



# Heet Shah

LLM Ops Engineer

✉️ shahheet85@gmail.com • 📞 +91 9104871425 • 🌐 https://www.linkedin.com/in/heet-shah-9ab2301b0 • 📍 Ahmedabad, Gujarat

LinkedIn: heet-shah-9ab2301b0 • GitHub: heetshah16

Engineer with deep experience fine-tuning LLMs (SFT, RLHF, PPO) and building deployment pipelines with gRPC, Redpanda, and Docker. Skilled at optimizing inference for production workloads.

## Work Experience

### LLM Ops Engineer

Aug 2025 - Present

PiHex Labs

Implementing scalable LLM deployment and optimization pipelines for client products.

### AI Developer

Jun 2023 - Aug 2025

Kookree.ai

Led fine-tuning of vision and language models and built high-throughput event streaming setup for multi-agent inference.

- Pretrained/fine-tuned models like Monoloco, Fairface, Detectron.
- Created RTSP → Redpanda pipeline with gRPC micro-services.

## Education

### Bachelor of Technology in Electrical Engineering

Aug 2020 - May 2024

IIT Jodhpur

CGPA 6.8

## Skills

### Model Training & Tuning (Advanced)

SFT, RLHF, PPO, TensorFlow, HuggingFace

### Backend & Infra (Intermediate)

gRPC, Redpanda, Docker, Django, Python

### Computer Vision (Intermediate)

Detectron, OpenCV

## Projects

### Speech Emotion Recognition

Real-time emotion classifier built in Streamlit.

Technologies: TensorFlow, Streamlit

Roles: Developer

## Languages

English (Fluent)