

DHRUVIL MISTRY

Mumbai, India — dhruvilmistry16@gmail.com — LinkedIn — GitHub

PROFESSIONAL SUMMARY

AI Engineer specialising in production-grade LLM systems, RAG pipelines, and agentic AI architectures. Experienced building multi-provider AI products using OpenAI, Anthropic, and Gemini APIs with LangChain, LlamaIndex, and FastAPI. Skilled in vector search, prompt engineering, LLM evaluation, and deploying scalable AI applications on AWS and Vercel across text, audio, and image workflows.

PROJECTS

Minutz – AI Meeting Intelligence Platform | FastAPI, Next.js, Whisper, GPT-4o, Supabase **Live** | **OpenAI × Outskill Hackathon – Selected**

- Engineered a browser-native meeting recorder intercepting WebRTC audio streams directly inside Chrome – no bot joins, no notifications – overcoming standard tab capture limitations.
- Built end-to-end async pipeline: 10s audio chunks → FastAPI → ffmpeg merge → Whisper transcription → GPT-4o extraction of summaries, action items, and sentiment; full turnaround in ~60s.
- Shipped full-stack product (Next.js, Supabase, Slack integration, 4 intelligence modes) on Vercel + Render; selected at OpenAI × Outskill Hackathon.

LegalMind – AI Contract Risk Analyzer | FastAPI, Python, Qdrant, OpenAI, asyncio **Completed**

- Designing a scalable agentic RAG pipeline for legal documents using 12 parallel async LLM risk-extraction tasks via FastAPI + asyncio, targeting <30s latency on 50-page contracts.
- Achieving 40% lower retrieval latency vs. naive chunking by implementing Qdrant vector search with recursive chunking and `text-embedding-3-small`; integrating LLM output evaluation to reduce hallucinations.

Clarity – AI Reading Assistant | JavaScript, OpenAI, Anthropic, Gemini APIs **Live**

- Shipped a Chrome extension delivering 1-click summarisation, entity extraction, and contextual Q&A via GPT-4o Mini, Claude Haiku, and Gemini 2.5 Flash through a zero-backend multi-provider LLM abstraction layer.

Skin Cure – Multimodal Healthcare AI | FastAPI, React, SwinV2, PyTorch **Live**

- Training 24-class SwinV2-Base classifier on 28,000+ dermatology images (~78% accuracy); building FastAPI + React app with AI predictions, confidence scores, and Google Maps dermatologist discovery.

EXPERIENCE

AI/ML Intern | Infosys Springboard, Remote **Sep 2025 – Nov 2025**

- Reduced model error rates by 20% across production ML pipelines by engineering end-to-end preprocessing, feature engineering, and evaluation workflows using Scikit-Learn and PyTorch on enterprise datasets.
- Accelerated NLP delivery by building sentiment analysis, text classification, and keyword extraction workflows with Hugging Face transformers; improved prompt engineering strategy to reduce LLM hallucinations on enterprise data.
- Improved pipeline reproducibility and CI/CD deployment consistency by containerising ML workflows with Docker, enabling repeatable model evaluation across environments.

Full Stack Web Designer | Contra Platform, Remote **Oct 2023 – Feb 2024**

- Delivered 10+ full-stack applications with 95% client satisfaction and zero post-launch critical bugs; reduced average delivery time by 30% through reusable design systems and component libraries.

TECHNICAL SKILLS

Languages: Python, JavaScript, HTML, CSS **Frameworks:** FastAPI, React, Next.js, LangChain, LlamaIndex

AI & Retrieval: LLMs, Agentic AI, RAG Pipelines, Prompt Engineering, LLM Evaluation, Embeddings, OpenAI APIs, Anthropic APIs, Gemini APIs, Hugging Face, Whisper, NLP

Vector Databases: Qdrant, Pinecone, ChromaDB, FAISS **ML:** PyTorch, Scikit-Learn, Computer Vision, Transformer Models

Infrastructure: PostgreSQL, MongoDB, Docker, AWS (EC2, S3), CI/CD, Git, Vercel, Render

CERTIFICATIONS

- Building LLM Applications with Prompt Engineering – NVIDIA — Generative AI with LLMs – DeepLearning.AI — Machine Learning Specialization – Stanford/Coursera

EDUCATION

B.E. Electronics & Telecommunications (Expected May 2026) | MCT Rajiv Gandhi Institute of Technology, Mumbai **CGPA:** 7.0/10

Diploma in Electronics & Telecommunications (2023) | Thakur Polytechnic, Mumbai **3rd Rank in Department**

HONORS & LEADERSHIP

- **Technical Lead, ISTE** – Organised and hosted multiple technical events across the college community. **LinkedIn Creator** – 3,000+ followers on AI, LLMs, and full-stack engineering.
- **Awards:** 1st Place Model Making (Thakur Polytechnic, 2022) | Runner-Up TechnoFest State-Level Competition (2023).