ANJALI PATEL

(+91) 9773299554 | anjalijpatel.45@gmail.com | LinkedIn | GitHub

EDUCATION

NEW LJ Institute of Engineering and Technology   Ahmedabad, Gujarat
Bachelor of Engineering in Computer Science, Minor in AI-ML   *Jul. 2022 – May. 2025*
R.C. Technical Institute   Ahmedabad, Gujarat
Diploma in Information Technology   *Aug. 2019 – May 2022*

TECHNICAL SKILLS

Languages: Python, Java, C/C++, SQL
AI/ML Frameworks: PyTorch, TensorFlow, Hugging Face Transformers, LangChain
Tools: GitHub, VS Code, Postman, Docker, Ollama, FastAPI
Domains: Deep Learning, NLP, Image Processing, LLMs
Other: FastAPI, Flask, REST APIs, Model Deployment

PROJECTS

RAG-Based Domain-Specific QA System | Python, PyTorch, FAISS, ChromaDB   Jan. 2025 – Apr. 2025
- Developed a finance-focused RAG system for domain-specific question answering using NLP techniques.
- Implemented document chunking, embeddings, and retrieval using FAISS and ChromaDB for improved accuracy.
- Fine-tuned LLMs to enhance contextual understanding of finance-related documents.
- Integrated the RAG-based QA system with a Streamlit interface for interactive, real-time querying and responses.

Deep Learning-Based Text Classification System | Python, PyTorch, Transformers   Dec. 2024 – Jan. 2025
- Developed a text classification model using PyTorch and Hugging Face Transformers.
- Trained the model to categorize various document types accurately.
- Enhanced model performance with knowledge distillation techniques.
- Integrated the model with a web API for real-time classification services.

EXPERIENCE

Artem Health Tech Pvt. Ltd.   Dec. 2024 – Present
                                                                                Ahmedabad, Gujarat
- **Associate AI-ML Engineer**   Jun. 2025 – Present
  - Fine-tuned LLMs using adapter-based methods (LoRA, PEFT) for efficient domain adaptation in healthcare NLP tasks.
  - Applied pruning, quantization, and knowledge distillation to optimize inference latency and model size.
  - Scaled modular AI pipelines with FastAPI and Docker, enabling reproducible and production-ready deployments.
- **AI-ML Intern**   Dec. 2024 – May 2025
  - Built a LangChain-based chatbot leveraging NER for structured medical data capture.
  - Developed a speech-enabled QA assistant using Whisper ASR, TTS, and RAG-based LLM integration.
  - Created retrieval pipelines using FAISS/ChromaDB with custom tokenizers and document chunking.

CERTIFICATES

Multi AI Agent Systems with crewAI (DeepLearning.AI)   Mar. 2025
- Learned how to design and implement multi-agent AI systems for task automation and collaborative workflows.

Machine Learning Specialization (Coursera, Stanford University)   Nov. 2024
- Gained in-depth knowledge in machine learning algorithms, supervised and unsupervised learning techniques.

Introduction to TensorFlow for AI, ML, and Deep Learning (Coursera)   Sep. 2023
- Developed hands-on experience with TensorFlow for implementing AI and ML models.