Jiaxuan (Gianna) Wang
New York, NY 11355 | (646) 431-2995 | Jianna0901@gmail.com
LinkedIn: linkedin.com/in/jiaxuanwang3896/ | Github: https://github.com/Jianna-tech

Experience

The Quality Life | New York City, NY   March 2021 – Present
Business Analyst, Full-Time
- Extracted 2 GB+ of online retail data with HubSpot, Amazon-Price API, Google Analytics and created ETL pipeline in Python.
- Analyzed key factors that influenced user spending, including customer browsing history/demographics with Logistic Regression, predicted spending with Random Forest, SVM, evaluated model performance with Cross-Validation.
- Identified the target customer by using KMeans and optimized the digital marketing campaign, improved profits by 12%.
- Monitored KPI performance with Logistic Regression and prepared monthly KPI reporting with 5+ business insights via Tableau.

Eric Adams 2021 | New York City, NY   October 2020 – April 2021
Data Analyst, Intern
- Developed the automated ETL pipeline to link with databases for 15 GB+ public datasets, saved 30% labor cost.
- Engineered 100k community data and extracted 15+ usable features including demographics, boroughs, median household income.
- Performed A/B tests with the webpage and conducted statistical analysis, increased Median Time Spent by 100%, CR by 8%.
- Improved pageviews by 300% by creating and publishing 6+ interactive dashboards using Plotly and R Shiny.

Columbia University | New York City, NY   March 2020 - August 2020
Data Science Consultant, Intern
- Analyzed 40k+ medical data to explore patient demographics and adherence to each medicine for a health insurance company.
- Built 5+ linear regression models to detect effects of patient demographics/symptoms on medication initiation, patient adherence.
- Identified which medications are more effective on different patients with T-Tests and Chi-Squared Tests for 4+ diseases.
- Generated an automated R reporting engine monthly updating actionable insights of products, improved working efficiency by 20%.

Bank of Nanjing | Shanghai, China   July 2018 – December 2018
Financial Data Analyst, Intern
- Identified 4+ key product metrics for business development and prepared the industry data with SQL queries.
- Conducted competitive analysis on the semiconductor industry to supported cross-team leadership in M&A decision making.
- Predicted the annual revenue with Time Series Models, increased the valuation of a listed semiconductor company by 10%.
- Promoted a $100M Assets-Backed-Securities program to a listed company with financial data analysis and analytical insights.

Projects

Model competition project in machine learning (SQL, R)   July 2020 - September 2020
- Performed feature selection on 100+ features from MIT's GOSSIS medical dataset. And reduced overfitting by analyzing feature distribution through the adversarial validation technique. Applied 20+ models to predict patient survival, SVM reached accuracy=98.3%.

Yelp User Analysis project (Python, R)   April 2020 – July 2020
- Extracted and transformed 3 GB+ of data from Yelp fusion API via Python, performed Text Mining and Sentiment Analysis on 10K+ user reviews to provide insights of user perceptions, created motion visualizations in R shiny and Tableau to report results.

SQL project for Library Management System (PostgreSQL)   February 2020 – March 2020
- Created and populated a relational database containing 4+ tables. And modified the database to encode borrowing rules.
- Wrote complex SQL queries using complex joins, grouping, having clause, window function, nested subqueries, etc.

Education

Columbia University, *M.S. in Applied Analytics*; GPA 3.7/4.0, New York, NY   September 2019 – February 2021
Relevant coursework: Python for Data Analysis, SQL & Relational Databases, Data Visualization & Storytelling, Machine Learning

Nanjing University of Finance & Economics, *BBA.in Accounting;* GPA 3.5/4.0, Nanjing, China   September 2015 – July 2019
Relevant coursework: Probability and Statistics, Finance, Accounting, Economics, Calculus, Corporate Finance, Excel

Skills

- **Data Analytics**: A/B Testing, Casual Inference, Tableau, PowerBI, SAS, Google Analytics, MS Excel, Access
- **Machine Learning**: Python (2 years +), R(2 years +), SPSS | Predictive Modeling, Hypotheses testing, Clustering Analysis
- **Database Processing**: SQL(4 years +), AWS S3, Google Cloud Big Query