Camen Piho

Data Scientist

Education

Brown University, *BA Classics and Mathematics (2015)*

Experience

Consultant *(1 year)*

University of Texas at Austin — *Researcher* *(June 2020 - Present)*

- Principal data scientist in a collaboration with UT Austin chimp researcher, Dr. Aaron Sandel
- Driving the development of a white paper by building Bayesian causal inference analyses to explore cortisol differences in wild chimpanzees

Reflective Earth — *Software Developer* *(May 2020 - August 2020)*

- Designed and built a NodeJS web app providing an interactive map tool to guide user insights into how albedo affects global warming
- Implemented a Flask backend using Xarray to quickly explore a dataset created from 30 years of ERA5 reanalysis data from the ECMWF
- Managed automatic deployment into Heroku-hosted backend and client's Webflow frontend

NASA — *Researcher* *(November 2019 - June 2020)*

- Improved wildfire detection algorithms using over 60TB of GOES-R series satellite data and proof of concept CNN models for transfer learning
- Authored an open-source python library for accessing and manipulating publicly available satellite data quickly and at-scale
- Used advanced parallelization and distribution techniques to perform model learning, inference, and evaluation on NASA's Pleiades Supercompute Cluster

Freebird *(3 years)* *acq. by Capital One in 2020*

Data Scientist *(January 2020 - March 2020)*
Jr. Data Scientist *(August 2018 - January 2020)*

- Developed well-tested Python libraries for recurring data analyses and risk assessments in order to provide reproducibility, maintainability, and stability of Freebird's risk estimates
- Performed Bayesian statistical testing to quantify the impact of proposed product changes and presented results to the CTO in order to inform product strategy
- Explored how Freebird's risk portfolio would change upon expansion to the global market
- Built Bayesian hierarchical models to predict product usage in order to increase the accuracy of our risk forecasts and deliver insights to the product and client teams to drive product strategy
- Developed Python, Scala, and SQL pipelines to ingest and process a wide variety of data sources to power Freebird's traveler notifications and risk modeling

Data Analyst *(July 2017 - August 2018)*

- Analyzed expected travel risk for Freebird's client pipeline to power pricing decisions
- Designed, authored, and maintained a web scraping data pipeline for ingesting monthly flight statuses from the *Bureau of Transportation Statistics*

Projects/Publications

Author, *"Prediction and Uncertainty Quantification of Daily Airport Flight Delays,"* with Thomas Vandal, Max Livingston, and Sam Zimmerman, Proceedings of Machine Learning Research (82), pp. 45-51. 2017.

Maintainer, *Bambi*. A high-level Bayesian model-building interface to fit mixed-effects models, built on top of the PyMC3 probabilistic programming framework.

Author, *Personal Website*. Built using VueJS and Flask to host projects, autobiographical information, and a nascent blog on science and wonder.

LOCATION

Chicago, IL

CONTACT INFO

camen.piho.r@gmail.com

PROFILES

github.com/camenpihor

camenpiho.com

linkedin.com/in/camen-piho

TECH

Python (Django, Flask, NumPy, Pandas)

SQL (PostgreSQL, Presto)

Scala

Javascript (VueJS, Node)

SYSTEMS

CI (GitHub Actions, Codeship, TravisCI)

DevOps (AWS, Docker, Papertrail, Git)

HPC (PBS, dask, mpi4py, ray)

CONCEPTS

ML (PyMC3, SciPy, Scikit-Learn)

Predictive Modeling (Bayes, MCMC)

Data Visualization (Jupyter, Matplotlib)