Camen Piho

camenpihor@gmail.com
github.com/camenpihor

Employment

Freebird

Data Scientist (Jan 2020 - March 2020)
Jr. Data Scientist (Aug 2018 - Jan 2020)
Data Analyst (Jul 2017 - Aug 2018)

- Built a new model to predict Freebird product usage across groups of travelers using Bayesian hierarchical modeling; developed it into a Python library for training, prediction, and evaluation.
- Developed Python and Scala services to ingest and process a wide variety of data sources, including weather forecasts and statuses of every flight in the world, to power Freebird's traveler notifications.
- Translated Jupyter notebooks for recurring data analyses into a well-tested Python library to provide reproducibility, maintainability, and stability for our risk estimates of a large contract with a major credit card issuer.
- Introduced pylint, black, and pydocstyle as part of the CI test suite across Python packages to improve code quality, readability, and standardization.
- Used Bayesian statistical testing to inform business decisions around the impact of a product change on customer rebooking behavior.

NASA

Research Consultant (Nov 2019 - Present)

- Wrote and launched highly parallel and distributed jobs on NASA's supercomputers to process over 60 TB of geospatial satellite images in order to detect wildfires over much of the Earth.
- Developed and maintained an open-source Python library to aid wildfire data exploration, and easily introduce other researchers to NOAA's publicly available satellite data and wildfire modelling (github.com/joyprojects/wildfire).
- Implemented a threshold model for wildfire detection to label wildfires for our data pipeline.

Publications

"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. Fall 2017.

Education

Brown University, Providence, RI
BA in Greek and Roman Classics (class of 2015)

Skills

Languages	Python, SQL, Scala
DevOps	AWS, Docker, Codeship, TravisCI, Papertrail, Git
Tools	PyMC3, SciPy, NumPy, Pandas, Matplotlib, HPC, multiprocessing, Django, Flask