Alan Flint

San Francisco, CA | 425-273-4884 | alan@flintagan.org
alanflint.com | linkedin.com/in/alan-flint/ | github.com/alan-flint

Education

University of San Francisco San Francisco, CA
Master of Science in Data Science *June 2020*
- Deep Learning Certificate
- Relevant Coursework: Python, Machine Learning, A/B Testing, Product Analytics, Deep Learning, Distributed Computing (Spark), Data Structures & Algorithms, Time Series, Data Visualization

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*
- Discovered the most influential website features for converting visitors to home buyers using an interpretable machine learning model and SHAP feature importance
- Built an internal power analysis tool in Python and 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
- 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

- Machine Learning Algorithms
  - Coded my own implementations of Random Forest, Decision Trees, k-means, Regularized Regression, and Gradient Boosting algorithms from scratch in Python
- Indigo Music Studio - indigomusicstudio.com
  - 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
- 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 - Code Repository
  - 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, matplotlib, Plotly, Flask, spaCy), SQL, Spark (PySpark, SparkSQL, SparkML), AWS (EC2, EMR, EBS), git, R (tidyverse, ggplot2)