

KAYAM PATHAN

AI Engineer

India | (+91) 9607263681 | kayampathan06@gmail.com | LinkedIn | GitHub

Summary

AI Engineer with an MSc in Artificial Intelligence and hands on experience building **Generative AI** and **Large Language Model (LLM)** systems, including **RAG** pipelines, **AI powered chatbots**, and autonomous **AI agents**. Proficient in **Python**, **SQL**, **LangChain**, **Vector Databases**, and **Prompt Engineering**, with a strong foundation in **NLP**, **Machine Learning**, and **Deep Learning**. Track record of shipping end to end GenAI solutions from data pipeline to deployment.

Education

Queen Mary University of London (Russell Group University)

MSc Artificial Intelligence – GPA 3.7/4.0

London, United Kingdom

2024 – 2025

Savitribai Phule Pune University

B.E. in Computer Science & Engineering, Distinction

Nashik, India

2020 – 2024

Professional Experience

BharatSkillz

Feb 2026 – Aug 2026

Generative AI Intern

Remote, India

- Independently engineered an **Intelligent Customer Support Chatbot** leveraging **Large Language Models (LLMs)** and **NLP**, automating responses to routine queries and cutting manual handling time by **30%**.
- Designed **prompt engineering** strategies and conversational flows that lifted intent recognition accuracy by **18%**, enhancing response quality and user experience.
- Integrated chatbot logic with backend **APIs** using **Python**, enabling real time, context aware query resolution across **500+** daily interactions and reducing support team workload.

Application Square Infotech

Mar 2022 – Aug 2022

AI/ML Research Intern & Data Analyst

Nashik, India

- Developed and evaluated **machine learning** classification models, achieving up to **88% accuracy**.
- Analyzed **1,000+** customer feedback entries using **NLP** techniques, surfacing insights that shaped key product improvements.
- Automated data preprocessing workflows with **Python** and **Pandas**, trimming data preparation time by **50%**.

Projects

Financial Document Intelligence Platform (RAG)

Independent Project

- Architected an end to end **Retrieval-Augmented Generation (RAG)** pipeline over real SEC filings (10-K/10-Q), combining a custom parser, **Hugging Face** sentence transformer embeddings, and a **FAISS vector database** to enable natural language querying with cited, source level answers across **8** public companies.
- Created a structured extraction and validation module using **LLM APIs** and **prompt engineering** to cross check every extracted financial line item against SEC's official structured (XBRL) data, achieving **96.6% exact match accuracy**.
- Implemented a resilient primary/fallback **LLM** architecture (Gemini API with a locally hosted Ollama model) to gracefully handle real world API rate limits, and deployed a live interactive demo on **Streamlit Cloud**, built entirely on free tier infrastructure.

Autonomous IT Support Agent (MCP)

Independent Project

- Launched an **AI agent** on the **Model Context Protocol (MCP)** to autonomously triage, categorize, and route IT support tickets, slashing triage time by **40%**.
- Established a **RAG** retrieval layer over internal runbooks/documentation with **LangChain** and **Hugging Face** models, improving semantic search accuracy for prior resolutions by **25%**.
- Applied **prompt engineering** to generate context aware ticket responses via **LLM APIs**, logging resolution data to **SQL** for SLA tracking; built **Power BI** dashboards spanning **5+** SLA metrics and containerized the service with **Docker** for **AWS** deployment.

AI Job Application Copilot (n8n)

Independent Project

- Engineered an autonomous **n8n** workflow integrating live job board **APIs** and **Google Gemini LLM** scoring to automatically match resume profiles against real time job postings, achieving a **43%** qualified match rate across **250+** listings per run.
- Implemented automated cover letter and interview question generation using structured **JSON** prompting, with rate limit aware throttling and **OAuth2** integrated **Google Sheets** logging for persistent application tracking.

Technical Skills

Languages: Python, SQL

Generative AI & LLMs: Large Language Models (LLM), Retrieval-Augmented Generation (RAG), LangChain, AI Agents, Model Context Protocol (MCP), Prompt Engineering, Hugging Face, Vector Databases, Chatbots

Machine Learning & Data Science: Machine Learning, Deep Learning, NLP, Feature Engineering, Pandas, NumPy

Cloud & DevOps: AWS, Docker, Git, GitHub

BI & Visualization: Power BI, SQL

AI-Assisted Development: Cursor, Claude AI, Google AI Studio, n8n, Supabase