

Rohit Kumar

+91 9798494187 - rohitcode005@gmail.com - [linkedin](#) - github.com/Rohitw3code - [kaggle](#)

TECHNICAL SKILLS

Programming Languages: Python, C++, C, Java, JavaScript

Libraries and Tools: TensorFlow, Transformers, PyTorch, PySpark (Databricks), Scikit-learn, Pandas, NumPy, Node.js, React.js, NLP, MySQL, Spark (ETL), Git, GitLab, Docker, MLOps, ZenML, MLflow, DVC, MongoDB, Linux

ML Architectures: CNN, RNN, GNN, YOLO, Transformers (BERT, LSTM, FLAN-T5, PEFT, LoRA, RLHF), VectorDB, Statistical Models, Reinforcement Learning, Unsupervised Learning

WORK EXPERIENCE

Software Engineer Intern

[Coreco Technologies](#), Pune, India

Internship

07/2024 - 11/2024

- Integrated **Yoda GPT** into a Slack bot within a CI/CD pipeline, streamlining deployment and updates.
- Designed and deployed an advanced **AI bot for Microsoft Teams** using Node.js and Azure App Services, actively used by Uptroop.io clients.

Data Scientist

[Enshrine Global Systems \(Gradstem\)](#), Michigan, USA

Remote Internship

06/2024 - Present

- Developed an automated resume parsing pipeline with **Selenium, Docker, and AWS**, improving application efficiency.
- Built GPT-driven models for **job matching and parsing**, analyzing 110k+ job listings and automating Workday entries.
- Boosted ATS scores by 30% through resume personalization and deployed scripts on AWS using Lambda functions.

Data Engineer

[Taiyo.ai](#), Silicon Valley Startup, USA

Remote Internship

03/2021 - 09/2021

- Engineered an OOP Python script for web scraping, automating data extraction and translation from global financial sources.
- Built scalable data pipelines using **Google Cloud, Docker, PySpark**, and Databricks for processing and visualization.

Prompt Engineer

Scale AI, Remote

09/2023 - 02/2024

- Optimized **LLM prompts** for generating code in data science tasks, improving accuracy and efficiency.
- Reviewed and refined AI-generated Python code to ensure adherence to best practices and performance standards.

PROJECTS

- SIH-Finalist (Smart Surveillance System for Large-Scale Events) (12/2024)** [Visit Live Project](#) Developed a system for facial recognition to locate missing persons and real-time emotion detection to prevent mishaps. Integrated multi-camera feeds for unified monitoring and tracked lost items via MAC addresses. Improved public safety with efficient live-feed analysis. Tech Stack: ReactJS, TailwindCSS, Flask, OpenCV.
- MLflow (CI/CD) (11/2023 - Recreated 11/2024)** [Visit Live Project](#) — [Video](#) — [Github](#) Built a no-code platform enabling streamlined data preprocessing, feature selection, and model training. Automated CI/CD pipelines for 37 algorithms, enhancing productivity and scalability. Used in academic ML projects by peers. Tech Stack: ReactJS, Flask, Pandas, NumPy.
- LLM Fine-Tuning Text Detoxification Using RL (09/2024)** [Github](#) Fine-tuned Flan-T5 for advanced summarization tasks with minimal computation using LoRA/PEFT. Applied RLHF for generating detoxified outputs while maintaining coherence. Optimized workflows for NLP applications. Tech Stack: Python, Transformers, RLHF.
- CCTV Video Violence Detection (MLOps CI/CD) (06/2024)** [Github](#) — [Deployed](#) — [Video](#) Achieved 97% accuracy in detecting violence in CCTV feeds using deep learning models. Deployed on AWS for scalability, with version control via DVC for reproducibility. Used in public area security pilot programs. Tech Stack: TensorFlow, AWS, Python, DVC.
- AI Multi-Agent System (AI Resume Creator) (05/2024)** [Github](#) Designed an intelligent system for resume generation optimized for job-specific requirements. Users could download resumes in LaTeX format and convert them to PDF. Improved job application efficiency. Tech Stack: CrewAI, Next.js, Flask, Python.
- AI ResumeRover (LLM LangChain) (09/2023)** [Video](#) — [Github](#) Built a platform enabling HR teams to filter resumes efficiently based on project requirements. Leveraged LangChain and FAISS for semantic search, saving hours of manual filtering. Tech Stack: LangChain, FAISS, Python, Streamlit.
- MedDiag: AI-Powered Disease Diagnosis (02/2024)** Developed a web application to classify medical images (e.g., fractures, skin diseases, tumors) using CNNs. Provided detailed condition explanations, symptoms, and precautions. Improved diagnosis accuracy and efficiency in healthcare pilots. Tech Stack: Python, TensorFlow, Flask.
- AI Maze Solver: Q-Learning (03/2024)** [Video](#) — [Github](#) Created a game-solving AI agent using Q-learning for optimal pathfinding in dynamic mazes. Introduced reinforcement learning concepts to students in workshops. Tech Stack: Python, Pygame, Q-Learning.

EDUCATION

Sarala Birla University

B.Tech in AI (Artificial Intelligence)

Kendriya Vidyalaya CRPF Ranchi

Class 12th CBSE: 92.2%

Ranchi, Jharkhand

11/2021 - 04/2025

Ranchi, Jharkhand

ACHIEVEMENTS

Samsung Solve for Tomorrow Top 30

CSR FITT (IIT Delhi), Onsite - Delhi

02/2023

- Ranked in the Top 30 out of 70,000+ participants nationwide for developing an innovative mobile application that provides instantaneous solutions and support to developers.

- [Laper App - Playstore](#)

Ranked Top 5 in Anveshana

Synopsys, Delhi

02/2024

- Developed a live IoT project utilizing ESP Cam, Arduino, and ML to prevent train accidents. Implemented obstacle and drowsiness detection for train drivers, winning cash prizes and goodies.