Nikhil Nageshwar Inturi

+1 (945)216-4026 | inturinikhilnageshwar@gmail.com | https://www.linkedin.com/in/nikhilinturi/

SUMMARY:

Generative AI and Machine Learning Engineer with 7+ years of experience in data-driven applications and scalable solutions. Proficient in Python, R, PyTorch, ETL pipelines, cloud(AWS, Azure), containerization (Docker, Kubernetes), and CI/CD workflows, with a strong track record in optimizing workflows and collaborating across teams. Passionate about solving complex problems with ML and AI.

SKILLS:

Programming: Python(NumPy, Pandas, SciPy), R, SQL, Shell Scripting, Java, Workflow languages (Cromwell and NextFlow), DQL
Machine Learning and AI: Deep Learning(Keras and PyTorch), NLP(RNN, LSTM, Transformers), Generative AI(LlamaIndex, LangChain)
Containerization and Deployment CI/CD: Git, Docker, Podman, Kubernetes, Jenkins, GitLab CI, and Docker Swarm
Databases and Cloud Tools: Redis, PostgreSQL, Snowflake, Redshift, Azure Cosmos DB, AWS DynamoDB, MongoDB, **AWS** and **Azure**
Certifications: Databricks GenAI Fundamentals, Graduate Certificate in Applied Machine Learning, Post Graduate Program in AI and Machine Learning, Amazon Cloud Computing Practitioner, Microsoft Certified: Azure AI/Data Fundamentals

EXPERIENCE:

Senior Data Scientist, The University of Texas at Dallas   Feb 2023 – Present
-   Led 3 cross-functional teams (15 scientists & 7 students) in image and sequencing analytics, and standardizing NGS pipelines.
-   Developed a GPT-powered RAG chatbot (LangChain) that unifies scRNA-seq, ATAC-seq, Bulk RNA-seq, Visium / Xenium spatial data related publications-cutting knowledge-retrieval time 70-90%.
-   Built an neuron-detection pipeline using image segmentation (Detectron2 | YOLOv11 | FastAI | SAM) that raised F1-score 0.78 → 0.89 (+15%) while slashing processing time 95%.
-   Deployed "Containerized Nextflow workflows" for Bulk RNA-seq, scRNA-seq, ATAC-seq, and spatial-omics; scaled to 50+ samples sets on Kubernetes and accelerated biomarker discovery 60%.
-   Developed ensemble models for Vagus Nerve Simulation data (XGBoost,LightGBM) achieving 93% accuracy, in predicting rat jaw-size.
-   Spearheaded global collaborations (McGill, UPenn, WashU and Univ. of Queensland) standardizing NGS and Spatial Transcriptomic pipelines and cutting turn-around from weeks-to-days.

Data Scientist, Aganitha Cognitive Solutions   Jun 2022 – Nov 2022
-   Developed clustering models(K-means, GMM) to identify AAV capsid sequence identification that cross the blood-brain-barrier, resulting in a 99.96% reduction in required in-vivo experiments.
-   Built an interactive mid-/downstream analytics suite (Python, Cromwell, Bash, R, REST APIs) that turns raw sequencing runs into dashboards in minutes, cutting analysis cycle time > 90% for five cross-functional teams.
-   Integrated structural & sequencing data into Snowflake, providing real-time tracking of capsid IDs and TM-scores across capsid sequences.
-   Fine-tuned SpliceAI-5 to detect novel human splice junctions, reducing false-positives ~30% accelerating biomarker discovery.
-   Developed AutoBLAST search algorithms that deliver 2× faster genome queries than standard BLAST, used in all internal genomic screens.

Data Scientist, Infosys Ltd   Sep 2018 – Jun 2022
-   Built 40+ universal DB connectors in Python for SQL (MySQL, Snowflake, Redshift) and NoSQL stores (Cosmos DB, DynamoDB, MongoDB, Cassandra), unlocking one-click data access in the Infosys Data Science & ML Platform (IDSMLP).
-   Automated ingestion & prep pipelines with Pandas, PySpark, and Airflow, cutting data-wrangling time 30% while maintaining 100% unit-test coverage and full lineage.
-   Integrated AutoML workflows-LightGBM, CatBoost, H2O, Keras-expanding IDSMLP's algorithm catalog 3× and halving model-development effort for client teams.
-   Containerized IDSMLP with Docker, CI/CD, and Nginx load balancing, reducing deployment cycles 75% & pushing platform uptime to 99.9%.
-   Built REST-based SAP CFIN reporting tool (Python+JavaScript) that ended manual data pulls & slashed generation time 40%.

EDUCATION:

The University of Texas at Dallas, *Master's in Business Analytics & Artificial Intelligence* | GPA: 3.9   Dec 2024
Purdue Global – Simplilearn, *Post Graduate Program in AI and Machine Learning* | GPA: 10   Nov 2021
Ramaiah Institute of Technology, *Bachelor of Engineering in Mechanical Engineering* | GPA: 9.8   June 2018