

Muhammad Haroon Rasheed K.M

haroonint144@gmail.com | +91 9087116254 | Chennai, Tamil Nadu

LinkedIn: linkedin.com/in/haroontrailblazer | GitHub: github.com/haroontrailblazer | Portfolio: haroon.pro

SUMMARY

Machine Learning & Data Engineer with hands-on experience in model development, LLM-based systems, MLOps, data processing pipelines, and real-time application deployment using Python, TensorFlow, Streamlit, SQL, and cloud platforms. Skilled in building agentic AI systems, RAG pipelines, multi-agent orchestration workflows, and MCP-integrated developer tools. Experienced in transforming analytical insights into scalable products and securing data workflows with privacy-first design.

EXPERIENCE

AI Product & Research Intern - CloseFuture (May 2026 - Present)

- Evaluate emerging AI tools and cutting-edge models at the intersection of research and product thinking, identifying where AI can create real, scalable impact and surfacing insights that inform strategic direction.
- Stress-test frameworks, experiment with frontier models, and build human-centered, production-ready AI solutions — applying a hands-on approach from initial exploration through to implementation.

Open Source Contributor - LangChain (2025 - Present)

- Contributing to LangChain, an open-source framework for building applications with LLMs, by implementing new features, fixing bugs, and improving documentation to enhance the developer experience and expand the capabilities of the framework.

Machine Learning Intern - The Mind Company (Jul 2025 - Aug 2025)

- Executed data preprocessing, cleaning, feature engineering, and visualization using NumPy, Pandas, Matplotlib, Seaborn & Scikit-learn for multiple structured mini-projects, improving data quality and modeling readiness.
- Built and evaluated baseline models (Linear/Multiple Regression, Logistic Regression, KNN, Decision Trees, Random Forest, Naive Bayes) in Scikit-learn with cross-validation.
- Assessed models using MSE, RMSE, and R^2 , documenting results and trade-offs to guide model selection.

Data Analyst Intern - Code From Basics, Chennai (Oct 2024 - Jan 2025)

- Built an interactive Power BI dashboard to analyze user engagement across major social media platforms.
- Designed KPIs, DAX measures, and trend visualizations to uncover insights on content performance, demographic behavior, and platform reach.
- Demonstrated strong analytical thinking and the ability to derive actionable marketing insights from raw data.

PROJECTS

Oreag – RAG & Memory as a Service (2026)

- Built a multi-tenant production RAG API and persistent memory layer that converts uploaded documents (29 formats) into a queryable REST endpoint with cited, grounded answers. Implemented a full indexing pipeline: file conversion via MarkItDown, recursive chunking with configurable overlap, vectorization through 6 BYOK/keyless providers (OpenAI, Gemini, Anthropic, Sarvam AI, Ollama, sentence-transformers), and pgvector cosine retrieval with an agentic multi-query loop.
- Engineered a memory graph with auto-detected cross-document links, conversation memory with multi-turn query rewriting, an MCP server (FastMCP) for Claude Code and Codex integration, and a self-serve dashboard with Playground, API, and Settings tabs. Secured using Postgres RLS, Fernet-encrypted provider keys, and SHA-256 API key hashing behind Supabase Auth.
- **Tech:** Next.js 16, React 19, TypeScript, FastAPI, SQLAlchemy 2, Supabase, pgvector, LangChain, FastMCP, Docker

EduOrchestrate – AI Learning Orchestrator (2026)

- Developed a local-first, dependency-free Node.js CLI that transforms any AI coding runtime (Claude Code, Codex, Gemini) into a daily job-role learning orchestrator. A single npx command profiles the target role, generates an adaptive 30-day minimum plan, resolves live learning resources via YouTube Data API and GitHub API, and ships daily tasks via SMTP and GitHub Actions automation.
- Architected a 6-agent multi-agent workflow (GoalMapper, CurriculumPlanner, Researcher, BuildMentor, FocusCoach, Evaluator) that converts passive learning into portfolio-ready artifacts and rendered SVG evidence cards. Vendor-neutral, no runtime lock-in across Claude Code, Codex, Gemini, and any MCP-capable coding agent.
- **Tech:** Node.js ≥18, Dependency-free CLI, Multi-Agent Workflow, GitHub Actions, SMTP, YouTube Data API, GitHub API, SVG Cards, MCP Runtimes

