Rohit Kumar

rohitcode005@gmail.com +91-9798494187 portfolio linkedin github kaggle

Technologies

- Programming and Query Languages: Python, SQL, JavaScript
- Machine Learning and Deep Learning: TensorFlow, OpenCV, CNN, Transformer, RLHF, LoRA, PEFT
- Big Data and Frameworks: PySpark, Databricks, Flask, Django, FastAPI
- Cloud and Tools: AWS lambda , AWS S3, Azure ML, Docker, Git, MLflow

Education

Sarala Birla University *Nov 2021 – May 2025*
B.Tech in AI (Artificial Intelligence)

Experience

AI Agent / Prompt Engineer *May 2025 – present*
PeptideWebMD (qubecare.ai)

- Building AI agents to automate end-to-end workflows for platforms like YouTube and Instagram. This includes content generation, video optimization, and engagement through comments and DMs. Also developed hiring agents for resume screening, job posting, and social media automation using Python, JavaScript, and n8n.

Data Extraction Engineer *Bhopal, India*
Relu Consultancy *Feb 2025 – Jun 2025*

- Developed a production-ready Chrome Extension using ReactJS for wishlist management and web scraping tool.
- Built Spring Boot APIs to process 300K+ influencer and financial data, using AI for stock/crypto predictions.
- Worked on Django HRMS backend, managed MySQL queries, and automated scraping tasks with cron jobs for seamless data workflows.

Data Scientist Intern *Oakland County*
Enshrine Global Systems (Gradstem) *June 2024 – April 2025*

- Built AI agents with persistent memory using FastAPI, integrating self-calling APIs and LangChain RAG systems with OpenAI function calling for context-aware document analysis and decision-making.
- Engineered a scalable ML pipeline for resume parsing using Python and AWS Lambda, enhancing talent acquisition efficiency and improving ATS match scores by 30% across 110K+ job applications.

Prompt Engineer *outlier.ai* *Remote Sept 2023 – Feb 2024*

- Review LLM model code geneartion Data science and ML training expert improved prompt generative AI model for better performance worked on google (bard i.e gemini) to review and train the ML model prompts

Data Engineer Intern *Taiyo.ai* *USA (Remote) Mar – Sept 2021*

- Designed a Big Data pipeline using PySpark on Databricks and Python to process World Bank datasets, integrating automated web scraping with Selenium to standardize multilingual data for ML models.

Projects

Built LLM from Scratch: GPT-2 Transformer Model *(Python, pytorch)* *GitHub*

- Built a GPT-2 based LLM in PyTorch with transformer blocks, multi-head self-attention, token/positional embeddings, and GELU-activated feed-forward layers.
- Used byte-pair encoding (5104 tokens) and a custom TikToken-based dataloader for efficient training.
- Pretrained for 10 epochs with loss-perplexity monitoring; applied temperature scaling, top-k sampling, and performed instruction/domain-specific fine-tuning.

AgenticHR - AI Agent *(Python, Solidity, AI Agent, Ethereum)*

- Built an AI-powered HR system with automated candidate extraction, summarization, and formatting. Created a custom vector DB and fine-tuned ”all-MiniLM-L6-v2” on resume–job datasets. Used a two-step selection: vector similarity search followed by an AI agent for top candidate ranking.

CCTV Video Violence Detection *(CNN, Transformer, OpenCV, TensorFlow)* *GitHub*

- Developed a deep learning system for real-time violence detection in CCTV feeds using CNN, ViT, and RNN/Transformer models, achieving 97% accuracy.
- Built a full pipeline with video preprocessing (OpenCV, NumPy), Flask-based deployment, and AWS hosting; generated timestamps and violence GIFs for easy review.
- Applied MLflow for tracking metrics and DVC for dataset/model versioning, showcasing strong computer vision

Fine-Tuning FLAN-T5 for Less-Toxic Text Summarization *(Flan-T5, python)* *GitHub*

- Fine-tuned FLAN-T5 with PPO and PEFT to create less-toxic summaries, using the DialogSum dataset and RoBERTa for toxicity evaluation. Achieved an 11.14% reduction in mean toxicity scores, leveraging Hugging Face Transformers and TRL libraries.

AI README.md Generator Live *(ReactJS, Flask, OpenAI APIs, Python)*

MLflow - Data Science and ML MLOps CI/CD Live (10/2024) *GitHub*

- Developed a full-stack web platform using ReactJS, Flask, and TailwindCSS that allows data scientists to perform data preprocessing, feature selection, and ML model training without writing code
- Enabled end-to-end workflows with automated code generation and support for 17 regression and 20 classification algorithms (e.g., XGBoost, SVM, Random Forest)
- **Tools:** ReactJS, Flask API, Pandas, NumPy, Scikit-learn, JavaScript, MLflow

TalkKAI – RAG-based Conversational AI Platform Live

YOLO from Scratch: Traffic Sign Object Detection *(TensorFlow,OpenCV)* *kaggle*

- Re-implemented the YOLO architecture from scratch using TensorFlow and NumPy by studying the original research paper and deriving all computations mathematically.
- Built a custom object detection pipeline for real-time traffic sign recognition using bounding box regression, anchor boxes, and IoU scoring.
- Gained in-depth understanding of computer vision fundamentals including grid-based detection, non-max suppression, and loss function design.

Air Drawing Math Solver & Invisibility Cloak *(OpenCV, HSV Masking)* *GitHub*

- Built an interactive air-drawing calculator using hand gesture and fingertip detection; supported draggable canvas, color selection, and open-hand gesture to move canvas in real-time.
- Integrated Google Gemini Vision API to extract and solve handwritten math expressions drawn in the air, with instant solution display on screen.
- Developed an invisibility cloak effect by detecting HSV color ranges and applying bitwise operations to blend background, creating real-time cloaking visuals.

Achievements

- **Samsung Solve for Tomorrow Top 30 Winner** - Ranked in the Top 30 of 70,000+ participants nationwide developed **Laper App**, an innovative mobile application that provides instant solutions for developers.
- **Winner , Anveshana Synopsys, Delhi (02/2024)** - Engineered an IoT project using ESP Cam, Arduino and ML to prevent train accidents with obstacle and drowsiness detection, winning cash prizes and goodies.
- **Amazon ML Challenge (09/2024)** - Secured All india Rank (AIR) 87 out of 75000+ Participants in Amazon ML challenge