PAIGE MCKENZIE

PaigeMcKenzie@utexas.edu • https://p-mckenzie.github.io/

EDUCATION

The University of Texas at Austin — Master of Science, Business Analytics — May 2018
Overall GPA: 3.71

The University of Texas at Austin — Bachelor of Science and Arts, Mathematics — May 2017
Certificate: Business Foundations Program (Highest Distinction)
Overall GPA: 3.96; summa cum laude; Phi Beta Kappa

EXPERIENCE

Cisco – *Software Engineer;* Research Triangle Park, NC — July 2018 – present
- Implement natural language processing methods to obtain customer insights

The University of Texas at Austin – *Workday Analyst Intern;* Austin, TX — May 2016 – May 2017
- Wrote Python program to check system data against user input, saving over 80 hours of validation on each data load
- Reviewed text of over 300 business process security policies using Python, discovering and correcting the 44% error rate
- Analyzed text of 42,000 security authorizations using Python, finding and eliminating 1,600 cases of inappropriate access
- Designed automated and executed manual end-to-end tests of business processes and integrations

PROJECTS

3M Accounts Receivable Capstone — Spring 2018
- Analyzed accounts receivable balances and invoices for all of 3M's US-based subsidiaries for 2013 through 2017
- Provided insights into customer-level changes in terms and made concrete recommendations for increasing AR turnover

Instacart Market Basket Analysis — Fall 2017
- Engineered 30 features for over 13 million user/product pairs to measure product affinity
- Implemented random forest, neural network, and XGBoost models to predict what set of products a user would re-order
- Achieved a 20% improvement in AUC over a naïve forecast

Personal Blog — Fall 2017 - present
- Discuss personal projects related to data science, including descriptive and predictive analyses
- Document code samples for creating visualizations, conducting analyses, and data scraping
- Available at https://p-mckenzie.github.io/

TECHNICAL EXPERTISE

Languages:	Python, working proficient in SQL, R
Machine learning:	Multivariate Regression, Clustering (k-means, agglomerative hierarchical), Binary and Multi-class Classification (logistic regression, naive bayes, decision trees, random forests, support vector machines, simple neural networks)
Data visualization:	matploblib, Tableau, bokeh, exposure to D3.js
Cloud computing:	Exposure to Amazon Web Services (MapReduce, S3), Google Cloud compute, Databricks (Spark/PySpark)

COMPETITIONS

USAA Hackchat – *2nd Place* — Spring 2018
- Analyzed customer service call transcripts in team of 4 students, over 12 hours of competition
- Used lemmatization and tf-idf to classify callers by the supporting document the call addressed

Citadel Data Open at Texas – *3rd Place* — Spring 2018
- Analyzed NYC taxi cab/Uber data for 2014-2015 in team of 4 students, over 8 hours of competition
- Identified and reported on demographics in neighborhoods where Uber created demand rather than cannibalizing taxi trips