

Siddhant Kasture

New York | siddhantkasture7@gmail.com | (680)-285-8077

github.com/sid12super | linkedin.com/in/siddhantkasture/ | sid12super.github.io

Summary

Data Scientist (May 2026 MS) focused on production model serving, ML pipelines, and agentic systems. Built CI/CD-driven data pipelines and model-serving integration for a client-validated forecasting platform, shipped an open-source production AI agent, and deployed an independent forecasting system end-to-end. 4.0 GPA, Phi Kappa Phi at Syracuse. Comfortable owning ML systems from data pipeline through deployment.

Education

Syracuse University, School of Information Studies, MS in Applied Data Science

May 2026

- **GPA: 4.0/4.0** | Phi Kappa Phi Honor Society (Inducted 2026)

- **Relevant Coursework:** Applied Machine Learning, Deep Learning, Natural Language Processing, Advanced Database Management, Data Warehouse, Text Mining

University of Mumbai, BE Electronics and Telecommunications

May 2023

Experience

AI Evaluation Specialist - Handshake AI (Remote)

Dec 2025 – Present

- Build evaluation harnesses and behavioral test suites for LLM and VLM outputs, contributing metrics that drove an 80% improvement in model accuracy and safety benchmarks.
- Reduced model hallucination rates by 15% by instrumenting drift monitoring across release cycles and translating noisy production telemetry into actionable engineering tickets.

Research Team Lead - HyperQuark Intelligence Labs (Remote)

Apr 2026 – Present

- Lead the LLM reasoning workstream within the lab's research group, coordinating experimental design and weekly research direction across collaborators.
- Drove team research on a DAG-based job-weighted skill-assessment framework, improving ISCO classification accuracy by 25% over keyword-based baselines.

Research AI Engineer - NEXIS Technology Lab

Sep 2024 – Dec 2025

- Built and fine-tuned a BERT transformer pipeline on 50,000 records via Hugging Face, achieving 84% accuracy and a 12% performance gain through systematic hyperparameter optimization.
- Engineered reproducible training and evaluation workflows, benchmarking against multiple baselines for robustness.

Projects

ProAg AI Market Research Tool - Pro-Ag Data Science Competition

Jan – May 2026

- Owned full-stack model-serving integration for a 3-person team: built the serving layer (confidence-tier, abstention, and SHAP-driver logic) connecting the team's XGBoost models into a live API and dashboard.
- Built the GitHub Actions CI/CD pipeline auto-refreshing USDA data daily (14 steps, 4-min verified runtime, auto-deploy to Render), serving models that scored 79.4% hog / 65.1% cattle high-conviction accuracy on a gated holdout.
- Ran the initial EDA surfacing the predictive signals (basis reversal, momentum, feed-cost drivers) that shaped the model feature sets.

MCP Deep Researcher - Production AI Agent

Mar 2026

- Designed and shipped an open-source production agent on LangGraph and GPT-4o with multi-turn memory, configurable tool use, and source credibility scoring for automated result ranking.
- Engineered SHA-256 content-hashed caching to eliminate redundant API calls; released as an MCP server integrated with Claude Desktop and Cursor IDE.

OpenAg Risk Twin - Commodity Forecasting Infrastructure

Jun 2026 – Present

- Solo-built and deployed a public forecasting system with GitHub Actions CI/CD, owning regime detection (dual-PELT, 7 validated regimes) and walk-forward evaluation (1,586 OOS predictions) against a random-walk baseline.
- Live at openag-risk-twin.vercel.app, demonstrating end-to-end ownership from data pipeline through model serving to deployment.

Betting Edge - Multi-Agent Decision System

Oct – Dec 2025

- Built a multi-agent application integrating a DQN-based predictive engine via LangChain orchestration with a ChromaDB retrieval pipeline; engineered a value algorithm identifying 15% positive expected value across 1,000+ matches (8.5% backtested ROI).
- Implemented responsible-AI guardrails via a fine-tuned DistilBERT classifier, mitigating algorithmic bias with audit logging for iteration.

Skills

Languages & Engineering: Python (primary), SQL, R, JavaScript / TypeScript, REST APIs, Git, FastAPI

ML & Deep Learning: PyTorch, TensorFlow, Scikit-learn, XGBoost, Hugging Face Transformers, BERT Fine-Tuning, Deep Q-Networks, CNNs

Agentic & LLM Systems: LangChain, LangGraph, MCP Server, Retrieval-Augmented Generation, Tool Use & Reasoning Loops, Multi-Agent Orchestration, LLM Evaluation

MLOps & Deployment: GitHub Actions (CI / CD), Docker, Model Serving, Walk-Forward Evaluation, Drift Monitoring, Render / Vercel Deployment

Cloud & Data: AWS SageMaker, Amazon S3, Snowflake, DuckDB, ChromaDB, Dask, Jupyter