

Vraj Patel

vjpatel4423@gmail.com — +1-631-780-2456 — Stony Brook, NY — LinkedIn — GitHub — Portfolio

Education

Stony Brook University

M.S. in Data Science

Stony Brook, NY

Expected May 2027

Pandit Deendayal Energy University

B.Tech in Computer Engineering – **CGPA: 8.72/10**

Gandhinagar, India

May 2025

Technical Skills

Languages: Python, SQL, R, C++, JavaScript/TypeScript, Java

Frameworks: FastAPI, React, Flask, Node.js, Streamlit, REST APIs, WebSockets

Cloud/DevOps: AWS (EC2, S3, Lambda), GCP (Cloud Run), Docker, Git, CI/CD, Cloudflare

Data/ML: PyTorch, TensorFlow, Pandas, NumPy, Scikit-learn, SpaCy, OpenCV, Tableau

Big Data: Spark, Hadoop, Kafka, Airflow — **Databases:** PostgreSQL, MySQL, MongoDB, Redis

Experience

FinTech Data Scientist Intern, Flits (India)

Feb 2025 – May 2025

- Built real-time options analytics engine processing **NSE/BSE tick data** for 200+ contracts, implementing async Websocket client with automatic reconnection; reduced manual monitoring overhead by 40%.
- Engineered ETL pipeline parsing 170K-row instrument master file, applying sector classification logic (SIC code mapping) and outputting normalized PostgreSQL schema with 30+ indexed industry categories.
- Developed sector rotation tracker correlating NIFTY index movements with constituent stock performance, implementing z-score normalization to identify top decile outperformers.
- Deployed **Streamlit + Plotly dashboards** with 30s polling intervals, implementing request throttling and connection pooling to maintain 100% uptime against exchange API rate limits.

Data Analyst Intern, YHONK – Noise Pollution Mitigation (India)

May 2024 – Jun 2024

- Built distributed web scraping pipeline using Selenium Grid (4-node cluster) with rotating proxies, extracting 50K+ school records and implementing exponential backoff retry logic.
- Executed geospatial analysis on **2.6M GPS-tagged events**, implementing ST_Buffer(100m) PostGIS queries and R-tree spatial indexing in QGIS to compute violation hotspots with $O(\log n)$ lookup.
- Applied ARIMA time-series decomposition on hourly violation counts, detecting 171% weekday anomaly and 130% monsoon seasonality using sliding window z-score thresholding.
- Automated violation detection with Python + PostGIS pipeline, implementing spatial joins (ST_Within) to flag vehicles in restricted zones for law enforcement reporting.

International Research Intern, AIT Brain Lab (Thailand) — Live Demo

Jul 2023

- Fine-tuned **T5-base and FactorSum** transformers on ParaSCI dataset (350K scientific paper pairs), achieving 0.42 ROUGE-L score for abstractive summarization via transfer learning on V100 GPUs.
- Built Flask REST API with Celery task queue for async inference, implementing request batching (batch_size=8) to reduce P95 latency from 4.2s to 1.8s under concurrent load testing.
- Integrated SpaCy NER pipelines and NLTK sentence tokenization for preprocessing, handling LaTeX math rendering via MathJax and PDF extraction through PyPDF2 with fallback OCR.

Featured Project

SBU VibeCheck – Real-Time Campus Event Platform — Live Demo

Nov 2025

React, TypeScript, FastAPI, PostgreSQL, WebSockets, Gemini API, Auth0, GCP

- Built full-stack event aggregation platform in **36 hours** (Hackathon Winner), implementing **WebSocket-based real-time synchronization** across clients with sub-200ms latency.
- Architected **OAuth 2.0 dual-token flow**: Auth0 for .edu email verification, then server-minted anonymous JWTs with RBAC claims for privacy-preserving user sessions.
- Integrated **Gemini 1.5 Pro with parallel function calling**, parsing natural language to SQL queries via custom tool schemas, reducing query formation time by 85% vs traditional search.
- Designed stateful sentiment reporting system with contextual prompts, broadcasting delta updates via Redis Pub/Sub to maintain real-time consistency across the campus network.
- Deployed React SPA on **Cloudflare Pages** with FastAPI backend on **GCP Cloud Run (auto-scaling)**, implementing TLS termination and CORS policies for production security.

Certifications

Azure Administrator Associate, AWS Certified Cloud Practitioner, Alteryx Designer Core