# ALEX HEILMAN, PH.D.    Data Scientist                    Seattle, WA

(859) 250–0470   |   AlexanderLHeilman@gmail.com   |   github.com/alexanderheilman   |   linkedin.com/in/alexheilman

## SUMMARY

Data Scientist with a Ph.D. in Chemical Engineering and a uniquely broad background spanning mathematics, programming, engineering, and design. A fast, avid learner with an aptitude for algorithmic thinking, a passion for tool-building, and an innate reluctance to step away from unsolved puzzles.

## TECHNICAL SKILLS

**Data Modeling**

- Linear / logistic regression
- NLP / Naïve Bayes
- Random Forest
- Gradient Boosting
- Clustering (DBSCAN, k-means)
- Dim. reduction (PCA, autoencoding)
- Recommender systems (SVD, NMF)
- Graph-based clustering / modeling

**Python**

- NumPy
- Scikit-Learn
- NLTK
- NetworkX
- PySpark
- Pandas
- Matplotlib
- PyMongo
- Selenium
- Flask

**Big Data**

- AWS
- Spark
- PostgreSQL
- MongoDB

**Statistical Methods**

- Frequentist, Bayesian, and probabilistic approaches
- Significance testing
- A / B testing
- Regression coef. analysis

**Web Skills**

- Scraping
- HTML

**Other**

- Git
- CRISP-DM

## DATA SCIENCE PROJECTS

**Online recipe aggregator**  |  Apr. 2019  |  github.com/alexanderheilman/recipe-vectorizer

Project goal: *Create a tool that simplifies and improves the online recipe search process by (i) collecting and comparing many recipes for a given dish and (ii) clustering and combining them into a few optimal recipes.*

- Automated scraping of online recipe data using Selenium-based Python scripts hosted on EC2 instances.
- Developed methodology for vectorizing recipes by parsing ingredient lists and standardizing quantities.
- Employed DBSCAN clustering and graph-based techniques (using NetworkX) to evaluate similarity of recipes and identify cluster centers representing the most authentic versions of several common dishes.

**Real-time fraud detection tool**  |  Mar. 2019

Project goal: *Create an interactive web app that uses a Random Forest Classifier to evaluate a live stream of event listings and allows internal fraud investigators to manually inspect potentially fraudulent listings.*

- Designed and built the front and back ends of the web app using HTML / Brython and Flask, respectively.
- Added interactive functionality that prompts user for input and updates the (Mongo) database accordingly.

## OTHER EXPERIENCE

**Exhibit Technician**  |  Pacific Science Center  |  Jan. 2018 – Jan. 2019

**Doctoral Student Researcher**  |  University of California, Santa Barbara  |  Jan. 2011 – May 2017

- Designed and constructed a custom microscopy system for nanoscale chemical interrogation of surfaces.
- Created software suite for instrument operation and automated data collection / analysis.
- Developed and published a methodology for scientific interpretation of novel experimental results.

## EDUCATION

**Certificate, Data Science**          Galvanize Data Science Immersive  |  Seattle, WA  |  2019
**Certificate, Machine Learning**    Instructor: Andrew Ng, Stanford University (via coursera.org)  |  2017
**Ph.D., Chemical Engineering**      University of California, Santa Barbara  |  Santa Barbara, CA  |  2017
**B.S., Chemical Engineering**        University of Kentucky  |  Lexington, KY  |  2010