

Shams Rupak

New York | 646-619-7873 | shamsrupak@gmail.com | [LinkedIn](#) | [GitHub](#)

Professional Summary

Driven and analytically minded aspiring Quantitative Trader with 3+ years of experience in applied mathematics, machine learning, and real-time data automation. Proven track record of building high-performance Python pipelines, improving search accuracy by 40%, and making fast, strategic decisions under pressure. Passionate about solving complex problems at the intersection of markets, math, and code.

Education

Stony Brook University, Stony Brook, NY

Master of Science, Engineering Artificial Intelligence

Aug 2025 - Dec 2026

Stony Brook University, Stony Brook, NY

Bachelor of Science, Applied Mathematics and Statistics

Aug 2021 - May 2025

- **Coursework:** Probability Theory, Numerical Analysis, Operations Research, Financial Mathematics, Data Mining, Object-Oriented Programming, Data Structures

Technical Skills

- **Programming:** Python, Java, R, MATLAB, SQL, VBA, C
- **Libraries/Tools:** NumPy, Pandas, Matplotlib, Seaborn, PyMuPDF, OCR, PyTorch, Git, Jupyter
- **Competencies:** Statistical Modeling, Algorithmic Thinking, Market Simulation, Machine Learning, Natural Language Processing, Retrieval-Augmented Generation (RAG), Smart Contract Analysis, Data Visualization

Experience

Extern

May 2025 - Jul 2025

AI and Automation Extern - Outamotion

Remote

- Engineered AI-powered document processing workflows that automated classification and data extraction using NLP, computer vision, PyMuPDF, and OCR, reducing manual handling by 60%
- Built a RAG-powered retrieval system with LlamaIndex to enhance search precision across unstructured mortgage data and evaluated transformer-based LLMs for model optimization
- Synthesized benchmarking results and deployment insights while evaluating top open-source AI models for document classification, with a focus on trade-offs between speed, context window, and accuracy

Stony Brook University

Jan 2025 - Mar 2025

Biomedical Informatics Bootcamp

- Applied ML techniques on biomedical datasets using Python and visualization libraries
- Collaborated on real-world healthcare problems using interdisciplinary, data-driven approaches

T-Mobile

Jul 2024 - May 2025

Retail Mobile Expert

Nesconset, NY

- Surpassed 100% of monthly sales quotas by integrating AI tools like ChatGPT and Apple Intelligence to streamline quoting and reduce interaction time
- Presented tech-integration strategy at a corporate roundtable, improving team-wide efficiency
- Recognized as a Values in Action Champion for mentoring peers and driving store performance

Extern

Sep 2024 - Nov 2024

Webacy Web3 Security Data Analytics Extern

Remote

- Conducted unsupervised ML analysis to identify smart contract vulnerabilities, improving risk categorization accuracy
- Peer-reviewed data labels to ensure 98% reliability across blockchain datasets
- Explored clustering methods for anomaly detection in transaction behavior

Excellence Community Schools

Sep 2020 - Sep 2023

Information Technology Intern

Bronx, NY

- Automated inventory tracking for 500+ IT assets and supported troubleshooting for over 200 users

Projects

Mortgage Document Classification with RAG & OCR

May 2025 - Jul 2025

<https://github.com/ShamsRupak/ai-doc-processing-suite.git>

AI & Automation Extern

- Created document pipeline using PyMuPDF and OCR to automate mortgage document classification
- Enabled semantic search across heterogeneous file types using Retrieval-Augmented Generation