Amaan Poonawala

amaanpoonawala05@gmail.com | +91 87792 76971
github.com/AmaanP314 | linkedin.com/in/amaan-poonawala | Portfolio

Skills

Languages: Python, JavaScript, TypeScript, SQL
AI Frameworks: LangChain, LangGraph, CrewAI
ML Libraries & Frameworks: Pandas, NumPy, Matplotlib, Seaborn, scikit-learn, OpenCV, MediaPipe, TensorFlow, PyTorch, Keras, Transformers
Databases: PostgreSQL, MongoDB, Redis, Pinecone, FAISS
AI/ML Concepts: Supervised & Unsupervised Learning, Large Language Models (LLMs), Fine-tuning, QLoRA, Agentic AI, Retrieval-Augmented Generation (RAG)
Technologies & Tools: Git, Docker, Streamlit, FastAPI, Jupyter, Hugging Face (Models & Transformers), AWS

Projects

Youtube AI Analyzer | Github | Live | Post — Jun 2025

- Build a full-stack web application that analyzes YouTube videos by summarizing video content and comments, performing sentiment analysis on comments, and visualizing comparative metrics across search results. It also supports natural language Q&A for both video and comment content.
- Built with Next.js (frontend) and FastAPI (backend), it utilizes large language models (LLMs) integrated via LangChain and FAISS for summarization and Q&A, with a model fine-tuned on YouTube comments to support sentiment analysis.
- Designed to reduce time spent on low-value content and help creators extract structured feedback via AI-powered summaries, interactive Q&As, and factual analysis that identifies accurate or potentially misleading statements in videos and comments.
- Python, Large Language Models (LLMs), LangChain, FAISS, YouTube-XLM-RoBERTa (comment sentiment), YouTube API, FastAPI, PostgreSQL, Next.js, Tailwind CSS

Fine-Tune XLM-RoBERTa | Github | Model — Feb 2025

- Fine-tuned a multilingual XLM-RoBERTa model for sentiment classification of YouTube comments using a dataset of over 1,000,000 labeled YouTube comments, improving accuracy from 69.3% to 80.17%.
- Adapted cardiffnlp/twitter-xlm-roberta-base-sentiment-multilingual using Hugging Face Transformers with PyTorch, applying label smoothing, early stopping, and custom preprocessing to handle comment-specific noise.
- Downloaded over 2,000 times by developers building content analysis and feedback systems, the model was designed to improve sentiment analysis on YouTube comments, which differ in tone, slang, and structure from other platforms.
- Python, PyTorch, Transformers, XLM-RoBERTa, Pandas, Scikit-learn, Datasets, Jupyter

Portfolio AI Assistant | Code | Chat — Mar 2025

- Developed a Agentic RAG based chatbot integrated into my portfolio designed to answer questions about my professional background, projects, skills, and learning journey
- Implemented a semantic chunking strategy, a custom hybrid search with a dense-only fallback, and conversational query rephrasing and a reranker. These methods improve retrieval of relevant information.
- Resolves challenges like context loss in large documents, unreliable search results, and disconnected conversational threads. It achieves this by autonomously adapting document splitting, search mechanisms, and user query processing.
- FastAPI, LangChain, Pinecone, BM25, Google Generative AI (for LLM and Embeddings), Agentic RAG (Retrieval-Augmented Generation), FlashRank, Python

Education

Thakur College of Engineering & Technology — Aug 2023 - Jun 2027 (Expected)
B.tech in Artificial intelligence & Data Science
Relevant Coursework: Databases, Discrete Maths, Data Structures and Algorithms, Operating Systems, Computer Networks

Thakur College of Science & Commerce — Sep 2021 - Apr 2023
HSC in Information Technology
Relevant Coursework: HTML, CSS, JavaScript, PHP

Courses & Certificates

Convolutional Neural Networks – Coursera: Built CNNs for image classification using PyTorch.
Complete Data Science Bootcamp – Udemy: Learned core Machine & Deep learning concepts.
LLM App Development – Deeplearning.ai: Learned LangChain and advanced RAG for building LLM-based applications.