Akash Pawar

Jersey City, NJ | akashpawar9619@gmail.com | (551) 328-5269 | LinkedIn | Github | Portfolio

WORK EXPERIENCE

Stevens Institute of Technology

Artificial Intelligence and Machine Learning Research Engineer — **June 2025 - Present**

- Architecting full-stack AI career chatbot (FastAPI, React, LangChain, GPT-4o) for personalized guidance and skills analysis.
- Prototyping Qdrant search over 300+ courses to recommend context-aware learning paths.
- Building resume parsing, career inference, and secure MongoDB/Auth0 backend for multi-platform chats.

Artificial Intelligence and Machine Learning Research Assistant — **April 2025 - June 2025**

- Designed a LangGraph-driven GPT evaluator that cut assignment-grading time 95% while preserving rubric fidelity.
- Engineered conditional routing to score Python scripts, Jupyter notebooks, LaTeX math, and cross-referenced analysis PDFs.
- Defined Pydantic state schemas so LLMs emit clean JSON, enabling one-click LMS feedback uploads.

ElevateMe

Data Analytics & Machine Learning Fellow Trainee — **March 2025 - Present**

- Engineered an end-to-end Azure SQL to Python ML pipeline spanning 300K+ contracts across 180 tables; feature-engineered genre, artist-experience, and commission signals that lifted gross-revenue R^2 from 0.44 to 0.80 and priced deals within ±$170.
- Mined cross-market pricing intelligence—discovering 2.3× multipliers on multi-artist events and commissions as the strongest revenue driver, powering precise contract decisions.
- Produced interactive data visualizations using Tableau highlighting artist pricing variations and multi-year revenue trends, enabling executives to make data-driven decisions for talent acquisition and market expansion.

PROJECTS

Deepseek for Advanced Mathematical Reasoning | Link — **February 2025**

- Fine-tuned DeepSeek-R1-Distill-Qwen-1.5B model that outperformed Claude-3.5 Sonnet on mathematical reasoning tasks.
- Reduced trainable parameters by 98.8% while doubling inference speed through LoRA adaptation and Unsloth framework.
- Prepared dual data preprocessing pipelines optimized for both low-RAM (16GB) and high-RAM environments using pandas and functional programming.

Protein Subcellular Localization Predictor using ESM2 | Link — **February 2025**

- Devised a high-accuracy protein localization system using Meta AI's ESM2-3B, achieving 84.79% top-3 and 92.09% top-5 accuracy across 12 cellular locations, demonstrating expertise in protein bioinformatics and deep learning.
- Optimized training performance by 1.8x through mixed precision and gradient checkpointing, reducing memory footprint by 50% while maintaining stability.
- Designed end-to-end pipeline from UniProt API data acquisition to model training and extensive evaluation.

Dependency Chain Analysis | Link — **December 2024**

- Processed 500,000+ relationships across 442,275 nodes in a large-scale dependency graph using Neo4j and NetworkX, extracting key structural patterns in under 25 seconds.
- Crafted 5 topological features and 20 semantic attributes, including centralities, local risk ratio, scope, and security metrics, contributing ~80% to classifier feature importance and enabling 100% accuracy in critical node identification.
- Implemented Node2Vec embeddings (128-dim) combined with handcrafted features, achieving 100% precision, recall, and F1-score on 88,455 nodes using Random Forest, compared to a baseline F1-score of 29% without custom features.

EDUCATION

Stevens Institute of Technology, Hoboken, NJ — **September 2023 - December 2024**

Master of Science, Machine Learning, (GPA: **3.93/4**)

Mumbai University, Mumbai, India — **August 2019 - May 2023**

Bachelor of Engineering, Computer Science, (GPA: **9.07/10**)

SKILLS & INTERESTS

Programming & Data Engineering: Python, SQL, pandas, numpy, Dask, DuckDB, subprocess, regex, JSON, Pydantic, ETL
Visualization & Analytics: seaborn, matplotlib, Plotly, Tableau, NetworkX, Neo4j, Streamlit
ML & NLP Frameworks: PyTorch, TensorFlow, scikit-learn, XGBoost, LightGBM, Hugging Face, LangChain, OpenAI, spaCy, LangGraph, LangSmith, Chroma, NLTK, FLAML, RLHF, LlamaIndex, PEFT (LoRA, bitsandbytes, Unsloth), Pinecone, Qdrant
MLOps & Cloud: Docker, AWS (S3, Lambda, SageMaker), Azure ML, Flask, FastAPI, Gradio, Git, MLflow, Weights & Biases
Interests: Lifelong learning, modern physics (especially relativity), exploring fascinating ideas, music, and recently crafting a second brain in Obsidian with Claude 3.7 Sonnet and MCP server integration.