MATT SHADISH

San Francisco, CA · (925) 487-1921 · m.shadish@engineering.ucla.edu · GitHub: mshadish

EDUCATION

Master of Science, Analytics **Expected June 2015**

UNIVERSITY OF SAN FRANCISCO, GPA 3.97

Relevant coursework: Advanced Machine Learning, Linear Regression Analysis, Time Series Analysis, Data Acquisition

Bachelor of Science, Mechanical Engineering **December 2013**

UNIVERSITY OF CALIFORNIA, LOS ANGELES, GPA 3.75

Honors**:** Phi Beta Kappa, Tau Beta Pi

Certifications: Base Programming for SAS 9

WORK EXPERIENCE

General Electric, San Ramon, CA **November 2014 – Present**

Data Analytics Intern

Parts Sourcing Project

- Integrated elasticsearch with a custom front end to improve usability for users searching through GE's parts database
- Created a Python Flask app to prototype an "Export to CSV" feature for our elasticsearch instance
- Implementing and deploying fuzzy matching of GE part descriptions as a MapReduce task to run on Hadoop

Company – Contact Mapping Project

- Doubled match rate of company names using Levenshtein distance metrics and custom word weightings in Python

Engage3, Davis, CA **December 2013 – August 2014**

Data Scientist, Category Management Algorithms

- Managed, designed, and implemented Oracle PL/SQL algorithms to categorize 800M+ retail product records
- Improved runtime of daily Oracle PL/SQL batch processes by 90% by developing incremental features
- Performed regression analysis of product prices to model how regular and promotional prices compare among retailers
- Collaborated with other members of the Data Science team using TortoiseSVN, Git, and Assembla
- Ran daily reports on user retention rates using PostgreSQL and weekly reports on data quality via Apache Solr queries
- Designed a Python script that relates similar retail products using image comparison to compensate for incomplete data
- Leveraged knowledge of pricing algorithms and worked with EIS team to begin migration to Apache Cassandra

ATAC, Sunnyvale, CA **June 2012 – June 2013**

Aviation Data Analyst

- Participated in airport and airspace noise and emissions analysis studies using Excel, Macros, and in-house software

RELEVANT TECHNICAL SKILLS

- Programming Languages
  - Proficient: Python, SQL (Oracle PL/SQL, PostgreSQL, MySQL), R, SAS
  - Familiar: Bash, LaTeX, JavaScript, C++
- Project management/version control: Git, TortoiseSVN, Assembla
- Other relevant technologies: elasticsearch, MongoDB, Apache Solr

PROJECTS

Kaggle: Driver Telematics Analysis **January 2015 – Present**

- Extracting and creating features from driver positional data in R to establish driver "fingerprints"

Kaggle: Sentiment Analysis on Movie Reviews **October 2014 – January 2015**

- Applied word vectorization in conjunction with a Support Vector Machine to classify movie review sentiment (on a discrete scale from 0 to 4) in Python – scored in the top half of contestants

Eccentric Exercise Machine Prototype **June 2013 – September 2013**

Founder, Lead Engineer

- Researched, designed, and fabricated a prototype exercise machine specifically made to facilitate eccentric exercise
- Drafted and submitted a provisional utility patent application
- Led a team of 4 in hand-fabricating the prototype, training team members in metal cutting, drilling & tapping holes