# VISHIT JIWANE

**AI/ML Engineer and Team Lead | LLMs, Agents, Evaluation and Voice AI**

+44 7887 361789 | vishitjiwane18@gmail.com | London, UK

vishitjiwane.com | linkedin.com/in/vishit-jiwane-992764244 | github.com/vishit18

## PROFESSIONAL SUMMARY

AI/ML engineer and team lead with experience building LLM products, agent workflows and voice AI for client platforms. Joined Encegen AI Labs as an intern in August 2025 and was leading its AI team by February 2026, owning planning, architecture and delivery. Built the evaluation and guardrail layer for client LLM features. Founder of Novessa, and built GulfSafe solo.

## TECHNICAL SKILLS

**Languages:** Python, SQL

**ML and Statistics:** PyTorch, scikit-learn, XGBoost, NLP, anomaly detection, sentiment analysis, speech recognition, voice AI

**LLMs and Agents:** RAG, multi-agent orchestration, LangChain, LangGraph, CrewAI, Google ADK, Microsoft Agent Framework, MCP, A2A, structured output (Pydantic), pgvector, self-hosted inference (Ollama)

**Evaluation:** Guardrails, Ragas, azure-ai-evaluation, LangSmith, Arize Phoenix, OpenTelemetry, hallucination testing, pytest

**Microsoft Azure:** Foundry, Container Apps, Event Hubs, Cosmos DB, AI Search, AI Services (Translator, Speech, Content Safety), Static Web Apps, Blob Storage, Container Registry, Application Insights, Managed Identity, Bicep, azd

**AWS and DevOps:** AWS (EC2, S3, SSM, Lambda), Docker, Linux, Git, GitHub Actions, CI/CD, MLOps

**Backend and Frontend:** FastAPI, Flask, REST APIs, PostgreSQL, SQLAlchemy, Supabase, React, Next.js, Tailwind

## EXPERIENCE

**AI/ML Engineer and Team Lead, Encegen AI Labs** Feb 2026 to Present

- Lead the AI team across several client products at once. I set sprint goals, own architecture and code review, report progress to stakeholders, and have run 10+ technical interviews that fed into hiring decisions.
- Designed and delivered a multi-tenant voice AI SaaS platform with agentic LLM workflows in LangChain and LangGraph, real-time voice pipelines, and RAG prompt infrastructure tuned for latency and accuracy under production load.
- Own the core AI pipeline on a document intelligence platform that processes 10,000+ scanned legal and property documents. Replaced OCR with multimodal LLM extraction, and added confidence-based validation, automated risk flagging and reviewer sign-off.
- Built the evaluation and guardrail layer across client LLM features: input, output and retrieval guardrails, regression suites, hallucination testing, and tracing through LangSmith and OpenTelemetry.
- Reworked CV screening to handle 100 resumes per batch instead of 20, and cut processing time from 200 to 30 minutes (85% less).

**AI/ML Engineer Intern, Encegen AI Labs** Aug 2025 to Feb 2026

- Built LLM pipelines, prompt workflows and multi-agent orchestration, with live AWS deployments across client environments.

**Founder, Novessa** novessa.tech | Aug 2026 to Present

- Founded an AI research company working on personality modelling and voice cloning to recreate deceased individuals, plus agentic and voice AI.

**Software Engineer (promoted from Intern), Dexter Intelligence** Oct 2024 to Aug 2025

- Contributed to two mobile apps: a children's educational app with 60,000 downloads and a road navigation app with 10,000.
- As an intern, built full-stack features for a real estate platform and a resort booking site, including search and data filtering.

**Software Engineer Summer Intern, Al Qoud Enterprise** Jul 2021 to Sep 2021

- Built features for a restaurant and food delivery app in a small team. First commercial work in a shared codebase.

## PROJECTS

**Automated Credit Memo Agent** github.com/vishit18/credit-memo-agent

- Pulls 10-K filings from SEC EDGAR, extracts financial line items, computes credit ratios and writes a structured credit memo. Extraction is scored against XBRL ground truth. The ratio calculations are plain deterministic code, and only the narrative uses an LLM.

**Multi-Agent Platform on Azure, Microsoft x Premier League Hackathon** github.com/vishit18/Novessa-Pitchside

- Built a Producer, Analyst, Verifier, Narrator and Gate workflow in Microsoft Agent Framework on two Foundry deployments (a small chat model and a reasoning model), scored with azure-ai-evaluation, with verified output cached in Cosmos DB.
- Shipped the API, my own MCP server and an A2A endpoint as three Container Apps from one image (1 to 3 replicas). Event Hubs carried 19,830 load-test messages with none lost. Deployed with Bicep, azd and GitHub Actions, using Managed Identity instead of keys.

**Agentic HR, Hiring Workflow System** github.com/vishit18/agentic-hr

- Turns CVs into deduplicated profiles and ranks candidates for each hiring request, with an explanation for every ranking. A named recruiter must approve each shortlist. Runs fully local (Ollama, pgvector) with schema-checked outputs, quotes verified against the CV, and a code fallback if the model is down.

**GulfSafe, Solo-built AI Intelligence Platform** gulf-safe.com

- Six-engine ML pipeline (sentiment, Z-score anomalies, trends, threat index, forecasting, event clustering) running every 15 minutes over live feeds from 12 Gulf countries. Shipped with Stripe billing and OTP auth to 25 users.

## EDUCATION

**MSc Artificial Intelligence and Robotics, University of Hertfordshire, UK (2:1)** Sep 2023 to Feb 2025

**BSc Computer Engineering, University of Bahrain** Sep 2018 to Jan 2023

Capstone: MBTI Speech Recognition, real-time personality classification from speech. First prize and national award nomination.