HARIHARAN CHANDRASEKAR

Boston, MA
hariharanchandrasekar2003@gmail.com || +1(617)708 – 6384 | + 917358489274
linkedin.com/in/hariharan-chandrasekar || github.com/Hariharan-afk

Education

Northeastern University, Boston, USA — 2024 - 2026
Master of Science in Artificial Intelligence — GPA: 3.42

SRM Institute of Science and Technology, Chennai, India — 2020 - 2024
B. Tech in Computer Science (AI & ML Specialization) — GPA: 9.14 / 10

Technical Skills

Languages: Python, SQL
Libraries & Frameworks: PyTorch, TensorFlow/Keras, Transformers, Scikit-learn, XGBoost, Hugging Face Transformers, LangGraph, Sentence Transformer, SpaCy, FAISS, YOLO, PaddleOCR, EasyOCR, FastAPI, Streamlit
Tools / Stack: Git/GitHub, Linux, Docker, Kubernetes (GKE), GitHub Actions (CI/CD), Airflow, DVC, Qdrant, PostgreSQL, Google Cloud: Cloud Run, Artifact Registry, Cloud Logging, Cloud Monitoring
Domain: Agentic AI / Multi-Agent Systems, Retrieval-Augmented Generation (RAG), Document AI (Invoice extraction), OCR/VLM pipelines, Information Retrieval (BM25 + Vector Search, reranking), Computer Vision (Object Detection), Cloud-native applications, microservices, containerization, MLOps / Production ML (CI/CD, deployment, observability)

Work Experience

Product Development Intern - Yantran LLC, Allen, TX — Jun 2025 – Aug 2025

- Built an Agentic AI system to dynamically extract and validate structured data for Invoice documents and documented development processes, system design, and deployment steps for future team use and maintainability.
- Prototyped a multi-stage OCR/VLM agent: Paddle OCR/Easy OCR → query-aware cropping (semantic re-ranking + Grounding DINO) → answer extraction with Donut / Florence-2, improving recall on key fields (totals, dates, vendor info).
- Achieved extraction accuracy of 95% in a test set of 100 invoice documents.

Projects

Automated Due Diligence & Market Intelligence Agent — Sep 2025 – Dec 2025

- Built an end-to-end market-intelligence system that generates automated due-diligence reports from heterogeneous sources (SEC filings, Wikipedia, NewsAPI).
- Implemented a production RAG pipeline: ingestion → cleaning/structuring → chunking (~800 tokens) → embeddings (BGE-large, 1024-dim) → validation → storage (Qdrant vectors + Postgres metadata).
- Designed a multi-agent workflow (Orchestrator → Planner → Researcher → Synthesizer → Evaluator) with loop-back quality control; combined BM25 + vector retrieval for higher-precision evidence gathering.
- Created an evaluation harness with synthetic test generation, ground-truth JSON test loading, bias checks, and metrics reporting (relevance/coverage/correctness) to gate releases.
- Productionized with CI/CD: tests/linting + threshold checks → Docker image build/push (Artifact Registry) → rolling deploy to GKE (API) + Cloud Run (UI).
- Added observability: deployment notifications, API/pod logging, health checks, latency and error monitoring (Cloud Logging/Monitoring).

RAG-Based QA System for Medical Drugs — Jan 2025 – April 2025

- Designed a retrieval-based QA system for drug-related queries using web-scraped data from MayoClinic.org.
- Built a full retrieval pipeline using MiniLM-V6 embedding, FAISS vector search, and Sentence-BioBERT reranking.
- Integrated Groq-hosted Llama-3.1 LLM for grounded answer generation based strictly on retrieved chunks.
- Achieved 87.5% retrieval accuracy and 60.96% F1@3, significantly outperforming traditional methods like BM25.

Multi-Label Classification of Thoracic Diseases — Jan 2025 – April 2025

- Developed a multi-label classifier for 15 thoracic conditions using the NIH ChestX-ray14 dataset, addressing the challenge of class imbalance.
- Implemented a custom CNN using DenseNet121 with focal loss and threshold tuning for imbalanced data.
- Personally proposed and built a centroid-based synthetic sampling method to augment rare label combinations.
- CNN achieved AUC > 0.94 for most labels, while XGBoost achieved AUC = 0.98 for rare diseases like Hernia.

Achievements & Publications

- Awarded with Scholarship amount of 25,000 USD from Northeastern University as a recognition of academic excellence
- Presented paper at ICESI 2024: “Probabilistic Eye-State Classification from EEG Signals”
- Gold Medal — Paper Presentation on “Automated Aeroponics System” at SRM Research Day 2023