

Full-Stack Data Scientist with 3+ years of experience building scalable ML and data pipelines in cloud environments with demonstrated success in predictive modeling, MLOps, deep learning, and Generative AI.

SKILLS

Languages	: Python, SQL, Java, Shell Scripting, JavaScript, R
Machine Learning	: PyTorch, TensorFlow, HuggingFace Transformers, Spacy, XGBoost, OpenCV, Sklearn, MLflow
Cloud	: Azure, Google Cloud, AWS, Docker, Kubernetes, Nginx, Redis, Apache Spark, Kafka
Generative AI	: LangChain, LlamaIndex, QLoRA, Faiss, Neo4j, RAG, MCP, Fine-tuning, Prompt Engineering

EXPERIENCE

Senior Data Scientist	April 2022 – Aug. 2023
App Orchid	Hyderabad, India
<ul style="list-style-type: none">Led a team of 3 data scientists to build automated multi-modal pipeline to parse legal documents, reducing manual processing by 100,000+ hours annuallyFine-tuned OpenAI’s GPT-3 davinci model for document classification, achieving 93% accuracyStreamlined data workflows using ETL tools like PySpark and Databricks, reducing processing times by 30%Applied Detectron2 and PyTorch for layout-aware image segmentation enhancing downstream OCR in the pipelineBuilt and deployed a NER engine using RoBERTa served via Flask API, achieving 94.5% precisionCreated CI/CD pipelines for Role Based Access Control, model versioning, and audit logging using MLflow, Docker & Azure DevOps to ensure secure, traceable model deployments	
Data Scientist I	July 2020 – April 2022
HDFC Life	Bengaluru, India
<ul style="list-style-type: none">Developed a risk classification pipeline for 50,000+ daily calls using an ensemble of BERT and PyTorch DNNs, achieving 98% accuracy; integrated with AWS S3 and AWS TranscribeUsed call telemetry and metadata to train deep autoencoder models for real-time fraud detectionTrained LSTM models in PyTorch for forecasting sales surges, improving business planning accuracyEnhanced face verification system accuracy to 98% by implementing ArcFace embeddings and Visual TransformersBuilt hyper-personalized recommendation system using XGBoost processing behavior data from 1M+ usersLeveraged SageMaker Pipelines and CloudWatch for monitoring and retraining of in-production models	

RESEARCH & PROJECTS

Graduate Research Assistant	August 2024 – Present
Prof. Pratik Sharma, Indiana University	Bloomington, Indiana
<ul style="list-style-type: none">Optimizing serverless latency of large language model (LLM) inference by GPU scheduling & KV-Cache reuseDeveloped a GPU queue forecasting model with Structured State Spaces (S4) for preemptive request schedulingDesigning a multi-agent LLM simulation framework using Model Context Protocol (MCP) to emulate GPU tasksSubmitted Publication: MQFQ-Sticky — Fair Queueing for Serverless GPU Functions, HPDC 2025	
Generative AI Research Intern	May 2024 – August 2024
Prof. Nalini Ratha, University at Buffalo	Buffalo, NY
<ul style="list-style-type: none">Designed an active RAG pipeline using LangChain & FAISS for dense embeddings and Selenium for web scrapingIntegrated Cohere for query embedding & reranking, leveraged GPT-4 for natural language answer generation	
Image Segmentation for Flood Inundation Mapping	Course Project: Computer Vision
<ul style="list-style-type: none">Implemented U-Net, LinkNet, Prithvi & SegFormer in PyTorch for flood pixel detection in satellite imagery, achieving 88% mIoU with optimal models on 7-channel (pre/post-flood RGB + elevation) dataset. Report	
Knowledge Graph RAG pipeline using Langchain and Neo4j	Course Project: Advanced NLP
<ul style="list-style-type: none">Extracted entity relationships from documents to populate Neo4j graph db using LangChain and LlamaIndexAnswered user queries by running Cypher queries over the knowledge graph and text generation by GPT-3.5	
Online Coding Platform	Personal Project
<ul style="list-style-type: none">Built a containerized coding platform using Azure Container Apps, Nginx, Serverless Functions, and Replit Agent; automated question generation with prompt engineering. Developing a mock interview chatbot	

EDUCATION

M.S. in Computer Science, Indiana University, Bloomington, IN — GPA: 3.7/4.0	Aug 2023 – Present
<ul style="list-style-type: none">Teaching Assistant: CSCI 555: Machine Learning, ENGR E 533: Deep Learning Systems	
B.Tech in Aerospace, Minor in Computer Science, Indian Institute of Technology Bombay	July 2016 – July 2020