Adam Wespiser, Resume, Fall 2016

adamwespiser [at] gmail [dot] com
(413)-281-4894

Objective:

Qualified data analyst and Haskell programmer with academic and start-up background seeking work developing business technologies. Six years developing, learning, and solving a diverse data problems while guiding my own multi-year research project. Experience co-founding a data-driven start up and extending machine learning based methodologies to complex business problems with messy data. Passionate self learner with a strong interest in data analysis software and previously employed building a programming language in Haskell for financial analysis.

Education:

University of Massachusetts Medical School, Ph.D. Program in Computational Biology, left after completing coursework, before degree.

University of Vermont. 2010. B.S. Degree in Biological Sciences.

Coursera.com courses in Database Design, Machine Learning, Algorithms, Statistics, Probabilistic Graphical Models

Work:

Writer, Software Engineering Communication *October 2016 - Present. Write You A Scheme, Version 2* Write You A Scheme, Version 2 is a public, open source book that is designed to propagate industry grade Haskell to a wide audience of software professionals. It does this by teaching the reader how to implement a programming language using Haskell best practices. Its main contribution to the field of engineering aims to be the demonstration that a subset of the Haskell language can be powerful and expressive, bringing research ideas into the production environment as simply as possible.

Algorithmic Consultant *June 2016 - July 2016. Althena Project* The Althena project is a technology group attempting to capitalize on emerging trends in distributed networking technologies to change the way consumers pay for internet. My role was to research issues including, but not limitted to, how algorithmic complexity would affect company growth, how other mesh networks solved these problems, and create engineering estimates of the effort required to overcome these obstacles.

Data Scientist *October 2015- May 2016. Elsen, co.* Originally hired to do data work, I was encouraged to learn Haskell and contribute to the development of an interpretered language written in Haskell. My role mostly consists of extending language features to include more statistical and data transformation features, as well as ensuring the accuracy and numerical stability of the system. Recent project include extending primitive language operators to work over timeseries data, creating aggregation functions over sets of timeseries, and building a symbolic algebra system for the dimensional analysis of units and currencies.

Co-Founder. *January 2015- September 2015. Ryzome Project.* Founded Ryzome to address price discrepancies, source obfuscation, and unreasonable mark up by finding customers the cheapest source of scientific resources. My role is to develop an algorithm capable of discerning groups of products across multiple vendors that are scientifically identical at extremely low false positive rate. Currently, we are both focusing on building automated tools for data collection while seeking funding to build a functioning protoype.

Rotation Student. *January 2013- December 2014, Weng Lab, University of Massachusetts Medical School.* Used machine learning, data visualization, statistics on projects with up

to terabytes data. Developed leadership skills through mentoring, and creation of a educational program. Focused free time on learning mathematical and statistical foundations of machine learning, and honed skill in R and software engineering techniques using Python.

Rotation Student. *August 2012 - January 2013. Zeldovich Lab, University of Massachusetts Medical School.* Designed, built and tested a 3d protein structure viewer with 'hands free' user interface using Microsoft Kinect. Presented the poster and live demo and 2012 NESS conference(New England Structural Symposium). Github Project Repo

Bioinformatician. *January 2011 - August 2012. Caffrey Lab, University of Massachusetts Medical school.*Prediction of protein-protein interactions, sequence features of RNA and genome annotation projects. Responsible for data management, analysis and visualization.

Work Projects:

Enhancer prediction: Given genomic regions, along with their quantitative epigenetic features, the objective is to find regions likely to be enhancers. I developed an platform to run multiple machine learning algorithms concurrently to find the best. Our results were compared to 8 other groups, and our test set performance had the highest correspondence to other groups. Github Project Repo

CTCF and splicing: Do some proteins near a gene influence how that gene is processed into RNA and thus a protein?. To answer this, I correlated metrics for protein binding and gene structure, then normalized by a false discovery rate. Github Project Repo

Analysis of LncRNA gene expression and prediction of functional lncRNA: Which lncRNA that are likely to have biological function? Are other lncRNA resulting from stochastic noise? During the project, I tried a multitude of approaches to train a classifier that could learn features of known functional lncRNA, including: PageRank to describe feature space, Principle Component Analysis for dimensional reduction, Information Theory statistics to describe exclusivity of expression, and modifications to logistic regression to account for the bias of unlabeled training set examples. Github Project Repo

Analysis of Large Dataset of RNA Sequencing Reads: I designed to better understand the quantitative properties of LncRNA using short read sequencing data, the raw data modern DNA sequencing. For one experiment of four total, up to 2 TB of data could be generated each run. Automating HPC jobs, file transfers, and data retention across systems solved my 'Big Data' problems without additional systems. Github Project Repo

Hobby Projects:

Write Yourself a Scheme in 48 Hours, version 2.0 I am working to update the classic Write Yourself a Scheme in 48 hours using modern Haskell idioms and providing the corresponding tutorial. For instance, replacing IORef with monad transformers. Still in the early stages, I am researching Scheme implementations and exploring the best pendantic abstractions capable of building a true Lisp-1 system. Github Project Repo

NFL Data Analysis by Play: A long time fan of football, my goal is to analyze whether football coaches are making rational decisions during the game. The dataset consists of play-by-play summaries over a 10 years, and required extensive cleaning. Github Project Repo

RTL-SDR signal processing with Raspberry Pi: A fun little project to make an RTL-SDR signal detector for collecting sources of radio broadcasts using a Raspberry Pi, GPS unit, and RTL-SDR(card for software defined radio). I have started writing the C code to decipher the radio input(IQ data), and am working on hardware integration issues with my raspberry pi. A pretty good excuse to learn about how information in encoded via radio. Github Project Repo

Professional Societies:

I am an activate member of the Boston Haskell Meetup group, and recently gave a talk on my experiences learning and using Haskell in a startup environment. YouTube Video.

Technical Skills:

Programming Haskell, Programming Language Design, SQL, relational tables, database design(BC-normal Form, 4NF, etc)..., Python, Perl, Bash, Ruby, C, C++, Java, Clojure, regular expressions, algorithms, data structures, git, vi/vim
Analysis machine learning, statistics and data analysis, web scraping, automated information extraction, R Programming Language, Rstudio, image recognition, probabilistic graphical models, Rcpp for integrating C++ code into R
Computing high performance computing with LSF scheduler, Google Compute Services, Linux server administration
Communication Data visualization with ggplot2, Edward Tufte Data Visualization Conference(2014), Dokuwiki, and Google Docs
Biology short read sequencing, RNA sequencing, genomics, epigenetics, functional discovery, lncRNA, protein interaction prediction

Awards:

Dean's Award for Outstanding Leadership, 2014 Umass Medical School. For creating an educational program designed to teach programming to students, post-docs and faculty on campus.Umass PR coverage More info on Bootstrappers

Eagle Scout Award, 2004

Volunteer:

IoT and Smart Cities Hackathon Taught a team of students how to use Arduino microcontrollers and answered questions about technology and business in Boston.

2016 Civic Tech Challenge Hackathon Helped High School students craft a website designed to increase community and police engagement.

Profiles:

LinkedIn: www.linkedin.com/in/adamwespiser
StackOverflow: http://stackoverflow.com/users/41932/wespisera
Twitter: @wespiser
Facebook: https://www.facebook.com/burlappsack
github: https://www.github.com/adamwespiser

Contact

Current Address: 43 Newbern Ave. Medford, Ma. 02155
Email: adamwespiser [at] gmail [dot] com
cell: (413)-281-4894
Date: November 2016