Alan Flint

alanflint.com | alan@flintagan.org | LinkedIn | 425-273-4884 | Github

Education

University of San Francisco San Francisco, CA

Master of Science in Data Science *Expected June 2020*

- Deep Learning Certificate
- Relevant Coursework: Machine Learning, A/B Testing, Product Analytics, Deep Learning, Distributed Computing (Spark), Data Structures & Algorithms, Time Series, Data Visualization (ggplot & plotly)

University of Maryland College Park, MD

Bachelor of Science in Mathematics and Economics *May 2019*

- University Honors College
- Relevant Coursework: Statistics & Probability, Linear Algebra, Numerical Analysis, Econometrics, R
- Studied abroad at the University of Sydney, Australia for one year

Work Experience

Trulia San Francisco, CA

Data Science Intern *November 2019 - Present*

- Built an internal power analysis tool in Python using Streamlit that calculates how long to run an A/B test for various in-house KPI's, providing key stakeholders a resource to efficiently and accurately run experiments
- Currently working on a machine learning propensity model to understand which user characteristics and website features are predictive in converting users to potential home buyers; using resampling techniques to alleviate the extreme class imbalance problem
- Deployed an ETL on AWS to automatically query for key business metrics and validate data sources using Presto and MySQL, saving hours of manual querying time

Projects

- Indigo Music Studio - Website
  - Developed a web application for users to generate new music with machine learning in collaboration with six other students; I focused on Google Analytics tracking and PyTests
- Machine Learning Algorithms
  - Coded my own implementations of Random Forest, Decision Trees, k-means, Regularized Regression, and Gradient Boosting algorithms from scratch in Python
- Article Recommender System
  - Constructed a system to recommend similar articles using word2vec centroids; deployed server on AWS to host Flask application
- Classified Damage to Buildings Hit by the Nepal Earthquake
  - Modeled damage level to buildings with Random Forest by engineering features from geographic and building structure variables, currently placed top 4% in the competition among 2300 participants

Skills and Technologies

- Python (scikit-learn, NumPy, pandas, PyTorch, spaCy, matplotlib, Plotly, Flask), SQL, Spark (PySpark, SparkSQL, SparkML), AWS (EC2, EMR, EBS), git, R (tidyverse, ggplot2)