

Pavan Tejavath

+91-7893936732 | hirepavan@thetejavath.com | LinkedIn | thetejavath.com

PROFESSIONAL SUMMARY

AI Engineer with expertise in **LLMs, RAG pipelines, embeddings, and vector databases (Pinecone, FAISS, Weaviate)**. Proven track record of delivering production-grade **AI assistants, chatbots, and scalable ML systems** using **Python, FastAPI, LangChain, and AWS**. Strong background in building guardrails, evaluation workflows, and full-stack integrations with React/Next.js.

EXPERIENCE

Optivus Technologies

Hyderabad, India

AI Engineer

Feb 2025 – Present

- Design and implement **LLM-powered features** using RAG pipelines and vector DBs, ensuring correctness, low latency, and cost efficiency.
- Ship **production-grade FastAPI microservices**, integrated with Next.js frontend, with full observability (structured logs, traces, metrics).
- Built guardrails with structured outputs, validation, caching, and prompt-safety to improve retrieval and answer faithfulness.
- Collaborate with founders and product teams, taking end-to-end ownership from design to deployment and documentation.

Suviko LLC

Remote

Software Developer

Jul 2024 – Feb 2025

- Developed cross-platform apps (**Fish Pond, CaféConnect**) using Flutter, Firebase, and FastAPI — scaled to 1500+ MAUs.
- Integrated APIs, push notifications, analytics dashboards, and real-time sync pipelines across event applications.

The Tejavath (Founder)

Remote

Founder / AI + Full Stack Engineer

Feb 2023 – Feb 2025

- Delivered AI chatbots and digital platforms for clients including **JarongMedia, Bismillahalalfoods, OurTimeJewelers, and SE Group Inc.**
- Built assistants with **LangChain, Pinecone, RAG**, enabling semantic search, memory-based reasoning, and multi-agent orchestration.
- Integrated AI features into React/Next.js + FastAPI stacks, deployed on AWS (Lambda, EC2) with CI/CD pipelines.

George Mason University

Fairfax, VA

AI Research Assistant

Sep 2022 – May 2024

- Designed NLP pipelines with HuggingFace Transformers for information extraction; built RAG prototypes improving accuracy by 92%.
- Benchmarked FAISS, Weaviate, Qdrant for semantic search and automated ingestion workflows using AWS Glue + Athena.

PROJECTS

PromptFlows Studio | LangChain, Pinecone, FastAPI, Streamlit

2024

- Built a platform for **RAG + agent workflows** with CoT prompting, LangGraph orchestration, and retrieval evaluation.

RecruitAI | FAISS, Flask, React

2024

- AI-powered recruitment assistant prototype: automated candidate screening via embeddings + semantic search, integrated into full-stack app.

myai.thetejavath.com | Pinecone, LangChain, Google Calendar API

2024

- Personal GenAI assistant for scheduling, semantic search, and conversations with TTS + RAG, deployed on cloud.

CERTIFICATIONS & STACK

Certifications: AWS Cloud Practitioner, HackerRank Advanced DSA, Python PCAP (Cisco)

Stack: Python, PyTorch, TensorFlow, HuggingFace, LangChain, RAG, Pinecone, FAISS, Weaviate, LangGraph, CrewAI, Prompt Engineering, FastAPI, Flask, React, Next.js, Docker, AWS (Lambda, S3, EC2), PostgreSQL, pgvector, CI/CD, Neo4j

EDUCATION

George Mason University

M.S. in Information Systems

Fairfax, VA

Aug 2022 – May 2024

KL University

B.Tech in Computer Science

Hyderabad, India

Jul 2018 – May 2022