

Matt Motoki

Kaggle Grandmaster · AI/ML Classification & Entity Matching

Remote | mattmotoki.github.io | [Kaggle](#) | [GitHub](#) | [Google Scholar](#)

Senior data scientist and machine-learning specialist with over a decade of experience building accurate, well-evaluated classification and prediction systems. Kaggle Competitions Grandmaster (peak global rank 18) with 13 national and international awards and **\$350,000+ in prize money**, including a prize-winning solution to the U.S. Department of Justice's **NIJ Recidivism Forecasting Challenge** — individual-level probabilistic classification judged on accuracy, calibration, and fairness. Strengths include **LLM-based classification and extraction**, rigorous evaluation on labeled data, probability calibration and confidence thresholds, and entity matching against reference databases.

EXPERIENCE

Shifted Energy — Staff Data Scientist | Honolulu, HI / Remote · Aug 2021 – Jun 2025

Led production machine-learning and optimization systems deployed across thousands of connected devices worldwide. Translated competition-grade modeling methods into reliable, monitored pipelines that retrained and scored on live data.

H2O.ai — Senior Data Scientist | Mountain View, CA / Remote · Jun 2020 – Aug 2021

Built algorithms and product features for **enterprise AutoML software** — automated model selection, tuning, validation, and NLP feature engineering — and partnered with customer-facing data scientists on Fortune 500 classification and prediction use cases.

Iterable — Senior Data Scientist | San Francisco, CA · Jun 2019 – May 2020

Developed machine-learning methods for **terabyte-scale customer engagement prediction** at a cross-channel marketing platform — modeling individual user behavior from large-scale event, message, and content-interaction data.

Independent — Data Science Consultant | Honolulu, HI · Jul 2017 – Jun 2019

Built predictive models, analyses, and dashboards for a remote-sensing startup, academic researchers, and local startups.

HiVet Solutions — Data Scientist | Honolulu, HI · Mar 2016 – May 2017

Developed data-driven methods for workforce planning, hiring, and compensation analysis with HR and Finance teams.

CLASSIFICATION & PREDICTION — SELECTED COMPETITION RESULTS

- **NASA Airport Throughput Prediction Challenge**, 2025 — **1st Place · \$30,000** — short-horizon prediction from large historical operations data
- **Net Load Forecasting Prize** (U.S. DOE), 2023 — **2nd Place · \$150,000** — large-scale prediction under high uncertainty
- **Digitizing Utilities Prize** (U.S. DOE), 2023 — **Phase 1 Winner · \$75,000** — analytics application turning messy operational data into decisions
- **NIJ Recidivism Forecasting Challenge** (U.S. DOJ), 2021 — **Prize Winner · \$48,500** — individual-level probabilistic classification, balancing accuracy, calibration, and demographic fairness
- **ASHRAE Great Energy Predictor III**, 2019 — **1st Place · \$10,000** — large-scale tabular prediction on Kaggle (3,600+ teams)

Also: eight further awards in forecasting and optimization, including the American-Made Solar Forecasting Prize (Runner-up · \$25,000) and two NeurIPS CityLearn Challenges (1st and 2nd). \$350,000+ total across 13 prediction, forecasting, and optimization awards (2017–2026).

EDUCATION

- **Doctoral Research, Electrical & Computer Engineering** — University of Washington · 2022 – 2025 — machine learning, game theory & optimization for large-scale prediction
- **M.S., Electrical Engineering** — University of Hawai'i at Mānoa · Dec 2015
- **B.S., Mathematics** — University of Hawai'i at Mānoa · Aug 2014

CORE SKILLS

LLM-based classification & extraction (prompting, few-shot, evals) · Text embeddings & entity matching / resolution · Gradient boosting & tabular ML · Class imbalance & probability calibration · Evaluation design & honest validation on labeled data · Confidence thresholds & human-review routing · Feature engineering · Model selection, tuning & ensembling · Reproducible ML pipelines · Python · PyTorch