Paul **Hein**

Data Scientist · Software Engineer

(+1) 520-289-9886 | paul.d.hein@gmail.com | pauldhein.github.io | pauldhein | pauldhein | @pauldhein

Summary

Experienced Research Programmer with a demonstrated history of working in the higher education industry. Skilled in Python (Programming Language), Data Science, Machine Learning, Software Development, and Continuous Integration and Continuous Delivery (CI/CD). Strong engineering professional with a Master of Science - MS focused in Computer Science from University of Arizona.

Education

University of Arizona — *Tucson, AZ*

M.S. in Computer Science — *2017 – 2019*

B.S. in Computer Science and Mathematics — *2013 – 2017*

Experience

Research Programmer — *University of Arizona*

ML4AI Laboratory — *May 2019 - Present*

- Built a pipeline to produce a computation graph from a program's abstract syntax tree
- Designed and implemented automated sensitivity analysis methods for the purpose of model selection among multiple competing computation graphs of the same real-world phenomena
- Assembled a parallel corpus of Python functions with their respective PEP standard docstrings from common Python libraries
- Constructed a classification model capable of determining whether a given docstring matched a given source code function
- Currently creating a stylistic domain adaptation tool for encoder/decoder methods

Graduate Research Assistant — *University of Arizona*

ML4AI Laboratory — *May 2017 - May 2019*

- Developed an experimental framework to properly test the effectiveness of various machine learning classifiers with varying linguistic feature sets; classifiers trained to identify biological context and associate the context with biochemical reactions
- Researched and tuned several machine learning classifiers including Logistic regression, SVMs (with various kernels), random forests, deep neural networks, and a gradient boosted decision tree classifiers.
- Developed a parallelized experimentation suite to allow our experiments to be run on the U of A's high-performance computing (HPC) cluster
- Developed visualizations of our classification results to aid in the explanation of our findings in our published manuscript

Undergraduate Research Assistant — *University of Arizona*

ML4AI Laboratory — *Aug 2013 - May 2017*

- Created time-series fluent representations of different features found in musical data
- Implemented Allen Interval Algebra to analyze relations between features in time-series data
- Constructed a Long-Short Term Recurrent Neural Network (LSTM) to generate time series data
- Created a geometric pattern finding algorithm tuned for detecting repeated patterns in music

Student Programmer — *University of Arizona*

Lunar Planetary Laboratory — *Apr 2015 - Jun 2016*

- Constructed a web-based data product organization tool to track a pedigree of all science products being created and revised during the length of the mission
- Gained experience working with large multi-user databases, distributed computing resources and fundamental concepts in graph theory

Skills

Programming Languages

Advanced	Python, Java, JavaScript/NodeJS, HTML/CSS, LaTeX
Intermediate	C/C++, Ruby, Scala, MATLAB, and MySQL
Familiar	Common Lisp, Haskell, Prolog, Verilog

Software

Systems	MacOS, Windows, and Debian Linux
Packages	NumPy/SciPy, Pandas, PyTorch, and scikit-learn
Etc.	Docker/Singularity, Git/SVN, and MS Office Suite

Awards

2017 **Full Year Academic Distinction**, University of Arizona

2013 **Wildcat Excellence Scholarship**, University of Arizona

2013 **AP Scholar with Distinction**, College Board

2009 **Eagle Scout**, Boy Scouts of America