# Nischal Chandur

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

## SUMMARY

Machine Learning Engineer with 2+ years of experience designing and deploying production-grade ML systems, with specialization in LLMs, RAG architectures, and geospatial data science. Skilled in transforming cutting-edge NLP and ML research into real-world, trustworthy products. Proven ability to lead end-to-end pipelines, from model development and cloud deployment to cross-functional collaboration with domain experts and business stakeholders. Strong focus on interpretability, ethical AI, and secure, scalable systems.

## EXPERIENCE

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

- Built a RAG assistant using Azure AI Studio to enable store managers to query KPIs from structured uploads (e.g., Excel).
- Designed an LLM-based support tool for field agents using embeddings of manuals and incident logs for contextual troubleshooting.
- Presented solution demos and architectural decisions to executives; worked with engineers to refine requirements.
- Created secure NLP interfaces for natural language interaction over mixed-format data.

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

- Proposed and integrated geospatial features into their ML pipeline, leveraging resulting in a significantly enhanced targeting strategy.
- Developed a Bayesian ensemble model to predict solar panel adoption likelihood, contributing to smarter customer acquisition and outreach strategies.
- Communicated value propositions to leadership through strategic insights and data-driven recommendations.

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

- Led development of a multilingual OCR-NLP pipeline for structured data extraction from public government documents, enabling regional demographic analytics for clients.
- Engineered geospatial risk models for NBFCs to assess borrower reliability by fusing behavioral and location-derived features into credit scoring.
- Delivered spatial analytics to a major automotive client, correlating product performance and sales patterns with demographic and location data.
- Presented findings and deployment options directly to product heads and senior business stakeholders.

## PROJECTS

***Lorekeeper – RAG-based Q&A System, University of Maryland*** **08/2024 – 12/2024**

- Developed a production-style Retrieval-Augmented Generation (RAG) system using Llama 3.2:1b, LangChain, and FAISS, enabling interpretable, real-time question answering over book corpora.
- Designed chunking, embedding, and search modules with modularity and scalability in mind.
- Implemented a front-end with source traceability and explanation features to enhance trust and transparency in LLM outputs. github.com/nchandur/lorekeeper

***NBA Game Outcome Predictor & Analytics Dashboard, University of Maryland*** **08/2023 – 12/2024**

- Built an end-to-end ML pipeline for predicting NBA game outcomes using ensemble models trained on 40+ years of historical data.
- Led team efforts in model design, REST API integration, and dashboard development using Flask.
- Focused on explainability and usability by surfacing key player metrics, model confidence intervals, and interactive visualizations. github.com/nchandur/NBA-prediction-model

## 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

## SKILLS

**Programming Languages**: Python, R, C/C++, Go, JavaScript, Bash, HTML/CSS
**LLMs & NLP**: LangChain, Hugging Face, ChromaDB, FAISS, Sentence Transformers, OCR, Prompt Engineering
**ML & Probabilistic Modeling**: PyTorch, TensorFlow, Scikit-learn, XGBoost, Keras, Bayesian Inference, Ensemble Learning
**Deployment & MLOps**: Docker, Azure AI Studio, AWS, Streamlit, Flask, REST APIs, CI/CD, Git
**Big Data & Infrastructure**: Spark, Hadoop, PostgreSQL, MongoDB, Redis, Dask, Structured & Unstructured Pipelines