

Priyansh Shah

917-680-5511 | New York | priyanshshah.2727@gmail.com | [linkedin.com/in/-priyansh-shah](https://www.linkedin.com/in/-priyansh-shah) | github.com/priyanshshahh

EDUCATION

Stony Brook University

Bachelor of Science in Applied Mathematics and Statistics

Stony Brook, NY

Aug. 2022 – May 2026

- **Awards:** Dean's List (2023–2025), Global Excellence Scholarship, Project Champion Award, Excellent Research Award
- **Relevant Coursework:** Quantitative Finance, Financial Mathematics, Data Mining, Probability & Statistics, Machine Learning, Data Analysis, Database Design, Linear Regression, Data Structures and Algorithms

TECHNICAL SKILLS

Languages: Python, SQL, R, MATLAB

Machine Learning: Scikit-Learn, XGBoost, LightGBM, Statsmodels, Pandas, Random Forest, TensorFlow, Keras, Predictive Modeling

Generative AI: LangChain, Hugging Face, Gemini API, Prompt Engineering, RAG Systems

Data Engineering: PostgreSQL, BeautifulSoup, Selenium, REST APIs, JSON/Parquet, DuckDB

Quant Finance: Algorithmic Trading, Backtesting, Financial Forecasting, Risk Modeling, Portfolio Management

Visualization & Tools: Tableau, Microsoft Excel, VS Code, HPC Cluster, Jupyter Notebook, PySpark, Git, Seaborn

EXPERIENCE

Stony Brook University (Virtually Integrated Projects)

Stony Brook, NY

Undergraduate Researcher (VIP)

Jan. 2026 – Present

- Extracted historic market and on-chain transaction data to identify Web3 adoption trends using Python and SQL, driving portfolio optimization insights through rigorous signal discovery and feature engineering.
- Led a 4-member team to forecast cross-sectional stock returns by estimating ML models on partitioned accounting and market features, utilizing **Seawulf** and **NVwulf HPC cloud infrastructure** to scale computations over massive financial datasets.

Blockchain Business Lab

Stony Brook, NY

Research Lead & Data Scientist

Aug. 2024 – Present

- Directed 4+ research initiatives and managed a team of 16 students using **Agile** workflows in **Notion**, culminating in streamlined lab operations and the authorship of a finance research paper.
- Built ensemble and deep learning forecasting models (**Random Forest**, **XGBoost**, **LSTM**) on large-scale financial datasets via systematic hyperparameter tuning, achieving high predictive accuracy in controlled experiments.
- Architected a fully automated agentic environment to extract actionable signals from financial text using **LLMs** and a custom agent architecture, enhancing simulated trading efficiency through vectorized backtesting and rigorous evaluation metrics.

Google

Remote

Google Software Engineer Program (G-SWEP) Fellow

Sept. 2024 – Jan 2026

- Selected from 2,000+ applicants to refine production-quality coding skills through 1:1 Google mentorship, engineering **Python** solutions for medium-hard algorithmic challenges using **Data Structures (Graphs/Trees)** and Dynamic Programming.

PROJECTS

Generative AI in Stock Portfolio Management | Python, LLMs, Perplexity API, WRDS

Oct. 2025

- Engineered a **SQL** pipeline extracting **WRDS** data and integrated **Perplexity** and **Gemini APIs** for portfolio construction, developing a Portfolio Validation and Cleaning framework ensuring complete index compliance.
- Executed Return Computation and Portfolio Rebalancing for simulated strategies, enforcing **risk management** rules to minimize drawdowns while capturing exceptional short-term total returns.
- Authored a URECA research paper on **LLM-based portfolio construction**, utilizing **data analysis** to evaluate prompt-engineered strategies, revealing systematic Tech sector concentration and peak Sharpe ratios.

Automated Crypto Hedge Fund | XGBoost, LSTM, vectorbt, Docker, APScheduler

Aug. 2025

- Engineered an end-to-end forecasting pipeline using **XGBoost** and **LSTM**, applying walk-forward validation and **Optuna** hyperparameter tuning to predict multi-horizon crypto returns across a Top-100 asset universe.
- Architected an autonomous agent framework via **APScheduler** to ingest market and on-chain data, implementing dynamic "Feature Decay" tracking to automatically down-weight declining predictive signals.
- Implemented Hierarchical Risk Parity and Top-K allocation algorithms within a **vectorbt** backtesting engine, minimizing drawdown risk while systematically evaluating strategies against realistic trading frictions.

Fraudulent Job Posting Detection | Python, NLP, Scikit-Learn

Oct. 2024

- Architected a fraud detection pipeline achieving **96.5% accuracy** using a **Multinomial Naive Bayes** classifier on text and metadata features.
- Engineered hybrid **TF-IDF** and **VADER** sentiment features, deploying a real-time **Dash** dashboard to visualize fraud probabilities.

Cryptocurrency Price Prediction & Analysis | Scikit-Learn, TensorFlow, Tableau, APIs

May 2024

- Engineered an end-to-end machine learning pipeline using **Scikit-Learn** and **TensorFlow** to forecast digital asset prices, implementing a suite of predictive models including **LSTM**, **ARIMA**, and regularized regressions.
- Automated historical market data ingestion via REST APIs and deployed an interactive **Tableau** dashboard to visualize model performance and unearth asset behavior patterns using **K-Means** clustering.