Venkatesh Shanmugam
+1 (703) 216-2540
svenkatesh.js@gmail.com
LinkedIn Profile

Hiring Team at the Hiring Company,

What if you could deploy **multi-modal AI models** on edge devices that not only predict outcomes with **78% ROC-AUC** but also slash deployment times by 94%? As a Machine Learning Engineer with a passion for optimizing **video and audio processing pipelines**, I've done exactly that—and I'm eager to bring this expertise to your team at Roku TV, where innovation in real-time, edge-compatible AI is paramount.

At ScriptChain Health, I architected a **multi-modal CNN/RNN system** for hospital readmission prediction, achieving **78% ROC-AUC** by refining **TensorFlow** pipelines for real-time clinical use. By leveraging **DeepSpeed** and multi-GPU clusters, I accelerated distributed training by 67%, ensuring scalability for high-volume data. My work on **CI/CD workflows** reduced deployment time from 4 hours to 15 minutes while maintaining sub-200ms latency—critical for edge devices like Jetson Nano, which aligns with Roku TV's focus on efficient, high-performance hardware.

Beyond healthcare, I engineered a **Graph-RAG** system for building codes, boosting query accuracy by 90% through **LangGraph** and conditional routing, all while optimizing for edge deployment. My background in **PyTorch** and **parameter-efficient fine-tuning (PEFT)** further underscores my ability to balance model performance with resource constraints—a must for delivering seamless, real-time video and audio experiences.

I'm excited by Roku TV's commitment to pushing the boundaries of edge AI, and I'm confident my blend of research rigor and production-tested skills can drive similar breakthroughs for your team. My full professional profile, including detailed project contributions, is available on LinkedIn and GitHub. I'd welcome the chance to discuss how my expertise aligns with your vision—let's connect.

Sincerely,

Venkatesh Shanmugam
+1 (703) 216-2540
svenkatesh.js@gmail.com
GitHub Portfolio
Personal Portfolio