

AAYUSH GAUTAM

LinkedIn: <https://www.linkedin.com/in/aayush-gautam-25b9a5247/>
GitHub: github.com/Epein5

Email: aayushgtm5@gmail.com
Mobile: +977 9869730965

EDUCATION AND FELLOWSHIP

Kathmandu University,
B.Tech in Artificial Intelligence; GPA: 3.65,
Fusemachines AI Fellowship
Fusemachines Inc. • Specialized training in AI, ML, and NLP

Dhulikhel, Nepal
2021 - December 2025
Kathmandu
Jan 2023 - Dec 2023

SKILLS SUMMARY

- **Languages:** Python, SQL
- **Frameworks:** PyTorch, TensorFlow, Scikit-Learn, Pandas, NumPy, Matplotlib
- **AI and NLP tools:** LangChain, LangGraph, Hugging Face, Vector Databases, RAG
- **Development & Deployment Tools:** FastAPI, Docker, Azure (CI/CD Pipelines)
- **Soft Skills:** People Management, Excellent communication

WORK EXPERIENCE

ARTIFICIAL INTELLIGENCE INTERN | Tecnvirons | [LINK](#)

- Built and deployed an AI-powered **voice agent system** with RAG search (Pinecone DB) and Plivo, improving first-call resolution by 40%.
- Created a WhatsApp onboarding bot with sentiment analysis and smart callbacks, boosting conversions by 35%.
- Developed a **FastAPI-Supabase** backend for logs, transcripts, and analytics, cutting manual processing by 30%.
- Implemented end-to-end **CI/CD pipelines** using GitHub Actions, automating container builds to Azure Container Registry and deployments to Azure Container Apps, accelerating release cycles.
- Automated customer insights dashboards to identify buying motivations and barriers, providing actionable data to guide product development and marketing strategies.

PROJECTS

Flash Notes | [LINK](#)

- A document assistant powered by Qwen 2.5 with RAG, enabling chat, summaries, and quizzes from PDFs, PPTX, and TXT files. Integrated Tesseract OCR for image-based text extraction
- It allows users to upload documents (PDF, PPTX, or TXT) and engage with them using Qwen 2.5. With Flash Notes, users can chat with their documents, generate summaries, and create quizzes based on the content, all designed to enhance document comprehension and learning.
- Tech Stack: Next.js, FastAPI, LangChain, Chromadb

FlightBuddy_AgenticRAG | [LINK](#)

- Built an airport assistant for Tribhuvan International Airport using a Query Router to handle RAG-based airport info via ChromaDB, flight bookings through an SQL agent with SQLite, and general queries using Gemini LLM.
- Agentic Framework handles Websearch, Vector Search, and SQL search
- Tech Stack: FastAPI, LangChain, Chromadb

Sahara (Ongoing) |

- Built a **Cluely-style AI assistant** powered by Google Gemini, supporting **text and audio inputs** with real-time streaming outputs.
- Engineered a **FastAPI + WebSocket backend** for low-latency, bidirectional communication, sustaining **1000+ concurrent sessions in testing**.
- Optimized **token usage and response streaming**, reducing API costs by ~30% while improving throughput and responsiveness.

AWARDS AND CERTIFICATES

KEC Hackathon(Most Business Viable Product)

February 2024

- Developed a Mobile and Web application for the detection, reporting, and management of fire and potholes.
- Acquired expertise in procedural programming paradigms and associated logical concepts, enhancing capabilities.

Khalisis “Bin There, Hacked That”

June 2025

- Created an end-to-end solution for the management of disposable and unrecyclable waste products.
- Designed a system to **reduce carbon emissions by ~25%** through optimized disposal workflows.
- Led data collection, identification, and cleaning, ensuring high-quality datasets for analysis and decision-making.

KUAIC Datathon’24 (2nd Runner-up)

January 2024

- Analyzed Nepal Census data to reveal gender disparities, showing only ~26% female labor force participation and <20% land ownership by women, highlighting the impact of a patriarchal society.