Adith Devakonda

GitHub: https://github.com/ASD02 | adithdevakonda@gmail.com | (404) 820-6011 | Suwanee, GA

PROFESSIONAL SUMMARY

AI/Data Engineer with ML, data science, and cloud experience. Proven track record at Codoxo developing ETL workflows, deploying ML models for healthcare fraud detection, and creating AI solutions. Winner of GenAI Innovation Challenge.

EDUCATION

Georgia Institute of Technology • M.S. in Computer Science (AI)
Georgia Institute of Technology • B.S. in Computer Science (Networks & ML) | May 2025

TECHNICAL SKILLS

Languages: Python, Java, SQL, JavaScript, C++
Technologies: Apache Spark, AWS (SageMaker, EMR, S3, EC2, Lambda), TensorFlow, PyTorch, LangChain, PostgreSQL, React.js, Node.js, Airflow, Django

HONORS/AWARDS

Innovation Challenge: Generative AI
Codoxo - 1st Place

Data Science @ GT Leadership Recognition
Director of Content

WORK EXPERIENCE

AI/Data Engineer at Codoxo | AI Solutions for Healthcare | Duluth, GA **May 2024 - Current**

- Developed LLM-based product using Bedrock API and Sonnet-4 for healthcare data validation, reducing client onboarding time by 20% and ensuring data integrity across multiple healthcare datasets
- Built ensemble learning model for document clone detection using 7-feature weighted classifier (MinHash, TF-IDF, SimHash) with threshold-based classification for large-scale record ingestion and fraud prevention
- Designed AWS Lambda functions as ETL middleware for claim audit ingestion, implementing REST APIs for seamless data flow and processing across multiple client environments
- Implemented PySpark processes for loading PostgreSQL tables from Parquet files in AWS S3, integrating into Airflow workflows using big data technologies for scalable data processing
- Researched and deployed multiple LLMs as endpoints on AWS SageMaker, demonstrating expertise in cloud-based model deployment and scaling for production environments
- Successfully onboarded multiple clients and created customized ETL workflows, ensuring data reliability and accuracy across diverse healthcare systems and compliance requirements

BI Analyst Intern at Billtrust | Remote **May 2023 - August 2023**

- Collaborated with engineers to develop LangChain tools, holding weekly meetings to ensure alignment and progress on research deliverables
- Developed application connecting Large Language Models to Snowflake Database, enabling users to generate SQL queries using natural language processing and improving query efficiency
- Utilized Python, SQL, Snowflake, SQLAlchemy, and LangChain for database integration and natural language query processing, contributing to enhanced user experience and system performance

Director of Content at Data Science @ GT | Atlanta, GA **January 2023 - January 2024**

- Managed 100+ member club as part of executive board, coordinating educational initiatives and member engagement to foster data science learning community
- Assigned tasks and mentorship teams to direct reports, tailoring assignments to individual skills to maximize productivity and professional development outcomes
- Organized and executed bootcamp meetings for 60+ students, including scheduling and logistics management for comprehensive educational experiences

Mentor and Workshop Leader at Data Science @ GT | Atlanta, GA **September 2022 - Present**

- Facilitate discussions between students to provide optimal Data Science learning experiences and career guidance
- Conduct workshops on Data Analysis and AI topics, employing various techniques to maximize attendance and engagement
- Plan and facilitate Data Preprocessing and ML workshops, utilizing Python, Pandas, NumPy, and PyTorch to deliver hands-on technical training

PROJECT EXPERIENCE

Wealth Inequalities in Motor Vehicle Collision Response Time

- Analyzed correlation between community wealth and emergency response times using PCA, Gaussian Mixture Models, Decision Trees, and CNNs to identify socioeconomic disparities in emergency services
- Utilized Python, Pandas, NumPy, Scikit-Learn, Streamlit, and GitHub for comprehensive data processing, statistical analysis, and interactive visualization development

ARC-AGI Agent

- Developed advanced neural network architecture for solving Abstraction and Reasoning Corpus for Artificial General Intelligence (ARC-AGI) problems, achieving 60%+ training accuracy through feature engineering, data augmentation, and ensemble learning techniques
- Implemented enhanced Multi-Layer Perceptron (MLP) with LayerNorm, dropout regularization, batch training, gradient clipping, and adaptive learning rate scheduling using PyTorch, along with statistical feature extraction and hyperparameter optimization for robust pattern recognition across diverse problem types