

Wonja Fairbrother

[Oakland, CA] [603.498.9504] [github.com/wonja] [<https://www.linkedin.com/in/wonja>] [wfairbrother@gmail.com]

Skills

Languages: Clojure, Go, Python, C/C++
Tools: Git, Kubernetes, Docker, Google Cloud Platform, AWS, MySQL, Protobuf, gRPC

Education

Tufts University · B.S. in Cognitive & Brain Sciences, Computer Science (2011)
GPA 3.7/4.0 · Magna Cum Laude

Experience

Mixpanel · Machine Learning Engineer / Tech Lead (Apr 2015 – Feb 2020)

- Served as Technical Lead of the Machine Learning team, responsible for developing, deploying, maintaining, and optimizing ML models to production, data pipeline design, stability & reliability, code quality and refactoring, technical designs, managing on-call rotations, and onboarding and mentoring junior engineers. Worked closely with product managers and designers to address user needs, providing technical context when necessary
- Led Technical Design for an internal ML platform enabling feature selection, model parameter specifications, and model performance comparison, and unifying several different pipelines that backed our ML products
- Productionized Mixpanel's first ML product, Predict, a tool which groups a customer's users by their likelihood to perform a behavior in the future by building a logistic regression model with limited-memory BFGS optimization on billions of event history data points using a distributed stratified sampling method
- Created an internal query endpoint for feature generation with enough flexibility to continue supporting Predict as well as propensity matching for a new feature, Causal Impact Analysis. Refactored the Predict C/C++ code to utilize internal data primitives, improving reliability and enabling cache reuse
- Built out an anomaly detection feature which sends mobile and web alerts to customers when anomalies are found in their time series data. First implemented with an ensemble model comprised of a Holt-Winters ETS forecaster and two modified versions of Netflix's Robust Anomaly Detection (RAD) with Robust Principal Component Analysis (RPCA), then later migrated to a more accurate and more flexible fleet-of-forecasters method using ARIMA, Seasonal ARIMA, ETS, moving average, and lagged predictors
- Managed the Machine Learning team for a year, responsible for fostering career development, delivering feedback to team members, and growing the team as hiring manager. During this time the team shipped many improvements and features including:
 - Interesting Segments - a feature that identifies subgroups of users whose funnel conversion or retention pattern differs from the overall population, and sends findings as alerts via the web UI
 - Anomaly Explains - an add-on to anomaly detection which finds specific users or user groups driving an anomaly, helping customers drill down on possible causes and fix issues faster
 - Smart Alert Config - allows customers to subscribe to ML alerts for the metrics they care about, reducing noise as well as unnecessary load on the anomaly and interesting segment pipelines

Moovweb · Software Engineer (Feb 2013 – Apr 2015)

- Created functionality to convert Google's XDM-style Gumbo parse tree to a custom DOM-like node class to enable developing a custom selector engine
- Implemented an interface between a Go service and the libxml C library, allowing for runtime selection of a dispatch table using specified libxml versions and improving modularity between system components
- Developed the backend daemon and worker pool queuing system of an application which detects and captures changes in upstream HTML and CSS of client websites

Autonomy HP · Technology Specialist (Aug 2012 – Feb 2013)

- Set up and maintained distributed networks of virtual machines for demonstration purposes
- Wrote scripts to ingest data from files into internal databases and to display the data effectively

UC Davis MIND Institute · Technical Research Assistant (Jul 2011 – Aug 2012)

- Improved, updated, annotated, and added several core features to the lab's MATLAB codebase for preprocessing and statistical analysis of fMRI data, including the ability to run modern whole-brain and functional connectivity analyses on outdated data sets and export them in a format compatible with contemporary statistical software
- Developed a user-thresholded Fourier filter to MRI scanner head motion signal data to detect and eliminate events surrounding motion spikes from data to be used in analysis; integrated script as a preprocessing step for all brain scan data in the lab