ROB ZHANG

DATA SCIENTIST

rzhang27@gmail.com
rzhang27.github.io
330-564-6532
Chicago, IL 60610
rzhang27
rzhang27

Education

Northwestern University
BA Economics 2015
Minors: Mathematics & Business

Skills

LANGUAGES AND TOOLS
Python
SQL
NoSQL
Git / Github
HTML / CSS
Shell / Bash
AWS

MACHINE LEARNING
Linear Regression
Logistic Regression
K-Nearest Neighbors
Naive Bayes
Gradient Boosting
Decision Trees
Random Forest
XGBoost
Support Vector Machines
K-Means Clustering
Principal Component Analysis
Singular Value Decomposition
Latent Dirichlet Allocation
Word2Vec
Sentiment Analysis

PYTHON LIBRARIES
Pandas
NumPy
Scikit-learn
StatsModels
BeautifulSoup
Scrapy
NLTK
Flask

DATA VISUALIZATION
Matplotlib
Seaborn
Plotly
Tableau
WordCloud

Experience

Metis — Chicago, IL
Data Science Teaching Assistant — Jan. 2019 to Current
- Provide personalized support and coaching to teams and individuals for various data science projects
- Lead course discussions and lectures
- Conduct technical reviews and team interviews

Data Scientist — July 2019 to Sept. 2019
- Metis is a full-time, ACCET accredited 12-week data science program with a project-based curriculum and coursework focused on Python, statistical modeling, machine learning, and big data tools
- Designed and executed five end-to-end projects requiring web scraping, database management, supervised and unsupervised learning, natural language processing, and interactive visualizations
- Presented results to technical and non-technical audiences

Gelber Group, LLC — Chicago, IL
Proprietary Trader — Aug. 2015 to Mar. 2019
- Directed a proprietary account with weekly turnover of ~$200 million
- Actively monitored and traded fixed income securities, stock indices, currencies, and commodities using fundamental and technical analysis, operating in both cash and derivative markets
- Conducted real-time evaluation of geopolitical events, macroeconomic developments, monetary policy, and relative pricing of domestic and international securities to form and execute trading strategies
- Spearheaded daily market update meetings with up to eight other team members
- Compiled trade activity and assessed performance metrics using Excel (i.e. pivot tables, regression analysis, etc.) to obtain maximum risk/return optimization
- Prepared weekly and monthly profit-and-loss statements for manager's review

Northwestern University — Evanston, IL
Instructor & Assistant — Sept. 2013 to June 2015
- Trained small groups of fellow undergraduate students in trading with Trading Technologies software
- Coached students in navigating the Bloomberg Terminal and obtaining Bloomberg Certifications
- Facilitated with organizing monthly lab sessions by coordinating with industry guest speakers

LiDa Corporation — Akron, OH
Business Intern — June 2013 to Sept. 2013
- Oversaw migration of company records to ERP system to increase daily operational efficiency by 40%
- Improved LiDa's accounting system by assisting with transition from Excel to QuickBooks Online
- Validated customer invoices and reconciled balance sheets to identify and correct discrepancies
- Supported upper management in handling customer complaints and resolving supplier disputes

Projects

Analyzing Adoption Speed of Pets Based on Online Listing — Sept. 2019
- Combined classification, NLP, and image processing to determine pet adoptability on Petfinder
- Implemented XGBoost as an optimum classifier to predict adoption speed based on a pet's listing
- Used OpenCV and Python Imaging Library (PIL) to extract image features for classification

Analyzing Game of Thrones Script with NLP — Aug. 2019
- Performed topic modeling using Natural Language Processing and unsupervised learning techniques
- Pre-processed unstructured text data before using dimensionality reduction (PCA, LDA, LSA, NMF) to analyze changes over the show's seasons
- Scraped all eight seasons of show transcript from Genius using Scrapy and stored data using MongoDB on AWS EC2

Predicting Customer Habits on Instacart — July 2019
- Used machine learning classification models (Logistic Regression, KNN, Random Forest, Naive Bayes) to determine which customers on Instacart are likely to repurchase within one week
- Engineered features from limited variables to maximize model's F1 score
- Created interactive Tableau visualizations to compare cart-makeup of frequent vs infrequent customers
- Combined 3.4 million rows of data with 5 additional tables using PostgreSQL on AWS

Predicting Earnings on the PGA Tour — July 2019
- Built and optimized regression models (linear, LASSO, ridge, polynomial) to predict a Tour player's annual tournament winnings from official PGA data
- Created separate models for traditional and shots-gained statistics to explore feature importance
- Web-scraped additional physical attributes of players to increase model complexity