Shakti Labhaniya

+91-9039302880 | shaktilabhaniya.work@gmail.com | shaktilabhaniya | S-Labs21

Education

Vellore Institute of Technology **Sep 2022 – Present**
Integrated MTech in Artificial Intelligence *CGPA: 9.04/10*

Technical Skills

Languages: Python, Java
Coursework: Data Structures & Algorithms, DBMS, Operating Systems, Computer Networks
AI/ML: Machine Learning, Deep learning, Generative AI, NLP, CNNs
Frameworks & Tools: PyTorch, TensorFlow, Keras, Scikit-learn, OpenCV, Hugging Face Transformers, Gradio
Databases, Cloud & Version Control: Milvus, Pinecone, AWS, Docker, Git, Github

Experience

Machine Learning Intern **Hyderabad**
Indian National Centre for Ocean Information Services ***Oct 2025***

- Developed a real-time multi-class fog detection system using Random Forest and a custom Pattern Recognition Filter, constructing a 10-class training dataset from 3+ satellite data.
- Operationalized a fully automated near-real-time pipeline integrating AVHRR and VIIRS data with SMTP/MIME alerting on an HPC server, achieving <4-minute latency and 100% per-file alert reliability.
- Achieved 95.97% overall classification accuracy, mIoU of 0.8703, False Alarm Rate (Fog) of 0.0178, and Critical Success Index (CSI) of 0.9814, indicating robustness in nighttime and daytime fog detection.

Projects

PageSense | **Ollama, JavaScript, Chrome Extension (MV3), Milvus (Zilliz)** **Nov 2025**

- Built a fully local, webpage-aware LLM Chrome extension using LLaMA-3.2-1B with Ollama (Manifest V3) and tested LLM context switching across 50+ webpages.
- Optimized real-time inference and streaming performance by using token-level generation controls, achieving sub-100 ms page ingestion latency and a 59.5% reduction in response time (6.33 ms → 2.56 ms).
- Enhanced retrieval performance by deploying a Milvus (Zilliz Cloud) vector database, achieving sub-50ms semantic search latency for semantic searches and ensuring context flow across multiple browsing sessions.

LatentHide | **Python, TensorFlow, Pytorch, StableDiffusion, DDPM, OpenCV** **June 2025**

- Architected a deep learning–based image steganography system using Stable Diffusion and DDPM, with a custom alpha-blending encoder–decoder for secure data hiding, trained to converge over 100 epochs.
- Standardized the data pipeline to 512×512 image resolution and created a custom Flickr8k dataset loader with batch-wise processing, enabling efficient evaluation across 8,000+ images.
- Obtained high image reconstruction quality with SSIM scores of 0.9644–0.9982 and low MSE.

Conversational AI Assistant | **Python, PyTorch, Hugging Face Transformers, Gradio** **May 2025**

- Designed a conversational AI system integrating STT, LLaMA-3.2-1B-Instruct (Hugging Face) backend, and ElevenLabs TTS, enabling real-time voice interaction with sub-2s response latency.
- Programmed session-based memory architecture using user IDs and vector-backed context storage, improving multi-turn contextual coherence by 40% across extended conversations.
- Integrated DistilBERT-based emotion detection to classify 6+ emotional states in real time, enhancing response personalization and increasing user engagement in conversational flow.

Research & Achievements

- Shakti Labhaniya, et al. "*A Survey of Machine Learning Techniques for Enhancing Scene Description in Assistive Systems*" Proceedings of the IEEE International Conference on Image Information Processing (ICIIP 2025), 2025. Available online: https://ieeexplore.ieee.org/document/11346259
- Contributed as second author to "Real-Time Fog Detection and Mapping Using Multi-Sensor AVHRR–VIIRS Satellite Data and Machine Learning Techniques", accepted and presented at AERIS-2026 National Conference.
- Semifinalist – NASSCOM Tech Developer Hackathon 2025