Yu Xiang

Bellevue, WA 98004

📞 518-833-4710 ✉ yux1991@gmail.com 🌐 yuxiang.info 💼 linkedin.com/in/yu-xiang-physics 🐙 github.com/yux1991

EDUCATION

Doctor of Philosophy in Physics, Rensselaer Polytechnic Institute — Aug. 2014 – Dec. 2019
GPA: *3.88* of *4.00* | Relevant Courses: *Quantum Mechanics, Statistical Mechanics* — Troy, NY

Bachelor of Science in Physics, Wuhan University — Sep. 2010 – Jun. 2014
GPA: *3.43* of *4.00* | Relevant Courses: *Advanced Mathematics, Linear Algebra, Probability* — Wuhan, China

SKILLS

Programming: Python, SQL, C/C++, Perl
Modeling: Machine Learning (Regression, Decision Trees, Clustering), Time Series (ARIMA, LSTM), Deep Learning (CNN, RNN, Transformer), Natural Language Processing (word2vec, seq2seq, BERT)
Packages: Scikit-learn, TensorFlow, PyTorch, Pandas, XGBoost, Qt, CUDA, NumPy, SciPy, Matplotlib
Techniques: AWS, Hadoop/Spark, Git, Docker

EXPERIENCE

Data Scientist, Amazon | Bellevue, WA — **May 2021 – Present**

- Built machine learning models to forecast critical quantities such as shuttle transactions and warehouse transfer arrivals for sales and operations planning at Amazon, saving at least several million dollars per year.
- Developed the Chatbot that translates human language into SQL queries to pull tables from the database, and make forecast for the target feature by learning from the historical data.
- Created a one-stop solution to the long-term forecasting problem with highly sparse inputs by designing a novel multivariate encoder-decoder attention model with two-head outputs.
- Setup the data pipeline and deployed the models on native Amazon web service (NAWS) to retrain and make predictions on any given cadence, while continuously monitoring the input data using a model drift detector.

Seismic Imager, CGG | Houston, TX — **Jan. 2020 –Feb. 2021**

- Provided optimal quality control for the input seismic raw data by analyzing the statistics from billions of seismic records using SQL and mapping out key metrics using Hadoop/Spark big-data tools.
- Processed petabytes of seismic data by detecting anomalies using sparse transformation, removing seismic echoes using adaptive subtraction, and imputing missing features with compressed sensing techniques.
- Helped the client save almost $30MM per production well by providing unbiased interpretation of the geological features using machine learning techniques including SVM and boosted trees.

PROJECTS

Open-Source Software Development for Electron Diffraction Image Processing — Jul. 2018 – Dec. 2019

- Designed the object-oriented architecture, implemented the software with Python (Github repository: PyRHEED), and provided technical support to users from several research groups across the world.
- Improved the average computational efficiency of the back-end processing module by about 100 times through NumPy vectorization and CUDA parallel computing.
- Built the data pipeline from raw reflection high energy electron diffraction (RHEED) images to processing-ready datasets through autonomous labeling, noise reduction, standardization and featurization.

RHEED Data Analysis With Machine Learning — May 2016 – Dec. 2019

- Extracted the three-dimensional probability density distribution of the diffracted electron waves from the preprocessed RHEED datasets by learning the parameters of a Gaussian mixture model.
- Simulated thousands of crystal domains based on the Voronoi tessellation using Monte Carlo methods, in order to be combined with the experimentally extracted features for model parameter estimation.
- Estimated the unknown statistics such as the lattice constant, grain size, and preferred orientations from the RHEED images with a Bayesian regression approach.

AWARDS

The Karen & Lester Gerhardt Prize in Science and Engineering at Rensselaer Polytechnic Institute — **May 2020**
Paul S. Ho '65 Prize in Physics at Rensselaer Polytechnic Institute — **May 2019**
Hillard B. Huntington Award (1976) at Rensselaer Polytechnic Institute — **May 2017**
Presidential Graduate Research Fellowship Award at Rensselaer Polytechnic Institute — **Oct. 2015**