AYAAN AHMED KHAN

(773)-937-2448 | ayaanahmedkhan12@gmail.com | in/ayaanahmedkhan/ | github.com/ayaan47-1 | Chicago, IL

EDUCATION

ILLINOIS INSTITUTE OF TECHNOLOGY | Chicago, Illinois **Expected May 2026**
Bachelor of Science in Artificial Intelligence; **Minor** in Architecture
Relevant Coursework: DSA, AI, ML, NLP, DBMS, Assembly, Data Mining, Discrete Math, Linear Algebra, Probability, Statistics, OOP

SUMMARY

AI engineer and backend developer building production LLM systems, including retrieval augmented generation, semantic search, and multi stage agent workflows. Experienced with asynchronous application programming interfaces, vector databases, Redis caching, job orchestration with retries, real time streaming, and observability.

WORK EXPERIENCE

AI & DIGITAL DEVELOPMENT INTERN | Chicago, Illinois **May 2025 - Present**
The Syntec Group

- Built an internal RAG chatbot over firm documents delivering cited, context grounded answers to reduce lookup time and improve response consistency.
- Implemented ingestion & retrieval workflow (chunking/embeddings/indexing) with emphasis on source traceability and hallucination reduction.
- Led an information architecture plus website redesign improving navigation and access to resources; used engagement analysis to iterate content performance.

PROJECTS

CLUTCH | Personal Project **January 2025 - Present**

- Built a staging deployed SaaS that generates research backed courses using a multi stage agent pipeline, orchestrating 5 stages with persisted job state, retry policies, and failure isolation to keep long running workflows reliable.
- Implemented an asynchronous FastAPI backend with PostgreSQL plus pgvector and Redis caching, and streamed real time job progress to clients using SSE backed by Redis to keep generation workflows responsive.
- Cut inference cost and improved output quality using LiteLLM routing and shared source deduplication; enforced structured outputs with PydanticAI validation and added observability via Sentry and PostHog.

SYNTEC AI CHATBOT | Professional Project **May 2025 - Present**

- Shipped a semantic search and retrieval augmented chatbot using GPT-4o and ChromaDB embeddings to deliver question answering across PDF documents, blog posts, and website content with automated source citations and context aware responses.
- Built a production oriented system with a React user interface and Flask application programming interface, integrated WordPress blog ingestion plus web scraping, and added Redis caching that reduced inference cost by 65%.
- Containerized the service with Docker for consistent deployments and enabled dual internal and client facing usage with security controls and access restrictions.

INVESTOCHAT | Personal Project **October 2025 - Present**

- Built a FastAPI service with PostgreSQL plus pgvector that powers a retrieval augmented investment chatbot over real estate brochures and documents.
- Implemented multi path retrieval with vector search plus SQL fallback and MMR to improve answer reliability.
- Integrating WhatsApp Business lead qualification with Airtable CRM sync, adding rate limiting and PII detection for safer client facing operation.

SENTIMENT ANALYZER | Academic Project **September 2025 - December 2025**

- Engineered a machine learning classifier using Logistic Regression with TF-IDF features (1-2 grams) to analyze over 50,000 Amazon fine food reviews, achieving high accuracy through L2 regularization and balanced class weighting on preprocessed text data.
- Built an interactive Streamlit web application with real time sentiment prediction and confidence scoring, implementing complete ML pipeline from data preprocessing (noise removal, text cleaning) to model deployment.

SPORTS AND METRICS TRACKER | Personal Project **January 2026**

- Built an end-to-end soccer video analysis pipeline using YOLOv8 detection and ByteTrack tracking, adding temporal smoothing for stable motion analytics and CPU only local processing with robust cross-platform video input and output.
- Developed export and visualization tooling, including CSV and JSON metrics with track filtering (minimum 15 frames), heatmap rendering with percentile transforms, and tiered error handling with debug mode progress reporting.

TREND ANALYZER FOR RAW MATERIALS | Personal Project **March 2025 - May 2025**

- Developed a cotton price forecasting system using Facebook Prophet and external regressors (crude oil, natural gas, soybeans), achieving 19.55% MAPE on 6 month forecasts with 755 days of real commodity futures data.
- Built a time series forecasting model with cross validation that improved 1 year prediction accuracy by 27.9% through multivariate regression analysis of correlated commodity markets (0.77 correlation with oil prices.)

SKILLS

Languages: Python, Java, SQL, JavaScript, R, Bash, C++, Svelte, TypeScript
Backend: FastAPI**,** Flask, SSE, WebSockets
Data: PostgreSQL, pgvector, Redis, ChromaDB
ML and LLM: RAG, embeddings, LangChain, Hugging Face, scikit-learn, PyTorch, TensorFlow
Infra and Observability: Docker, Git, CI and CD, Sentry, PostHog