

M. KAIF KOHARI

+44-7587754976 · kaifkohari10@gmail.com · [LinkedIn](#) · [GitHub](#) · [Medium](#)
London, UK

PROFILE SUMMARY

ML Engineer with 3+ years of experience and a proven track record of translating cutting-edge research into end-to-end robust, scalable solutions for commercial products, with a focus on LLMs, fine-tuning techniques and MLOps. Passionate about improving LLM reasoning, achieving SOTA results with small models and leveraging RL techniques.

KEY SKILLS

Python Javascript C++	RL, LLM finetuning	AWS
Machine Learning (Pytorch, Keras)	ML infra (Docker, Kubernetes, kubeflow)	REST APIs
Generative AI	MLflow, DVC, CI/CD	Linux
Dev (Django, Flask)	Advanced ETL & Data pipelines	SQL Databases
		Vector retrieval Databases

PROFESSIONAL EXPERIENCE

Syntagma Labs, London

MAY 2024 - NOV 2024

Machine Learning Intern

- Built ML algorithms for crypto transaction analysis to support AML investigations, using Pytorch and **networkx** to predict illicit transactions and detect anomalies with 97% accuracy.
- Conducted AI-driven research on terrorism financing and scam transactions with bitcoins using crypto APIs, LSTMs, and graph data visualization, providing critical insights into crypto risks

ZCLAP, Texas

AUG 2023 - DEC 2023

Machine Learning Engineer

- Deployed multiple advanced AI systems like video AI meeting assistant, official government document extractor, video summarizer using Large Language models, GANs and SOTA hugging-face stable diffusion models for text to image generation.
- Successfully commercialized these AI solutions, leading to their sales on the AWS Marketplace, enhancing client engagement and revenue streams for the company

Vsualthree60

AUG 2022 - NOV 2022

Computer Vision Intern

- Collaborated on the development of Computer Vision models to precisely perform emotion detection, people count, age/gender detection in videoframes, achieving seamless synchronization.
- Models were successfully integrated in CCTV cams across various supermarkets.

Tungsten Studios

JUN 2021 - JULY 2021

Data science Intern

- Created an algorithm to automate detection of elements (doors, windows) via pixel intensity tools with Python in architectural drawings, accelerating layout mapping by 40%.

ML Intern

- Executed critical safety Computer Vision systems, including Driver facial Drowsiness Detection system, an automated license plate recognition system for identifying vehicles violating traffic laws for a government NGO

PROJECTS**Finance Lab – Agentic research assistant for hedge funds and analysts**

Demo - <http://financelab.click:8000/>

- Engineered a **RAG** pipeline with advanced retrieval (rerankers, Qdrant vector DB and a finance-finetuned LLM) that extracts KPIs, earnings, revenue, etc from enormous reports like 10-Ks, 10-Qs in seconds, cutting analyst research time by ~90%
- Built a Dockerized FastAPI app and deployed on AWS ECS Fargate enabling natural language querying and structured outputs for analysts, traders, and portfolio managers.

Favourite Video Game Describer

<https://pht5symp5h.us-east-1.awsapprunner.com/>

- A fun agent that transforms your story descriptions into personalized game recommendations by analyzing gameplay and genres using MCP tool calling functionality and GPT 4o.
- A Dockerized Flask app, deployed on AWS cloud App Runner for scalability.

Pneumonia Diagnosis on X-rays with CNNs

<https://github.com/Kaif10/Pneumonia-Diagnosis>

- Created and trained 3 different CNN architectures with Keras achieving 90% sensitivity and 85% specificity for X-ray pneumonia detection, leveraging transfer learning and finetuning methods to handle a large, imbalanced medical dataset.

Emotion Detection from twitter profiles

https://github.com/Kaif10/Emotion_Detection_in_Tweets

- Trained a **bi-directional LSTM network** to do sentiment analysis on COVID-era Twitter data using BeautifulSoup and Selenium for real-time scraping of tweets from Twitter.
- Performed comprehensive ETL processes, including data cleaning, transformation, and structuring, to convert raw scraped tweets into a high-quality dataset optimized for ML model training.

EDUCATION & CERTIFICATIONS**Bachelor of Engineering**

University of Mumbai, SFIT (8.0/10)

Masters in Applied Data science

Royal Holloway, University of London (Distinction)

- Implemented Video Action-Recognition for Master's thesis, using (ResNet-152)CNN, LSTM, and Attention-Mechanism for crowd surveillance, achieving 97% accuracy.

COURSES & CERTIFICATIONS

- Advanced Computer Vision (Deeplearning.ai)
- Cloud Computing-AWS (IBM)
- Agile And Scrum (IBM)

EXTRACURRICULAR ACTIVITIES**AI/ML Intern (Hackathon Selection)**

Selected at DJSCE college hackathon for an AI internship with an NGO, applying ML to impactful social causes.

Treasurer and Technical Head, Computer Society of India (CSI)

Led a team of 26 in organizing workshops on data science, communication, and leadership. Hosted a club talk show, interviewing prominent tech professionals. - The CSI Show

Machine Learning Content Creator

Created ML content for Medium, Orchids International School and UpGrad KnowledgeHut simplifying complex tech topics.