

Matthew Rui

(720) 965-1646 | matthewrui0905@gmail.com | linkedin.com/in/matthewrui | github.com/mrui0905

EDUCATION

Duke University

May 2026

Bachelor of Science, Computer Science and Mathematics, Statistics Minor (GPA: 3.99)

Durham, NC

- **Related Coursework:** Data Structures and Algorithms, Computer Systems, Advanced Algorithm Design, Randomized Algorithms, Database Systems, Deep Learning, Machine Learning, Engineering Design
- **Thesis:** Polynomial Convergence of MCMC Phylogenetic Tree Sampling Algorithms
- **Awards:** Citadel Datathon, 1st Place - \$10,000 Prize, USC Big Data Case Competition, 3rd Place - \$2,000 Prize

EXPERIENCE

Bridgewater Associates

June 2025 – Present

Investment Engineer Intern

New York City, NY

Castleton Commodities International

June 2024 – Aug. 2024

Quantitative Trading Intern

Stamford, CT

- Engineered a predictive model for power futures pricing, leveraging encrypted market data and distributed processing with Apache Spark. Resulted in a 13% improvement in pricing accuracy, leading to its adoption as the core model across all trading desks and integration into algorithmic strategies managing \$50M+.
- Spearheaded real-time data ingestion from 100k+ sensors into Snowflake leveraging Python, Apache Kafka, and REST API integrations. Developed 97% accurate anomaly detection XGBoost model for pipeline monitoring.
- Deployed a real-time disaster/outage intelligence platform using AWS, Snowflake, Next.js, and React.

Crestone Capital

May 2023 – Aug. 2023

Software Engineer Intern

Boulder, CO

- Created a Python/Flask application on AWS featuring a Llama 2 chatbot for internal use and integrated with a LangChain RAG system to automate quarterly client report generation, saving 200+ hours per week.
- Developed a full-stack application with Python, Flask, PostgreSQL, and AWS to centralize internal access to web-scraped (using BeautifulSoup) SEC regulatory data. Implemented a similarity ranking feature to sort and compare firms using VectorDB.
- Built a Go-based private blockchain prototype featuring SHA256 hashing, private wallet functionality, and transaction management using BadgerDB. Presented to firm to evaluate potential for a tokenized investment fund.

Michaud Capital Management

June 2022 – Aug. 2022

Machine Learning Intern

Denver, CO

- Engineered a real-time data pipeline with TwitterAPI and Python to monitor online sentiment during Federal Reserve speeches through NLP. Developed and tuned a 91% accurate Random Forest model to predict real-time treasury yields.

PROJECTS

Discrete Diffusion Language Models

Jan. 2025 – Present

- Designed theoretical discrete diffusion model architectures, including the derivation of closed-form solutions to Markov processes, and tested models for improvements in LLM text & music generation (Paper submitted).
- Utilized PyTorch, CUDA, and Hydra to implement Transformer-based Neural Network and graph-based diffusion processes, incorporating custom noise schedules, gradient accumulation, mixed precision training, and distributed training via Slurm.

SKILLS & INTERESTS

Languages: Python, Java, Go, C, SQL, Javascript, Typescript, HTML/CSS, R, MATLAB, x86 Assembly, Latex

Libraries/Frameworks: React, Node.js, Next.js, Flask, Dash, Streamlit, PostgreSQL, VectorDB, BadgerDB, REST API, Hadoop, Spark, PyTorch, Scikit-learn, BeautifulSoup, LangChain, Pandas

Tools/Platforms: Git, Kafka, Hydra, Slurm, AWS, Snowflake, VS Code, PyCharm, Jira, Confluence

Interests: Candle collecting (200+ and counting!), Tennis, Poker, Spicy Ramen, Skiing, Latin Literature, Sci-Fi Novels