Rishi Guptha Mankala

Willing to relocate | [LinkedIn](#) | [Portfolio](#) | 934-263-3087 | rishiguptha.mankala@gmail.com

WORK EXPERIENCE

Nevara AI New York City, NY
AI/ML Engineering Intern *Jun 2025 - Aug 2025*

- Built an AI sales-call auditing MVP (Next.js/TypeScript, Supabase, OpenAI, Deepgram) for a coach overseeing 100+ reps. Automated notes-to-report in 90–140s, cut per-rep review time ~75%, and achieved 3.5–3.75/5 satisfaction in a pilot with 3 reps × 3 calls.
- Owned CI/CD and serverless delivery with GitHub Actions to Vercel and FastAPI on AWS Lambda (Mangum, Docker). Tuned memory and reserved/provisioned concurrency to stabilize p95 latency at 30s (transcribe), 30–50s (notes), and 60–90s (report).
- Shipped timestamp-synced media player, transcript, and notes (HLS.js, SSE, React Query). Added click-to-seek and auto-highlighting that improved audit navigation throughput by 2x and reduced context switching.
- Authored Next.js APIs (/transcription, /ai-notes, /gong-scrape, /fathom) with SSE progress, XML validation, and versioned Supabase storage (Auth/Postgres/Storage via Prisma). Integrated Gong and Fathom and led AI-report prompt and context engineering (token cap <128K) to keep cost < $0.75/call and cut retries by ~30%.

Stony Brook University Stony Brook, NY
Data Science Researcher *Jan 2025 - May 2025*

- Built Python-SQL pipelines for S&P 500 time-series data and trained LSTM, CNN, and SVM models using TensorFlow/PyTorch, improving model accuracy by 20% and risk scoring by 15% through Bayesian inference and backtesting.
- Delivered dashboards (Tableau, Python) integrating predictive outputs and risk metrics, enabling 30% faster portfolio decisions and surfacing volatility trends via SVM optimization and feature extraction.

Kimberly Clark Bengaluru, India
Data Engineering Intern *Jan 2023 - July 2023*

- Automated metadata tagging and validation in Collibra using Python/PostgreSQL workflows, earning 7 certifications and boosting governance team productivity by 25% while reducing manual effort by 30%.
- Migrated 10K+ assets into Collibra via Excel-based ETL with 100% data integrity, resolved 15% pipeline inconsistencies, and built Power BI/Tableau dashboards to visualize KPIs and reduce reporting errors by 18%.
- Designed and deployed 5+ advanced Collibra workflows to support real-time data quality checks and lineage validation, enabling faster stakeholder reporting and improved oversight across governance domains.

Matchday AI Hyderabad, India
Data Scientist Intern *Jan 2022 - Dec 2022*

- Built computer vision pipelines with OpenCV and deep learning to detect field lines, ball position, and player movement from broadcast footage, achieving 90%+ event tracking accuracy.
- Applied Homography Transformation to convert dynamic camera feeds into 2D field projections, powering spatial analytics for tactical coaching insights.
- Developed a full-stack annotation platform (Flask, JS, SQLite), reducing manual labeling time by 40% and enabling deployment of data-driven dashboards for 6+ ISL teams via Star Sports.

PROJECTS

NutriSmart: RAG-Driven Insights for Fitness and Nutrition

- Delivered 30% faster fitness insights and 40% lower latency by deploying a RAG service (Sentence Transformers + FAISS) containerized with Docker on AWS EC2/FastAPI, indexing 1.2K+ papers.

NYC Taxi Insights

- Improved fare/duration accuracy 30% and query performance 35% by engineering a GCP ETL for 10M+ rides, modeling star-schema marts, and training LightGBM with BERT-derived features for dynamic pricing.

Sales Insights Dashboard

- Raised sales-forecast accuracy 30% and cut prediction turnaround 15% by building a Power BI hub over CRM/ERP using SQL + DAX, enabling faster pipeline planning and territory targeting.

Spotify Playlist Analytics

- Enabled daily KPIs by designing a serverless ETL (AWS Lambda/Glue/S3) for 18K+ playlist logs and optimizing Athena SQL to reduce processing time 40%, informing genre strategy and retention.

EDUCATION

Stony Brook University Stony Brook, NY
Masters of Science, Data Science

SKILLS

Programming Languages: Python, Java, SQL, TypeScript/JavaScript, R
Data Science & Machine Learning: PyTorch, TensorFlow, scikit-learn, Pandas, NumPy, NLP, ASR (Deepgram), Large Language Models (OpenAI API), Prompt & Context Engineering, Statistical Analysis/Modeling, Deep Learning
Cloud Platforms: AWS (Lambda, S3, EC2, Glue, Redshift), GCP (BigQuery, Cloud Storage, Vertex AI), Vercel, Supabase (Auth/Postgres/Storage)
Data Platforms & Pipelines: Snowflake, Apache Spark, Airflow (Astronomer), dbt, Data Modeling, PostgreSQL
Backend & Web: Next.js (React), FastAPI, RESTful APIs, Server-Sent Events (SSE)
Data Visualization: Tableau, Power BI, Looker Studio, AWS QuickSight, Matplotlib
DevOps & Tools: Docker, Linux, Git/GitHub, GitHub Actions (CI/CD), Microsoft Excel, Google Sheets, Collibra