RISHI MOHAN

San Francisco, CA | linkedin.com/in/mohan-rishi | rishimo2012@gmail.com | rishimo.github.io

EDUCATION

UNIVERSITY OF SAN FRANCISCO   *San Francisco, CA*
Master of Science in Data Science   *December 2024*
Relevant Coursework: Machine Learning, Linear Regression, Time-series, Distributed Systems, Databases

UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN   *Champaign, IL*
Bachelor of Science in Industrial Engineering, minor in Statistics   *May 2023*

RELEVANT EXPERIENCE

MACHINE LEARNING ENGINEERING INTERN   *San Francisco, CA*
Metaphor Data   *October 2023 – Present*
- Developed open-source Python-based connectors for document sources and corresponding TypeScript ingestion backend to extend data crawlers and index client knowledge-bases.
- Built and tested document vector embedding pipelines using Azure OpenAI text embedding, large language models, and MongoDB vector search to improve and tune multi-source search capabilities.

DIRECTOR OF OPERATIONS   *Urbana, IL*
Illini Solar Car   *August 2019 – May 2023*
- Managed network of 30+ team alumni to act as cross-functional advisory board for vehicle design & operational decisions while collaborating with the leadership team.
- Managed $100k+/yr. financials and materials through 501c3, negotiating support agreements with 60+ sponsors valued between $5,000-$35,000.

OPERATIONS MANAGEMENT INTERN   *Kearney, NE*
Eaton Vehicle Group   *May 2022 – August 2022*
- Analyzed 9 years of hourly output data across 6 lines for 25+ parts, utilizing time-series modeling to track theoretical capacity and forecast production, enhancing accuracy in setting production goals by 30% through a multiple linear regression model.
- Streamlined database maintenance and report generation for process owners, saving 10 hours weekly.
- Visualized performance indicators to pinpoint production shortfalls, guiding decisions on capital investments, machine rebuilds, and operator retraining for efficiency improvements.

SELECT PROJECTS

DISTRIBUTED SYSTEMS PROJECT – UNIVERSITY OF SAN FRANCISCO   *February 2024*
- Developed an automated data pipeline using Apache Airflow to orchestrate interactions between MongoDB, Spark, and BigQuery, handling daily ingestion of 500-2000 rows and weekly processing of 5000-15000 rows.
- Implemented aggregation pipelines in MongoDB to optimize data processing and runtime.

A/B TESTING PROJECT – UNIVERSITY OF SAN FRANCISCO   *January 2024*
- Designed an online A/B testing procedure with a simulated response surface to optimize 4 experiment parameters and predict minimum target variable.
- Employed Response Surface Methodology (RSM) to optimize parameter search space with projected objective coming within .03% of the optimal value.
- Performed statistical testing to verify hypotheses and provide appropriate statistical power for reporting.

SENIOR ENGINEERING CAPSTONE – JULIAN ELECTRIC   *December 2022*
- Collaborated with an industry partner to reduce production waste by 5% by replacing destructive tests with non-destructive techniques for ultrasonic welding of dissimilar metal cables.
- Developed a Python OpenCV algorithm for weld quality analysis with over 99% accuracy and compiled detailed cost-error statistical evaluation

SKILLS / TOOLS

Python (matplotlib, numpy, pandas, plotly, prophet, pytorch, scikit-learn, scipy, seaborn), R, Git, PostgreSQL, MongoDB, Spark, AWS, Airflow, Databricks, Azure OpenAI, A/B testing