Jun Yu

Redmond, WA || (541) 678-8666 || zariable@gmail.com || linkedin.com/in/zariable

QUALIFICATION

A proven track record of leading teams of scientists and engineers to build end-to-end machine learning (ML) products to solve real world problems and deliver business value.

- ❏ Tech lead working in a wide range of ML-driven applications, including recommendation system and ranking, causal model, forecasting, advertising, fraud and risk management.
- ❏ CS Ph.D. majored in Machine Learning research with 10+ years experience in developing ML algorithms and applying ML techniques to solve large scale real world problems.

EXPERIENCE

Amazon - Senior Applied Scientist and Tech Lead October 2017 – Now

- ❏ [Recommendation] Lead the team in building the recommendation engine to rank and personalize content on Amazon Sell Central Homepage and other push delivery channels (emails, push notifications and SMS) in order to maximize user engagement.
- ❏ [Downstream Impact] build Amazon's first seller downstream impact causal model to estimate the incremental long-term economic impact of a single seller-initiated action. Outcomes are used in budget allocation, content recommendation and A/B test reporting.
- ❏ [Forecasting] build ML models to forecast products' sales velocity in realtime to source high-quality deals and optimize the deal scheduling for Amazon Prime Day 2018 & 2019.

University of Washington - Adjunct Professor January 2019 – Now

I have a passion for teaching and helping more people into the field of machine learning and data science. I currently teach several graduate-level courses in the Foster School of Business, including Advanced Machine Learning, Deep Learning, and Natural Language Processing.

eBay - Staff Applied Science Tech Lead December 2013 – October 2017

- ❏ [Marketing] Led a team of applied researchers to power eBay's paid internet marketing by improving the bidding strategies on Google and Facebook to maximize ROI. Our models improved ROI by a total of 30+%, leading to tens of millions more GMB annually.
- ❏ [Risk Management] Built a large-scale machine learning model to predict seller risk and reduce the number of defective transactions on eBay via search ranking demotion, which translates to tens of millions GMB lift annually.
- ❏ [Search] Implemented and published a topic model based approach to retrieve diverse items based on user buying intents and improved the user satisfaction by 6+%.
- ❏ Award: eBay Critical Talent Award & eBay Seattle Technical Achievement Award.

Oregon State University - Ph.D. Researcher September 2006 – December 2013

- ❏ Developed probabilistic graphical models to predict species distribution with citizen science data and significantly improved the accuracy of state-of-the-art species distribution models.
- ❏ Developed a generative mixture model in the eBird human/computer learning network to quantify the skill level of citizen scientists in the eBird project.

EDUCATION

Ph.D. in Computer Science September 2006 – December 2013
Oregon State University, OR GPA: 3.98
Thesis: *Machine Learning For Improving The Quality of Citizen Science Data.*

B.S. in Computer Science September 2002 – June 2006
Wuhan University, China

SKILL SETS

Amazon Web Service for building ML products using various AWS products.
Python and R for building machine learning models, data analytics and prototyping.
Scala and Java for building and deploying end-to-end application in production.
Spark and Scalding for building data pipelines and larges-scale ML models.

PUBLICATIONS

[1] *Bayesian and empirical Bayesian forests*. ICML. 2015.

[2] *Can observation skills of citizen scientists be estimated using species accumulation curves?* PLOS-ONE, 2015.

[3] *The eBird enterprise: an integrated approach to the development and application of citizen science*. Biological Conservation, 2014.

[4] *Modeling misidentification of bird species by citizen scientists*. AAAI. 2014.

[5] *HC-Search for Multi-Label Prediction: An Empirical Study*. AAAI. 2014.

[6] *Clustering species accumulation curves to identify groups of citizen scientists with similar skill levels*. IAAI. 2014.

[7] *Latent Dirichlet Allocation based diversified retrieval for e-commerce search*. WSDM. 2014.

[8] *Clustering species accumulation curves to identify groups of citizen scientists with similar skill levels*. NIPS workshop. 2013.

[9] *Modeling misidentification of bird species by citizen scientists*. NIPS workshop, 2013.

[10] *Automated data verification in a large-scale citizen science project: a case study*. eScience, 2013.

[11] *eBird: a human/computer learning network for biodiversity conservation and research*. AI magazine, 2013.

[12] *Crowdsourcing citizen science data quality with a human-computer learning network*. NIPS workshop. 2013.

[13] *eBird: a human/computer learning network for biodiversity conservation and research*. IAAI. 2012.

[14] *The implementation of automated data verification processes in a large-scale citizen science project*. eScience workshop. 2012.

[15] *Multi-label classification for species distribution modeling*. ICML workshop. 2011.

[16] *Modeling experts and novices in citizen science data for species distribution modeling*. ICDM. 2010.

[17] *Learning algorithms for link prediction based on chance-constraints*. ECML. 2010.

[18] *Chance-constrained programs for link prediction*. NIPS workshop. 2009.