XINYAO HAN

+1 (617) 821-8462 | Cambridge, MA | xinyao1@mit.edu | LinkedIn: xinyao-han

EDUCATION

MIT SLOAN SCHOOL OF MANAGEMENT — Cambridge, MA
Candidate for Master of Business Analytics, Operations Research Center, August 2023 — 2022 - Present

- Coursework: Machine Learning, Optimization Methods, Analytics Edge, Analytics Lab
- ML Project: Social media ads CPA prediction on marketing strategy (OCT & XGB), with 26% cost reduction (Python, Julia)
- Optimization Project: Optimized delivery routes with random customer locations & stochastic travel times, reducing delivery cost by 60% (Julia)

NEW YORK UNIVERSITY — Shanghai, China
Bachelor of Science in Computer Science, Minor in Mathematics, GPA: 3.9/4.0 — 2018 - 2022

- Coursework: Natural Language Processing, Algorithms, Probability and Statistics, Econometrics
- NLP Project: Implemented BERT multi-task model on criminal legal dataset to answer paragraph questions
- Honors: NYU Honors Scholar, Global Futures Scholar, Dean's Honors List 2018-2022, NYU Excellence Award
- Leadership: Career Development Center Intern (Promoted to Leader), Undergraduate Business Association Vice President

TECHNICAL SKILLS

- Python (PyTorch, NLTK, scikit-learn, Selenium, Pandas), R, SQL, Julia, ArcGIS, C, Java, Git

EXPERIENCE

MIT OPERATIONS RESEARCH CENTER | LIBERTY MUTUAL — Cambridge, MA
Research Assistant to Professor. Vivek Farias — Fall 2022 - Present

- Designed and implemented scalable imputation algorithm to imputed 2 high-dimensional matrices on 500k insurance data with $1.2B revenue, analyzed imputation mechanisms, presented results to Data Science teams to improve pricing (Python)
- Performed tree models and neural network on insurance loss prediction, enhancing baseline model performance (Tensorflow)

MIT SLOAN | UNITED NATIONS — Cambridge, MA
Analytics Lab Team Member, UN Mine Action Service, Awarded 2nd Place out of 22 teams — Fall 2022

- Feature-engineered satellite imagery data into patches (ArcGIS), classified buildings in rural Colombia from pixel-level (CART, XGBoost, random forest), achieving 98% AUC (PyTorch)
- Integrated classification models to UNMAS global analytics team to evaluate effects of Colombian mine clearance project, providing insights into subsequent urban developments of Colombia

BLOOMBERG — Shanghai, China
Global Data 2021 Summer Intern — Summer 2021

- Crawled 100+ reports (Selenium), extracted key points and established a data repository in Bloomberg, transformed data with statistical correlations, visualized data to analyze macro & micro finance trends (Python)
- Developed product template with data mining on domestic policy research, presented product to senior management (40+ European & APAC team leads, managers) and used as Bloomberg prototype product

EBAY INC — Shanghai, China
Data Scientist Intern (Honor: Intern Top Performer 1/40+) — Spring 2021

- Led team of 2 interns to conduct clustering-based cohort analysis and regression-based incrementally analysis utilizing 5-year data for 200B users to evaluate effectiveness of major promotion events, resulting in 1.5% increase in user retention
- Designed and proposed 15+ decision, guardrail, and learning metrics to measure promotion impact, constructed dashboards to analyze business metrics performance and user behaviors (Python, Tableau), presented data-driven promotion strategies to business and operations teams bi-weekly

NYU CENTER FOR DATA SCIENCE AND ARTIFICIAL INTELLIGENCE — Shanghai, China
Research Assistant to Professor. Guodong Chen — 2020 - 2021

- Conducted text analysis on 70M HSBC payment transactions, preprocessed features via word embeddings (NLTK)
- Segregated customers using text data on consumption behavior with Multinomial Naive Bayes & NN models, achieving 88% accuracy (PyTorch), introduced customer segmentation strategies to HSBC Data Science team

ADDITIONAL INFORMATION

- Volunteer: Researched Chinese female maternity policy & economic effects, volunteer teacher for migrant children in China
- Interests: amateur pianist, avid yogi, adrenaline junkie, figure skating learner