ZACHARY HORTON

704-219-8999 | zhhorton@mit.edu | LinkedIn | Project Portfolio

EDUCATION

MASSACHUSETTS INSTITUTE OF TECHNOLOGY Cambridge, MA
Master of Business Analytics, Operations Research Center, GPA 5.00/5.00 2023 - 2024

- Coursework: Machine Learning, Optimization, Deep Learning, Product Analytics, and Software Systems for Data Science
- Deep Learning: Built a transformer model to generate queries and analyze S&P 500 forecasts from LSTMs [Demo] (Python)
- Realtime Analytics: Optimal rideshare service decision-making using machine and reinforcement learning (Python)
- Machine Learning Project: Improving K-means and K-medoids clustering using mixed-integer optimization (Julia)

NORTH CAROLINA STATE UNIVERSITY Raleigh, NC
Bachelor of Science in Industrial & Systems Engineering, Minor in Statistics, GPA 4.00/4.00 2019 - 2023

TECHNICAL SKILLS

- Programming: SQL | Python (Pandas, NumPy, DuckDB, Scikit-Learn, TensorFlow) | R | Julia | SAS | OOP
- Software: Databricks | Microsoft Fabric | MySQL | Spark | Salesforce | Tableau | PowerBI | Git | Streamlit
- Certificates: Cognitir Advanced SQL | Cognitir Intro to SQL

EXPERIENCE

BOOTH & ASSOCIATES Raleigh, NC
Data Sciences and Engineering Specialist II Fall 2024 - Present

- Leading the Data Services division, focusing on expanding current offerings and developing new technologies for clients
- Developed internal library to identify design errors in electrical components on utility-scale system model (Python, Git)
- Designed and implemented scalable data processing to automate ingestion and analysis of 575M+ energy usage records, enabling optimal rate design informed by industry trends, competitive intelligence, and user behavior (Python, Git)
- Architected a full-stack framework for a utility client to identify $300K+ in annual revenue leakage and drive strategic customer engagement via a targeted marketing campaign (Python, MySQL, Streamlit, Git)

MIT | THERMO FISHER SCIENTIFIC Cambridge, MA
Data Science Intern, Capstone Project Spring 2024 - Fall 2024

- Developed ensemble forecasting model, boosting revenue accuracy by 40% and cutting planning costs by 70% (Python, Git)
- Built a fully automated ETL pipeline to create and deliver new monthly forecasts, reducing senior staff time by 10% (Fabric)

MIT | NATIONAL INSURANCE CRIME BUREAU Cambridge, MA
Data Science Intern, Analytics Lab Project Fall 2023

- Engineered machine learning system with an AUC of 0.83 to predict fraudulent claims and identify high-fraud events from 127 tropical storms to enhance workload forecasting and reduce $200K in annual investigative expenses (Julia, Python, R)
- Examined 690K fraudulent claims spanning 18 years to discern distinguishing factors of high-risk financial events (Tableau)

CORNING INCORPORATED Charlotte, NC
Manufacturing Data Analytics Intern 2022 - 2023

- Automated ETL pipeline for monthly sales data from 2018 to increase velocity and availability, leveraging natural language processing with Meta's LLM (BART) to analyze 7.3M rows, adding customer-level granularity to analytics (PySpark)
- Generated 83 process quality metrics utilizing a convolutional neural network computer vision model in Databricks, earning finalist recognition for the Manufacturing Leadership Council: Engineering & Production Technology award (Python, SQL)
- Showcased project outcomes and strategic impact to CDIO and senior IT leaders across Corning’s 5 central business units

UNC REX HOSPITAL Raleigh, NC
Spring Senior Design Project Team Member Spring 2023

- Accomplished goal of decreasing the daily backlog of surgical trays from 600 to 75 by developing an optimized scheduling model; reduced required FTEs from 15 to 10 and weekly hours from 584 to 405 compared to the previous schedule (R)
- Presented findings, experiences, and future recommendations to senior clinicians, nurses, and hospital managers (Tableau)

NC STATE UNDERGRADUATE RESEARCH Raleigh, NC
SimOpt GUI Developer Summer 2021 - Spring 2023

- Created graphical user interface (GUI) for open-source simulation and optimization library and was acknowledged for contributions in a research paper published in the INFORMS Journal on Computing 35(2):495-508 (Python, Git)

Diabetic Retinopathy (DR) and Macular Edema (ME) Researcher Spring 2022

- Applied linear models and continuous time Markov chains to model the progression of DR and ME for 400K+ patients (R)

ADDITIONAL INFORMATION

- Interests: lifelong baseball player and skier; drafting recipe book for young adults and new cooks