

Satoru Shibata | Data Scientist

Location: 19-1 Kodai, Odawara City, Kanagawa Prefecture, Japan

Telephone: +81 080-6708-5390 **Email:** satorushibatad4dr@gmail.com

Birth Date: June 27, 1991 (29 years old)



Professional profile

I experienced over 3 years development as a Data Scientist with Python3/R/SQL after [Sophia University Graduate School](#) and Technical Advisor on Data Science for PM at [NTT DATA Corporation](#) for 6 months.

In addition to My Development Experience, I acquired many technical skills applied to Data Science business:

- Machine Learning: LightGBM, XGBoost, CatBoost, t-SNE, k-means, Deep Learning
- Empirical Analysis: Dynamic Panel Data Analysis, Time Series Analysis, Generalized Linear Model, Significance Test
- Development Environment: Python3/R in Jupyter Notebook and SQL (MySQL/BigQuery/Redshift SQL)

Career summary

September 1, 2020 - February 28, 2021 EDIX, Inc, Tokyo, Japan
Lead Data Scientist and Technical Advisor

Outline

I was dispatched from EDIX, Inc to NTT DATA Corporation and served the PM as a Technical Advisor on Data Science.

I also advised executives at EDIX, Inc on Data Scientist positions as a recruitment assistant.

Key responsibilities

- I advised the PM on feasibility and complementary elements for projects especially during meetings;
- For example, during the project for measures against food loss, as for AI algorithm on recommendation, is this project theoretically or developmentally feasible as a result of using data provided from the contractor?
- If not possible, what exactly should be supplemented?
- I asked job seekers questions about the theory and implementation of Data Science:
- For example, “Would you be able to explain us about SVM (Support Vector Machine) and the difference between Inner Join and Left Join?”
- However, the job seeker could not provide any specific answers, and I recommended officers at EDIX, Inc not to hire.
- As a result, the job seeker was rejected.

Key achievements/projects

- Complementing PM and the members with Data Science Development/Food loss and recommendations for foreign tourists
- My unit price: over 1 million yen (around 9579.65 US \$) / month
- Recruitment assistant for Data Scientist positions at EDIX: Officers at EDIX, Inc explained to me that they would like to decide to hire or not hire job seekers on the basis of my advice, and that was the case.

April 1, 2017 – July 31, 2020

**SERAKU Co.,Ltd., Tokyo, Japan
Data Scientist**

Outline

I joined as a new graduate in a Data Scientist position after completing my Master Degree in Environmental Science majoring Environmental Economics and Energy Policy at Sophia University Graduate School. I was dispatched from SERAKU Co.,Ltd. to customers to develop and report on Data Science.

Key responsibilities

- Machine Learning and Empirical Analysis:
- LightGBM, Random Forest, Neural Network Regression, SARIMA Model, and Unit Root Test with Python3/R in Jupyter Notebook;
- Sometimes, I create required SQL on MySQL/BigQuery/Redshift SQL;
- In reporting, I adopted PowerPoint/Excel/Word depending on demand from the clients.

Key achievements/projects

- My source code and reports as deliverables on Data Science Development for the clients. / An each project
- My business at SERAKU Co.,Ltd. was SES (System Engineering Service);
So, I could not know my key achievements but the clients only would be able to check it;
- However, IT industry in Japan and Korea adapt this SES;
- I did not like the system of SES and remove SERAKU Co.,Ltd. to step up my career as a Data Scientist such as PM/PL.

Education & qualifications

- [Kaggle 3 Experts](#) – [Kaggle](#) - March 2021
 - [Code Highest Rank: Top < 0.2% \(317/161898\)](#) - Kaggle Code - March 2021
 - [Code Solo 3 Silver + 13 Bronze Medals](#) - Kaggle Code - March 2021
 - [Discussion Highest Rank: Top 0.3% \(574/188433\)](#) - Kaggle Discussion - March 2021
 - [Discussion 103 Bronze Medals](#) - Kaggle Discussion - March 2021
 - [Datasets Highlight Rank: Top 10% \(354/3463\)](#) - Kaggle Datasets - March 2021
 - [Datasets Solo 3 Bronze Medals](#) - Kaggle Datasets - March 2021
- [Master of Arts in Environmental Studies](#) with [Total GPA3.74/4.00 \(93.5%\)](#) - Sophia University Graduate School - March 2017
- [Bachelor of Economics](#) - Hosei University - March 2015
- [Statistical Data Analyst](#) - The Institute of Statistical Science - February 2017
- [RPA Developer Foundation](#) - UiPath Academy - April 2019

References available on request

Kaggle Abstract:

Competitions	Top	Teams	Notebooks
Predicting Molecular Properties	21%	553/1636	
ALASKA2 Image Steganalysis	25%	272/1095	
M5 Forecasting - Uncertainty	26%	237/912	
TReNDS Neuroimaging	32%	336/1051	
OpenVaccine: COVID-19 mRNA Vaccine Degradation Prediction	33%	532/1636	1 Bronze Medal
Riiid Answer Correctness Prediction	34%	1133/3406	3 Bronze Medals
Jane Street Market Prediction			3 Silver + 6 Bronze Medals
Lyft Motion Prediction for Autonomous Vehicles			1 Bronze Medal
HuBMAP - Hacking the Kidney			1 Bronze Medal
Rainforest Connection Species Audio Detection			1 Bronze Medal

Kaggle Awarded Code:

- 3 Silver Medals:
 - [“LightGBM Classifier & Logistic Regression + Report”](#)
 - [“Optimize LightGBM HyperParameter with Optuna & GPU”](#)
 - [“Optimized LightGBM with Optuna adding SAKT Model”](#)
- 13 Bronze Medals:
 - [“RFCX: Residual Network with TPU Customized”](#)
 - [“Optimized Logit LightGBM Classifier & CNN Model”](#)
 - [“Predict Optimized Logit LightGBM Classifier & CNN”](#)
 - [“Riiid!: Optimized LightGBM with Optuna”](#)
 - [“LightGBM with the Inference & Empirical analysis”](#)
 - [“Optimize CatBoost HyperParameter with Optuna & GPU”](#)
 - [“LGBM on Lyft Tabular Data \[Inference\] + Tuning”](#)
 - [“Visualization Model with Customized Sharp Weighted”](#)
 - [“Research with Customized Sharp Weighted”](#)
 - [“Visualization & Empirical analysis of Submission”](#)
 - [“Introduction of LightGBM Tuner with Optuna”](#)
 - [“COVID-19 with H2OAutoML Baseline”](#)
 - [“【日本語】HuBMAP - Exploratory Data Analysis”](#)