GREGORY TOMY

Gregory.Tomy@colorado.edu | 917-428-6003 | linkedin.com/in/gregorytomy/ | github.com/GregoryTomy

Professional Experience

GlobalFoundries – Machine Learning Intern — Oct 2023 – Jan 2024

- Implemented an XGBoost digital twin of the Technology Computer-Aided Design (TCAD) process, achieving over 50% reduction in training time relative to existing deep learning model, while maintaining accuracy.
- Developed and optimized data generation scripts for TCAD, cutting iteration time by 80% and significantly reducing errors.

Sandia National Labs – Graduate Student Researcher — Sep 2023 – Dec 2023

- Built a deep learning model to predict the impact of atmospheric SO_2 injections on temperature using a 15-year dataset with over 5.4 million rows and PyTorch.

CVS Health – Data Scientist Intern — May 2023 – Aug 2023

- Developed a business case for Aetna's smoking cessation program, projecting $3-5M annual revenue increase.
- Utilized Google BigQuery for data cleaning and analysis of the Clinical Data Repository (CDR) and Aetna databases, identifying a target population of 666,000 smokers.
- Analyzed 6+ million patient records with 50+ features in the CDR to identify individuals at risk of heart disease.

Brean Strategic Advisors – Valuations Analyst — July 2018 – July 2021

- Compiled SQL data and built data models for RMBS loans over $3 billion, identifying breaches of contract and quantifying material impact.
- Led a team in forecasting financial performance for low-income credit unions, successfully securing National Credit Union Administration Secondary Capital program funding for 6 clients.
- Developed Python data cleaning pipeline and forecasted mortgage lender portfolio prices, reducing data preparation time by 98% and automating error handling with 95% accuracy.

Projects

Sommelier LLM Chatbot:

- Developed an AI-powered chatbot that recommends wine using Microsoft Azure, OpenAI, LangChain, and Retrieval-Augmented Generation (RAG), and deployed using FastAPI, Streamlit and GitHub Actions. (link).

Automated Lung Cancer Detection:

- Built an end-to-end deep learning pipeline with PyTorch for the identification and classification of malignant lung nodules from CT scans using Convolutional Neural Networks (link).
- Achieved 80% nodules detection and 58% precision in identifying malignant cases, across 228K instances.

Optimizing Discount Strategy with Causal Inference:

- Utilized synthetic control, regression discontinuity, and double/debiased machine learning to assess an e-commerce company's discounting strategy, finding a $0.25 loss per dollar increase in discounts (link).
- Created a causal model for targeted discounts, recommending discounts for the top 55% of customers based on a custom score.

Education

University of Colorado Boulder — Aug 2022 – May 2024

Masters in Applied Mathematics | **GPA: 4.0**

Courses: Machine learning, Bayesian statistics, Applied deep learning, Data mining, Causal inference.

New York University — Aug 2014 – May 2018

Bachelors in Economics and Business Studies

Certificates

Google TensorFlow Developer – March 2024

Skills

Python	PyTorch	Google AutoML	Flask
SQL	Prompt Engineering	Microsoft Azure	FastAPI
R	LangChain	AWS Lambda	Git/GitHub Actions
TensorFlow	Google BigQuery	Docker	Presentations