Ren Zhang

https://ryanzhang.info

Massachusetts 617-893-5778
ryanzjlib@gmail.com

Experience

Goolge - Search — Cambridge, MA
Software Engineer — *Jun. 2022 - Present*

† **Search Query Understanding**: Better understanding of the intent behind every user query.
Member of the search ranking team. Machine Learning for query expansion.

NielsenIQ - BASES — Needham, MA
Lead Data Scientist — *Jul. 2021 - Jun. 2022*

† **Technical Lead on Deep Learning**:
Training new data scientists and software engineers on building, training and serving deep learning models with PyTorch.

† **Line and Price Optimizer**: Discrete choice based solution for product innovation decision making.
Managed the Backend service for BASES LPO solution. Lead the team paying down technical debt and migrate this legacy solution to cloud.

Senior Data Scientist — *Apr. 2019 - Jul. 2021*

† **Forecasting models**: A suite of Gradient Boosting Tree models and RNN models trained to solve various forecasting problems the team is facing.
Designed overall model architecture. Implemented ETL pipeline, automated model tuning and deployment.

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

† **Explorys**: Big data platform for storing longitudinal EHR data with analytic tools on top.
Designed Avro data schema for new record types, implemented the ETL pipeline.
Created a React web app for displaying and editing Avro data schemas.

† **Watson Health Cloud**: HIPPA and GXP compliant cloud solution.
Worked on developing an internal dashboard aggregates and displays the status of applications running on the cloud.

† **Watson for Genomics**: Research tool uses NLP and knowledge base to link clinical trials to the genomic sequence.
Implemented graph traversal strategies on knowledge graph to enhance entity extraction results.
Trained recurrent neural network models to classify documents into genetic mutation types.

Projects — Softwares

Denoise AutoEncoder for Tabular Data: A python package that prepossesses 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.

DCASE 2019 Freesound Audio Tagging Challenge: 19th place solution using convolutional neural networks to tag audio clips. Models were trained with SpecAugment and Mixup to enhance generalizability and combat label noise.

Publication

Zhou L., Zhang R., Chakraborty P., Farooq F., Hensley Alford S., Comparative Effectiveness of Edoxaban and Warfarin in Prevention of Stroke and Systemic Embolism in Non-valvular Atrial Fibrillation Using Observational Healthcare Data, in ISPOR 2018.

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

Education

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

Coursework: Programming, Databases, Hadoop, Time Series, Optimization, Statistical Learning, Data Mining

Skills

Experties: Deep Learning, Machine Learning, Statistical Analysis, Optimization, Software Development
Languages: C++, Python, R, Javascript, SQL
Libraries / Frameworks: JAX, Pytorch, Tensorflow, Sklearn, Optuna, Shap, Pandas, PySpark, Numpy, Scipy, Django, Streamlit, Matplotlib, Seabor
Systems—Tools: Linux, Shell Script, Git, Docker, Jenkins, Hadoop, Nginx, Vim, Tmux