Jun Yu

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

- ❏ Engineering leader with strong ownership, proven people and team management, problem solving and analytics skills, and extensive software development experience.
- ❏ Single-threaded owner of business critical products leading teams of client, backend and machine learning engineers and managers across multiple Geo locations, and in highly cross-functional settings.
- ❏ Domain expertise in Machine Learning (CS Ph.D. with 20+ publications) and 10+ YoE in industry building large-scale recommendation systems, forecasting, fraud detection, and risk management.

EXPERIENCE

Snap inc. - Senior Engineering Manager October 2019 – Now

I am the single-threaded owner of a business critical area "*Friending*" at Snap and led teams of mobile, backend, and ML engineers and managers across multiple geo locations. We build, iterate, and own products, such as Quick Add, notification/sms, spam prevention, that help users make virtuous friends on Snapchat so that they can better experience our value proposition. Our innovations (e.g. EBR, GNN) helped boost topline Friending OKRs by 2-3x since my tenure here.

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.

Amazon - Senior Applied Scientist and Tech Lead October 2017 – October 2019

[Forecasting] built 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.
[Recommendation] Led the team in building the recommendation system to personalize and rank content on Amazon Seller Central Homepage and other push delivery channels (emails, push notifications and SMS) in order to maximize user engagement.
[Downstream Impact] built 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.

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 ML model to predict seller risk and reduce defective transactions in search result pages, leading 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.

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

SKILLS

AWS and Google Cloud for building softwares and products using various cloud products.
Scala and Java for building and deploying end-to-end applications in production.
Spark for building data pipelines and large-scale ML models.
Python for building machine learning models, data analytics and prototyping.

PUBLICATIONS

[1] *Embedding Based Retrieval in Friend Recommendation.* SIGIR. 2023.

[2] *Graph Neural Networks for Friend Ranking in Large-scale Social Platforms*. WWW. 2021.

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

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

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

[6] *A latent variable model for discovering bird species commonly misidentified by citizen scientists*. AAAI. 2014.

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

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

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

[10] *Machine Learning For Improving The Quality of Citizen Science Data*. PhD Dissertation. 2013.

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

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

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

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

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

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

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

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

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

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

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