

KATIE CAO

New York, NY - Lcao18@fordham.edu
Linkedin.com/in/katiecao - (814) 218-7569 – Katiecao1.github.io

EDUCATION

FORDHAM UNIVERSITY

New York, NY

Master of Science in Business Analytics, Data Science Track, GPA: 3.80, GMAT 690/800

August 2018 – December 2019

- Relevant Coursework: Advanced SQL for Data Scientists, Data Mining, Web Analytics, Big Data Analytics, Text Analytics, Advanced Computational Finance, Database Management

MERCYHURST UNIVERSITY

Erie, PA

Bachelor of Arts - Double Major in Accounting and Finance, Summa Cum Laude, GPA: 3.92

September 2013 - May 2017

PROFESSIONAL EXPERIENCE

ZENON AI (FinTech Consulting Firm)

New York, NY

Data Science Intern

July 2019 – September 2019

- Revised a customer segmentation model by training an Artificial Neural Network and XGBoost to predict customer churn for the marketing department of the client ETRADE - an online trading platform
- Developed Python code using Natural Language Processing packages to extract payment terms from sentences
- Wrote SQL queries to extract data points to perform analysis on client's database to target customer segmentation

IDEAS42 (Behavioral Economics Consulting Firm)

New York, NY

Senior Associate

March 2018 – July 2018

Associate

August 2017 – March 2018

- Analyzed and forecasted project data used in capital budgeting for ongoing projects across multiple domains for internal reporting to the CFO and middle management
- Queried data and files for delivery to clients and loading onto internal databases
- Coordinated with cross-functional teams (operations, human resources) in the NY, Boston, San Francisco, and DC offices to allocate human capital and financial resources to projects and evaluate project performance through dashboards
- Automated tools and scripts to streamline expense reports and repeatable tasks, increasing 50% efficiency between teams
- Led a team of two employees by providing technical guidance and engagement opportunities

AWARDS & ACHIEVEMENTS

1st Prize – 2018 Fordham Accounting and Tax Society Consulting Case Competition – sponsored by PwC & Baker Tilly

- Provided a strategic recommendation on changes to management structure, customer journey, and product strategy shift for the IPO of MongoDB - a NoSQL database company; presented to industry professionals against 14 teams

Honorable Mention – 2019 Fordham March Data Crunch Madness Competition – sponsored by Deloitte

SKILLS & OTHER TECHNICAL PROJECTS

Languages: Python, R, SQL

Tools: Tableau, Alteryx, Power BI, Hadoop, Jupyter Notebook, SPSS Modeler, MS Excel / VBA, Adobe Photoshop

Applying Machine Learning to Predict 2019 NCAA March Madness Bracket (Python, Tableau, SPSS Modeler, Excel)

- Feature-engineered new variables and applied machine learning algorithms including Neural Network, Random Forest, Logistic Regression to build a predictive model for NCAA Men's Basketball Tournament with 72% accuracy and 0.49 log loss.
- Achieved honorable mention after competing against 36 teams at the 2019 Fordham March Data Crunch Madness Competition sponsored by Deloitte

Grubhub vs. Uber Eats Twitter Sentiment Analysis (Python, Excel, Tableau)

- Performed sentiment analysis by using Python code (TwitterAPI & TextBlob packages) to crawl tweets from Twitter to compare the popularity between different services in the NY Tri-State area, thereby giving recommendation to improve their business
- Increased the social media marketing conversion rate by analyzing competitors, major customer complaints by geographic location, most popular foods, influencers, and visualizing analytics results using Tableau

Pharmacy Prescription Tracking Database (SQL)

- Built a relational pharmacy prescription tracking database and schema from scratch using Oracle Data Modeler and Oracle APEX
- Generated reports and analyzed drug price, drug order, diseases, physicians, patient visits using SQL queries

Predicting Citi Bike Trip Duration Using Data Mining Algorithms (R, Excel, SPSS Modeler, Tableau)

- Constructed Decision Tree, Neural Network, Logistic Regression algorithms to build and compare statistical models using R and SPSS Modeler to predict riders' trip duration with an 80% of accuracy, thereby giving pricing strategy of surcharging and reducing the free period for subscribers to fall closer to trip duration

Dimension Reduction and Visualization for High Frequency Trading and Implied Volatility in Finance (Python)

- Identifying pattern in stock options and high frequency trading stocks by employing and visualizing dimension reduction methods (Principal Component Analysis, Kernel PCA, Sparse PCA, t-SNE, LLE, ISOMAP) from different industry sectors using Python packages (sklearn, matplotlib, scipy, pandas, seaborn).