

Suraj Pathak

Kathmandu, Nepal

☎ 9864662199 | ✉ surajpathakcs@gmail.com | 🌐 github.com/surajpathakcs | surajpathak.com

Professional Summary

Software engineer specializing in AI systems integration, designing production-oriented RAG pipelines, LLM orchestration graphs, and grounding/verification layers. Full stack engineer experience in building backend services, APIs, and web platforms end-to-end; recent work includes a support assistant with a stateful agentic workflow, hallucination-grading retry loop, and guardrails, alongside a full-stack e-commerce platform built with NestJS, PostgreSQL, and Next.js.

Skills

AI/LLM Engineering	RAG pipelines, LangGraph, vector search, embeddings, Qdrant, ChromaDB, guardrails, observability.
Languages	Python, TypeScript, JavaScript, SQL, C#, YAML.
Backend Engineering	NestJS, FastAPI, Node.js, REST API architecture, authentication systems.
Frontend Engineering	Next.js, React, Zod, Redux Toolkit, Tailwind, Streamlit.
Databases	MySQL, MongoDB, PostgreSQL, ChromaDB, Qdrant, MinIO (S3-compatible object storage).
Infra & Tools	GIT, Postman, Docker/Docker Compose, NVIDIA Nemo guardrails, OpenAPI/Swagger, Logfire, Render, Linux environments.
Miscellaneous	Testing, Debugging, System Design, Problem-solving, Research, Collaboration.

Personal Projects

RAG Support Assistant for Fonepay Merchant Support (independent project)

2026

- Built a multi-stage retrieval pipeline (top-k vector search → semantic reranking → context assembly) with character-based prompt budgeting to stay within context limits; developed a Streamlit interface.
- Designed a 4 node stateful LangGraph workflow (planner → retriever → responder → grader) with a conditional hallucination-grading retry loop and fallback handling.
- Implemented input-side guardrails using NVIDIA NeMo Guardrails and full-pipeline tracing using Logfire decoupled from user-facing error messages.

Academic Projects

Student Report Card Extraction Using Document Understanding Transformer (DONUT)

TU 2025

- Developed an OCR-free document understanding system using the **DONUT Vision-Encoder-Decoder model** to extract structured academic information from student report card images.
- Built an end-to-end fine-tuning pipeline including dataset preparation, JSON schema conversion, train/validation/test splitting, custom **PyTorch** dataset handling, mixed-precision training, gradient checkpointing, and checkpoint management on **Kaggle** GPU infrastructure.
- **Fine-tuned naver-clova-ix/donut-base** to generate structured JSON outputs containing student details, academic records, and subject-level grades; implemented inference and evaluation workflows for batch document processing.

Work Experience

Ashva Tech – Frontend Developer - Educampaign Consultancy

Jul 2025

- Engineered responsive user interfaces using **Next.js**, implementing modular component architecture and structured routing.
- Designed scalable UI components supporting maintainable feature expansion and consistent design patterns across the application.

Ashva Tech – Full Stack Developer - E-Commerce Platform Development

Jan 2026

- Built a modular NestJS backend organized into **18 domain and infrastructure modules** covering authentication, commerce workflows, bookings, inventory, notifications, storage, and administrative operations using DTO validation and layered architecture.

- Designed PostgreSQL schemas with indexing and query optimization; implemented geocoding-based address autocomplete, S3-compatible image storage using MinIO, and service booking with calendar synchronization.
- Developed Next.js storefront and admin interfaces using App Router, server-side rendering, client components, Tailwind CSS, and REST API integration.

Education

Tribhuvan University

Kathmandu, Nepal

BSc. Computer Science and Information Technology (BSc.CSIT)

Graduated June 2026

- **Relevant Courses:** Artificial Intelligence, Database Management Systems, Operating Systems, Cryptography, Computer Graphics, Design and Analysis of Algorithms

Languages

English Full Professional Proficiency

Nepali Native Proficiency