

YUVAL MEHTA

+91 7045983199 — yuvalmehta.728@gmail.com — linkedin.com/in/yuvalmehta728 — github.com/Yuval728

Summary

Versatile AI/ML Engineer with hands-on experience in agentic systems, LLM orchestration, and scalable ML deployments. Specialized in LangGraph-based multi-agent coordination, MLOps, and deep learning pipelines. Proven ability to drive end-to-end AI development in high-compliance, production-grade environments like GxP.

Education

NMIMS - Mukesh Patel School of Technology, Management and Engineering

June 2025

Bachelor of Technology in Computer Engineering — CGPA: 3.84/4

Mumbai, India

Experience

xLM - Continuous Intelligence

June 2025 – October 2025

AI/ML Engineer

Mumbai, Maharashtra

- Engineered AI agents within **cIV**, automating GxP compliance workflows and expanding task coverage by **65%**.
- Orchestrated LangGraph-based multi-agent systems with retry, memory, and control flows, slashing execution time by **30%** and raising success rate by **40%**.
- Collaborated with engineering and QA to embed traceable logic, enhancing audit-readiness across workflows.

xLM - Continuous Intelligence

January 2025 – May 2025

AI/ML Intern

Mumbai, Maharashtra

- Constructed a traceability matrix generator, reducing manual overhead by **60%** and boosting consistency by **45%**.
- Prototyped 3 AI-driven document intelligence solutions, accelerating internal validation cycles by **50%**.
- Deployed real-time logging and QA pipelines, achieving **100%** traceability in early-stage toolchains.

IIT Kharagpur

July 2024 – May 2025

Machine Learning Intern

Kharagpur, West Bengal

- Devised a video feature extractor using autoencoders and GNNs, improving frame processing efficiency by **30%**.
- Optimized hyperparameters to cut training duration by **25%** and raise validation accuracy by **18%**.

JM Financial Ltd

July 2024 – November 2024

Data Science Intern

Mumbai, Maharashtra

- Streamlined KYC document verification with OCR and deep learning, reducing processing time by **40%**.
- Synthesized dashboards and data pipelines, enhancing insight delivery throughput by **3×**.

Projects

AI Therapist | Python, LangGraph, LLMs, Supabase, Vector DB, FastAPI

- Architected an LLM-powered therapy agent with emotion control and memory, improving session consistency by **70%**.
- Refined empathy alignment by **52%** through feedback-driven tuning in simulated conversation loops.

LipReader AI | Python, PyTorch, OpenCV, MLflow, TorchServe, Docker

- Engineered a lip-to-text model with **87%** character accuracy and **25%** faster inference via TorchServe.
- Streamlined experiment cycles using MLflow, trimming iteration time by **10%**.

AQI Prediction | Python, PyTorch, GCP, FastAPI, MLOps, WandB

- Devised a dual-view attention model to forecast AQI and 6 pollutants, achieving **93% R²** accuracy.
- Minimized cloud training expenditure by **35%** through targeted optimization and automated sweeps.

WS Buddy | PythonLangGraph, QDrant, Whisper, LLM, Diffusion, Docker

- Developed a multimodal chatbot integrating voice, image, and context memory, reaching **85%** task success.
- Elevated engagement by **30%** through dynamic activity updates and real-time interaction.

Technical Skills

Programming Languages: Python, JavaScript, SQL, C, C++, Java

Libraries: Scikit-learn, NumPy, Pandas, Matplotlib, OpenCV, NLTK, SpaCy, SciPy, HuggingFace, Weights & Biases

Frameworks: PyTorch, TensorFlow, Keras, MLFlow, Langgraph, Langchain, FastAPI, Django, Flask, CrewAI

Databases: MySQL, PostgreSQL, Supabase, SQLite, MongoDB, ChromaDB, Redis, FAISS, Qdrant

Tools: VS Code, Jupyter, Kaggle, Git, GitHub, GitHub Actions, DVC, Docker, Nginx, Postman

Cloud Platforms: AWS, GCP

Achievements

Achieved **274th out of 74,824 participants** (Top 1%) in the Amazon ML Challenge 2024, with a team of 3-4 members.

Published a paper at IEEE APCIT 2024: Examining ML Approaches for Early Diabetes Prediction.