Preston Hinkle

Seattle, WA | 614-208-4506 | tphinkle@gmail.com | tphinkle | tphinkle

Skills

Languages	Python · C++ · SQL
Tools	Python scientific computing stack (pandas, sklearn, etc.) · Flask · Git · LaTeX · Qt
Data Science & Engineering	Statistics · Modeling · Computer vision · Data visualization · GUI + web app development

Experience

Data Science Fellow *Jan. 2018 –*

Insight Data Science *Seattle, WA*

- Built *shelfie*, an interactive web app that automatically prices and catalogues book collections based off images shot on cell phone.
- Images are uploaded to the Google Cloud Vision API for annotation and post-processed using a custom image processing pipeline in order to group words by spine.
- Book tokens are matched to real book candidates using Amazon API calls and the match is compared against a corpus comprised of 50 M titles, authors, and publishers contained in a PostgreSQL database.
- Used techniques from natural language processing such as tokenization, fuzzy string comparison, and inverse document frequency weighting to validate the book identifications.
- Involved using **Python**, **Flask**, **Javascript**, **SQL**, **web scraping**, **cross-validation**, and **computer vision** and **natural language processing** techniques.
  App repository: shelfie repo
  App website: shelfie app

Experimental Biophysicist *Sep. 2012 – Sep. 2017*

University of California, Irvine *Irvine, CA*

- Devised and ran experiments on custom fabricated nano- and microfluidic devices with potential applications for biotechnology; designed a microfluidic apparatus capable of measuring the stiffness of human cell lines, an important biomarker for cell malignancy.
- Awarded a fellowship from the UCI Data Science Initiative for writing an open-source Python software package to analyze imaging data from an ultra-high speed camera capable of generating $\sim$250 GB/experiment.
- Wrote software which enabled remote operation of instruments in the lab, including a C++ GUI program which allowed asynchronous control of an ultra-high speed camera (communication via TCP/IP), a high-speed picoammeter (NIDAQmx library), and a remotely controlled pump (RS-232 serial commands). The software enabled simultaneous recording of independent measurement devices, and allowed for the first time simultaneous electrical and optical recording of cell passages through microfluidic channels with high time resolution.
- Involved **experimental design**, **software development and data analysis in Python**, and **C++**, **computer vision**, **time-series analysis**, and required **effective communication**, both in paper authorship and in talks delivered at conferences.
  Software repositories: pore_stats · cell_controller · nanoIV

Data Science Workshop Instructor *Sep. 2016 – Nov. 2017*

UCI Data Science Initiative *Irvine, CA*

- Devised curriculum for and instructed quarterly graduate-level machine learning workshops in Python.
- Responsible for teaching subjects including **generalized linear models**, **scientific computing in Python**, **machine learning with scikit-learn**, **data visualization**, and **model cross-validation**.
- Taught 50 students Python and machine learning in a two day workshop as part of a program ran jointly by NASA and California State University, Los Angeles. Independently devised the curriculum and taught the workshops.
  Workshop repositories: Predictive Modeling in Python · Introduction to Python for Data Analysis

Education

Ph.D., Physics *Sep. 2017*

University of California, Irvine *Irvine, CA*

B.S., Physics and B.S., Astronomy *Jun. 2011*

The Ohio State University *Columbus, OH*