Roy Ma

Machine Learning Engineer

[**My Codes**] · [**roy.ma9@gmail.com**] · [**Kaggle**] · [**Linkedin**] · Manassas, VA 20110

Experienced Machine Learning Engineer with 4+ years of expertise in designing and optimizing ML/DL solutions. Possesses a strong background in finance and statistics and ranks top 0.8% on Kaggle. Demonstrated ability to create and deploy scalable, real-world solutions for complex business models. A native bilingual speaker of English and Mandarin, consistently working effectively with cross-functional teams.

Technical Projects & Competitions

Competitions Expert, Kaggle rank top 0.8%

LLMs – You Can't Please Them All (Silver Medal, Ranked Top 3%)

- Collaborated with ML experts to develop adversarial strategies that improved the robustness and security of Large Language Models (LLMs).
- Engineered template-, heuristic-, and optimization-based injection attacks using iterative Bayesian chaining to infer LLM behaviors without explicit parameter access.
- **Technologies:** `Pytorch`, `Transformers`, `scikit-learn`, `XGBoost`, `Numpy`, `Pandas`

Jane Street Real-Time Market Data Forecasting (Bronze Medal, Ranked Top 6%)

- Designed and trained an RNN architecture with online fine-tuning to enhance predictive accuracy using real-time market data.
- Developed robust data and feature engineering pipelines that reduced noise and improved evaluation metrics.
- Deployed cloud-based training workflows with quality assurance and hyperparameter optimization.
- **Technologies:** `Pytorch`, `Tensorflow`, `XGBoost`, `CatBoost`, `scikit-learn`, `Polars`, `Numpy`, `Pandas`, `SQL`, `AWS EC2`, `AWS S3`, `Docker`, `Git`

Santa 2024 - The Perplexity Permutation Puzzle (Bronze Medal, Ranked Top 5%)

- Developed simulated annealing and evolutionary algorithms to solve large-scale black-box optimization problems with a search space of 9e157.
- Trained a deep neural network to evaluate candidates, reducing computational resource requirements for LLM forward passes by over 75%.
- **Technologies used:** `Pytorch`, `scikit-learn`, `Simulated Annealing`, `Genetic Algorithm`, `Numpy`, `Pandas`

Micro-Masters in Statistics and Data Science, MITx
Feb 2024 to Feb 2025

- Successfully completed 4 MIT graduate-level courses covering probability, statistics, machine learning, and computational data analysis with 93% in overall grades. Passed 4 practored exams covering each course.
- Built transferable ML engineering skills, including data/feature engineering, model development and deployment, testing, and QA.

Professional Experience

Machine Learning Engineer, SSR Tax & Legal
Mar 2021 to Jan 2025, Beijing, China

- Developed an interactive client-intake AI system incorporating intent classification, semantic analysis, and retrieval-augmented generation (RAG), cutting manual document triage time by 55%.
- Refactored document processing pipelines using state-of-the-art NLP/LLMs and distributed computing to reduce inference latency by 20%.
- Built ML-driven risk detection workflows that cut document review time by 35%.

- Collaborated in an Agile, cross-functional team with developers, legal, and finance experts.

Business English Instructor, Freelance
Dec 2019 to Present, Beijing, China

- Delivered over 800 hour-long live sessions and managed 59 tailored student programs, including courses for university professors, executives, and legal professionals.
- Independently developed and authored curricula during a challenging period, showcasing adaptability and cross-domain expertise.

Director of Business Development, Cleveland International Fund
Nov 2017 to Oct 2019, Cleveland, OH, U.S.

- Led a team to close six private placement offerings, raising over $370M for commercial real estate developments in the U.S. Midwest.
- Managed relationships with 300+ FINRA-accredited investors, negotiating strategic partnerships and securing new business opportunities.

Head of International, Multibank Exchange Group
Mar 2016 to Sept 2017, Beijing, China

- Supervised a cross-functional team of 30+ professionals, overseeing brokerage operations, hedge fund products, and market research.
- Exceeded strategic targets and optimized international business operations.

Skills

Machine Learning: Deep Learning , Artificial Intelligence (AI) , Neural Networks , Applied Machine Learning , Predictive Modeling , Statistical Modeling , Model Development/Selection , Pattern Recognition , TensorFlow , PyTorch , LangChain
Data Engineering: Data Science , Statistics , Feature Engineering , Data Visualization , Tableau , Pandas , Polars , NumPy
Domains: Natural Language Processing (NLP) , Large Language Models (LLM) , Prompt Engineering , Conversational AI , Text Mining , Retrieval-Augmented Generation (RAG) , Image Processing , 3D Rendering/Reconstruction , Generative AI , Transformers , Diffusion , Image Generation , Fine Tuning
SDE Fundamentals: Python , Javascript , Algorithms , Data Structures , Computer Science , Problem Solving , Full-stack Web Development , Troubleshooting , Maintenance
MLOps: Amazon Web Services (AWS) , Kubernetes , Cloud Computing , Distributed Training , High Performance Computing (HPC) , Distributed Computing

Certificates and Languages

English - Native, **Mandarin** - Native, **Japanese** - Business Conversational
FINRA, Series 63 - Uniform Securities State Law; **FINRA**, Series 82 - Private Securities Offerings Representative

Education

Micro-Master's, Statistics and Data Science, MITx
Bachelor of Economics, University of Maryland