

# HARDIK GULATI

+91 9667563485 | hardikglt2507@gmail.com | [hardikgulati.me](https://hardikgulati.me) | [LinkedIn](#) | [GitHub](#)

Third-year B.Tech student in Artificial Intelligence and Machine Learning with a strong foundation in Python, data analysis, and statistics. I build practical AI and data-driven systems, with recent work spanning a retrieval-augmented research assistant, an accessibility automation tool, and a study of text extraction errors in government PDFs. Interested in roles where I can work across machine learning, backend systems, data, and problem-solving.

## SKILLS

**Languages:** Python (proficient), SQL (proficient), Java (decent), C (decent), C++ (basic)

**Backend:** FastAPI, SQLAlchemy / SQLAlchem, Pydantic, REST API design

**Frontend:** React

**Data / ML:** NumPy, Pandas, FAISS, sentence-transformers, TensorFlow/Keras (coursework)

**Databases:** SQLite, PostgreSQL

**Tools:** Git, GitHub Actions (CI), Playwright, pytest, Docker (basic), Excel, Power BI

## PROJECTS

### A11y GitHub Issue Logger — FastAPI, React, Playwright, axe-core

Jul 2025 – Present

Live: [a11y-github-issue-logger.vercel.app](https://a11y-github-issue-logger.vercel.app) · Repo: [github.com/Hardik25-iso/A11y-Github-Issue-Logger](https://github.com/Hardik25-iso/A11y-Github-Issue-Logger)

- Built a tool that scans a public URL with axe-core, checks the target repository for duplicate issues, drafts a structured accessibility (WCAG) issue, and logs it to GitHub after review.
- Used a bring-your-own-token model so no user credentials are stored server-side.
- Set up GitHub Actions CI running lint, type checks and a pytest/Playwright test suite on every pull request; deployed the backend to Railway and frontend to Vercel.
- Rate-limited the API with slowapi and worked across feature branches with reviewed pull requests.

### Scholar — Retrieval-Augmented Research Assistant — FastAPI, React, FAISS

Present

Live: [scholar-production-480b.up.railway.app](https://scholar-production-480b.up.railway.app) · Repo: [github.com/Hardik25-iso/Scholar](https://github.com/Hardik25-iso/Scholar)

- Built a document Q&A system that ingests PDFs/DOCX/PPTX, chunks and embeds them with sentence-transformers, and answers questions using FAISS vector search combined with BM25 keyword search.
- Implemented a citation layer so answers point back to the source passage they came from.
- Wrote an evaluation script measuring retrieval recall on a small test set of questions, and used it to compare retrieval configurations rather than assuming defaults were best.
- Data model in SQLAlchemy/SQLite; background jobs handled with arq.

### DevAudit / SILT — Text-Layer Corruption in Government PDFs (in progress) — Python

Present

Repo: [github.com/Hardik25-iso/DevAudit](https://github.com/Hardik25-iso/DevAudit)

- Measuring how often born-digital Indian government PDFs carry a Devanagari text layer that extracts without error but is silently wrong, caused by legacy non-Unicode fonts.
- Audited 1,602 documents from 8 issuing bodies across 3 states using a five-bucket classification, identifying 141 distinct legacy font variants.
- Pre-registered the decision threshold and random seed before running the audit.

## EDUCATION

### B.Tech, Artificial Intelligence & Machine Learning — Symbiosis Institute of Technology, Pune

2024 – 2028

CGPA: 7.78

### Class XII (CBSE), Vishwa Bharti Public School

2024

88.8%

### Class X (CBSE), Vishwa Bharti Public School

2022

91.8%

## CERTIFICATIONS & ACTIVITIES

- Supervised Machine Learning — Infosys Springboard
- Artificial Intelligence — LaunchED Global
- Hashgraph Developer — The Hashgraph Association
- Machine Learning Specialization — DeepLearning.AI
- Deep Learning coursework — regression, classification and regularized neural networks (TensorFlow/Keras), including a from-scratch MNIST FNN
- Technical Volunteer — Cybersecurity & Blockchain Club, SIT Pune

Jan 2026  
Feb 2026  
Jul 2025  
Completed