Hrishi Mehta

Mumbai, Maharashtra

+91-9892090398 mehtahrishi45@gmail.com linkedin.com/in/hrishi-mehta-889732256 github.com/mehtahrishi
portfolio-theta-lemon-33.vercel.app

SUMMARY

AI/ML Engineer specializing in LLM-powered systems, Retrieval-Augmented Generation (RAG), and intelligent automation.

Experienced in building production-ready AI applications integrating semantic search, vector databases, and scalable APIs. Strong ability to transform unstructured data into deployable AI features that improve user engagement and decision-making.

EDUCATION

University of Mumbai   **2022 – 2025**
*Bachelor of Science in Information Technology (B.Sc IT) | CGPA: **8.7/10***   *Mumbai, Maharashtra*

TECHNICAL SKILLS

ML & AI: PyTorch, TensorFlow, Scikit-learn, NLP, LLMs, RAG, Embeddings
Programming: Python, SQL, Flask, Django
Cloud Platforms: Google Cloud, Oracle Cloud, AWS Cloud
Data: Pandas, NumPy, Matplotlib, Feature Engineering, EDA
Databases: Neo4j, MongoDB, SQL, PostgreSQL
LLM Tools: Gemini, LangChain, Groq, Grok, Mistral, Openrouter, MCP Servers, Claude
Developer Productivity Tools: Copilot, AI CLI, Cursor, Kiro, AntiGravity

EXPERIENCE

Web Developer   **July 2025 – Present**
Freelancer   *Remote*

- Integrated AI-powered features such as recommendation systems and chat interfaces into web products.
- Developed AI automation agents and Machine Learning models for intelligent feature delivery.
- Collaborated cross-functionally to translate business problems into data-driven solutions.

Python Developer Intern   **Mar 2025 – Jun 2025**
Codotech and VaultOfCodes   *Remote*

- Built machine learning models and integrated them into web applications.
- Implemented Machine Learning models and AI Agents using Python libraries such as Pandas, NumPy, and NLTK, integrated into responsive front-ends with HTML and CSS.
- Conducted data preprocessing, feature engineering, and model evaluation.
- Developed APIs to serve ML predictions to frontend systems.

Artificial Intelligence Intern   **Mar 2025 – Apr 2025**
INeuBytes and Edunet Foundation   *Remote*

- Built AI-powered solutions using LLMs, RAG pipelines, and NLP techniques for real-world problem solving.
- Utilized Python libraries like Torch, TensorFlow, Pandas, and NLTK for AI model development and data analysis.
- Processed and analyzed datasets using Pandas and NumPy to generate insights.
- Designed intelligent systems for Q&A, summarization, and semantic search.

PROJECTS

Catalyst Flow ML – ML Experimentation Platform ⤢ | Next.js, FastAPI, Docker   **January 2025**

- Built a full-stack machine learning experimentation platform enabling users to train, compare, and visualize multiple ML models in real time.
- Engineered live training visualization pipelines for 9 algorithms (Random Forest, Gradient Boosting, SVM, Neural Networks) with continuous metric tracking for overfitting detection.
- Developed an automated ML preprocessing pipeline including encoding, scaling, imputation, and train/validation/test splitting for structured datasets.

- Architected a Dockerized microservice system with FastAPI backend, Next.js frontend, background ML workers, and persistent experiment tracking using SQLAlchemy.

AI Wallet – Multi-Provider LLM Interaction Platform ⧉ | Next.js, TypeScript, Python   Jan 2025
- Developed a full-stack platform enabling users to discover and interact with AI models from multiple providers including OpenAI, Gemini, Anthropic, Groq, Mistral, and OpenRouter.
- Implemented real-time streaming AI responses using Server-Sent Events (SSE) with FastAPI async generators and dynamic frontend rendering.
- Engineered a model discovery engine that fetches, categorizes, and displays model capabilities (chat, vision, embeddings) along with provider rate limits.
- Designed a secure API key architecture where credentials remain client-side, ensuring privacy while enabling direct provider communication.

Chat with PPT and PDF - Curr AI Agent ⧉ | JavaScript, Gemini AI SDK, Redis Cloud   May 2025
- Architected Curr AI Agent, a Generative AI platform enabling context-aware interrogation of unstructured datasets (PDFs, .pptx) via a sophisticated Retrieval-Augmented Generation (RAG) pipeline.
- Engineered a semantic search ecosystem utilizing Google Gemini-, implementing high-dimensional vector embeddings and cosine similarity ranking to optimize token usage and response accuracy.
- Orchestrated a resilient Python (Flask) backend infrastructure, fortifying RESTful endpoints with secure server-side session persistence and state management protocols.

Movie Recommendation System ⧉ | Python, ML, NLTK, Torch   March 2025
- Orchestrated a monolithic AI-driven web platform utilizing TensorFlow and PyTorch (Torch) for backend predictive modeling, integrated seamlessly with a responsive HTML5/CSS3 frontend.
- Designed a robust RESTful API ecosystem in Flask, acting as the middleware between the frontend client and the computational logic, enabling real-time asynchronous data fetching.
- Implemented advanced Dimensionality Reduction and data normalization techniques to refine the recommendation algorithm, significantly enhancing prediction accuracy and user engagement metrics.

CERTIFICATIONS

- ISO/IEC 20000 IT Service Management Associate - Skillfront
- ISO/IEC 27001 Information Security Associate - Skillfront
- Gemini Certified University Student - Google
- Neo4j Certified Professional – Neo4j GraphAcademy
- ACCELQ Automation Engineer – ACCELQ
- AWS Databricks Platform Architect – Databricks

VOLUNTEER

Google Cloud Skills Boost   March 2025 – Present
Innovator   *Remote*
- Focusing on personal growth through innovative problem-solving, learning cutting-edge cloud technologies, and fostering engagement within the tech community.

AWS Community   June 2025 – Present
Community Builder   *Remote*
- Drive community engagement by creating in-depth technical content (blog posts, videos, and tutorials) focused on AWS cloud technologies, sharing expertise, and providing direct feedback to AWS product teams.

DLLE, Mumbai University   June 2024 – May 2025
Member – Department of Lifelong Learning and Extension   *Mumbai, India*
- Supporting lifelong learning programs by mentoring others and advancing educational accessibility and community development.