Harsh Sanjaysingh Kachhaway

7249246230 | harshkachhaway8001@gmail.com

LinkedIn: https://www.linkedin.com/in/harsh-kachhaway

GitHub: https://github.com/Harsh-Kachhaway

Hugging Face: https://huggingface.co/harsh8001

Objective

A motivated Computer Science student with a focus on Generative AI, intelligent systems, and machine learning. Experienced in building AI-powered applications and working with data-driven workflows, with a strong foundation in data science and deep learning. Interested in developing scalable AI solutions, leveraging LLMs, and contributing to innovative, real-world applications.

Education

B.Tech in Computer Science Engineering
Govindrao Wanjari College of Engineering & Technology, Salai Gadhani, India
Affiliated to: Dr. Babasaheb Ambedkar Technological University
G**raduated :** 2025

Data Science & Artificial Intelligence Program
Naresh i Technologies, Hyderabad
Duration: 6 Months
2025 – 2026

Internship Experience

Python Developer Intern — *Codemore* (2024)

- Developed and tested Python applications and scripts.
- Collaborated on automation and data processing projects.

AI/ML Intern — *Next24Tech* (2023)

- Worked on machine learning models for predictive analytics and image processing.
- Involved in model training, data preprocessing, and evaluation.

Tech Intern — *EduSkills Foundation, AICTE & Google* (2023)

- Participated in Google-sponsored workshops on cloud computing, AI, and web technologies.
- Contributed to collaborative projects and learned industry-grade tools

Harsh S Kachhaway
7249246230
Harshkachhaway8001@gmail.com

Projects

Local Text-to-SQL Analytics Engine

Developed a privacy-focused, local-running Text-to-SQL agent that translates natural language queries into complex SQL statements for relational databases. Integrated a local **Qwen 1.5-Instruct** model, fine-tuned/optimized via **Unsloth** for accelerated inference on consumer hardware. Orchestrated agentic logic using **LangGraph** for multi-step query verification and **LangChain** for database schema retrieval and execution.
Tools: Python, LangChain, LangGraph, Qwen 1.5-Instruct, Unsloth, Transformers, SQLite/PostgreSQL.

Autonomous Local AI Workflow & RAG Platform

Engineered a comprehensive, offline-capable platform featuring 10+ modular AI workflows for end-to-end media production and knowledge retrieval. Implemented **RAG (Retrieval-Augmented Generation)** to allow the system to reason over local documentation using Qwen models. The pipeline supports local image synthesis, audio cloning, and enhancement, all managed through a unified interface.
Tools: Python, Gradio, ComfyUI, Stable Diffusion, Flux, RAG (FAISS/ChromaDB), Qwen.

LLM from Scratch 105M

Architected and trained a custom Language Model from scratch, scaled to **105 million parameters**, trained on curated Wikipedia datasets. Implemented advanced architectural components including **RoPE (Rotary Positional Embeddings)** for superior context handling and a standard **Attention Decoder-only architecture**. Managed a **32,000 token vocabulary** and implemented efficient training loops to ensure model convergence from initialization.
Tools: PyTorch, Python, Hugging Face `transformers` (modeling primitives), Tokenizers, CUDA.

Skills

Generative AI & LLMs

- **Core Skills:** Prompt Engineering, LLM Integration, RAG Systems, AI Workflow Automation.
- **Model Techniques:** PEFT/LoRA Fine-tuning, Quantization, Diffusion Models (Stable Diffusion, Flux).

Machine Learning & Deep Learning

- **Frameworks & Libraries:** PyTorch, TensorFlow, Keras, Scikit-learn, Transformers.
- **Applied AI:** Computer Vision, Model Training & Evaluation, Custom LLM Architecture (RoPE, Attention Mechanisms).

Data Science & Analytics

- **Processing:** Data Cleaning, Feature Engineering, Exploratory Data Analysis (EDA).
- **Tools:** Pandas, NumPy, Matplotlib, Seaborn, Statistical Analysis.

Infrastructure & Systems

- **Development:** FastAPI, REST API Design, Event-driven Systems, Automation Pipelines.
- **Data Handling:** SQL, SQLite, MongoDB, Vector Database Management.
- **Tools:** Git, Docker/Containerization, VS Code, Jupyter Notebooks.