Lance Fernando

LinkedIn: in/ljfernando | Email: Ljfernando@usfca.edu
Personal Site: ljfernando.github.io | Phone: 510-557-2623
Blog: medium.com/@jonaven96 | San Francisco, California

Experience

Eventbrite **Nov 2018 - Present**
Data Scientist Intern

- **Customer LTV model**: Built a multi-class random forest model to predict the lifetime value of our customers with 96% accuracy and a Micro-Avg. F1 score of 0.96.
- **Venue data enrichment**: Deduplicated over 5 million venue addresses using a HMM and fuzzy string similarity metrics resulting in venue data assets to be used in predictive models and other analyses. Implemented the search ranking methodology for our venue marketplace using these assets.
- **Customer insights**: Collaborated with the Product team to define metrics and performed statistical analyses that uncovered product features that are highly correlated to the success of our customers. Presented insights to the company and impacted the product roadmap.

ABC News|USF Data Institute **Sep 2018 - Nov 2018**
Graduate Research Scientist

- Created a real-time web-app for the Decision Desk at ABC News in election forecasting. First team to call multiple 2018 US midterm races by building a proprietary mathematical model.

CA Technologies **Aug 2017 - May 2018**
Data Visualization Research Intern

- Developed a rule-based anomaly detection dashboard that identified important features and the optimal way to visualize each feature.

USF Visualization and Graphics Lab **Jun 2017 - May 2018**
Research Assistant

- Researched the most effective aesthetic choices in building scatter plots. Conducted a MTurk study that emulated scatter plots created by five common visualization tools and asked users to perform various interactive tasks.

Class/Personal Projects

NewsPhi
A web-app that allows for contextual news consumption by providing a news feed with a dashboard conveying controversy of a given article's topic and impact of the author for over 100K articles. Utilized LDA for topic creation and assessed the variability of sentiment scores of articles within a topic to compute controversy.

Predicting Parking Turnover (github.com/Ljfernando/TurnoverPrediction)
Leveraged 92m parking meter readings to forecast the availability of parking throughout San Francisco. Preprocessed data and built a random forest model using PySpark and increased efficiency using an EMR cluster.

D3.js Visualization Portfolio (github.com/Ljfernando/DataVizPortfolio)
Compilation of interactive d3.js visualizations from a dashboard visualizing the world's ecological footprint to a chloropleth map showing expensive and less expensive regions of San Francisco Craigslist rent.

Markovian Blues (github.com/Ljfernando/MarkovBlues)
Artificially generated 12-bar blues melodies using a variable-order Markov model that captures relationships between successive notes and the global structure of the 12-bar blues songs fed into the model.

Education

M.S. in Data Science|University of San Francisco **Jun 2018 - Jun 2019**
Coursework: Advance ML|Deep Learning|Experimental Design|Distributed Computing|Time Series Analysis|Linear Regression|Bayesian Statistics

B.S in Data Science|University of San Francisco **Sep 2014 - May 2018**

Skills & Technologies

Languages: Python, R, Javascript
Database & Visualization Technologies: PostgreSQL, MySQL, Tableau, ggplot2, D3.js, matplotlib, plotly
Distributed Computing: Apache Spark (SparkML, SparkSQL), AWS (EC2, EMR, S3), Hive, Presto