

Siva Ganesh

AI/ML Engineer

siva.ganesh@email.com

+91 76543 21098

Mumbai, Maharashtra

SUMMARY

Innovative AI/ML Engineer with a deep focus on Generative AI, Large Language Models (LLMs), and Computer Vision. 4 years of experience building, fine-tuning, and deploying state-of-the-art machine learning models. Expert in Python, PyTorch, and cloud-based ML pipelines.

EXPERIENCE

Machine Learning Engineer

2023 - Present

DataMinds AI

- Led the development of a domain-specific customer support chatbot by fine-tuning Llama-2 7B models on internal datasets, reducing support ticket volume by 30%.
- Designed and implemented a RAG (Retrieval-Augmented Generation) pipeline using Pinecone vector database and LangChain for accurate document retrieval.
- Optimized model inference latency by quantization and deploying models using ONNX Runtime and Triton Inference Server.

Data Scientist

2020 - 2023

Visionary Tech

- Developed high-accuracy object detection models for retail inventory management using YOLO v5 and Faster R-CNN architectures.
- Deployed computer vision models to edge devices (NVIDIA Jetson) using TensorFlow Lite, ensuring real-time performance.
- Conducted extensive data analysis and preprocessing to improve dataset quality and model generalization.

PROJECTS

DocQuery AI

An intelligent document search engine that allows users to query PDFs using natural language.

VisionGuard

Real-time safety monitoring system for manufacturing plants using CCTV feeds and AI analysis.

SKILLS

Python | PyTorch | TensorFlow | Generative AI | LLMs (Llama, GPT) | LangChain | Computer Vision (YOLO) | NLP | FastAPI | MLOps

EDUCATION

M.Tech in Artificial Intelligence

2020

IIIT Hyderabad