Nischal Chandur

+1(858)-241-6448 | chandur.nischal2@gmail.com

nchandur.github.io/portfolio | linkedin.com/in/nchandur | github.com/nchandur

Professional Experience

Data Science Graduate Intern, Ecolab – Naperville, IL, USA — Jun 2024 – Aug 2024

- Developed and implemented a multilingual RAG pipeline for assisting global field agents in diagnosing and resolving machine malfunctions, offering step-by-step guidance and product recommendations based on historical logs and training manuals.
- Established an ML model validation framework and synthetic data generator to enhance the reliability of predictive maintenance systems by balancing datasets and improving accuracy in predicting rare anomalous events without depending heavily on infrequent real-world failures.
- Communicated project vision and outcomes to directors and executive leaders, facilitating alignment across different teams and supporting well-informed decision-making.

Machine Learning Engineer, Reworked.ai – Miami, FL, USA — Apr 2024 – May 2024

- Developed and deployed a statistically-backed ML lead generation algorithm, optimizing customer acquisition for solar installations by integrating demographic, geospatial, and solar-specific features to predict homeowner adoption likelihood.
- Collaborated closely with leadership to refine the lead generation process, delivering a production-ready ML predictor for seamless system integration, enhancing high-potential customer identification and outreach strategies.

Data Scientist, Latlong (ONZE Technologies Pvt. Ltd.) – Bengaluru, KA, India — Sep 2022 – Jun 2023

- Led a team in developing data extraction and ML pipelines, including a multilingual OCR-NLP pipeline for structured demographic data extraction and geospatial database augmentation.
- Developed and deployed geospatial risk models for an NBFC, integrating location data to predict loan repayment and visualize regional defaulting rates for refined strategies.
- Delivered advanced spatial analytics and strategic insights to an automotive client, optimizing sales and marketing through analysis of geospatial and demographic data. Presented findings to executives for decision-making support.

Academic Projects

GoReads! – Scalable Content- Based Book Recommendation System — May 2025 - July 2025

- Built a Go-powered book recommendation system, featuring a data pipeline that gathers book details, stores them in MongoDB, creates Ollama API vector embeddings, and loads them into Qdrant for semantic search.
- Built a Gin web application in Go with REST API endpoints for health checks, data and author retrieval, and various recommendation functionalities (by book, author, or genre).
- Containerized the full system, including Qdrant and the Go web application, using Docker Compose for streamlined deployment and isolated service management. github.com/nchandur/go-reads

Lorekeeper – RAG-based Q&A System — Aug 2024 - Dec 2024

- Developed an end-to-end RAG system using Llama3.2:1b, LangChain, and FAISS for real-time, interpretable Q&A on The Lord of the Rings literary corpora.
- Engineered a data pipeline involving intelligent PDF parsing, content-based text chunking, and optimal vector embedding selection to build the knowledge base.
- Implemented an interactive Streamlit frontend that enhances transparency by displaying retrieved source chunks alongside LLM responses, ensuring interpretability. github.com/nchandur/lorekeeper

Education

University of Maryland, College Park, MD, USA — Aug 2023 – May 2025

Master of Science in Data Science

Coursework: Natural Language Processing, Computer Vision, Big Data Systems, Algorithms for Data Science, Cloud Computing

Technical Skills

Programming Languages: Python, R, C/C++, Go, Node.js, Bash
Machine Learning & AI: Scikit-learn, Sentence Transformers, Large Language Models, RAG Architectures, Exploratory Data Analysis, Feature Engineering, Statistical Testing, Bayesian Inference, Time Series Analysis, Data Modeling
Data Management & Databases: PostgreSQL, MongoDB, Redis, ChromaDB, Neo4j, Data Cleaning, ETL/ELT
Big Data & Distributed Systems: Spark, Hadoop, Kafka, Airflow, RabbitMQ, Dask, Snowflake
Cloud Plaforms & MLOps: Microsoft Azure Databricks, Azure ML Studio, Amazon Web Services (AWS), Docker, Git, REST APIs