MindScope – AI Mental Health Analytics (2025)

- Designed and implemented an AI-driven machine learning platform for analysis of human cognitive retention and focus patterns, surfacing real-time behavioral insights from cognitive signal data using ML models and visualization pipelines.
- **Tech:** Python, TensorFlow, Scikit-learn, Streamlit, Pandas, Data Engineering

StrengthX – Password Security and Breach Detection (Sep 2025 - Nov 2025)

- Developed and deployed a privacy-first password security platform used by 100+ test users, featuring LLM-powered password generation, SHA-1 hashing, and real-time breach checks via the HIBP API. Built a Streamlit interface and deployed on Render for public access.
- **Tech:** Python, Streamlit, SHA-1, HIBP API, Custom LLM, Render

LANGUAGES, SKILLS AND CERTIFICATES

Technical Skills:

- Programming & AI – Python, TensorFlow, PyTorch, Scikit-learn, ML Algorithms, LLMs, Feature Engineering, LLM-Based Agents, LangChain, LangGraph, LangSmith, Agentic-AI Orchestration, AI Workflow Orchestration
- Data & Visualization – Pandas, NumPy, SQL, MongoDB, MySQL, Power BI, Tableau, EDA, Data Engineering
- MLOps & Deployment – Streamlit, Docker, GitHub Actions, REST API, FastAPI, Data Pipelines, Model Deployment, Cloud Deployment, Model Registry, Experiment Tracking
- Web & Agentic Systems – Next.js 16, React, TypeScript, FastAPI, Supabase, pgvector, LangChain, FastMCP, MCP Runtimes, Node.js, RAG Pipelines, Multi-Agent Orchestration
- Mastered AI Tools – Claude.ai, Claude Code, Codex, Claude Cowork, IBM BOB, Antigravity, Deep Agents

Achievements:

- Built and deployed Oreag – a production RAG API with agent memory, MCP server, and BYOK multi-provider support with full multi-tenant security (Postgres RLS, Fernet encryption).
- Built and deployed EduOrchestrate – a vendor-neutral multi-agent CLI orchestrator running across Claude Code, Codex, and Gemini with automated GitHub Actions delivery.
- Built & deployed 4+ AI/ML applications spanning real-time monitoring, breach detection, predictive analytics, and agentic retrieval.
- Implemented privacy-first security architectures using client-side hashing and zero data retention.
- Developed and optimized ML pipelines with feature engineering, class imbalance handling, and performance evaluation.
- Created automated dashboards and production-ready applications using Python, Streamlit, and Render.
- Hosted a MongoDB seminar for 60+ students with 90% positive feedback.
- Maintained an open-source GitHub portfolio with 7+ AI/ML, data science, and Python projects.

Certifications:

- Certified Power BI Analyst
- Oracle Certified AI & ML Engineer
- LangChain & LangGraph
- LangChain Deep Agents
- Cisco Certified Data Analyst
- Cisco Certified Python Programmer
- IBM Certified AI Professional
- IBM Certified WatsonX Orchestrator
- Google Certified Cybersecurity Professional

Soft Skills:

- Demonstrated strong time management while handling multiple academic and project deadlines.
- Highly adaptable, quickly learning new tools and technologies during project development.
- Skilled in problem-solving and debugging complex ML/model issues.
- Strong critical thinking and analytical reasoning applied in data-driven decision making.
- Effective team collaborator, experienced in open-source contributions and seminar delivery.

EDUCATION

Bachelor of Technology in Artificial Intelligence (Current CGPA: 8.7 / 10)

Anna University, Chennai - 600025

Sep 2023 – Oct 2027

- Course Works: Exploratory Data Analysis, Machine Learning, Database Systems, Deep Learning, AI-Agents Orchestration (LangChain, LangGraph, LangSmith)