Aaron Deich

San Francisco, CA | 831.421.2907 | deichaaron@gmail.com
linkedin.com/in/aaron-deich | adeich.github.io | github.com/adeich

SKILLS

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

EXPERIENCE

Data Science Fellow, *Insight Data Science,* San Francisco May 2020 - Present

- Consulting for retail computer-vision analytics startup Pathr.ai, I designed a movement-analysis pipeline that classified hundreds of human 2D paths by behavior ('needs assistance'; 'just browsing'; etc).
- Optimized dimensionality-reduction via a pipeline in Pandas, numpy, and scikit-learn that embedded movement via customizable stacks of 15+ transforms (e.g. curvature, spatial frequencies, acceleration, Kalman filter), obtaining 95% accuracy on proof-of-concept human motion data.
- Achieved first useful behavior classification on company datasets.

Data Science Consultant, *Lick Observatory,* CA Dec 2018 - May 2019

- Located primary causes of an $8M robotic telescope's high error rate via analysis of 5 years (20GB) of data from 16 sensor telemetry channels (thermometers, windspeed, motor positions, etc).
- Discovered, using multivariate correlation analysis combined with cleaning, regression, and PCA in Python/Jupyter/Pandas, that the error was due to the thermal control system—my work has since guided the observatory to reduce the telescope's error by 3x, to theoretical minimum levels.
- Predicted the telescope's real-time error by training a CNN in TensorFlow over historical telemetry data, achieving a 13x mean error-accounting reduction for starlight measurements.

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 - Aug 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 big astronomy data to classify galaxies.

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