Nijgururaj Ashtagi

+1(872)258-2580 | github.com/nijgururajofficial | linkedin.com/in/nijgururaj15/ | nijgururaja12@gmail.com

Skills

Languages: Python, SQL
Data Engineering: Spark, Delta Live Tables, Azure Data Factory, Azure Synapse, BigQuery, Autoloader
Machine Learning/AI: TensorFlow, Scikit-learn, OpenCV, Google Gemini, LangChain, LLM, SLM, Agents
Cloud Platforms: Azure, AWS, Google Cloud
Tools: Docker, Git, GitHub, Power BI, Looker Studio, Databricks
Certifications: Google Cloud Big Data and ML Fundamentals, Modernizing Data Lakes with GCP

Experience

AI Engineer January 2025 – Present
S3CURA *Chicago, IL*

- Architected and deployed a machine learning-powered Video Forensics System for VigilAI using **AWS Lambda, API Gateway, S3, and DynamoDB**, enabling serverless, scalable, and low-latency **video analytics** and **AI-driven surveillance.**
- Integrated Google Gemini Vision API with **knowledge graph embeddings** to extract rich semantic insights from surveillance video, improving natural language query accuracy by **40%**
- Designed and developed over **20 + production-grade REST APIs** including endpoints for user authentication, chat management (create, rename, delete), **video uploads, media processing**, and **system status monitoring**
- Implemented a WebSocket-based real-time communication system for instant event streaming and user interaction within the VigilAI interface.
- Utilized **AWS DynamoDB** for low-latency storage of user metadata, chat history, and inference logs; optimized data schema and access patterns to reduce read/write costs by 30%.
- Engineered asynchronous ML inference pipelines using **asyncio and aioboto3**, resulting in a **2x improvement** in response time for video feature extraction and GenAI inference.
- Containerized inference tools using **Docker** and deployed **MLOps** workflows via **Lambda Layers** and **ECR**, ensuring modular and scalable deployment across environments

Projects

Demand Forecasting and ETL for an Online Marketplace | *SQL, Google Cloud, BigQuery, Pandas, Matplotlib, Looker Studio*

- Applied time-series modeling and statistical analysis (Python, Pandas) to forecast ride demand, a key business problem requiring optimization of resource allocation.
- Engineered a scalable ETL pipeline on Google Cloud (SQL, BigQuery) to process and model transactional data, enabling real-time analytics for a fast-paced environment.
- Designed and implemented cost-optimization techniques for analytical queries, including partitioning and clustering, to ensure efficient data processing.
- Developed interactive dashboards (Looker Studio) to deliver actionable insights on demand patterns to cross-functional stakeholders.

Deep Learning for Automated Image Captioning (Published Research) | *NLP, CNN, RNN, LSTM, TensorFlow, Python*

- Developed and trained a sophisticated deep learning model (Python, TensorFlow) combining a CNN and LSTM to perform a complex sequence-to-sequence task.
- Rigorously evaluated model performance using BLEU scores and utilized attention heatmaps to interpret results, demonstrating an understanding of systems with complex trade-offs.
- Applied advanced NLP techniques for text preprocessing and embedding, a skill applicable to understanding user-generated content on media platforms.
- Published findings in the peer-reviewed journal *Springer – Data Science and Applications*, showcasing self-driven and impact-oriented research.

Real-Time ETL and Data Modeling for Scalable Analytics | *Python, Spark Streaming, Azure Databricks, Delta Live Tables, SQL*

- Built a robust, real-time ETL pipeline using Python, Spark Streaming, and Azure Databricks to handle high-velocity data for immediate analysis.
- Designed a multi-layered star schema (Bronze/Silver/Gold) using SQL and Delta Live Tables to optimize data structures for downstream analytics and modeling.
- Implemented automated data quality checks and schema evolution to ensure a reliable and scalable data foundation for

analytical teams.

- Orchestrated CI/CD workflows for the data pipeline, demonstrating the ability to build and maintain production-grade systems.

Education

Master of Science in Computer Science *2024 - Present*
Illinois Institute of Technology, Chicago, IL
Coursework - *Machine Learning, Advanced Database Organization, Data Preparation and Analysis*

Bachelor of Engineering in Computer Engineering *2020 - 2024*
Savitribai Phule Pune University, Pune, India
Coursework – *Big Data, Deep Learning, DBMS, Design and Analysis of Algorithms*