Aaron Deich

San Francisco, CA

linkedin.com/in/aaron-deich | github.com/adeich | adeich.github.io

SKILLS

Programming Languages: Python, SQL, Bash, C, Mathematica
Toolkits: Numpy, Scipy, Pandas, Scikit-Learn, Matplotlib, Seaborn, Jupyter, TensorFlow, Git, Conda
Software Engineering: ETL Pipelines, Deployment Environments, OO Data Structures & Algorithms
Data Modeling: PCA, CNN, Fourier Analysis, Causal Inference, Regression, Signal Processing
Geospatial Data: Simulation & Forecasting; Trajectory Embedding & Classification

EXPERIENCE

Data Science Fellow, *Insight,* San Francisco   May 2020 - Present
- Consulted for geospatial intelligence company, Pathr, where I established a **trajectory embedding** that classifies human movement by behavior, achieving 98% accuracy on proof-of-concept test data.
- Created a **processing pipeline** in Pandas that runs each trajectory through customizable stacks of 10+ transforms (e.g. **curvature, FFT, Kalman filter**), scored by **clustering** quality metrics.
- Achieved clear visual sorting of the 300 test trajectories via the embedding + **PCA**.

Consultant, *Lick Observatory,* CA   Dec 2018 - Mar 2019
- Analyzed 5 years (20GB) of data from 16 sensor telemetry channels (thermometers, windspeed, motor positions, etc) to discover the **root cause** of high stellar measurement error for an $8M robotic planet-hunting telescope.
- Trained a **CNN** in **TensorFlow** to predict the error from the telescope's 16-sensor state-vector and achieved a 13x mean reduction in live error for historical validation data.
- Discovered, using **correlation analysis** combined with **cleaning, regression, and PCA** in Python/Jupyter/Pandas, that the error was due to the thermal control system—the observatory has since reduced the telescope's error by 3x, to theoretical minimum levels.

Physics & Calculus Teacher, *Jewish Community High School,* San Francisco CA   Aug 2018 - Jun 2019
- Taught AP Physics (calculus-based Mechanics) and AP Calculus to 55 students.
- Designed lessons around **experimental design** and discrete, programming-based calculus methods.

Software Engineer, *Markit on Demand,* Boulder CO   Feb 2012 - May 2013
- Built and maintained 40+ custom **ETL pipelines** in **Python/SQL** for ingesting and storing json & XML data from financial clients' FTP servers, adding 4+ pipelines every month.
- Designed scheduled jobs in Python for high-frequency file transfer and transformation; SQL error-checking to maintain our production environment; and logic to manage redundancy across 3 company data centers.

EDUCATION

San Jose State University   2018
MS, Physics—Applied Python-based ML to classify galaxies.

Reed College   2011
BA, Physics—Thesis: Developed an MC orbital simulation of planets around exploding stars.