Lance J. Fernando

Cell: 510-557-2626 | LinkedIn: in/ljfernando
Email: ljfernando@usfca.edu | Personal Site: ljfernando.github.io

Education

B.S. in Data Science | Concentration in Computational Analytics **Fall 2014 - expected Spring 2018**
University of San Francisco
Cumulative GPA: 3.81 | Major GPA: 3.66

Skills

- Data Analysis **(R/Python)**
- Machine Learning **(R)**
- Visualization **(R**, **D3.js**, **Tableau)**
- Querying **(SQL)**
- Web-Scraping **(R, Python, Java)**
- OOP **(Java)**

Certifications

Designing, Running, and Analyzing Experiments | UC San Diego, Coursera **Spring 2018**
https://www.coursera.org/account/accomplishments/certificate/KVEDKCDACQ68

Data Science Ethics | University of Michigan, Coursera **Spring 2018**
https://www.coursera.org/account/accomplishments/certificate/ND5A4UVRH5W4

Projects

Predicting Article Popularity w/ Ensemble Methods **Spring 2018**
https://ljfernando.github.io/PredictingArticlePopularity/ Group Project
Used decision trees, random forests and boosted trees to predict the popularity of *Mashable.com* articles. Hypothesis testing, PCA, parameter tuning and ROC assessment was used to aid models achieving an **AUC: 0.70**

Shiny Clustering App **Spring 2018**
https://ljfernando.shinyapps.io/Clustering_Pipeline/ Individual Project
Created an **RShiny** dashboard that visualizes the results of various clustering algorithms and performs consensus clustering. Contains functionality to import a dataset, tune algorithm parameters and export clustering assignments

Spam Detection Using Naive Bayes **Spring 2018**
https://ljfernando.github.io/SpamDetectionNaiveBayes/ Group Project
Implemented a Naive Bayes Classifier from scratch in **R** to detect spam emails based on their content. Extracted over 170 features from our corpus of raw emails and achieved a cross-validated **misclassification error rate of 5%**

One-Stop Shop Regression Function **Spring 2017**
https://ljfernando.github.io/Regressience/ Group Project
Programmed a reusable function that runs linear regression, shrinkage methods and regression trees in **R**. It then outputs cross-validated results with visualizations to assess each algorithm's success

Visualizing Ecological Footprint **Spring 2017**
https://ljfernando.github.io/project-Ljfernando/ Individual Project
Created an interactive dashboard that joins a mercator map with various other plots to express the proportion of impact each region has on our global footprint. Visualizations produced using **Javascript** and **D3.js**

Movie Recommender **Fall 2016**
https://ljfernando.github.io/MovieRecommendation/ Individual Project
Developed movie recommendations using Python that takes in a user's inputted movie ratings and outputs movies based on user-user collaborative filtering using 100k movie ratings and multiple distance metrics

Experience

Research Intern **August 2017 - Present**
CA Technologies (Santa Clara, CA)
- Program a backend visualization recommendation system in **Python, R** and **RShiny**
- Create interactive visualizations using **plotly.js** and **d3.js** with **react.js+redux** as well as **angular.js**
- Conduct literature review to aid in the development of the visualization recommendation system

Data Intern **December 2016 - Present**
The Climate Music Project http://www.theclimatemusicproject.org
- Analyzed and aggregated historic and future climate data using **R** and visualize data using **ggplot2**
- Scrape google scholar search results using **R**
- Assisting in developing an open-sourced climate-music tool

Undergraduate Research Assistant **June 2017 - Present**
Visualization and Graphics Lab http://vgl.cs.usfca.edu/
- Design user studies deployed on **MTurk** using **JS**, **Python** and **R** and provide analyses of results
- Present research poster at on-campus student research fair