

M. KAIF KOHARI

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

PROFILE SUMMARY

ML Research Engineer with multiple years of experience, often as the founding ML engineer at growing and scaling teams. Proven track record of translating cutting-edge research into real world, robust and scalable infrastructure for commercial products. Passionate about LLM reasoning, Multi-Agent systems and low latency programming.

KEY SKILLS

Python C++ JavaScript	CI/CD automation	AWS
Machine Learning (Pytorch, Tensorflow)	Infra (Docker, Kubernetes, Kubeflow)	REST APIs
Agentic AI	Airflow, Data Version Control	Linux
Dev (Django, Flask)	Advanced ETL/Data Pipelines	SQL Databases
		Vector Databases

PROFESSIONAL EXPERIENCE

Stealth mode, London

MARCH 2025 - PRESENT

ML Consultant

- Advising stealth startups on integrating multi-agent systems in their workflows.
- Built an agentic finance model for a stealth firm to be used by top film producers that converts film production requirements into structured and deterministic cross-territory funding plans.

Syntagma Labs, London

MAY 2024 - NOV 2024

Machine Learning Research 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, scam transactions with bitcoins using gradient boosting, LSTMs, and graph data visualization, providing 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 with OpenCV, Yolo to precisely perform emotion detection, people count, age/gender detection in videoframes, achieving seamless synchronization which were 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**

- Engineered an Agentic RAG pipeline with langchain and advanced retrieval (rerankers, Qdrant vector DB and a finance-finetuned LLM) that extracts and calculates KPIs, earnings, etc from enormous SEC filings like 10-Ks, 10-Qs in seconds, cutting 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**Demo**

- 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 with Convolutional Neural Networks**Github**

- 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 on twitter profiles**Github**

- 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 using pandas, NLTK, 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 (2023)**

University of Mumbai, SFIT (8.0/10)

Masters in Applied Data science (2025)

University of London (Distinction)

Research - Spatiotemporal Action Recognition in Videos Using ConvLSTM with Attention: A Comparative Analysis and Implementation

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

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