Greg Faletto

Data Scientist · Ph.D. Student

| gregory.faletto@marshall.usc.edu | gregoryfaletto.com/ | github.com/gregfaletto | https://www.linkedin.com/in/gregfaletto/

Employment History (Summary)

[Undisclosed Tech Startup]

Los Angeles, CA

Research Consultant — *Oct. 2019 -*

- Developing an algorithm for a core business objective (can't disclose details under agreement).
- Working with adviser and communicating regularly with key startup data scientists.

ZipRecruiter

Santa Monica, CA

Data Analytics Research Intern — *Jul. 2017 - Jan. 2018*

- Developed a model to predict whether individual job seekers apply to a given job listing to infer preferences for forms of compensation.
- Chose observational study design and model form. Accessed data via SQL. Conceived outcome metric, constructed from raw data.
- Coded in R utilizing `glmnet` and other packages to preprocess data and train model. Created visualizations and tables with `ggplot2`. Created a more flexible (but less interpretable) random forest model withl `randomforest` and `caret` packages to compare fit.
- Wrote a paper analyzing results, describing possible applications and next steps for project.
- Worked mostly independently under light supervision from ZipRecruiter and Dr. Sanjog Misra (Prof. of Marketing, U. of Chicago).

Live Nation

Los Angeles, CA

Machine Learning Intern — *Feb. 2017 - May 2017*

- Working remotely, independently created and implemented an algorithm in R utilizing JSON files from the setlist.fm API to predict future concert set lists of bands based on past set list data. Code and brief white paper: https://github.com/gregfaletto/setlistpredictor.

Education

University of Southern California Marshall School of Business

Los Angeles, CA

Ph.D. in Data Sciences and Operations (Statistics Track) — *Aug. 2018 -*

- Advisor: Dr. Jacob Bien
- Current research projects: developing a novel feature selection method; solving applied business problem with Opendoor (see above).
- Coursework includes: Mathematics of Statistical Learning, Generalized Linear Models, Convex Optimization, Economic and Financial Time Series, Probability, Stochastic Processes, Mathematical Statistics, High Dimensional Statistics and Big Data Problems (machine learning survey course)

Washington University in St. Louis

St. Louis, MO

B.A. in Physics — *Aug. 2006 - Dec. 2009*

- Graduated Phi Beta Kappa and with College Honors in Arts and Sciences. Dean's List every semester.

Online Coursework (MOOCs)

Jan. 2016 - Dec. 2017

- **Coursera:** "Machine Learning" (Andrew Ng/Stanford), "Econometrics: Methods and Applications" (Erasmus University Rotterdam).
- **Stanford Lagunita:** "Statistical Learning" (Trevor Hastie and Rob Tibshirani), "Databases: SQL," "Databases: Relational Algebra."
- **Datacamp:** "Intro to Python for Data Science," "Data Visualization with ggplot2," "Intermediate R," "Introduction to R."

Personal Projects

Valuing a Digital Firm Using a Customer Retention Model — *Jul. 2018 - Jul. 2019*

- Used a methodology developed by scholars from Wharton to valuate Buffer, an online firm, using publicly disclosed data.
- Developed a model for customer retention and forecasted residual customer lifetime. More information at gregoryfaletto.com/blog.

Princeton University

Fragile Families Challenge — *Jun. 2017 - Aug. 2017*

- Cleaned data and created models to predict three different outcome variables in a training set with only 2,121 observations, nearly 13,000 covariates, and a great deal of missing data. Code and details: github.com/gregfaletto/fragilefamilies.

Skills

- **Languages:** R (proficient), SQL (proficient), Matlab (intermediate), Python (some experience), Java (some experience).
  Other software: Google Sheets (proficient), LaTeX (proficient), Excel (intermediate), Stata (some experience), Tableau (some experience), SPSS (some experience), Minitabs (some experience).