HARRISON ZHAO

Data Scientist | Bayer | Cornell

@ zzhharrison@gmail.com 949-463-6067 in www.linkedin.com/in/harrison-zihao-zhao web: harrisonzzh.github.io

EXPERIENCE

Data Scientist

Bayer

June 2018 – present St Louis, MO

- Built a two-stage (prediction and optimization) model for plant field-testing allocation.
  - Developed a recurrent neural network (Bi-LSTM) with PyTorch to predict allocation preference. Used the predicted preference to guide optimization model to the optimal solution and improved performance by 5x.
  - Built a large-scale (160k decision variables) mixed integer optimization model with CPLEX to optimize allocations according to business rules and field-testing objectives.
  - Delivered allocations of 1.2 million plots across North America. Improved key metrics on maturity matching, environmental diversity by 12%. Automated the entire application through Domino and AWS.
- Implemented a prediction model that uses genetic data to predict plant phenotypes (eg. yield, plant height).
  - Built an ensemble model of gradient boosting decision tree and ridge regression. Tuned hyperparameters through AWS Sagemaker.
  - Achieved 0.3 – 0.5 correlations on the major phenotypes, which beats currently production model by 3%.

Lead Optimization Scientist

OptiFunder

Dec 2018 – July 2019 St Louis, MO

- Designed a smart loan funding model from scratch that provides optimal strategies to warehouse lenders.
  - Worked with the founder to understand loan business. Built an optimization model that helped clients reduce funding cost by 30% - 40%.

Graduate Research Assistant

Cornell Tech

Jan 2017 – May 2017 New York, NY

- Worked with three graduate students to enable natural language navigation for simulated drone.
  - Built a recurrent neural network with attention mechanism for language embedding. Enabled the system to better capture key words in the natural language queries. (report link)

PROJECTS

Simulated Self-Driving Car

- Built an autopilot program to drive in a simulator. Implemented a convolution neural network with TensorFlow. Trained with 40,000 driving images on AWS EC2 to predict steering angle. (demo)

Plant Disease Detection IOS App

- Developed an IOS app that can predict plant disease by taking pictures of plant leaves. Trained a convolution neural network model with 20k images across three different plants. Integrated the trained model into an IOS app that can predict plant disease in real time. (demo)

EDUCATION

M.Eng. in Operations Research

Cornell University

May 2018 New York, NY

B.S. in Mathematics

University of California, Irvine

March 2017 Irvine, CA

COURSEWORK

Applied Machine Learning
Optimization Methods
Computer Vision
Natural Language Processing
Linear Algebra
Statistics

SKILLS

Programming

Python SQL R Matlab

Machine Learning

Scikit-Learn LightGBM XGBoost PyTorch Keras TensorFlow CoreML

Implementation

Domino Data Lab AWS Sagemaker AWS S3 AWS EC2 Docker

Optimization

CPLEX Gurobi CBC

OTHER PROJECTS

- Airline Pricing
- Image Search Engine
- MNIST
- Sentiment Analysis
- House Prices