

# Raj Kapadia

✉ Email: [raajforyou@gmail.com](mailto:raajforyou@gmail.com)  
☎ Phone: 9558515995  
🌐 GitHub: [View Profile](#)  
🌐 Portfolio: [View Portfolio](#)

## Professional Summary

AI & Generative AI Engineer with **6+ years of experience** specializing in **LLM fine-tuning, NLP, AI-powered chatbots, and deep learning models**. Strong expertise in **GPT-based applications, text generation, vector databases, and AI model inference optimization**. Experienced in **cloud-based AI deployments (AWS, GCP), API development, and production-ready AI architectures**. Passionate about **building scalable, intelligent AI systems** to enhance automation and decision-making.

## Skills

### Generative AI & NLP

- Large Language Models (GPT-4, LLaMA, Claude, Falcon)
- Fine-tuning & Inference Optimization (Hugging Face, LoRA, PEFT)
- Text Generation, Summarization, Question Answering

### AI/ML/DL Frameworks

- TensorFlow, PyTorch, LangChain, OpenAI API, Hugging Face Transformers
- Vector Databases (Pinecone, FAISS, Weaviate, ChromaDB)

### AI-Powered Chatbots & Agents

- Google Dialogflow ES/CX, RASA, LangChain Agents
- Conversational AI, Multi-Turn Dialog Systems

### Backend & Cloud for AI

- Python (Flask, FastAPI), Node.js (Express)
- AWS (SageMaker, Lambda, S3), GCP (Vertex AI, Cloud Functions)

## Work Experience

### Let's Enkindle | *Lead AI & Generative AI Engineer*

📅 Sep 2023 - Present

- **Designed & fine-tuned custom LLMs**, optimizing inference time by **50%** using **LoRA & Hugging Face Transformers**.
- **Developed AI-powered chatbots** using **Google Dialogflow ES/CX**, integrated with **WhatsApp & enterprise workflows**, increasing automation by **60%**.
- **Built a Text-to-SQL NLP model**, allowing users to **query databases in natural language**, reducing manual SQL queries by **75%**.
- **Created an image similarity search engine** leveraging **vector embeddings (FAISS, Pinecone)**, improving retrieval accuracy by **95%**.

### The Other Fruit/Nes Tech Ltd | *AI/ML Engineer*

📅 Jan 2022 - Aug 2023

- **Developed an AI-based document summarization system** using **GPT-3, BERT**, and **LangChain**, reducing manual report analysis time by **80%**.
- **Implemented a multimodal AI system** combining **text & image inputs**, improving recommendation accuracy by **75%**.
- **Optimized AI model inference pipelines**, reducing processing time from **25 sec to 10 sec** using **TensorRT & ONNX**.
- **Built a real-time speech-to-text assistant** using **Whisper AI & OpenAI API**, enhancing accessibility for **10K+ users**.

### Freelance / Fiverr / Upwork | *AI & Chatbot Developer*

📅 2018 - Present

- **Developed 100+ AI-powered chatbots** using **Google Dialogflow, RASA, OpenAI API**, deployed for **business automation & customer support**.
- **Fine-tuned LLaMA & GPT-4 models** for **custom AI applications**, optimizing **inference latency by 45%**.
- **Integrated AI agents into enterprise applications** using **LangChain, OpenAI GPT APIs**, automating decision-making tasks.

## Education

🎓 **Master of Engineering (Power Systems)**  
Sankalchand Patel College of Engineering, Gujarat (2011-2013)  
CPI: 8.3

🎓 **Bachelor of Engineering (Electrical Engineering)**  
Global Institute of Technology, Jaipur (2006-2009)

Projects

- **LLM-Powered AI Chatbot** – Built an AI chatbot using **GPT-4, LangChain, and FAISS**, handling **10K+ queries daily** with real-time responses.
  - **AI-Powered Resume Screener** – Developed a **job candidate ranking system** using **BERT & OpenAI embeddings**, reducing HR screening time by **X%**.
  - **Voice-to-Text AI Assistant** – Created an **NLP-powered transcription tool** using **OpenAI Whisper API**, enhancing accuracy for **speech recognition tasks**.
- 

Keywords for ATS Optimization

AI Engineer, Generative AI, LLMs, GPT-4, OpenAI API, LangChain, Hugging Face, NLP, Chatbot Development, Google Dialogflow, AI Model Fine-Tuning, PyTorch, TensorFlow, Vector Databases, API Development, AWS SageMaker, GCP Vertex AI