

HARSHIT WALDIA

Delhi Cantt, Delhi, India 110010

📞 7060546501 ✉️ harshitwaldia112@gmail.com [in linkedin.com/in/harshit-waldia](https://www.linkedin.com/in/harshit-waldia) github.com/harshitwaldia

Education

Graphic Era Hill University

Oct 2021 – Aug 2025

Bachelor of Technology in CSE with Machine Learning and AI

8.17

Relevant Coursework

- Data Structures
- Algorithms Analysis
- Artificial Intelligence
- Systems Programming
- Software Methodology
- Database Management
- Machine Learning
- Computer Architecture

Experience

Omnie Solutions (I) Pvt. Ltd.

Sep 2025 – Jan 2026

AI/ML Software Developer Intern

Noida, India

- Built machine learning models and time-series forecasting pipelines, using feature engineering and evaluation techniques to improve prediction accuracy.
- Worked on neural networks and gained a strong understanding of the mathematical concepts behind machine learning algorithms.
- Integrated machine learning models into production workflows to support data-driven decision making.
- Developed and deployed RESTful APIs using FastAPI for an enterprise e-commerce platform, enabling smooth backend integration.

Renix Informatics

Nov 2024 – Feb 2025

AI Research Intern

Tamil Nadu, India

- Assisted in developing DocX, an AI-powered document processing system, working on OCR, key-value extraction, and PII redaction modules, improving overall processing efficiency by approximately 20%.
- Analyzed document processing workflows and contributed to solutions that enhanced automation, reducing manual effort by around 15%.
- Helped improve data extraction accuracy by refining approaches for key-value detection and preprocessing techniques.
- Supported product demonstrations and internal reviews, contributing to better presentation of features and receiving positive feedback from the team.

Projects

DocTalk – LLM Powered PDF Chatbot | Python, LangChain, FAISS, Streamlit, Groq API

GitHub

- Built a Retrieval-Augmented Generation (RAG) based chatbot to interact with PDF documents using LLMs.
- Implemented semantic search using FAISS vector database and embeddings for context-aware responses.
- Integrated LangChain pipelines to process user queries and generate accurate answers from document context.
- Developed an interactive UI using Streamlit for real-time document querying and response visualization.

FastAPI RBAC Microservice | Python, FastAPI, JWT, PostgreSQL

GitHub

- Designed a scalable microservice implementing Role-Based Access Control (RBAC) for secure API authorization.
- Developed RESTful APIs with FastAPI and implemented JWT-based authentication for user sessions.
- Created role-permission mapping to restrict endpoint access based on user roles (admin, user, etc.).

Synthiverse: Stable Diffusion-Based Creative Synthesis System | Python, PyTorch, Diffusers, OpenCV

GitHub

- Developed a modular text-to-image generation pipeline using Stable Diffusion for creative visual synthesis.
- Integrated Hugging Face Diffusers and PyTorch to generate high-quality, prompt-driven images.
- Engineered and fine-tuned prompts to improve image quality, diversity, and stylistic control.

Technical Skills

Languages: Python, C++, Java, SQL, JavaScript, HTML/CSS, R

Frameworks/Libraries: FastAPI, LangChain, Transformers, TensorFlow, Scikit-learn, OpenCV, YOLO, RAG

Databases: PostgreSQL, MySQL

Developer Tools: Git, GitHub, Jupyter Notebook, Visual Studio Code, Slack

Certifications

OCI Generative AI Certified Professional, Oracle Certification (1Z0-1127-24)

Aug 2024

Microsoft Certified: Azure Data Fundamentals, Microsoft Certification (DP-900)

July 2024

GitHub Foundations Certification, GitHub Certification

June 2025