

# PRATIK JADHAV

AI/ML/LLM Engineer | Software Engineer | Full-Stack/Backend Engineer  
San Francisco, CA | +1 (312) 678-6684 | pratik.jh2017@gmail.com

Portfolio: [pratikjh.com](https://pratikjh.com)  
GitHub: [github.com/PratikJH153](https://github.com/PratikJH153)  
LinkedIn: [linkedin.com/in/pratikjh](https://linkedin.com/in/pratikjh)

## WORK EXPERIENCE

### VIBEO INC.

Chicago, IL

Co-founder & Founding Engineer (Qualitative Research - B2B) [Website](#)

10/2025 – Present

- Pivoted Vibeo to B2B qualitative research, building a multi-agent orchestration platform that automates session analysis, insight extraction, and knowledge graph integration, cutting research timelines from **3 weeks to 3 days**.
- Built a custom analysis pipeline per study with shareable outputs delivered within **24 hours**, using Python (FastAPI), LangGraph, Agentic RAG with Neo4j knowledge graphs, Redis, and multi-model LLM orchestration.
- Led **paid pilot programs** with **Dyson, Visa, Fitness First, and Kraft Heinz**; generating **\$5K in revenue** and demonstrating clear ROI, while producing over **30,000+ insights** and **100+ shareable reports and decks**.
- Accelerated real-time AI inference, processing **20,000+ transcripts per study**, enabling multilingual transcription, semantic embeddings, summarization, cross-session comparison, and trend detection with integrated chat functionality.

Co-founder & Founding Engineer (Nightlife Video Discovery App - B2C) [Demo](#)

02/2024 – 09/2025

- Led a 4-person engineering team to spearhead a B2C AI-powered video discovery app using Flutter; secured a **\$25,000 Google Cloud grant** and scaled backend to support 25+ concurrent users with <100ms response time.
- Built a NestJS backend supporting end-to-end logic from ingestion to real-time processing for live recommendations.
- Constructed a real-time vision pipeline analyzing **50+ daily** video uploads and **150+ ambience parameters** with vision models, Pinecone for vector retrieval, Langchain for orchestration, and Neo4j for graph-based recommendations.

### WAJOBA AI LLC.

Atlanta, GA

Backend Developer Remote Intern

02/2023 – 06/2023

- Improved backend modules using Node.js and optimized complex MongoDB aggregation pipelines to facilitate the automated generation of **200+ monthly customized reports**, improving data processing throughput by 20%.
- Streamlined data pipelines by integrating Airbyte, managing the ingestion of **50k+ records**, developing custom Airbyte source connectors, and automating high-volume user communications via SendGrid.

## TECHNICAL SKILLS

**Languages & AI/ML:** Python, TypeScript, C++, Dart, RAG, LangGraph, QLoRA, FAISS, Pinecone, Pandas, NumPy

**Full-Stack:** React, Next.js, Flutter, BLoC, Redux, Node.js, FastAPI, Express, NestJS

**Backend & Data:** PostgreSQL, MongoDB, Redis, Neo4j, Kafka, RabbitMQ, REST/GQL, Microservices, System Design

**Cloud & DevOps:** AWS/GCP, Docker, Terraform, CI/CD, Prometheus, Grafana, OpenTelemetry, Langfuse

**Testing & Tools:** PyTest, Jest, Clerk

## PROJECTS

**GITAS – AI-POWERED GITHUB CODE INSIGHTS PLATFORM** [GitHub](#)

05/2025 – 06/2025

- Coded a full-stack AI platform (Next.js, tRPC, Prisma) integrating GitHub and Gemini to enable **automated commit-level code understanding** via diff-based analysis and structured summaries (features, bugs, risks).
- With a RAG-based **“Ask the Codebase”** feature using LangChain and pgvector, embedding repository files and enabling semantic search with context-aware, streaming Q&A over large codebases.
- Engineered a scalable backend handling **100,000+ lines of code**, including repository indexing, and query handling; integrated secure GitHub access for private repos with persistent data models for commits, embeddings, and Q&A.

**ENTHEM – GEOSPATIAL SOCIAL DISCOVERY APP** [GitHub](#) | [Demo](#)

06/2024 – 09/2024

- Developed a cross-platform Flutter application utilizing a Node.js, TypeScript backend and Neo4j, enabling **real-time** encrypted chat, location-based clustering, and **graph-based** personalized social recommendations.
- Enhanced query performance by implementing **Spatial Indexing** and **Index-free Adjacency** to traverse nodes, enabling proximity-based user grouping and **hyper-local** discovery.

## EDUCATION

### ILLINOIS INSTITUTE OF TECHNOLOGY

Chicago, IL

Master's in Computer Science (Concentration in AI & ML) | GPA: 3.8/4.0

08/2023 – 05/2025

Relevant Coursework: DBMS, Distributed Computing, Deep Learning, Machine Learning, DS/Algorithms, Big Data

Served as Graduate Teaching Assistant for Mobile Application Development (CS 442)

### D.Y. PATIL INTERNATIONAL UNIVERSITY

Pune, India

Bachelor of Technology in Computer Science (Specialization in AI & ML) | GPA: 3.85/4.0

08/2019 – 06/2023

## LEADERSHIP & INNOVATION

**SHAPESHIFT CONFERENCE** [Conference Details](#)

Chicago, IL

Project Builder & Presenter

05/2025

- Designed an RAG-driven “Live Professor” utilizing 3D avatar to bridge physical and virtual learning environments.
- Showcased technical implementation to Fortune 500 design executives, earning recognition for immersive education.

**MULTIDISCIPLINARY INNOVATION WORKSHOP** [GitHub](#) | [Demo](#)

Chicago, IL

Team Leader (Institute of Design)

01/2025 – 05/2025

- Collaborated with Steelcase Inc. to prototype AI-integrated physical spaces, optimizing student-university interaction.
- Architected a multi-modal AI study room featuring agent-based interaction and a gamified UX to enhance learning.