https://github.com/probinso

Personal Statement

I have 7+ years experience, solving increasingly difficult problems with software. My diverse work history spans both products and research, including data pipelines/management, medical informatics, programming languages, and computer securities, as well as audio, image, and text processing. I have developed strong interests in data analysis/exploration, information retrieval, machine learning, and data engineering. I look forward to understanding and helping to solve your domain specific problems.

Persistent Experience :

C, C++ Python Linux TDD

R, dplyr ET_EX Git. Ticketing Systems

Work Experience _

- Marine Acoustics Engineer Graduate Research with ACO

(continuing)

Researched deep unsupervised anomaly detectors, for sane indexing of 10 year audio track Developed ACOio, an intuitive ACO track explorer, to rejuvenate data access for researchers Developed principled audio processing and noise-suppression, focused on marine bioacoustics

Technologies: Python, ACOio, tensorflow, scipy.signal, Jupyter

- Data Science Technical Expert Jan. 2019 - present Thinkful (Part Time) Advised career transitioners in mastery of professional and data science topics Provided safe environment for 1:1 training and instruction in a remote-first setting

- Graduate Data Science Intern June 2018 - Sept. 2018 NASA Jet Propulsion Lab Designed & developed Expert Modeling/Recommender System, by extending the Author-Topic-Model Developed and specified program from use-case meetings with top NASA/JPL employees Implemented stable/principled text normalization, tokenization, and model evaluation Open Source contributions to the gensim's natural language processing library

Technologies: Python, nltk, gensim, pyLDAvis, pandas, Jupyter

OHSU Fair Neuroimaging Lab - Research Assistant

Oct. 2017 - June 2018

Contributed to microbiome population analytics tools, to study female reproductive system Supported research on developing brains, including ABCD, and several ADHD/ASD studies Developed processing pipeline and audit tools for reported data and FMRI images

Technologies: Python, Bash, R, neo4j, ponyorm, stan, GitLab, Docker

Contractor/Consultant

RGB Optics / C&W Energy USA (Part Time)

Nov. 2015 - Aug. 2018

Developed custom photo image processing tools for low cost spectral profile analysis Authored educational material in light pollution's effect on economy, biology, and technology Provided live technology tutorials and consulting on optimization and image processing

June 2016 - Sept. 2016 ComScore

Worked to support large, custom, memory mapped, data store for demographic analysis

Melinae

Setup infrastructure in AWS to enable secure & sustainable remote-first workflow Provided hands on training in Python and R to industry professionals

Technologies: Python, numpy, skimage, R, AWS, PostgreSQL, Perl, C++, Jira, ZenHub

- Research Engineer Galois Inc. April 2014 - Dec 2015

Developed processing pipelines and workflows to enable evaluator work for DARPA programs Produced biannual quantitative and qualitative reports on for DARPA and language developers Participated in programs sharing new technologies to research and industry professionals

Technologies: Python, SLURM, Scala, Figaro, Chimpy, Docker, Jira, Basecamp

${\tt EMC}^2$ Isilon Storage - Software Development Engineer Dec. 2012 - July 2013

Brought to schedule a lagging anchor release feature in approximately 5 months Designed and developed password manager to support Data At Rest Encryption Wrote unit tests using libcheck to attain > 80% code coverage

Technologies : C, C++, Python, SQLite, Subversion, FreeBSD, OpenSSL

Neato Projects _

Multiple Myeloma Clinical Trials, custom named-entity boosted topic model numpy, Python Topic modeling and applications, a presentation for non-statisticians beamer, MTEX Workshop collaborative introduction to GitHub and slides markdown, bash, github.api Morphological Watersheding Algorithms pyspark, numpy, ndimage, Python Relevance Vector Machine Information Retrieval Cluster/Rank Demo Harness flask, numpy, nltk, sklearn, Python Gene Data Breast Cancer Drug Predictor R, caret Distributed Fully Homomorphic Encryption System Hadoop, Sagemath Concurrent Elliptic Curve Cryptography Module Sagemath, Erlang Multilingual Analysis of Subordinating and Coordinating Conjunctions R, Perl AdaRailz Concurrent Model Train Control System Ada

Education

Oregon Health Science University

CSLU

Computer Science MSc 2016 - present

Expected Graduation: Sept. 2019

Courses :

* Advanced Topics in Machine Learning
Deep Learning
Machine Learning
Artificial Intelligence
Statistical Methods
Univariate Statistical Analysis
Analysis of Sequences

- \star Natural Language Processing
- * Advanced Topics in Signal Processing Speech Signal Processing Problem Solving with Large Clusters Image Processing Information Retrieval Computing Ethics

Western Washington University

Unix Software Development

Computer Science BS, Mathematics Minor June 2012

Computer Science

Courses :

Windows Software Development
Operating Systems
Computer Networks
Software Project Requirements Analysis
Software Project Design
Software Project Implementation
Analysis of Algorithms

Linear/Non-Linear Data Structures
Discrete Structures
Formal Languages/Automata
Programming Languages
Concurrent Programming
Computer Organization I/II
Probability and Statistical Inference
Object-Oriented Programming in C++

Electives :

Homomorphic Encryption Systems Cryptography & Elliptic Curves Artificial Intelligence Natural Language Processing Functional Programming Computer Graphics Computer Architecture
Number Theory
Elementary Real Analysis
Abstract Algebra
Linear Algebra I/II
Ordinary Differential Equations