KEVIN (JING MING) LI

Evanston, IL | kevinli2024@u.northwestern.edu | 224-714-9081| linkedin.com/in/kevinjmli

EDUCATION

Northwestern University, M.S in Machine Learning and Data Science *December 2024*

Northwestern University, B.S in Industrial Engineering, Computer Science Minor (GPA: 3.96/4.00) *June 2023*

- **Relevant Coursework:** Machine Learning, Stochastic Optimization, Database System, NLP, Operations Management, Statistics
- **Awards:** Magna Cum Laude, Industrial Engineering Department Award (2 graduates out of 80), Research Scholarship ($5,000)

WORK EXPERIENCE

The Blackstone Group **New York, NY**

Data Science Intern, Private Equity Team (Python, Snowflake) *June 2023 - August 2023*

- Designed and prototyped a new store location selection tool for 3 quick-service-restaurant clients, using 6 alternative data sources to engineer 35 features on demographics, cannibalization, and accessibility while achieving 3x faster dataframe creation, trained and tuned a random forest model on 3.8M row dataset, generating 8.8% revenue uplift compared to baseline model.
- Determined 5 drivers of post-COVID revenue growth for a large aviation company by leveraging controlled variable analysis and executed A/B test with private equity team to optimize non-fuel pricing, resulting in 5% increase in customer reactivation.
- Reduced investment memo generation time by 98%, leveraging Langchain to simultaneously parse 10 structured dataframes.

Dunkin' Donuts **Chicago, IL**

Data Science Consultant, Operations Team (Python, Simio, R, Excel) *March 2023 - June 2023*

- Analyzed 80 weeks and 300K rows of product sales to optimize ordering and storage for 24 remodeled stores in Chicago.
- Addressed storage misordering by building a sales forecasting model using ARIMA, achieving 13% MAE through backtesting.
- Quantified misordering severity with Euclidean Distancing, leading to $30K annual cost reduction per franchise.

Chicago Board Options Exchange (CBOE) **Chicago, IL**

Quantitative Research Intern, Data Analytic Team (SQL, Python, Tableau, Hadoop, Snowflake, HUE) *June 2022 – August 2022*

- Analyzed 30M trading transactions over 4 months on customer behavior of NANOS (newly-launched options product) to identify short-expiration trends and customer churn contributing to 3 major deliverables for C-suite.
- Utilized data manipulation techniques (recursive CTE, LOD Functions) to uncover trends in product volume and retention rate, achieving 15% increased monthly revenue and driving tailored product designs for retail customers.
- Reduced feature span from 1K+ to 6 with PCA on 350K account data, improving cluster visualizations. Built Gaussian Mixture Model which accurately identified high-frequency algorithmic traders, enhancing trade monitoring accuracy by 75%.

Illinois Attorney General's Office **Chicago, IL**

Data Analytics Intern, Civil Rights Bureau (Python, SQL, R, Excel) *January 2022 – May 2022*

- Leveraged Python and SQL to analyze 12 years & 400K rows of data and discovered 9 potential patterns of racial discrimination including increased police misuse of force in minority communities, informing policy changes for 10k residents.
- Developed Lasso and Logistic Regression models on disciplinary data, achieving an 86% precision rate in predicting suspension-related incidents, thereby aiding Civil Rights Bureau in identifying potential bias in school disciplinary actions.

RESEARCH EXPERIENCE

The Morton Group (CDC, NIH Funded), Northwestern University **Evanston, IL**

Data Research Assistant, COVID-19 Staged Alert System (Python, JSON) *October 2021 – May 2023*

- Implemented a staged-alert system in Austin, Texas using stochastic optimization (SEIR) to minimize ICU intake.
- Achieved 300% increase in data input speed by implementing web-scraping and automated seed generation to enhance data collection efficiency for vaccine doses, hospitalization rates, and ICU capacities across 5 age groups.
- Validated predictive alert model for COVID-19 by simulating 15 different transmission scenarios, increasing accuracy (R-value) of future predictions by 11.6% compared to original model, aiding in effective resource allocation.
- Applied statistical analysis to identify 7 high-risk population outcomes for COVID-19, influencing socioeconomic policies for 2.3M people. Presented findings to Head of Austin Public Health, contributing to city's low per capita COVID-19 death rate.

SONIC Lab (NASA & NSF Funded), Northwestern University **Evanston, IL**

Data Science Research Intern, Performance Metric for MultiTeam Systems (Python, AMPL) *March 2022 – June 2022*

- Leveraged mixed-integer programming to evaluate astronauts' collaboration for space tasks, with optimal solution projecting a 20% improvement in task efficiency and proposed model adopted by NASA for future missions.
- Delivered 40% decrease in computation time by spearheading migration process from AMPL to Python, devised intricate variable-to-parameter guidelines and introduced wrapper to solve objective functions across 248 inter-team simulations.

PROJECTS

Amazon Review Classification: Applied TF-IDF and sentiment analysis on 500K pre-processed retail reviews, built a late fusion algorithm on tuned SVM and Neural Network models to classify 'positive' and 'negative' reviews, achieving 0.91 F1-score on test set.
Recipe Recommender System: Leveraged LSA to create 7-dimensional user and recipe profile to recommend 10 new recipes for 3000 test users similar to previously downloaded recipes, achieving 32% precision@10 score.

SKILLS & INVOLVEMENTS

Tools: Python, R, AMPL, SQL, Tableau, Snowflake, MATLAB, Excel | Machine Learning, Product Development, AB Testing
Leadership: Quant Analyst & Golf Lead at Sports Analytics Group | Outreach Chair at Industrial Engineering Undergrad Board