Alan Flint

alan@flintagan.org | linkedin.com/in/alan-flint/ | 425-273-4884

Education

University of San Francisco San Francisco, CA
Master of Science in Data Science *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), Data Ethics, Python, SQL, pandas, PyTorch

University of Maryland College Park, MD
Bachelor of Science in Mathematics and Economics *May 2019*

- Honors College Citation
- Relevant Coursework: Statistics & Probability, Linear Algebra, Numerical Analysis, Econometrics, R

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

Predicted Short Term Outcomes for Critically Ill Patients
In-Class Kaggle Competition *January 2020 - February 2020*

- Predicted patients' heart rate and mean arterial pressure using categorical and time series data for each patient; increased model performance with a stacked ensemble of tree based models
- Finished 4th out of 42 pairs of other data science students in my program

Machine Learning Algorithms
Class Projects, University of San Francisco *October 2019 - February 2020*

- Coded my own implementations of Random Forest, Decision Trees, Naïve Bayes, k-means, and Regularized Regression algorithms from scratch in Python
- Compressed grayscale and color images using k-means++ by clustering the RGB values and re-assigning each pixel to the value of its centroid

Classified Damage to Buildings Hit by the Nepal Earthquake
Machine Learning Competition, drivendata.org *October - December 2019*

- Modeled the level of damage 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, SQL, Spark, scikit-learn, NumPy, pandas, PyTorch, AWS, git, R, ggplot2, matplotlib