Shams Rupak

New York, NY — 646-619-7873 — shamsrupak@gmail.com
LinkedIn: linkedin.com/in/shams-rupak-262906272 — GitHub: github.com/ShamsRupak

Professional Summary

M.S. Engineering Artificial Intelligence candidate (Dec 2026) with strong foundations in Machine Learning, C++ systems programming, software engineering, and scalable systems design. Experienced designing, deploying, and optimizing end-to-end ML pipelines, Retrieval-Augmented Generation (RAG) systems, and OCR-based document automation for production-style workflows. Seeking internship or full-time Software Engineering / Machine Learning roles.

Technical Skills

Programming Languages: Python, C++, Java, SQL, JavaScript, C, R
Machine Learning & Artificial Intelligence: PyTorch, scikit-learn, XGBoost, LLMs, RAG, LlamaIndex, NLP, OCR, embeddings, semantic search, clustering, regression, model evaluation, feature engineering
Systems & Data: Data Structures & Algorithms, OOP, STL, memory management, NumPy, Pandas, REST APIs, ETL pipelines, Git, Linux, performance optimization, debugging
Web: React, HTML/CSS, Chrome Extensions

Education

Stony Brook University — M.S. Engineering Artificial Intelligence Aug 2025 – Dec 2026
Machine Learning, Deep Learning, AI for Robotics
Stony Brook University — B.S. Applied Mathematics & Statistics Aug 2021 – May 2025
Data Structures & Algorithms, Object-Oriented Programming, Probability Theory

Experience

Stony Brook University — Teaching Assistant (C++ OOP) Jan 2026 – Present

- Support 30+ students in core C++ systems course covering OOP design, STL containers, dynamic memory, RAII, and debugging (segmentation faults, memory leaks)
- Lead code reviews emphasizing modular architecture, abstraction, algorithmic efficiency, and runtime performance optimization

AI Engineering & Automation Extern May 2025 – Jul 2025

- Designed and deployed modular AI document processing pipeline in Python (PyMuPDF, Tesseract OCR, NLP, LLM classification) across 1,000+ financial documents, reducing manual processing time 60%
- Implemented Retrieval-Augmented Generation using LlamaIndex, contextual chunking, and embedding-based semantic search to improve structured data extraction accuracy
- Benchmarked transformer-based models across latency, context window, and precision trade-offs to inform production-level deployment decisions

Webacy — Web3 Security Data Analytics Extern Sep 2024 – Nov 2024

- Applied unsupervised learning and clustering to detect anomalous blockchain transactions and smart contract vulnerabilities; improved labeled dataset reliability 95%+

Projects

Financial Document Classification with RAG & OCR May 2025 – Jul 2025
github.com/ShamsRupak/ai-doc-processing-suite

- Built scalable document ingestion and classification system using PyMuPDF and OCR for structured financial data extraction
- Developed embedding-based semantic retrieval pipeline enabling contextual document Q&A and modular evaluation framework for maintainability

Student Buddy — AI Chrome Extension Sep 2025 – Present
github.com/ShamsRupak/student-buddy-extension

- Engineered AI-powered Chrome extension using JavaScript and Gemini API with modular frontend architecture and scalable API integration for real-time academic assistance