Ren Zhang

Massachusetts (617)893-5778
ryanzjlib@gmail.com

Experience

Google - Search Intelligence Core Modeling — Cambridge, MA
Senior Software Engineer — *Jun. 2022 - Present*

† **Search Relevance and Ranking**
Developing fine-tuned Gemini LLM models in JAX to optimize matching between user queries and web search results.
Leading model development initiatives to enhance search result relevance using state-of-the-art large language models.

† **Search Query Understanding**
Built and deployed BERT and RNN models in TensorFlow to generate signals that help search ranking better understand user intent.
Designed end-to-end experimentation and serving infrastructure for production ML systems serving Google-scale traffic at tight latency requirement.

NielsenIQ - BASES — Needham, MA
Lead Data Scientist — *Apr. 2019 - Jun. 2022*

† **End-to-End ML Solutions for Consumer Research**
Led technical initiatives across forecasting, optimization, and deep learning capabilities for product innovation decision-making.
Built and deployed gradient boosting and RNN forecasting models, managed backend services for discrete choice optimization platform, led cloud migration from legacy systems, and established PyTorch training programs for data scientists and engineers.

† **Forecasting Platform Development**
Designed and implemented automated ML pipeline for business forecasting solutions.
Built model architecture, ETL pipelines, automated hyperparameter tuning, and deployment systems for gradient boosting and RNN models addressing various forecasting challenges.

IBM - Watson Health — Cambridge, MA
Data Scientist — *Apr. 2017 - Mar. 2019*

† **Watson Health Platform Development**
Built end-to-end data solutions across multiple Watson Health products including Explorys EHR platform, cloud infrastructure, and genomics research tools.
Designed Avro schemas and ETL pipelines for longitudinal health data, developed React dashboards for cloud application monitoring, and implemented ML models for genomic mutation classification with enhanced entity extraction using graph traversal algorithms.

Skills

Expertise: Large-Scale Machine Learning, Deep Learning, MLOps, Distributed Systems, Statistical Analysis
Languages: Python, C++, SQL, JavaScript
ML/AI Frameworks: JAX, TensorFlow, PyTorch, Transformers, Optuna, Scikit-learn
Infrastructure: Cloud Platform, Kubernetes, Docker, BigQuery, Apache Beam
Data Analytics: Pandas, NumPy, PySpark, Matplotlib, Seaborn, SHAP

Education

Bentley University — Waltham, MA
Master of Science in Business Analytics / Data Science - GPA: 4.00 — *Sept. 2015 – Dec. 2016*
Focus: Machine Learning, Statistical Analysis, Optimization

Publication — Open Source

Denoise AutoEncoder for Tabular Data: A python package that preprocesses arbitrary tabular formatted data and train autoencoder networks in denoise setup to learn a transformation that results robust representation for supervised or unsupervised downstream tasks.

Babaian T., Zhang R., Lucas W. DTMi – a New Interface for Informed Navigation, in proceedings HCI International 2017.