Yu Zhao

Senior Data Scientist at Veritone, Inc.
Orange County, California Area

Summary

Background:
- Solid background in Math, Statistics and Computer Science, with
fluency in analytic tools (Python, SQL, Matlab).
- 5 years academic experience in developing and applying advanced
statistical modeling techniques to solve estimation and prediction
problems.
- 5 years industrial experience in applying machine learning/deep
learning approaches for product design and development (speech
transcription, object detection and tracking, etc).

Skills and specialties:
- Programming Languages: Python (Tensorflow, Keras, PyTorch,
scikit-learn), shell scripts, C#, Java, C/C++, VBA
- Scientific Software: R, SQL, Matlab/Simulink, Tableau
- Machine Learning: Linear/Logistic Regression, Ridge, LASSO,
Neural Networks, SVM, Naïve Bayesian, LDA/QDA, Decision
Tree, PCA, Nearest-Neighborhood, Ensemble, Clustering, NLP,
Collaborative Filtering
- Deep Learning: face recognition, object detection, object tracking,
speech recognition
- Statistics: Hypothesis Testing, Maximum Likelihood Estimation
- Time Series Analysis: ARIMA, State Space Model, Causality
Analysis
- Big Data Techniques: MapReduce, Hadoop, PIG/HIVE, NoSQL

———

Experience

Veritone, Inc.
Senior Data Scientist
June 2018 - Present (1 year 10 months)
Orange County, California Area

AmeriHome Mortgage Company, LLC
Data Scientist/Financial Engineer
March 2017 - June 2018 (1 year 4 months)
Woodland Hills, CA

WD, a Western Digital company
Staff Engineer
May 2013 - February 2017 (3 years 10 months)
Irvine, CA

• Obtained drive manufacturing/processing data via Hadoop based product data warehouse, performed pre-processing.
• Built linear regression model to quantify the relationship between recording density and magnetic signal quality of disk drive systems, including feature engineering and model training/scoring.
• Improved 2% product yield by optimizing recording format based on recording quality prediction.
• 2 Invention Disclosure Award, 1 Trade Secret Award.

University of Southern California
Research Assistant
August 2008 - May 2013 (4 years 10 months)

• Developed a data-horizon robust system identification algorithm that improved parameter estimation accuracy and variance in data insufficiency scenarios.
• Resolved input-error collinearity in closed-loop data using banks of time series models.
• Proposed data-driven disturbance modeling approach to boost the prediction performance of industrial model predictive controllers and achieved the improved prediction accuracy on the Tennessee Eastman Process.
• Implemented Principle Component Analysis (PCA) based anomaly detection technique to monitor the wafer etching process and detected abnormal events and faults.
• Side projects: Dynamic PCA based anomaly detection, subspace identification based anomaly detection.
• 10 journal/conference publications with more than 110 worldwide citations.

WD, a Western Digital company
Intern
May 2012 - August 2012 (4 months)
San Jose

• Modeled dual-stage mechatronics system via statistical modeling techniques using complex mechanical data. This modeling module has been integrated into WD product design platform for 2.5/3.5 inch disk drive products.

Education

University of Southern California
Ph.D, Electrical Engineering · (2008 - 2014)

University of Southern California
Master of Science (M.S.), Financial Engineering · (2011 - 2013)

Harbin Institute of Technology
Bachelor's degree, Automation · (2004 - 2008)