

Samuel Davis

Data Scientist



+91 8652289114



samuel.davis.aiml@gmail.com



github.com/sam99dave

SKILLS

Language - Python, SQL

Deep Learning Frameworks - PyTorch, Tensorflow, LangChain

Libraries - OpenCV, FAISS, Opensearch, ChromaDB, Transformers, Distributed Training, Accelerate, Lightning, Ray, PEFT, Pydantic, MLFlow

API - Flask, FastAPI

Web App - Streamlit

Cloud Vendors - AWS

Serverless Applications - AWS Lambda

Source Code Management - GitHub, GitLab, CodeCommit

Other Skills - Computer Vision, NLP, Classical ML, MLOPs, LLMs

ACHIEVEMENTS

Contributed in setting up sports capability DEMO for AWS Re-Invent which attracted many clients.

Developed critical solutions that fast tracked production deal

CERTIFICATIONS

- Nvidia Fundamentals of Accelerated Computing with CUDA
- AWS Certified Cloud Practitioner
- MLExpert Certificate Of Completion
- Deep Learning Specialization
- Machine Learning Course (Stanford)

ABOUT ME

Experienced in designing and building Machine Learning and Deep Learning solutions, with a passion for developing AI solutions that solve real-world problems.

EXPERIENCE

Machine Learning Engineer

Quantiphi | Mumbai

QDox Product Development & POCs

Aug 2022 - Present

- Spearheaded the execution of 5 Proof of Concepts (POCs), showcasing exceptional problem-solving skills and technical acumen. Remarkably, 2 of these POCs converted successfully for full-scale production systems.
- Designed and implemented cutting-edge Language Model (LLM) solutions for chatbot, document, and information retrieval applications using tools such as Vector Database, RAG and LangChain. Prioritized innovation and efficiency in every aspect of project development.
- Successfully executed an internal POC displaying capabilities for generating SQL queries from Language Models such as Flan-T5 using LangChain and AWS RDS.
- Elevated the platform's capabilities by introducing game-changing features like Document Comparison and Document Indexing, enhancing overall user satisfaction and usability.
- Spearheaded the exploration and fine-tuning of cutting-edge Question Answering models, including DeBERTav3 and Donut.
- Led a major overhaul of the OCR codebase, preparing it for production. Seamlessly integrated the Textract APIs, significantly enhancing document processing capabilities. Diversified OCR options by incorporating PyTesseract and EasyOCR, broadening the codebase's utility.

Sports Analytics - CoachFirst

Mar 2022 - Jul 2022

- Played a key role in the development of performance tracking system for coaches.
- Optimized the Jersey Number Recognition module by implementing batching, and parallel image loading using multiprocessing.
- Engineered Action Recognition using a Spatial Temporal Model (ResNet-LSTM) combined with DenseNet, achieving faster training and lower latency while maintaining high accuracy.

Sports Media - SportCast

Nov 2021 - Mar 2022

- Spearheaded the creation of a cutting-edge player archival system for Bundesliga, revolutionizing the way football highlights are managed and analyzed.
- Implemented a highly efficient Object Detection system using YOLOv5, optimizing image processing through batching and parallel image loading, significantly improving system performance by reducing problems of trackID destabilization, breaking, and switching.
- Implemented Cosine Similarity with FAISS to achieve a 3x reduction in latency, enabling rapid player identification through accelerated Face Recognition.
- Orchestrated seamless end-to-end ML pipeline integration, ensuring smooth data flow and processing from frame extraction to player curation.

Machine Learning Engineer Intern

Quantiphi | Mumbai

R&D

Jul 2021 - Nov 2021

- Created real-world scenarios utilizing Unreal Engine.
- Conducted training and assessment of Object Detection, Instance Segmentation, and Object Tracking Models using synthetic data.
- Demonstrated the potential of training with synthetic data to the R&D team.

EDUCATION

Bachelor of Engineering - Computer Engineer

July 2017 - June 2021

GPA: 8.97