Stacy Lee

stacy-lee.github.io
linkedin.com/in/lee-stacy
slee422@illinois.edu

Education

University of Illinois at Urbana-Champaign
M.S. in Statistics *Analytics Concentration, GPA: 3.65/4.00* — Expected: Dec 2018
B.S. in Civil and Environmental Engineering *Systems Concentration* — May 2017
Minor in Mathematical Statistics

Skills

Languages: *Python, R, SQL, SAS, C++, MATLAB* — Web Design: *HTML/CSS*
Python Libraries: *Pandas, Scikit-Learn, Numpy, Regex* — R Libraries: *Tidyverse*
Tools: *Amazon EC2, Apache Spark, UNIX/Bash, Hive, Tableau*

Experience

Ameren — Champaign, IL
Data Science Innovation Intern *Data Analytics Team* — Jan 2018 - Present

- Lead a 5-person team for the project in identifying individuals out of 1.4 million customers with the highest likelihood of enrolling in two energy savings programs administered by Elevate Energy
  - Clean customer demographic data with 140 features of mixed data types for ML
  - Perform feature engineering (correlation plots, frequencies, RFE, feature importance)
  - Implement machine learning methods (Random Forest, XGBoost) using top 20 features
  - Present and communicate weekly analytics updates with Elevate Energy
- Support innovative project to replace the traditional utility pole assessment procedure for a more passive system check
  - Write Python scripts to reformat data collected from accelerometers for efficient analytics
  - Brainstorm analytical methods to classify utility pole health based on physical movements

RailTEC, University of Illinois at Urbana-Champaign — Champaign, IL
Research Assistant *Train Safety Analytics Group* — Jan 2016 - May 2017

- Created a geospatial visual that predicted train derailment severity based on train speed
- Investigated the effect of train speed on estimated derailment severity on millions of trains
- Fixed SQL queries; tabulated data on hazardous materials release from derailed tank cars
- Applied regression methods on data regarding train casualties using R programming language
- Utilized imputation methods for missing values for statistical analysis

WSP — Newark, NJ
Safety and Security Intern *Transit Rail Operations Planning* — May 2017 - Aug 2017

- Created a simplified track signal diagram to convey complex data to government stakeholders
- Collaborated on presentations by building upon the data visual with senior consultants

Projects

Software Engineering CS 412: Intro to Data Warehousing and Data Mining, *Fall 2017*

- Implementation of the Random Forest Algorithm in Python from Scratch
  - Gini index attribute split; majority voting classification; runtime under 3 min.

Involvement

Civil Engineering Undergraduate Advisory Board — Champaign, IL
President — 2016 - 2017

- Established the first undergraduate advisory board in the civil engineering department
- Organized social events of ~50 attendees to increase student and professor interactions

Awards

Data Science: **Top 10** Team with Best MSE Score - Synchrony Financial Datathon 2018
Top 20% (Round 1); **Top 50%** (Round 2) - TEXATA 2017
Energy: **Finalist** Team - U.S. Dept. of Energy's Race-To-Zero Sustainable Design 2015
Entrepreneurial: **Semi-Finalist** Team - Cozad New Venture Competition 2014