

HARSH MANKODIYA

• +1 (602)-517-6750 • hankodi@asu.edu • linkedin.com/in/harshmankodiya • github.com/hmankodiya

Education

Arizona State University

August 2023 - May 2025

Master of Science, Computer Science: GPA 3.70

Tempe, USA

Courses: NLP, Statistical Learning, Artificial Intelligence, Data Mining

Institute of Technology, Nirma University

August 2019 - May 2023

Bachelor of Technology, Computer Science Engineering

Ahmedabad, India

Courses: Machine Learning, Deep Learning, Data Structures, Linear Algebra, Calculus, Probability and Statistics

Professional Experience

Cellino Biotech

May 2024 - August 2024

Machine Learning Intern

Cambridge, USA

- Developed a proof of concept for a central embedding model for patch selection, anomaly detection, cell segmentation and cell classification.
- Fine-tuned **DinoV2** using **Vision Transformer** based heads for downstream segmentation tasks, achieving average F1-Score of **82%**. Utilized **Weights & Biases** for experiment tracking, artifact logging and hyperparameter sweep.
- Integrated **GCP** API calls with **PyTorch Dataset** utilities to streamline **Zarr** to **Tensor** conversion. Added a local caching mechanism improving throughput.
- Automated the retrieval of artifact metadata from a **PostgreSQL** database and integrated it into the pipeline for validating clustering efficacy and artifact detection workflows.
- Containerized the inference pipeline with **Docker**, enabling real-time data processing and easy integration with cloud-based services.

Lens Lab, Arizona State University

August 2023 - May 2024

Research Assistant

Tempe, USA

- Integrated eXplainable AI techniques with RL agents in **Gymnasium** environments to enhance decision explainability.
- Trained **Proximal Policy Optimization (PPO)** RL Agents using **StableBaselines3**. Intergrated **VAE**-based feature extraction for image stream processing.
- Leveraged pre-trained **CLIP** models for **zero-shot segmentation** and concept sampling for policy rollouts.
- Published findings at **NeurIPS 2024 SATA Workshop**, focusing on explainability in robotic decision-making.

Bosch

January 2023 - May 2023

Research Intern

Bangalore, India

- Formulated working principal for **GradCAM** and **GradCAM++** based **gray-box adversarial training** for **image segmentation models**.
- Utilized **PyTorch Lightning** to automate **data-processing, model training, evaluation, and inference** and implemented experiment tracking using **MLFlow**.
- Trained SegNet and U-Net segmentation models on **NVIDIA DGX A100** systems, achieving high relative **IoU** scores exceeding **85%** across multiple datasets.

Relevant Projects

Multilingual Sentiment Classification using LLM | Python, PyTorch, HuggingFace

Dec 2024

- Conducted **PEFT** on **Llama2-7B**, utilizing **Quantized Low-Rank Adaptation (Q-LoRA)** to achieve **4-bit quantization**, reducing trainable parameters by approximately **0.60%**.
- Fine-tuned **Llama2-7B** on just **2%** of a multilingual sentiment dataset spanning **12 languages**, categorized into three classes: positive, neutral, and negative.
- Witnessed **30%** increase in **test AUC** and a **20%** increase in **test accuracy**.
- Performed a comparative analysis by fine-tuning **GPT2** and **BERT**, highlighting their relative performance.

Reinforcement Learning for Algorithmic Trading | Python, stable-baselines3

Nov 2024

- Engineered an RL-driven trading system utilizing **DQN**, **PPO**, and **A2C** to optimize trading strategies.
- Designed a multi-reward function incorporating technical indicators, implemented with **MLP** and **LSTM**, boosting decision-making accuracy by **23%**.
- Developed an **LLM-based reward** function with real-time human feedback for adaptive learning, achieving an initial **21%** performance improvement.

Technical Skills

Languages - Python, C++, Shell, Docker, Git

ML Frameworks - PyTorch, HuggingFace, Jax, TensorFlow, scikit-learn, XGBoost, Stable-Baseline3, Gym, LangChain, LangGraph, Ollama

Python Libraries - NumPy, SciPy, Pandas, OpenCV, Pillow, Zarr, Dask, Seaborn, Matplotlib, Plotly, W&B, MLFlow, PySpark

ML Techniques - LLMs, RAG, Knowledge Distillation, Reinforcement Learning, CLIP, Image Captioning, Image Classification, Image Segmentation, VAE, GANs, Style Transfer, GradCAM, TCAV