Researched spoken dialogue system, NLP, search system, and big data system by building machine learning model with big data and collecting web-scaled text data.

Key Accomplishments:

- Created a TV scene classifier by genres from closed caption program searches. Achieved over 80% precision through the prediction engine, which was released as a feature of a production system. The work resulted in a credit for developing a new patent.
- Designed and built a distributed crawler with Apache Nutch and Hadoop cluster to support NLP and speech recognition research. Employed the crawler to collect two hundred million websites for Text Corpus. Created the full-text search system utilizing Solr for the high-accuracy speech recognition system.
- Constructed large-scale regression trees employed to identify anomalies with factory sensor data. Employed MapReduce to create a model drawing on 100GB of training data.

EDUCATION & CREDENTIALS

Master of Engineering in Electrical Engineering and Computer Science (2008)

Nagoya University School of Engineering, Nagoya, Japan

Thesis: "Mental Tension Detection in the Speech Based on Physiological Monitoring."

Bachelor of Engineering in Electrical Engineering and Computer Science (2006)

Nagoya University School of Engineering, Nagoya, Japan

Languages: Japanese (Native), English (Professional proficiency; IELTS General Training Overall 7.0)

Publication

An introduction of MLOps, M. Ariga, N-monthly Lambda Note, Vol. 1, No. 1, Lambda Note, 2019

Machine Learning at Work, M. Ariga, S. Nakayama, T. Nishibayashi, O'Reilly Japan, 2016

Machine Learning at Work 2nd Edition, M. Ariga, S. Nakayama, T. Nishibayashi, O'Reilly Japan, 2021

Selected Papers

"Development of TV Program Searching System for Natural Sentences"

NLP Meetup for Young People, the 5th Symposium, 2010

"Detecting a Genre of TV Programs using a Non-Semantic Frame Keyword on TV Program Searching by Natural Language"
the 23rd meeting, The Japanese Society for Artificial Intelligence, 2009

"Mental Tension Detection in the Speech based on Physiological Monitoring"

IEEE International Conference on Systems, Man and Cybernetics, 2007

Distinctions

Recipient of 14 Patents (10 in Japan and 4 in the United States)

Cloudera APJ SE Award Grow with Cloudera (Dec 2017)

Cookpad Staff Award as a Member of Research and Development Team (Jan 2016)

The Founder of Meetups: kawasaki.rb (<http://kawasakirb.connpass.com/>), Machine Learning Casual Talks
(<http://mlct.connpass.com/>)

Affiliation

Public Relations Committee Member (2011-2013), The Japanese Society for Artificial Intelligence

Co-organizer of ML Infrastructure and Productionization Working Group(2018-2021), [MLSE: SIG Machine Learning Systems Engineering](#)