

Andrew V. Franco

702 756 1720 | And.Fra.97@gmail.com | [LinkedIn](#) | [GitHub](#)

PROFESSIONAL SUMMARY

Machine Learning Engineer leveraging 8+ years of clinical data domain experience to safely build robust, full-lifecycle AI solutions. Specialized in generative AI (RAG, LLMs), medical image segmentation, and time-series classification. Adept at driving models from architecture and training to containerized cloud and on-device edge deployment.

TECHNICAL SKILLS

Generative AI & NLP | LLMs, RAG, Pinecone, LangGraph, LangChain, BioBERT, DeBERTa, Prompt Engineering
Deep Learning & ML | 3D/1D CNNs, U-Net Segmentation, Time-Series Classification, Optimization, Quantization
Languages & Libraries | Python, PyTorch, Keras, TensorFlow Lite, MONAI, NumPy, Scikit-Learn, Pandas, C, C++
MLOps & Deployment | MLflow, FastAPI, Docker, Hugging Face Spaces, Mixed Precision Training, Experiment Tracking
Edge AI & Hardware | STM32 (Cortex-M), FreeRTOS, DMA, UART, ADC, SPI, STM32-X-Cube-AI, Edge Inference

PROJECTS

SentinelMD - Clinical RAG System | LangGraph, Pinecone, BioBERT, DeBERTa, OpenFDA, FHIR [Live Demo](#)

- **Agentic Pipeline:** Built a LangGraph RAG pipeline ingesting clinical queries and FHIR R4 resources, embedding PubMed abstracts via BioBERT into namespaced Pinecone indexes with OpenFDA drug label enrichment.
- **Hallucination Detection:** DeBERTa-v3 NLI scoring for claim-level entailment verification against PubMed literature and FDA drug labels achieving 0.774 RAGAS faithfulness, with per-response reliability scoring, inline evidence citations, and FHIR R4 input support via HAPI server integration.
- **Production Deployment:** FastAPI backend with SSE streaming on Hugging Face Spaces, React frontend on Vercel, MLflow monitoring tracking confidence scores and claim distributions across production queries.

NeuralSeg - 3D Brain Tumor Segmentation | PyTorch, MONAI, FastAPI, Docker [Live Demo](#)

- **Architecture & Training:** Engineered a 3D U-Net segmentation model on BraTS 2023 GLI dataset achieving Dice scores of 0.930 (WT), 0.910 (TC), and 0.848 (ET) with Hausdorff95 distances of 5.94mm, 3.76mm, and 3.23mm. Built a distributed training pipeline utilizing mixed precision and MLflow experiment tracking.
- **MLOps Deployment:** Developed a FastAPI inference service with automated N4 bias correction and sliding window inference; containerized via Docker and deployed to Hugging Face Spaces with live 3D visualization.

Real-Time ECG Monitor With Arrhythmia Detection | Python, C, C++, Keras, Tflite, STM32, FreeRTOS, DMA

- **AI/ML Architecture:** Engineered a 1D Convolutional Neural Network (CNN) for classifying Normal Sinus Rhythm, A-Fib, and First Degree Heart Block based on the PTB-XL dataset. Achieved 96% accuracy with 0.92 macro recall and macro-F1 of 0.90, validated against stratified clinical ground-truth labels.
- **Edge Deployment:** Applied post-training quantization to reduce model size by ~90% (23KB .tflite), achieving measured ~380ms inference on Cortex-M4 hardware, while maintaining complete pre-quantization accuracy.
- **Integration Strategy:** Authored a 37-page proposal detailing a 7-phase integration plan, \$74k-\$124k/unit annual ROI, SaMD-aligned HIPAA-compliant deployment, and HW-SW handshakes for hospital adoption

WORK EXPERIENCE

Cardiac Telemetry Specialist (Monitor Technician) / ICU Unit Coordinator **June 2015 – October 2023**
Fundamental LTC / Universal Health Services Las Vegas, NV

- **Clinical Data Workflow:** Managed ICU unit operations ensuring 100% uptime for 80+ networked telemetry nodes in a 24/7 ICU, executing critical care protocols and maintaining Class II medical device documentation.
- **HITL & Signal Integrity:** Served as the "human-in-the-loop" (HITL) validation layer for automated detection algorithms, interpreting raw ECG waveforms to mitigate alarm fatigue for nursing staff.

EDUCATION

Bachelor of Science, Computer Science | Western Governors University (ABET) **March 2026**
Awards: Excellence Award in Systems Thinking and Applications