Kirtan Amrutiya

kirtanamrutiya07@gmail.com | +91 9510941315 | Linkedin | Github | Portfolio

WORK EXPERIENCE

Physical Research Laboratory (PRL), Dept. of Space, Govt. of India — Ahmedabad, Gujarat
Research Intern — Sep 2025 - Jan 2026

- Designed and implemented Factorized Vision Transformer (FAViT) with global-local attention split (1:7 ratio), reducing quadratic complexity from $O(N^2)$ to $O(N \log N)$ while achieving 90.8% top-1 accuracy on CIFAR-10 vs. 80.3% baseline.
- Developed Window-Based ProbSparse attention mechanism, enabling 1.28x training speedup and 41% GPU memory reduction (9.2GB → 5.4GB) on NVIDIA T4 via contiguous window sampling over random points.
- Engineered Point-Based ProbSparse model with Informer-inspired M(max-mean) sparsity, selecting top-u VIP patches to hit peak 90.82% accuracy on upscaled CIFAR-10 (224x224), outperforming vanilla by 10+ points.
- Optimized hierarchical ViT stages (2-2-6-2 blocks) with Mixup/CutMix/EMA training, conducting ablation studies on factorization ratios and visualizing attention maps to validate locality preservation.

SLTL Group (Sahajanand Laser Technology Ltd.) — Gandhinagar, Gujarat
C++ and Java Intern — Jun 2022 - Sep 2022

- Integrated NLP into chatbots and deployed to website, reducing customer query response time by 30% for e-commerce platform
- Collaborated with 10-developer team to resolve 150+ bugs, cutting load times 40% across platforms
- Contributed to UI redesign, increasing user engagement 35% and daily active users from 5K to 6.75K in 3 months.
- Optimized C++/Java backend processes, achieving 50% average speed improvements for 20K+ monthly visitors

Projects

Factorized Vision Transformer (FaViT) with ProbSparse Attention

- Architected Factorized ViT with ProbSparse self-attention, reducing computational complexity from $O(N^2)$ to $O(N \log N)$ while achieving 90.8% top-1 accuracy on CIFAR-10 (224×224) compared to standard ViT baseline at 80.3%
- Developed Point-Based ProbSparse model with Informer-inspired sparsity achieving 90.82% peak accuracy, conducting ablation studies on Mixup/CutMix/EMA and attention map analysis

VocalNavigator AI - Voice-Controlled Desktop Assistant

- Architected modular desktop assistant with Porcupine hotword detection, Whisper transcription, and Voice Activity Detection, executing real-time system actions via PyAutoGUI across audio, STT, NLP, and actions modules
- Engineered hybrid NLP parser combining keyword-based recognition (open, close, click, scroll) with Google Gemini LLM fallback for complex queries, optimizing speed and flexibility.

Education

Swaminarayan University Ahmedabad, Gujarat
Diploma, Information Technology - CGPA - 8.86, CPI - 9.21

Ahmedabad institute of Technology Ahmedabad, Gujarat
Computer Science in Artificial Intelligence and Machine Learning CGPA -7.32, CPI - 7.27

SKILLS & INTERESTS

Skills: Vision Transformers, Attention Mechanisms, Deep Learning, Computer Vision, Natural Language Processing, Python, C++, Java, JavaScript, SQL, Docker, Git/GitHub,

Courses: CS50: Introduction to Programming with Python, Harvard University, Build Real-World AI Applications with Gemini and Imagen, Google Cloud Platform