

# Shiva Kharbanda

Delhi, India | sh\*\*\*\*1@gmail.com | +91 844-XXXX-744 | linkedin.com/in/s\*\*\*\*-\*\*\*\*a | github.com/s\*\*\*\*\*nda

## SUMMARY

Seasoned software engineer with 3 years of experience building **production data pipelines, low-latency APIs, and enterprise data apps** using **Databricks, FastAPI, Django, and modern cloud tooling**. Focused on data platforms and applied AI, delivering scalable, high-impact solutions in production.

## PROFESSIONAL EXPERIENCE

### Software Development Engineer

Sciqiq

Nov 2023 – Present

#### GA Call Transcripts Pipeline

- Automated processing of **6,000–8,000 call transcripts/day** with Databricks + Azure Data Factory, enabling intent detection, sentiment analysis, QC, and complaint checks.
- Scaled NLP inference via **Azure OpenAI Batch API** with **Pydantic-based validation** to handle high-volume workloads reliably with consistent throughput.
- Orchestrated pipelines with **Automic** and coordinated stakeholder requirements for seamless production rollout.

#### Master Data Management Tool

- Delivered a Django/DRF web app with **role-based access control**, enabling secure, multi-team usage.
- Built **approval workflows, dynamic item tracking, and ingestion pipelines** validating thousands of records with full audit trails.
- Launched **SLA/TAT dashboards** to surface bottlenecks and cycle times, helping ops leads reduce delays and improve throughput.

#### Fare Rules Extraction

- Built a FastAPI service powered by an **open-source LLM deployed with vLLM** to ingest and validate airline fare rules from XML, cutting end-to-end parse time from **2–4 minutes to ~13 seconds (~95% faster)**
- Used async I/O to parallelize category processing; added retries and dead-letter handling for bad inputs
- Designed MongoDB schemas and Django dashboards for **run health, error drill-downs, and rule coverage**
- Onboarded **50+ airlines (~90% market coverage)** and shipped to production with monitoring and alerts

#### Sciqiq Platform (AutoML Module)

- Owned the front end for an AutoML module: multi-sheet experiments, dataset/version selection, and run tracking
- Partnered with stakeholders to design workflows and acceptance criteria; wrote specs and UI prototypes
- Fixed **200+ bugs** and UX issues, reducing support tickets and improving session stability

### Data Engineer

Daaslabs

Nov 2022 – Oct 2023

#### Assessment Tool

- Built an enterprise data-maturity assessment app (Django, DRF, HTML/CSS/JS) with authentication and admin tooling
- Created an adaptive questionnaire engine that calibrates difficulty from prior answers to increase scoring fidelity

#### ONDC (Open Network for Digital Commerce)

- Led a team of three to develop a Selenium-based collector for **50+ vendor dashboards**, loading results into PostgreSQL
- Productionized with Jenkins pipelines and schedules; added checkpointing and idempotent loads for consistency

## TECHNICAL SKILLS

Languages:	Python, JavaScript, Go (intro)
Frameworks & Libraries:	Django, Django REST Framework, FastAPI, Flask, React, Next.js, LangChain, LangGraph, VLLM, PydanticAI
Data & Infrastructure:	Databricks, Azure Data Factory, Azure Data Lake, Amazon S3, ETL/ELT, SageMaker
Tools & Platforms:	GitHub, Azure DevOps, Docker, Docker Compose, Jenkins, PostgreSQL, MySQL, MSSQL Server, MongoDB

## PERSONAL PROJECTS

### AnalytiQ [GitHub]

Streamlit app for data quality, preprocessing, and AutoML with GenAI assist (Streamlit, AutoML, Generative AI).

### Prompt ACE [GitHub]

Chrome extension for prompt versioning with secure API key storage and AI enhancements (Manifest V3, JavaScript, LangChain.js, Google Gemini).

### sfinance & SFinance MCP Server [PyPI] [Docs] [GitHub] [MCP GitHub]

Published Python package on PyPI for extracting financial tables into pandas DataFrames, fully documented on ReadTheDocs, plus a separate MCP server exposing Indian stock-market data programmatically (Selenium, BeautifulSoup, Pandas, Model Context Protocol).

## EDUCATION

### BTech in Computer Science and Engineering – MDU University

2018 – 2022