

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,Transformer,PyTorch,Pyspark (databricks), Sklearn, Pandas, Numpy, LLM, nodejs,reactjs, 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 , Reinforcement Learning , Unsupervised

WORK EXPERIENCE

Software engineer

[Coreco Technologies](#), Pune, India

Internship

07/2024 - 11/2024

- Integrated Yoda GPT into the Slack bot within a CI/CD pipeline for seamless deployment and updates.
- Designed and developed an advanced AI bot for Microsoft Teams using Node.js, managing the full development lifecycle.
- Successfully Deployed the Teams bot on **Azure App Services within a resource group**, now actively used by Uptroop.io clients.

Data Scientist

[Enshrine Global Systems \(Gradstem\)](#), Oakland County, Michigan, United States

Remote-Internship

06/2024 - 08/2024

- Developed an automated resume parsing and job application pipeline with Selenium, Docker, and AWS, increasing application efficiency and accuracy.
- Built GPT-driven models for job matching and resume parsing; performed data analysis on 110k+ job listings for model improvement. Created a job classification pipeline, categorizing 10k+ job listings, and automated form filling on Workday and Greenhouse using OpenAI API for personalized entries.
- Improved ATS scores by 30p.c through resume personalization and led the deployment of automation scripts on AWS with Lambda functions.

Data Engineer

[Taiyo.ai](#), San Francisco california, USA -Silicon Valley startup

Remote-internship

03/2021 - 09/2021

- **Web Scraping:** Engineered an OOP Python script for web scraping, extracting and standardizing data from World Bank and IADB websites with automated translation for consistency.
- Architected a scalable data pipeline using Google Cloud and Docker for efficient data transfer. Utilized PySpark on Databricks, along with Pandas and NumPy, for data preprocessing and employed Matplotlib and Seaborn for exploratory data analysis and visualization.

Prompt Engineer

Scale AI, remote

09/2023 - 02/2024

- **Enhanced prompt-based AI model :** Reviewed and optimized a language model (LLM) for generating code in data science and machine learning tasks
- **Code review :** Examined code generated by the LLM model, ensuring accuracy, efficiency, and adherence to best practices in Python and other programming languages.

PROJECTS

- **SIH-Finalist (Smart Surveillance System for Large-Scale Events) (12/2024)** [Visit Live Project](#) Facial Recognition: Located missing persons and children by matching faces. Multi-Camera Integration: Unify multiple webcams into a single monitoring system. Emotion Detection Fight Alerts: Detected distress, fights, and violence with real-time notifications to police.
Item Phone Tracking: Tracked lost items and detected stolen phones using MAC address tracking. Phone Snatching Detection: Identified snatching incidents through live camera feeds. Tech Stack: Built with ReactJS, TailwindCSS, and Flask for scalability and usability. Impact: Practical solution for improving safety at large events, ideal for MP Police implementation.
- **MLflow(CI/CD) (11/2023)-(Recreated 11/2024)** [Visit Live Project](#)— [Project Video](#) [Github](#) Developed a web tool for a data scientist using Flask,React,Tailwind CSS,Pandas,and NumPy.This platform enables data scientists to perform all data preprocessing task,feature selection, and machine learning model training tasks without writing code. It supports 17 regression and 20 classification algorithms and code generation tech stack : reactjs,Flask api,pandas,numpy,sklearn,javascript

- **LLM Fine-Tuning and Text Detoxification Using Reinforcement Learning : 9/2024** [Github](#)
 - - Fine-tuned Google **Flan-T5 model** using **PEFT/LoRA** on the DialogSum dataset for dialogue summarization. Applied Reinforcement Learning from Human Feedback (**RLHF**) for text detoxification using the Roberta-Hate-Speech-Dynabench dataset. Combined detoxification and summarization to generate clean, concise outputs.
- **CCTV Video Violence Detection Deep Learning (MLOps CI/CD pipeline) : 06/2024** [Github](#) —[Deployed-Visit Video](#)
 - * Developed a deep learning model for violence detection in real-time CCTV feeds, achieving 97p.c accuracy. Integrated an end-to-end pipeline using Flask, **Keras**, TensorFlow, NumPy for video preprocessing, and deployed on AWS, displaying timestamps and GIFs of detected violence for review. Utilized DVC for version control and **MLflow** to track accuracy and hyperparameters.
- **AI Multi-Agent System (AI Resume Creator): 05/2024** [Github](#) Developed advanced AI system that enables users to quickly and easily create optimized resumes tailored to specific job postings, increasing their chances of securing the job. This system generates LaTeX code and converts it into a PDF file. Tech : CrewAI, nextjs ,flask
- **AI ResumeRover (LLM langchain) (09/2023)** Functionality: [Project Video](#) [GitHub](#) A platform for HR to upload and manage hundreds of resumes simultaneously. Enabled HR to filter resumes efficiently based on project needs and employee. Technology Stack: Langchain and fine-tuned with OpenAI and FAISS , Python , Streamlit,vectorDB
- **MedDiag: AI-Powered Disease Diagnosis and Doctor Report Summary (02/2024)** Developed a web app that predicts disease types from images using deep learning with CNN. It supports three types of images: bone fractures, skin diseases, and CT scans of brain tumors. The site classifies the disease, displays symptoms, describes the condition, and suggests precautions.it also summarize doctor report cards for easier understanding.
- **(Reinforcement Learning) AI Maze Solver: Q-Learning Project (03/2024)** Created a maze game where an AI agent navigates towards the goal using the Q-Learning Algorithm.Developed entirely from scratch in Python pygame,problem solving skill [video show case](#) [Github](#)

EDUCATION

Sarala Birla University
B.Tech in AI (Artificial Intelligence)

Ranchi, Jharkhand
11/2021 - 04/2025

Kendriya Vidyalaya CRPF Ranchi
Class 12th CBSE 92.2%

Ranchi, Jharkhand

ACHIEVEMENTS

Samsung Solve for Tomorrow Top 30

CSR FITT (Corporate Relations - IIT Delhi), onsite - Delhi 02/2023

- Achieved a remarkable ranking in the Top 30 among more than 70,000 participants across india for developing an innovative mobile application that provide instantaneous solutions and support to a developer
[Laper App -playstore](#)

Ranked Top 5 in Anveshana

Synopsys , Delhi 02/2024

- Developed a live IoT project utilizing ESP Cam, Arduino, and machine learning for train accident prevention Implemented obstacle detection and drowsiness detection for the train driver, resulting in winning cash prize and goodies in a competition