

Xiangyu Zeng

608-471-2060 | xzeng3@andrew.cmu.edu | Pages.cs.wisc.edu/~xiangyu | Pittsburgh, PA

Education

-
- Carnegie Mellon University** | Pittsburgh, PA Aug. 2019 – Dec. 2020
- M.S. in Information Systems Management (Business Intelligence & Data Analytic) | GPA:3.5
- University of Wisconsin Madison** | Madison, WI Sep. 2015 – Dec. 2018
- Bachelor of Arts in Statistics; Minors in Math, Business & Computer Science (Honor Program)
 - GPA: 3.77/4.0 | Major GPA:3.85/4.0 | Dean's List 2016-2018

Skills

Programming Languages: Python (sklearn, pandas, tensorflow, prophet, etc.), SQL (regulars, analytic functions, etc.), Matlab, Java (oop, collection, data structure, etc.), R (dplyr, ggplot2, prophet, etc.), C++(OOP), Excel, Shell scripting, Git

Front end: JavaScript, AJAX, Flask, Dash, Scrapy | **Back end:** MongoDB, NoSQL, Hive, Apache, REST, Spark, Hadoop

Coursework: OOP in Java, Intro to Machine Learning, Data-Focused Python, Database Management, Unstructured Data Analytics, Intro to Artificial Intelligence, Data Structure and Algorithm, Advanced Business Analytic, Computer engineering

Certifications & Competition: Google analytics certification, Udacity data & business Analyst nanodegree, Keith block entrepreneurship competition, MadHack hackathon team 3rd prize 2018, Huskie hack 2017

Bootcamp Completion: Laioffer Fullstack Bootcamp (2019), Bittiger DS bootcamp (2018), Touchdown BA bootcamp (2018)

Experience

-
- Data Scientist Part-Time Intern** | Soothsayer Analytics | Livonia, MI Sep. 2019 – Present
- Participated in the development of a machine learning-based M&A platform for small and medium-sized companies that enables successful deal searching process for business buyers, business sellers, and financial intermediaries
 - Built, modeled and maintained databases on buyer and seller information with ER modeling and normalization techniques
 - Conducted exploratory data analysis in Python (Seaborn, matplotlib, etc.) to draw operational insights from trading data
- Data Scientist Intern** | Soothsayer Analytics | Livonia, MI Jan. 2019 – Aug. 2019
- Refined client E-commerce website's search relevance mechanism and final model had improved search relevancy by 23%
 - Created twitter monitoring web app and presented weekly feature-specific analytic reports to management group
 - Web app reduced data workflow processing time by 40% and enhanced MAU performances by 31% for the next 3 months
 - Analyzed E-commerce website's clickstream data and forecasted future web traffic page views across multiple time series
- Data visualization Developer** | Metascience Design Lab | Madison, WI Sep. 2018 – Dec. 2018
- Led 4-person team on front and back end development with Statscope (an interactive data visualization web application) to help researchers by eliciting user predictions and offering multiple visual representations of research data
 - Used Python (D3.js&Plotly) and REST API to add interactive web-based components (buttons, sliders, etc.) on app interface
 - Released the internal test version of Statscope, having been adopted by 6 research groups for research data visualizations
- Data Analytic Intern** | The Ronin Society | Austin, Texas May. 2018 – Aug. 2018
- Consulted five clients to maximize their sales efficiency by 30% on average through revenue optimization, demand forecasting and productivity analysis projects to support high-priority enterprise initiatives
 - Built a predictive model to offer comprehensive loan data assessments for investor clients to limit risks and avoid defaults
 - Designed a fraud detection program in Python base on 20GB transactional data with over 64% accuracy
- Financial Analyst Intern** | China Asset Management | Beijing, China May. 2017 – Aug. 2017
- Extracted data from financial terminals (Choice, CSMAR and Wind) to study regional market business landscape
 - Developed comprehensive financial reports to present market data, make forecasts and guide investment strategies

Project

-
- Real Time Twitter Visualization Dashboard – Topic Tracking and Brand Improvement [Soothsayer]**
- Automated collection process of tweets streaming from Twitter API and store them in **Relational & NoSQL** databases
 - Built an interactive analytical dashboard on the web using **Python** (Dash & Plotly) and **AJAX** tool frames
 - Tracked and analyzed trending topics in tweet texts with **NLP** (NLTK&RE) and **Sentiment Analysis** (TextBlob) techniques
- Lending Risk Analytics – Default Rate Predictions [The Ronin Society]**
- Processed raw loan data with 887,379 customers, 79 features and 8 label responses via EDA, encoding and standardizations
 - Built various **ML** models in **Apache Spark** (MLlib) and **Python** to predict the loan default likelihood with 59% accuracy
- Java Web Service Development – Event Search and Ticket Recommendation [Laioffer Full Stack Bootcamp]**
- Built a content-based recommendation algorithm to match similar events based on categories in **Java**
 - Implemented an interactive single-page webapp using **AJAX** and deployed server on **AWS** to handle 150 queries per second
- IOS Store Monetization Experiment Design [Bittiger Business Analytic Bootcamp]**
- Identified potential monetization opportunity to improve buyer conversion by 2x through exploratory analysis in **R**
 - Made recommendations to incentivize user to make purchase and to measure impact of **A/B** test in online purchase flow