Aaron Mangum

SKILLS

COMPUTER
Python • pandas • numpy • scipy
scikit-learn • statsmodels • gensim
D3.js • matplotlib • NLTK • flask
BeautifulSoup • Selenium • mongoDB
SQL • C • Objective-C • ADA • PHP
Django • HTML • CSS • MapReduce
Hadoop • Hive • Spark • Git/Github
Bash • AWS

STATISTICS & MACHINE LEARNING
Hypothesis testing • Maximum
Likelihood Estimation • Bayesian
Inference • Linear Models • k-Nearest
Neighbors • Support Vector Machines
Random Forests • Decision Trees
Neural Networks • Naive Bayes
k-Means and other clustering algorithms
Topic Modeling

LANGUAGE
Japanese (Fluent)
Spanish and German (Basic)

EXPERIENCE

DATA SCIENTIST, CONSULTANT 2015-PRESENT

➡ Used named entity recognition and an SVM classifier to automatically tag a document for companies, agencies and geography.

➡ Wrote a twitter-bot that finds the sentence in a text corpus most similar to a user's tweets using a NoSQL database and doc2vec.

➡ Working with SumAll.org on various projects including predicting war crimes in the Syrian Civil War.

➡ Created an interactive map and chart in D3 based on Ashley Madison user data. **http://gifsandgraphs.com/Dear-Ashley-Madison/**

DATA SCIENTIST, METIS 2015

➡ Scraped website data and used text mining in order to show the effect of violent media consumption on violent crime using linear regression. Created a script to help BeautifulSoup spider entire websites offline, negating the need to worry about an IP address being blocked for a DoS attack.

➡ Brute-force analyzed the contributing factors for heart disease using multiple supervised machine learning models across differing hyper-parameters. Created a D3 visualization illustrating the chance in accuracy resulting from different model-feature combinations. **http://gifsandgraphs.com/Heart-Disease-D3/**

➡ Predicted the passage of federal legislation using natural language processing, logistic regression and clustering. Used unsupervised machine learning to train custom sentiment analysis.

➡ Used text processing and dimensionality reduction with word2vec to predict restaurant health code violations with Yelp data. Ensembled multiple models to increase accuracy. Model tested and performed well in a data science competition. **http://gifsandgraphs.com/Winning-Kaggle/**

ATTORNEY, CONSULTANT 2008-2014

Worked for various law firms, including Skadden Arps, Clifford Chance, Ropes & Gray, and Hogan Lovells. Prepared Japanese written materials for antitrust litigation with the Department of Justice.

TECHNOLOGY COORDINATOR, GW CENTER FOR THE STUDY OF GLOBALIZATION 2003-2004

Designed and built interactive organization website. Created internal web-based applications in PHP, including a content management system for the organization staff.

COMPUTER CONSULTANT, DOLSEY LTD 1999-2000

Created company website. Troubleshooted company logistics software written in an unfamiliar programming language.

EDUCATION

SYRACUSE UNIVERSITY SCHOOL OF LAW; SYRACUSE, NY	JURIS DOCTOR (NYS BAR ADMITTED)
THE GEORGE WASHINGTON UNIVERSITY; WASHINGTON, DC	INTERNATIONAL ECONOMICS
KANSAI UNIVERSITY; OSAKA, JAPAN	JAPAN STUDIES
THE GEORGE WASHINGTON UNIVERSITY; WASHINGTON, DC	COMPUTER ENGINEERING (BS - PARTIAL)