Harsh Goyal

+91-8368203499 | harshgoyal2004@gmail.com | linkedin.com/in/harshgoyal2004 | github.com/Harshgoyal2004

Education

Vellore Institute of Technology **2022 – 2026**
B.Tech in Computer Science and Engineering, CGPA: 8.86 *Vellore, India*

Central Board of Secondary Education (CBSE) **2022**
Class 12, Percentage: 88.6 *New Delhi, India*

Central Board of Secondary Education (CBSE) **2020**
Class 10, Percentage: 92.4 *New Delhi, India*

Experience

Feynn Labs Services **March 2025 – May 2025**
Machine Learning Intern *Remote*

- Defined the EV market problem statement and conducted structured research on adoption drivers, barriers, and competitive landscape.
- Analyzed 1,123+ survey responses using K-Means and Random Forest, identifying key customer segments and validating a Rs 43.5 Cr opportunity ($R^2 = 0.7482$).
- Built predictive models and an EV market dashboard to support targeted pricing and GTM recommendations.
- Delivered concise insights including estimates and segment-wise strategy recommendations.

Projects

AI Trip Planner – Agentic Workflow System GitHub

- Reduced response latency by 60% by building an LLM-driven travel planner (LangChain, LangGraph, FastAPI).
- Achieved 98% tool-call accuracy for weather / places / currency APIs via structured tools and ReAct-style prompting.
- Cut hallucinations by 90% through strict system prompts and controlled agent–tool routing.

Chicken Disease Classification Pipeline GitHub

- Developed a modular ML pipeline, reducing experimentation and iteration time by 40%.
- Used DVC for dataset and model versioning, ensuring fully reproducible and traceable experiments.
- Containerized and deployed a Flask inference service on AWS EC2 using Docker, ECR, and GitHub Actions CI/CD.
- Trained a VGG16 transfer-learning model achieving 88.7% accuracy with inference latency under 120 ms.

Text Summarization Pipeline GitHub

- Built a 5-stage NLP pipeline for dialogue summarization, improving preprocessing throughput by nearly 3 times.
- Fine-tuned the Pegasus model on the SAMSum dataset with improved ROUGE-L over baseline.
- Developed a FastAPI inference engine generating summaries with latency under 150 ms.
- Deployed the system on AWS EC2 using Docker, ECR, and automated CI/CD workflows, reducing release cycles by 70%.

Technical Skills

Languages: Python, SQL, C++, Java, R
ML/Agentic AI: TensorFlow, PyTorch, Scikit-learn, Transformers, LangChain, LangGraph, LLM Tool Calling
MLOps/LLMOps: DVC, Docker, CI/CD, Experiment Tracking, Model Deployment
Backend/Cloud: FastAPI, Flask, REST APIs, Microservices, System Design Basics, AWS (EC2, ECR, S3), Azure
Data/Tools: ETL Pipelines, Feature Engineering, Pandas, NumPy, Git, Linux, MySQL
Problem Solving: DSA/Algorithms, LeetCode (200+), HackerRank 5-star (Rank: 55k)
Profiles: leetcode.com/harshgoyal2004 | hackerrank.com/harshgoyal2004

Certifications

Amazon ML Summer School (MLSS) Certificate August 2025
AWS Certified Solutions Architect – Associate Certificate July 2025
Microsoft Certified: Azure AI Engineer Associate Certificate June 2025
Artificial Intelligence – Nanyang Technological University (NTU), Singapore Certificate June 2024

Extracurricular Activities

IEEE Professional Communications Society **February 2024 – December 2024**
Senior Core Member *VIT Vellore*

- Organized hands-on machine learning workshops for more than 100 participants.
- Mentored over 20 juniors in coding and data science activities.
- Led technical events that increased ML community engagement by 35%.