Nischal Chandur

858-241-6448 | chandur.nischal2@gmail.com
nchandur.github.io/portfolio | linkedin.com/in/nchandur | github.com/nchandur

EXPERIENCE

Data Science Graduate Intern, Ecolab – Naperville, IL, USA **06/2024 – 08/2024**

- Developed and deployed a multilingual RAG pipeline to assist global field agents with machine malfunction diagnosis and resolution, providing step-by-step guidance and product recommendations from historical logs and training manuals.
- Created an ML model validation framework and synthetic data generator for predictive maintenance systems, improving robustness and accurately predicting rare anomalous events by balancing datasets and reducing reliance on infrequent real-world failures.
- Presented project vision and impact to directors and executive leaders, fostering cross-functional alignment and informed decision-making across diverse stakeholders.

Machine Learning Engineer, Reworked.ai – Miami, FL, USA **04/2024 – 05/2024**

- 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.

Geospatial Data Scientist, Latlong – Bengaluru, KA, India **09/2022 – 06/2023**

- Led the team involved in developing data extraction and ML pipelines for location-based solutions, 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.
- Provided advanced spatial analytics and strategic insights to an automotive client, optimizing sales and marketing by analyzing geospatial and demographic data, identifying performance drivers, and presenting findings to the executives to facilitate decision making.

PROJECTS

GoReads! – Scalable Content- Based Book Recommendation System **05/2025 – 07/2024**

- Engineered an end-to-end book recommendation system in Go, leveraging its efficiency for ML applications. Developed a comprehensive data pipeline that ingested data from internet sources, stored data in MongoDB, generated vector embeddings via Ollama API, and ingested them into a Qdrant vector database 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 **08/2024 – 12/2024**

- 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, United States **08/2023 – 05/2025**

Master of Science in Data Science

GPA: 3.9

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

SKILLS

Python, R, C/C++, Go, Node.js, Bash, PyTorch, Scikit-learn, Keras, Sentence Transformers, Large Language Models, Prompt Engineering, RAG Architectures, ChromaDB, Pinecone, Weaviate, Neo4j, Exploratory Data Analysis, Feature Engineering, Statistical Testing, Time Series Analysis, Data Cleaning, Data Modeling, ETL/ELT, Spark, Hadoop, Kafka, Airflow, RabbitMQ, Dask, Snowflake, PostgreSQL, MongoDB, Redis, Microsoft Azure Databricks, Azure ML Studio, Amazon Web Services (AWS), Docker, Git, REST APIs, Bayesian Inference, Optical Character Recognition, Flask, Django, FastAPI