Ashutosh Patra

 Pune, India   ashutosh2211@cc.cc   +91 9967556873   **in** ashutosh   ashutosh

Summary

AI Engineer with ~2 years of experience in **ML/DL, LLM fine-tuning, OCR/NLP pipelines, and MLOps lifecycle management**. Delivered AI solutions including **diffusion-based VFX automation, OCR-based document data extraction, and LoRA-driven LLM fine-tuning**, reducing manual effort by 70%. Skilled in **Python, TensorFlow, PyTorch, Hugging Face, and OpenAI APIs**, with experience driving **cross-team AI adoption in ethical and scalable ways**.

Work Experience

AI Data Engineer   *Pune, India*
Ancrew Global Services PVT LTD   *Jun 2025 – Present*

- Implemented Multimodal Model OCR pipelines (Anthropic Haiku, Langchain, Hugging Face) to automate invoice processing and reducing cost by **98%**
- Fine-tuned LLaMa-3-8b with LoRA and Unsloth for a Healthcare client and **increased efficiency by 30%**.

AI Architect   *Mumbai, India*
Blowfish Media Technology   *Oct 2024 - june 2025*

- Designed agentic workflows to detect and flag inappropriate content, scaling to **300+ sources/hour**.
- Designed a platform for LoRA training with security and lease-out with expiration timeline.
- Built an automated AI VFX diffusion pipeline for de-aging in feature films, **reducing turnaround by 70%**.
- Built an agentic workflow for an application to prepare storyboard from pdf script using Deepseek-R1, Flux pipeline **reducing time by 98%**
- Collaborated with cross-functional teams to deploy scalable APIs and ensure production stability.

Cloud Engineer Trainee   *Pune, India*
Teleglobal International PVT LTD.   *Jan 2024 - Sep 2024*

- Automated the PII discovery pipeline before ETL jobs for **Data Security**.
- MongoDB tables to ETL based transformation and automated pipeline for data discovery and updation.
- PoC on Agentic RAG based solution for a client on their Splunk UAT system.

Projects

VerveAI
Technologies: Langchain, AWS Bedrock, OpenSearch, Titan Text Embedding, Sentence Transformer, Docker

- Implemented text embedding with Titan Text Embedding.
- Developed a RAG chatbot using AWS Bedrock LLaMa 2 (13B parameters).
- Integrated OpenSearch to enhance vector search performance.

AzureVector
Technologies: Langchain, Streamlit, AzureML, LLaMa, Docker

- Built a RAG bot using LLaMa 3 (70B parameters) from AzureML model catalog.
- Implemented Conversation Buffer Memory for stateful interactions.

SiameseID
Technologies: Tensorflow, OpenCV, Siamese Network

- Developed a Siamese network for face verification, achieving precision comparable to modern mobile authentication systems.

- ○ Implemented triplet loss function from the Siamese network paper.
- ○ Achieved nearly 100% precision with a single anchor face image.

Education

Master of Science in Data Science	*Bhubaneswar, India*
Silicon University	*2022 - 2024*
Bachelor of Science in Mathematics	*Bhubaneswar, India*
Utkal University	*2018 - 2021*

Technologies

Languages: Python (Proeficient), Go (Beginner)
ML/DL: TensorFlow, PyTorch, Scikit-learn, Hugging Face Transformers
OCR/NLP: OpenCV, EasyOCR, Tesseract, spaCy, NLTK
Tools: AWS, Docker, Kubernetes, Git, GitHub Actions, MongoDB, PostgreSQL