

SAMYAK NAHAR

📞 +91 8080319065 ✉ samyaknabar004@gmail.com 🌐 github.com/samyak197 🔗 linkedin.com/in/samyaknabar/ 🌐 samyaknabar.com

Summary

AI Engineer with 2+ years of hands-on experience shipping production LLM and agentic systems end-to-end — data pipelines, RAG, tool-using agents, and AWS deployment. Currently driving the technical direction of Retail Agent, a revenue-intelligence agentic system at Propheus used by enterprise US retail clients.

Experience

Propheus

Jul 2025 – Present

AI-ML Engineer

On-Site

- Building **Retail Agent**, an agentic system giving retailers a **30-day** forward view of nearby events, competitor promotions, and weather; on the project since inception and now driving its technical direction.
- Architected a three-stage pipeline: **Serper** and events/weather API integrations capture raw signals; **OpenAI** and **Gemini** agents with custom **LlamaIndex** tools and **ChromaDB**-backed retrieval classify store-relevant signals; a deterministic (non-LLM) model computes exact USD revenue impact.
- Generated counter-promotion recommendations that helped retailers offset day-to-day revenue swings of **20–40%**, driving an average **8%** per-store revenue lift for a leading US retail brand.
- Led scraping and market-analysis engagements for enterprise clients, including **Gomart**'s new dark-store launch: analyzed competitor dark stores, local demographics, and buyer preferences to design a premium-area product assortment that drove strong sell-through and margins post-launch.

Syngenta

Jan 2025 – Jul 2025

AI-ML Intern

On-Site

- Engineered an ML pipeline using Python, SageMaker, S3, and Redshift to forecast monthly sales for **40+** products across **4** countries, achieving **5–14%** MAPE.
- Implemented an **XGBoost** demand model for stock management, reducing over- and under-stocking of inventory across country-level warehouses.
- Developed AI agents using open-source LLMs and AWS Bedrock to translate natural language into SQL, giving **10k+** non-technical users near-instant access to a **23GB** database.
- Leveraged AWS SageMaker Pipelines to automate data preprocessing, model training, and evaluation, triggered on a schedule or new data commits.

EduPlus

Aug 2024 – Dec 2024

GenAI Intern

Remote

- Engineered a Q&A chatbot using **Llama 3**, FastAPI, and ReactJS, serving **1,500+** student inquiries annually with low-latency cached retrieval for common questions.
- Deployed Llama 3 on AWS SageMaker with a similarity-threshold gate that filters low-relevance retrievals before generation, cutting off-topic responses.
- Built a CI/CD pipeline with GitHub Actions for automated AWS EC2 deployments.

Millennium Semiconductors

Feb 2024 – Jul 2024

AI-ML Intern

On-Site

- Built a cross-brand component-matching engine: when a customer requests a part (MOSFETs, ICs, etc.) from a brand outside Millennium's supplier line card, the system finds the closest equivalent from internal-provider brands by comparing electrical specs — voltage, current, temperature range, topology, and more.
- Aggregated component data from **50+** sources via scraping pipelines and distributor APIs (specs not readily available in structured form), normalizing heterogeneous datasheet parameters into a comparable schema; where structured parameters were unavailable, fell back to embedding-based semantic similarity over product descriptions.
- Shipped it as a semantic search engine for the business team, replacing manual datasheet-by-datasheet comparison; received commendation for organization-wide productivity gains.

Projects

NyayaBodh – GenAI Legal Research Engine | LLM, NER, NLP, RAG | [GitHub](#)

- Developed a platform enabling judges and legal researchers to search case precedents via natural language queries, eliminating manual journal searches.
- Implemented advanced NER algorithms achieving **96%** accuracy extracting entities such as places, courts, judges, lawyers, and witnesses from case PDFs.
- Automated legal document summaries using GenAI, eliminating manual summarization and potential human error.
- Received positive feedback from **2** practicing lawyers for reducing case research time from weeks to hours.

AI-Powered Video Navigation System | *Whisper, LLaMA, Open-Source* | [Colab](#)

- Built an open-source system for intelligent video navigation, processing **300+** hours of technical training content for a leading automation company.
- Implemented Whisper for speech-to-text with timestamp-based navigation, and LLaMA for natural language queries that jump users directly to the relevant moment — eliminating manual video scanning.

Text-Generation Transformer From Scratch | *PyTorch, Self-Attention, Tokenization* | [GitHub](#)

- Implemented a Transformer in pure PyTorch — multi-head self-attention, positional encoding, and a custom BPE tokenizer following Karpathy's approach — with no high-level libraries.
- Trained on TV-series scripts with Adam on an NVIDIA RTX 4090 (**24GB** VRAM) and cloud GPUs to generate coherent script-style text.

Skills & Interests

Languages: Python, SQL, C++

GenAI & Agents: LlamaIndex (agents & custom tools), RAG, function calling / tool use, prompt engineering, LLM evals; OpenAI, Gemini, Llama, Whisper, AWS Bedrock, Serper

Vector Stores: ChromaDB, pgvector, embedding models, semantic search

ML: PyTorch, TensorFlow, XGBoost, HuggingFace, Transformers, NLP, NER, CNNs, time-series forecasting

Cloud & MLOps: AWS SageMaker & SageMaker Pipelines, EC2, S3, Redshift, Docker, GitHub Actions CI/CD

Web & Databases: FastAPI, Django, React, PostgreSQL, MySQL, MongoDB

Interests: Chess, Football, Music

Education

VIIT College of Engineering

B-Tech in Artificial Intelligence & Data Science

Pune, Maharashtra

2021 – 2025

Leadership Experience

Runner-up (2nd Place) – Smart India Hackathon 2024 Grand Finale

- Placed 2nd at the national Grand Finale after selection among **380** teams; built an AI-driven Legal Research Engine prototype to expedite legal research.

Runner-up (2nd Place) – DataHacks III Hackathon

- Developed a chatbot that ingests IT infrastructure documentation, asks targeted follow-up questions to fill information gaps, and generates a comprehensive cybersecurity report.

Winner (1st Place) – Vortexa Tech Hackathon

- Won first prize among over **250** participants, developing a Legal Document Generation Engine in **12** hours using LaTeX compilers and LLMs.