

SIDDHESH SATHE

Senior AI Engineer

London, UK | thesidsat@gmail.com | [linkedin.com/in/thesidsat](https://www.linkedin.com/in/thesidsat) | github.com/thesidsat | [kaggle.com/sidcodegg](https://www.kaggle.com/sidcodegg)

UK Right to Work: ILR (Global Talent Visa)

SUMMARY

AI and ML engineer with 7+ years building production ML and LLM systems in fintech, quantitative trading and wealth tech. Author of Tine, an open source branching notebook runtime in Rust (9K+ PyPI installs) with a native MCP interface for AI agents. Built LangChain and Azure OpenAI analyst tools, credit default models on trade finance data, and live quantitative trading strategies. Kaggle Notebooks Expert (Top 1.7% worldwide).

TECHNICAL SKILLS

Languages: Python, Rust, SQL, Scala

GenAI and LLM: LLM application development, RAG and semantic retrieval, LangChain, LlamaIndex, agentic workflows, Model Context Protocol (MCP), Azure OpenAI, vector search, NLP (NER)

ML and Data Science: PyTorch, sklearn, Bayesian modeling, recommender systems, gradient boosting, time series forecasting

Data and Infrastructure: AWS, Azure, GCP, Databricks, Spark, PostgreSQL, Terraform, Airbyte, Docker, Kubernetes

MLOps and Backend: MLflow, DVC, FastAPI, Flask, CI/CD, Git

EXPERIENCE

Senior Data Scientist

Dec 2025 to Present

Serve Legal

London, UK

- Lead data science for performance intelligence products covering ~**4,500+ UK retail sites** (Lidl, M&S, Tesco, Asda) under a no win no fee model, reporting to the Director of Technology.
- Built a **Bayesian risk scoring model** that ranks stores by compliance and operational risk to prioritise audits and target revenue; early trials lifted client compliance rates by **5 to 10%**.
- Designed a PostgreSQL star schema over audit data, enabling multidimensional analysis across store, region, audit type and time to power reporting and modelling.

Founding Data Scientist and AI Engineer (Contract)

Apr 2025 to Nov 2025

Wealthly Ltd.

London, UK

- Sole data and AI hire (contract) reporting to the CEO; built the wealth tech data and ML foundations and defined the first engineering plan.
- Integrated the **Intelliflo** adviser platform API for client and portfolio data.
- Built a recommender and a proprietary client and advisor scoring system; piloted with independent financial advisors.

Data Scientist and Data Engineer

May 2022 to Mar 2025

TradeTeq Ltd.

London, UK

- Built ML models predicting repayment and flagging likely **defaults** on trade finance transactions for largely unrated SMEs, reporting to the Chief AI Officer.
- Modelled repayment probability across buckets and lifted the primary credit model's **AUC from ~0.58 to 0.70**, using features including NACE industry codes, repayment history, invoice amount, issue and expiry dates and interest rate.
- Built a LangChain and Azure OpenAI data analyst agent for natural language querying over trade finance data.
- Reengineered vector embeddings for semantic search, cutting ~**£45k** in annual infrastructure cost while preserving retrieval quality; built Scala and Spark ETL on Databricks (medallion lakehouse) with Terraform and Airbyte.

AI Engineer, Quantitative Trading

Jan 2020 to May 2022

Redev Ltd.

Leeds, UK

- Developed and backtested profit taking and risk management strategies deployed to the live trading engine across **FX and crypto books (multimillion notional)**, delivering a backtest validated ~**15% increase in gross profit**.
- Cut nightly processing runtime by **40% (from 5h to 3h)** through backend and AWS optimisation.
- Built an MLOps pipeline (MLflow, DVC, Kubernetes) and containerised Flask REST inference APIs.

Data Scientist (Independent Consultant)

Jul 2017 to Sep 2019

Upwork (Top Rated)

Remote

- Delivered ML solutions to **20+ international clients** across media, mobility and healthcare; achieved Top Rated status.
- For **Al Jazeera**, built an NER pipeline quantifying audience time spent per topic to inform editorial and content strategy; built demand forecasting and allocation models rebalancing a bike share fleet across time slots.

OPEN SOURCE AND SELECTED PROJECTS

- Tine: Branching Notebook Runtime for AI and Humans** | *Rust, Python, MCP, WebSockets* | [tine.sh](#)
- Open source (Apache 2.0) runtime with **9K+ PyPI installs** where experiments branch like code, letting humans and AI agents attach to one local runtime and share state, logs and outputs.
 - Built the Rust execution engine: projects each branch into a **DAG** (petgraph) with topological scheduling, plus a SQLite content cache keyed on code, input and environment lockfile hashes that recomputes only changed cells and rehydrates cached artifacts.
 - **MCP native**: ships a stdio Model Context Protocol adapter (VS Code, Cursor, Claude) alongside browser and desktop clients over a shared HTTP and WebSocket backend.
- Northtara.ai: Consumer LLM and Voice Application** | *Python, LLMs, RAG, Skyfield*
- Solo built a consumer LLM and voice product over a low resource historical **Sanskrit** text corpus, pairing retrieval based interpretation with an open source astronomical compute engine (NASA JPL ephemeris via Skyfield).
- Kaggle Notebooks Expert: Top 1.7% Worldwide**
- Global rank **870 of 51,474**; notebooks upvoted by Kaggle Grandmasters.

EDUCATION

University of Leeds, UK <i>MSc, Advanced Computer Science (AI)</i> • <i>Merit</i>	<i>2019 to 2020</i>
Savitribai Phule Pune University, India <i>B.E., Computer Engineering</i> • <i>First Class</i>	<i>2013 to 2017</i>