

XINYU LI

(929) 442-4917 | x13228@columbia.edu

<https://www.linkedin.com/in/xinyu-l-206831227/> | Green Card Holder/US Permanent Resident

EDUCATION

Columbia University, Columbia Engineering – Operation Research & Columbia Business School
MS in Business Analytics (Track: Data Analytics Algorithm)

New York, NY
Expected Dec 2023

GPA: 3.979/4.00.

Award: Business Analytics Graduate Fellowship for Academic Excellence (\$1,000 Scholarship)

Coursework: Data Analytics (Python/SQL/Web Scraping/ML), Statistics & Probability, MS ML, Optimization Models, Applied Machine Learning, Intro to Database.

Baruch College, CUNY

New York, NY

BBA in Accounting with Intensive Minor in Math

Dec 2020

GPA: 3.985/4.00.

Coursework: Multi-variable Calculus, Inferential Statistics, C++, Java, Computer Algorithm, Stochastic Process Models.

WORKING EXPERIENCE

Columbia Business School

New York, NY

Teaching Assistant II (Part-Time)

Sep 2023 – Expected Dec 2023

- Teaching Assistant for marketing analytics course, which covers popular data science techniques used in marketing area.
- Providing academic and administrative support for faculty, including organizing classes, grading HWs and exams & scheduling events.
- Acting as a liaison between students and faculty to address logistical student concerns or meet individually with almost students to discuss a challenging topic.

QuidelOrtho (NASDAQ: QDEL)

Rochester, NY

Data Science Analyst Intern (Full-Time)

May 2023 – Aug 2023

- Developed and maintained 500 lines of Python code and 100 lines of SQL code within Azure Databricks, spanning data collection, data profiling, exploratory data analysis, feature engineering, statistical analysis, and machine learning model development.
- Orchestrated machine learning experiments utilizing Auto ML. Developed and fine-tuned multiple ML models manually and logged models into Auto ML experiments using python, achieving accuracy of 86% in the use case.
- Collaborated closely with the commercial team to conceptualize data-driven solutions for addressing commercial challenges. This resulted in expedited resolutions for unplanned equipment fixes, elevating client satisfaction by 50% and driving monthly savings of \$120k.
- Analyzed SAP HANA View and Power BI dataflow to decipher the underlying business logic used in formulating key business metrics from raw data.
- Presented the final project on ML models for addressing unplanned equipment fixes to the Data & Analytics Manager, earning an "excellence" performance review.
- Effectively planned and executed a data science project roadmap within Azure DevOps, ensuring a seamless progression from concept to implementation. Proficiently managed code integration, version control, and facilitated collaboration between Azure Databricks and Azure DevOps.

Apple (NASDAQ: AAPL)

New York, NY

Student Data Scientist Intern (Part-Time)

Feb 2023 – May 2023

- One semester long capstone project between Columbia University and Apple to apply machine learning to solve supply chain challenges.
- Collaborated with Apple's analytics team to address apple's current challenges in supply chain and cybersecurity.
- Met with Apple's supply chain manager on a weekly basis to deliver reports and business insights, and to modify plans to align with manager's goals and feedback.
- Conducted market research to rank the importance of supplier features and assigned percentage importance score to each feature and developed a supplier scoring model using Python to assess individual supplier's importance.
- Built KNN models, Gradient Boosting and Random Forest Model using Python's Scikit-learn package to group suppliers and identify the key supplier group.
- Applied sentiment analysis to analyze supplier's surveys to gain insights into their level of confidence about their cybersecurity system.

UHY Advisors LLP

New York, NY

Audit & Assurance – Asian Practice (Full-Time)

Feb 2022 - Jul 2022

- Delivered over 20 audit engagements including 10-Q quarterly interim review, 10-K year end audit, SPAC entity IPO closing audit, S-4 form for SPAC entity & 8-K form (unscheduled material event) over CCH ProSystem fx Engagement.
- Audited each account in the financial statement by analyzing large set of accounting data through Excel. Identified and corrected over 5 million misclassification of accounting items to ensure no material mistakes and compliance with US GAAP.

GEN II Fund Services

New York, NY

Experienced Private Equity Fund Accountant (Full-Time)

Feb 2021 – Feb 2022

- Oversaw accounting for a startup private equity fund with \$500M AUM and provided fund administration support for three legacy funds with \$1B AUM.
- Developed data pipelines between Sage Intacct and Qlik to automate financial and accounting reporting.
- Mentored interns in PE accounting and software usage, resulting in an 80% retention rate for full-time positions.
- Established fund structures based on investment documents and partnership agreements to ensure correct accounting data flow.
- Analyzed large financial data set using Excel (VLOOKUP, PivotTable, Aggregation Functions, etc.) to generate ad-hoc reports.
- Imported accounting data to Sage Intacct by making daily accounting entries, reorganized large accounting datasets by compiling accounting work papers, and audited financial statements in compliance with US GAAP to ensure data reliability and accuracy.

PROJECT

Tahoe Health Care Analytics – Reduce Risk of Readmission | Python, Machine Learning, Scikit-Learn

- Developed logistic regression (lasso/ridge), random forest, GBM, XGBoost and SVM models using Python's scikit-learn library to predict the likelihood of patient readmission within three months, based on patient demographics, medical history, and other relevant variables.
- Utilized the predictive models to suggest a proactive care plan for patients predicted to be readmitted within three months, resulting in significant cost savings of nearly \$2 million by avoiding government Medicare penalties.

Spotify – Exploratory Data Analysis | SQL, Big Query, Looker Studio, Google Cloud Computing

- Analyzed Spotify's user ratings and reviews data to evaluate the performance of the Spotify music streaming app over time.
- Defined metrics to measure the performance of the Spotify music streaming app and compiled SQL queries using Big Query to aggregate and organize the data according to the defined matrix.
- Utilized Looker Studio to generate business reports and data visualizations that provided actionable insights into the performance of the app.

Amazon - E-commerce XP: "Add to cart" buttons | SQL, Big Query, Probability & Statistics, A/B Testing

- Conducted A/B testing to assess the impact of adding extra "Add to Cart" buttons on the website, defining metrics, forming treatment and control groups, and determining an optimal sample size.
- Analyzed the outcomes of the A/B test by computing the p-value associated with the observed disparities between the control and treatment groups with respect to the predefined metrics. This comprehensive analysis provided valuable insights into the effectiveness of the proposed interface modification.

SKILLS

Programming: SQL, Python (Statsmodels, Scikit-learn, Pandas & Numpy, matplotlib), Web Scraping (Beautiful Soup), API Keys (Twitter/Yelp), Auto ML.

Math: Probability, Statistics, Machine Learning (Regression, RF, GBM, SVM, XGBoost, PCA, Clustering Methods, Recommendation System), A/B Testing.

Software: Azure Databricks & DevOps, Google Cloud Computing (Big Query/Looker Studio/Colab), Qlik, Excel/Word/PPT, Jupyter Notebook, Sage Intacct.