SUMMARY:

As an experienced AI/ML Engineer, I specialize in designing, training, and scaling deep-learning solutions for NLP and computer vision applications using Python, PyTorch, and TensorFlow. Proven ability to build end-to-end pipelines, optimize model performance, and integrate large-language-model services into production applications.

HIGHLIGHTS:

- Designed and developed an intelligent conversational system combining LLM and Stable Diffusion for multilingual dialog and local image generation on consumer GPUs

- Implemented token-aware memory management for contextual understanding and intelligent conversation trimming

- Architected CORS-enabled FastAPI backend with asynchronous endpoints and VRAM-optimized inference

- Developed responsive Streamlit frontend featuring real-time status monitoring and automatic language detection

- Deployed ML models for payment-risk scoring and anomaly detection in healthcare-fintech datasets, adhering to HIPAA-compliant data-governance standards

- Collaborated in Agile sprints, performing code reviews and mentoring junior team members

KEYWORDS:

Python, PyTorch, TensorFlow, LLM, Stable Diffusion, FastAPI, Docker, AWS, Git, MLOps, CI/CD, Agile, HIPAA-compliant data-governance standards.