

Asfand Yar

Islamabad, Pakistan

+92 331 0329489 • asfandyar273263@gmail.com • in asfandyar100 • Asfand-Yar-dev

Gen AI Backend Developer with hands-on experience building intelligent systems, backend services, and scalable APIs using Python and FastAPI. Skilled in developing RAG pipelines, working with Open-source LLMs, and applying prompt engineering to optimize model outputs for enterprise applications. Experienced in containerizing and deploying AI-Powered applications using Docker across virtual environments. Focused on building efficient, low-latency backend infrastructure that bridges generative AI with scalable, production-grade systems.

Experience

Emumba

GenAI Engineer Intern

On-site

Jan 2026 – Mar 2026

- Designed scalable backend systems using Python and applied prompt engineering to optimize LLM outputs for enterprise applications.
- Engineered scalable Retrieval-Augmented Generation (RAG) pipelines to enhance model accuracy and enable contextual answering across applications.
- Collaborated with cross-functional teams to deploy scalable AI features, ensuring low-latency data processing and API integration.

Developers Hub Corporation

AI Intern

Remote

Mar 2025 – May 2025

- Engineered Python scripts to extract, clean, and format raw data, creating high-quality datasets for testing internal machine learning models.
- Conducted foundational prompt engineering and benchmarked various Open-source language models to document performance and response accuracy.
- Built internal UI dashboards and automated scripts to visualize AI outputs and test API endpoints, streamlining day-to-day development tasks.

Technical Skills

- AI & Machine Learning: Generative AI, Large Language Models (LLM), RAG Pipelines, Prompt Engineering, AI Agents, AI Automation
- Frameworks & Tools: FastAPI, LangChain, OpenAI API, n8n, Pipecat, Cartesia, Github, Git
- Languages & Development: Python, SQL, Node.js, REST APIs, HTML5, CSS3, React.js, Tailwind CSS
- Data & Infrastructure: Redis, Docker, Containerization, Virtual Environments, Vector Databases

Key Projects

Real-Time Audio & RAG

AI-Powered Voice Bot

- Implemented a real-time conversational voice agent using a local LLM, Pipecat, and Cartesia for natural text-to-speech interactions.
- Engineered a RAG pipeline with a vector database to enable the voice bot to answer complex, context-specific queries accurately.
- Managed the project in a virtual environment, containerized the application using Docker, and successfully published the image to Docker Hub for public access.

FYP Project

Intervexa - AI Web-Based Mock Interview System

- Developed an AI-driven interview platform that dynamically tailors technical questions based on the candidate's specific job role and skill set.
- Integrated voice recognition to analyze verbal responses and facial expression capture to monitor candidate confidence during the interview.
- Provided a comprehensive feedback dashboard to help users prepare for real-world interviews effectively.

National AI Hackathon

Decision Debater AI

- Built an AI-Powered debate platform using LangChain and Retrieval-Augmented Generation (RAG) to provide evidence-based arguments.
- Integrated vector search with LLM to retrieve relevant information and generate context-aware responses from multiple knowledge sources.
- Created backend APIs using FastAPI to support real-time debate generation and seamless integration with frontend applications.

Education

BUIITEMS

Bachelor of Science in Computer Science, GPA: 3.83/4.00

Quetta

Oct 2022 – Dec 2026 (Expected)

Tameer-e-Nau Public College

Intermediate in Pre-Engineering, Grade: A+

Quetta

Mar 2019 – Apr 2021

Certifications & Workshops

- ACT AI National Training Programme (2026)
- Microsoft Learn Student Ambassador (2024)
- Generative AI Application Developer Certificate
- ChatGPT for Everyone (Learn Prompting)
- Google AI Essentials (2024)
- National AI Hackathon (atomcamp)
- Anthropic Claude Code in Action
- WEB3 Masterclass (Algorand)

Leadership & Achievements

BUIITEMS Entrepreneurial Society

Executive Lead

Quetta

Aug 2024 – Present

Mobile Phone Repairing & Maintenance

Gold Medalist

EME Center, Quetta Cantt

2020

Languages

English: Advanced | Urdu: Professional | Punjabi: Fluent | Pashto: Intermediate