Nithin Sameer Yerramilli

/in/nithin-sameer-yerramilli | /Nithinsameer | ynsameer@gmail.com | Portfolio | +1 240 927 7584

Summary

Data professional combining machine learning expertise with business analytics acumen. Skilled in developing AI solutions, analyzing complex datasets, and communicating insights through visualization and stakeholder collaboration.

Education

Master of Science Data Analytics Engineering, George Mason University CGPA: 3.96/4.0 Jan 2023 - Dec 2024
Bachelor of Technology in Computer Science, Dayananda Sagar University CGPA: 3.50/4.0 Aug 2018 - Jun 2022

Skills

Languages and Tools: Python, R, SQL, Tableau, Git, Latex, Databricks, Next.js, Excel
Frameworks and Libraries: TensorFlow, PyTorch, Scikit-learn, Pandas, NumPy, Seaborn, Matplotlib, NLTK
Cloud Technologies: AWS

Experience

Data Analyst Intern Aug 2024 - Present
Graduate Studies Dept. (GMU)
- Implementing machine learning models to analyze student admission data, improving prediction accuracy by 23%.
- Creating interactive dashboards using Tableau (Salesforce) to visualize key metrics, increasing data accessibility and enabling data-driven student engagement strategies.
- Executed 5 interactive Salesforce dashboards, improving stakeholder data accessibility and enabling more informed decision-making through visualized key metrics and trends.

Data Scientist Aug 2024 - Dec 2024
Erasmus.AI
- Built ClimateGPT's core function calling architecture using Python, enabling the AI model to execute 60+ specialized climate analysis functions through a unified tool interface.
- Developed an intelligent query router and flexible tool calling system that allows ClimateGPT to analyze climate data from NOAA, NASA and World Bank databases through natural language queries.
- Led technical design decisions as Product Owner in an Agile team of 4, implementing modular data processing pipelines that power ClimateGPT's automated climate impact analysis capabilities.

Machine Learning Research Associate Aug 2023 - Aug 2024
Costello College of Business (GMU)
- Led advanced statistical analyses on India's largest household survey dataset (174,000+ households across 27 states) using R and Python, uncovering a 12% decrease in spending patterns post-app bans through CEM, difference-in-difference analysis, and regression with survey weighting.
- Designed and implemented machine learning models achieving 85% accuracy in forecasting consumer behavior trends. Optimized big data preprocessing workflows using PySpark, enhancing efficiency for downstream ML tasks.
- Validated results from Research papers and Journals. Communicated findings to stakeholders detailing KPI's.

Projects

AI-Powered Multilingual Chatbot - Headstarter Fellowship Jul 2024 - Aug 2024
(Generative AI, NextJS, Tailwind, RAG)
- Engineered a sophisticated ML-based chatbot using Next.js and executed Retrieval Augmented Generation (RAG), increasing query response accuracy by 22% across multiple languages.
- Implemented RAG architecture using Langchain and Pinecone, significantly expanding the chatbot's knowledge base and improving its ability to provide contextually relevant responses.
- Composed a custom LLM routing mechanism using Groq's Mixtral-8x7B and OpenAI's GPT-4, optimizing response times and expanding the knowledge base for contextually relevant responses.

Demographic Bias in Recidivism Prediction - College Project Jan 2023 - May 2023
(R, Chi-Square Analysis, Stepwise Selection)
- Developed a predictive model achieving 76.5% accuracy in identifying recidivism risk factors using R, employing Stepwise Selection and Chi-Square analysis.
- Identified key insights into the correlation between work history and recidivism using a Decision tree, shedding light on racial disparities in post-release supervision.
- Implemented Decision Tree analysis using R, identifying critical employment thresholds that differentiate recidivism probabilities, with employment status influencing recidivism predictions at a key threshold of 64.3% days employed.