DEEPAK KUMAR RAUTA
AI Developer | LLM Engineer | RAG & Agentic Systems Enthusiast
Phone: 6372757074 | Location: Hyderabad, Telangana
E-Mail: kdeepak8250@gmail.com | LinkedIn

PROFESSIONAL SUMMARY

AI/ML Developer specializing in Generative AI, RAG architectures, and multi-Agent systems. Leverages deep technical expertise in RLHF and model alignment from the Google Gemini Project to build robust, hallucination-free applications. Proficient in deploying scalable AI solutions using Python, FAISS, and Streamlit.

WORK EXPERIENCE

Associate – Google Gemini AI Project | Wipro **(April 2025 – Present)**

- Optimized Model Alignment via RLHF: Generated high-quality ground-truth data for Reinforcement Learning from Human Feedback (RLHF) loops, identifying complex logic gaps and edge cases to refine model reasoning.
- Technical Failure Analysis: Conducted root cause analysis on multimodal model outputs to detect hallucinations and safety vulnerabilities, ensuring robust adherence to Responsible AI frameworks.
- Data Integrity & Quality Assurance: Achieved 100% precision and recall in data annotation pipelines, minimizing dataset noise to provide high-fidelity training signals for model fine-tuning.

SKILL SUMMARY

- **Programming Languages:** Python, SQL
- **Artificial Intelligence & Machine Learning:** Scikit-learn, TensorFlow, PyTorch
- **Generative AI & Large Language Models:** LLaMA, Falcon, Mistral, Gemini AI, Prompt Engineering
- **Agentic AI & Retrieval Systems:** RAG Pipelines, Vector Databases (FAISS, Pinecone)
- **Tools & Frameworks:** Streamlit, GitHub, RESTful APIs, Microservices Architecture
- **Version Control:** Git, GitHub Actions

PROJECTS

1. **RAG Document Q&A Chatbot | GitHub Link**
   - Developed a context-aware chatbot capable of answering queries from uploaded documents and live web sources using a Retrieval-Augmented Generation (RAG) framework.
   - Engineered a pipeline integrating PDF parsing, vector embeddings (Sentence Transformers), and semantic retrieval via FAISS for precise and fast context matching.
   - Enhanced information accessibility and user interaction through real-time query processing and relevant web content integration.
   - Tech Stack: Python, Streamlit, FAISS, Sentence Transformers, Google Gemini AI, GitHub
2. **Vehicle Number Plate Detection | GitHub Link**
   - Developed an AI-powered web application for automatic vehicle number plate detection and text extraction using OpenCV and Tesseract OCR.
   - Implemented Haarcascade classifiers for plate detection and optical character recognition (OCR) for accurate number extraction.
   - Outcome: Delivered a reliable end-to-end system for real-time vehicle plate recognition with automated text extraction.
   - Tech Stack: Python (OpenCV, NumPy, PIL, Tesseract OCR), Streamlit

EDUCATION

- **Master of Computer Applications (MCA)** **(November 2022 - June 2024)**

  National Institute of Science & Technologies, Berhampur, Odisha

CERTIFICATION

- **Full Stack Data Science & AI** **(July 2024 – March 2025)**
  Naresh IT, Hyderabad, Telangana