

# THOMAS OU

(203) 243-2852 | thomas96@gmail.com | linkedin.com/in/thomas-ou

Mathematics and CS student specializing in ML, full-stack development, and algorithmic optimization. Proven track record building scalable systems and applying data-driven solutions across finance, legal tech, and nonprofit sectors.

## EDUCATION

### University of Pennsylvania

May 2027

BA in Mathematics, Minor in Statistics; Accelerated MSE in Computer Science

**Coursework:** Algorithms, Machine Learning, Big Data Analytics, Game Theory, Statistical Modeling, Cryptography

## TECHNICAL SKILLS

**Languages:** Python, Java, OCaml, R, SQL, JavaScript, HTML/CSS

**Frameworks & Libraries:** PyTorch, TensorFlow, Scikit-learn, Pandas, NumPy, SciPy, React, Flask, Firebase, spaCy, NLTK, Three.js, GSAP, quantmod, tidyverse, Shiny

**Tools & Technologies:** Git, Docker, Excel VBA, Tesseract OCR

**Specializations:** Machine Learning, NLP, Algorithm Design, Game Theory, Bayesian Statistics, Data Visualization, Time Series Analysis, Full-Stack Development, Data Engineering

## PROJECTS

### Bayesian Opponent Modeling Engine for Poker AI

Sep 2023 – May 2024

- Designed Bayesian inference system analyzing 10,000+ hands with 85% bluff prediction accuracy; built poker simulation engine
- Applied game-theoretic EV optimization improving performance by 12% against exploitable opponents
- Tech:** Python, NumPy, Pandas, SciPy, Bayesian Statistics, Game Theory

### Full-Stack Tutor Matching Platform – Trumbull Tutors

Mar 2020 – May 2025

- Architected web app serving 200+ students with intelligent matching algorithm; built with Python/Flask and Firebase
- Developed analytics dashboard visualizing 5,000+ hours, reducing administrative overhead by 60%
- Tech:** Python, Flask, Firebase, JavaScript, React, HTML/CSS

### Quantitative Trading Analytics Platform – Wharton Investment & Trading Group

Oct 2024 – May 2025

- Engineered real-time data pipeline processing 50+ equities; implemented ARIMA/GARCH models for volatility forecasting
- Built interactive Shiny dashboard for portfolio optimization and risk-adjusted return visualization
- Tech:** R, quantmod, tidyverse, Shiny, Time Series Analysis

## EXPERIENCE

### Research Assistant – Defense Innovation & Policy, Penn Center for Undergraduate Research

Jul 2025 – Present

- Developing Python OCR pipeline digitizing 50+ years of military appropriations data for Prof. Michael C. Horowitz
- Designing SQL database and applying NLP for automated document classification and entity extraction

### Research Assistant – Computational Physics, Princeton Plasma Physics Lab

Jul 2024 – Aug 2024

- Implemented Monte Carlo simulations modeling particle trajectories in fusion reactors; achieved  $R^2 \geq 0.90$  prediction accuracy
- Automated pipeline processing 100GB+ simulation data, reducing analysis time by 75%

### Data Engineering & Analytics Intern, Flushing CPA Tax Center

Feb 2024 – May 2024

- Built ETL pipelines automating 50,000+ financial records, improving tax projection accuracy by 15%
- Developed regression models forecasting quarterly tax liabilities with 92% accuracy; created VBA anomaly detection macros

### Legal Tech & NLP Intern, Law Offices of Jonathan Spodnick

Sep 2023 – Dec 2023

- Engineered NLP classification system categorizing 1,000+ legal documents, reducing manual review time by 50%
- Conducted statistical analysis on settlement data using regression and clustering techniques

## LEADERSHIP

### Co-Founder & CTO, Trumbull Tutors (501c3)

Mar 2020 – May 2025

- Co-founded nonprofit serving 200+ FGLI students; led technical strategy and managed team of 40+ volunteers
- Developed matching platform and analytics dashboards; secured \$15,000+ in grants through data-driven reporting