RAJEEV NANDAN DAMARLA

GITAM, Visakhapatnam – 530045
rajeevnandan382@gmail.com +91 9481509488 Portfolio
LinkedIn GitHub

B.Tech CSE candidate (GPA: 9.24/10) specializing in Generative AI, High-Performance ML serving, and Elite Full-Stack Product Engineering. Developed PyTorch diffusion models and optimized FAISS RAG databases at IIT Kanpur. Expert in building premium, responsive user experiences (Next.js, Framer Motion, WebSockets) to serve low-latency AI models. Seeking an AI / Full-Stack engineering internship to build state-of-the-art intelligent products.

EDUCATION

GITAM University – School of Technology *2023 – Present*
B.Tech in Computer Science & Engineering — CGPA: 9.24/10 (Latest Sem GPA: 9.73/10)

RELEVANT COURSEWORK

AI/ML & Core CS: Artificial Intelligence (Applications), Natural Language Processing, Compiler Design, Database Management Systems, Design & Analysis of Algorithms, Operating Systems, Computer Networks, Data Structures, Automata Theory
Mathematics & Security: Linear Algebra, Discrete Mathematics, Probability & Statistics, Graph Theory, Number Theory, Cryptography & Network Security

EXPERIENCE

Generative AI Intern — Certificate *June 2025 – August 2025*
Indian Institute of Technology, Kanpur
- Developed diffusion model pipelines in PyTorch, improving image generation quality (FID score reduction of 15%)
- Reduced inference latency by 22% through batching and mixed precision in PyTorch diffusion model pipelines
- Implemented LLM experimentation framework comparing 3 architectures, documenting results that improved baseline output quality by 18% on internal evaluation prompts
- Collaborated with 2 team members to present weekly progress, incorporating feedback to refine model evaluation protocols and prompt sets

PROJECTS

Image Generation using Diffusion Model — Check it out! *Jan 2024 – Mar 2024*
Research Project
- Developed text-to-image diffusion model generating 512x512px images from prompts, achieving CLIP score of 0.78 for alignment and creativity
- Processed and filtered 10K+ synthetic training samples, improving dataset quality and reducing training time by 30%
- Deployed interactive Gradio interface on Hugging Face Spaces, serving 500+ image generations with ¡2 s average latency

Insight GPT — Check it out! *Oct 2025 – Dec 2025*
Research Project
- Developed domain-specific LLM using Retrieval-Augmented Generation (RAG) with FAISS, processing 500+ PDFs (80K pages) for semantic search
- Achieved 0.82 top-1 retrieval accuracy on 200-query evaluation set, reducing answer latency by 35% through caching and optimized indexing
- Implemented PDF ingestion pipeline handling 50 documents/hour, deployed on Hugging Face Spaces with real-time query interface

Real-Time Object Detection — Check it out! *Jan 2026 – Mar 2026*
Research Project
- Implemented client-side real-time object detection using DETR-ResNet-50, processing webcam feed at 15 FPS with ¡100 ms latency
- Achieved mAP@0.5 of 0.68 on COCO val subset, displaying bounding boxes with confidence scores in-browser
- Optimized model loading and inference, reducing initialization time by 40% and memory footprint by 25% through quantization

TECHNICAL SKILLS

Programming Languages: Python, C, C++, Java, JavaScript, SQL, HTML/CSS
AI/ML & Data Science: TensorFlow, PyTorch, Scikit-learn, Hugging Face, Pandas, NumPy, Jupyter Notebook, Google Colab
Full-Stack & UI/UX Design: Next.js, React, Node.js, Vercel, Tailwind CSS, Framer Motion, GSAP, WebSockets, MySQL, SQLite, Figma
Cyber Security: Wireshark, Splunk, Kali Linux, Nessus, Metasploit, Snort, ELK Stack
Skills Summary: RAG, Diffusion Models, FAISS, PyTorch, Next.js, Framer Motion, UI/UX Design, MLOps

CERTIFICATIONS

Certifications available at: https://github.com/Rajeev91691/Certifications