

EDUCATION

University of California, Berkeley
Masters of Information and Data Science
Focus: Natural Language Processing

Duke University School of Law
Juris Doctor – May 2012
Focus: Entrepreneurship & IP Law

University of California, Berkeley
Double Bachelor's Degree – May 2009
Focus: Law and Economics

SUMMARY

I am currently a data science graduate student and a licensed attorney. I have 2 years of professional experience in product development for a YC-backed legal tech startup, as a business immigration and human trafficking attorney working on petition strategy, and as an economic litigation analyst working on case strategy. My work at the junction of law, economics, and data analytics has motivated me to find broad applications of my unique skill set. My areas of interest:

- Combining data munging and statistical analysis skills to process and mine massive structured & unstructured datasets
- Applying machine learning algorithms to find novel patterns and derive business knowledge for practical application
- Translating findings into actionable insights through data visualizations and an understanding of business objectives

SKILLS AND TECHNIQUES

- Platform Experience: Spark, Hadoop, SQLite, MongoDB, Softlayer, Elasticsearch, and AWS
- Languages: Python (proficient), R (proficient), SQL (proficient), and CSS/HTML/JavaScript/D3 (prior experience)
- Analysis Tools: SAS, pandas (Python), numpy (Python), scikit-learn (Python), NLTK (Python), and Excel
- Visualization Tools: Tableau, pyplot (Python), ggplotlib (R), d3 (Javascript), dc (Javascript), and Powerpoint

PROJECTS & COMPETITIONS

Daily Fantasy Sports Lineup Optimizer

- Used a genetic algorithm with a 500 population pool and 10 generations to optimize multiple Draft Kings lineups
- Iteratively improved both lineup optimization and prediction-optimization through dry-run tests over 6 months
- Achieved a 55% in-the-money success rate after adjusting genetic algorithm parameters and prediction formula

NOAA Weather Prediction System

- Built a scalable weather prediction system based off 10TB of hourly climate data collected from 34,000 U.S. stations
- Wrote bash and Python scripts to parallelize the download and extraction of object storage files into a Hadoop cluster
- Wrote Spark data aggregation code for specific location queries to be fed into random forest classifiers for prediction
- Iteratively refined classifier parameters using error analysis and loss function plotting to improve prediction accuracy

San Francisco Crime Classification Kaggle Competition (Currently 3rd Place)

- Wrote Python code for feature vector creation, preprocessing, and classifier training using numpy and pandas libraries
- Tested and applied various learning algorithms and found optimal components for PCA dimensionality reduction
- Conducted error analysis on classifiers to identify common misclassification patterns and areas for improvement

Yelp Business Analytics Dashboard

- Wrote JavaScript code using d3, dc, crossfilter, and bootstrap to create a Yelp data dashboard for business owners
- Conducted usability and A/B tests with prospective users to optimize user interaction and satisfaction with the product

Enron Email Sentiment Analysis

- Distributed computing project using EMR, EC2, MapReduce, and NLTK to analyze a 210 GB Enron dataset
- Wrote Python scripts to preprocess raw email data and a MapReduce script to classify entire Enron email dataset

PROFESSIONAL

Teleborder | San Francisco, CA

June 2013 – October 2013

Product Development and Legal Data Research

- Led legal product analysis for a YC-backed startup developing intelligent tools to manage business immigration
- Extracted predictive information from government spreadsheets containing over 100,000+ entries and 10 subfields
- Designed and executed A/B testing for multiple website designs to measure user behavior patterns
- Compiled summary statistics and created data visualizations to demonstrate probability of success for various visas

Asian Pacific Islander Legal Outreach | San Francisco/Oakland, CA

September 2012 – April 2013

Immigration Law Bridge Fellow

- Awarded a six-month fellowship at a community-based, social justice public interest law group
- Prepared T visa applications for human trafficking victims seeking legal status for themselves and their families

OSKR, LLC | Emeryville, CA

May 2008 - September 2008

Economic Litigation Analyst

- Performed data analysis, financial modeling and research for antitrust and intellectual property litigation
- *Semiconductor Antitrust*: Pinpointed collusive behavior of direct competitors and estimated monetary damages
- *Credit Card Competitive Practices*: Proved damages claimed against our industry-leading client were negligible
- *Athlete Union Licensing*: Analyzed players' claims against NFL for unpaid royalties and unlicensed use of images