

# Harsh Kumar

☎ 9341045191 | ✉ harsh.k22@iiits.in | in [Linkedin](#) | GitHub [GitHub](#) | Portfolio [Portfolio](#)

## Education

### Indian Institute of Information Technology, Sri City

Bachelor of Technology - Electronics & Communication Engineering; CGPA: 7.5

Chittoor, India

Nov 2022 - June 2026

### D.A.V Public School, Aurangabad

CBSE(X); Percentage: 91.4

India

Mar 2009 - May 2019

## Experience

### Research Assistant — IIIT Sri City

Aug 2025 – Nov 2025

Data Scientist — Hybrid Intrusion & Anomaly Detection

- Deployed FastAPI production APIs for IoT anomaly detection with CNN/attention-based models, 10K+ daily requests.
- Built scalable data pipelines processing high-volume sensor data using NumPy, Pandas, feature engineering.
- Integrated XGBoost and deep learning models into backend workflows with optimized inference and error handling.
- Configured MLflow tracking, Docker containerization, Git version control for reproducible model deployments.

### CodeClause

Dec 2025 – Jan 2026

Data Science Intern

- Configured AWS S3 buckets (raw/labeled/processed data), SageMaker Ground Truth labeling, 30+ annotated images.
- Deployed Lambda functions for preprocessing, configured IAM roles, orchestrated workflows via Step Functions.
- Built production-ready ML pipeline integrating S3, Lambda, Step Functions for end-to-end automation.

## Projects

### Waste Detection System [↗](#)

(GitHub [↗](#))

End-to-End Computer Vision Pipeline

- Built waste detection system using **YOLOv5** with dataset annotation and augmentation, achieving 18% accuracy improvement.
- Designed modular Python pipeline with **OpenCV** for image processing, inference, bounding-box filtering, confidence scoring.
- Containerized using **Docker**, deployed on **AWS/Azure** with **CI/CD** pipelines for scalable real-time inference.
- Technologies:** YOLOv5, Python, OpenCV, Docker, AWS, Azure, CI/CD, Git

### Book Recommender System [↗](#)

(GitHub [↗](#))

Collaborative Filtering Based Recommendation Platform

- Built **collaborative filtering recommender** with **Scikit-learn** nearest-neighbor, analyzed user-item rating matrices.
- Preprocessed sparse matrices with **feature engineering** and **data normalization**, filtered 200+ books and 50+ user ratings.
- Deployed **Streamlit** app on **AWS** with **Docker** for real-time recommendations and scalable infrastructure.
- Technologies:** Python, Scikit-learn, Streamlit, Docker, AWS, Kaggle, GitHub

### AI-Powered Medical Chatbot [↗](#)

(GitHub [↗](#))

LLM-Driven Healthcare Assistant with RAG and API Deployment

- Developed LLM medical assistant integrating Gemini via LangChain with RAG architecture on Pinecone vector database.
- Built document ingestion and semantic retrieval pipelines achieving sub-second response times for context-aware queries.
- Deployed Flask APIs with Redis caching, structured logging, and error handling for scalable real-time inference.
- Technologies:** Gemini, GPT-4, LangChain, Pinecone, Hugging Face, Flask, AWS, GitHub Actions

## Technical Skills

**Programming & Data:** Python, SQL, NumPy, Pandas, C++

**Data Science:** EDA, Feature Engineering, Data Preprocessing, Statistical Analysis, Data Visualization

**Machine Learning:** Classification, Regression, Clustering, Ensemble Methods, Hyperparameter Tuning, Scikit-learn

**Deep Learning & LLMs:** Transformers, Hugging Face, LangChain, RAG, GPT-4, Gemini, Azure OpenAI

**MLOps & Deployment:** FastAPI, REST APIs, Docker, MLflow, CI/CD, AWS, Azure, Model Deployment

## Achievements & Certifications

- Oracle Cloud Infrastructure 2025 Certified Generative AI Professional [Link](#)
- Complete Data Science, Machine Learning, Deep Learning, NLP Bootcamp — Krish Naik [Link](#)
- Solved 250+ problems on LeetCode/GfG with focus on algorithms, data structures, and problem-solving.

## Leadership & Responsibilities

### IIIT SriCity Cricket Team Captain

September 2024 - nov 2025

IIIT Sri City Intercollegiate Team

- Captained the IIIT Sri City cricket team in intercollegiate tournaments, leading match strategy and on-field performance.

### Events Coordinator at Abhisarga

March 2024 - December 2024

Annual College Technical Festival

- Managed technical event logistics for 1,500+ participants across 15+ competitions, ensuring 100% on-time execution.