LIKHITHA PARUCHURI

San Jose, California • +1 4085606262 •likhiparuchuri132@gmail.com • https://www.linkedin.com/in/likhitha-paruchuri/

EDUCATION

San Jose State University (SJSU) | *Master of Science in Artificial Intelligence* | ***GPA - 3.7*** **Aug 2023 - May 2025**
ICFAI Foundation for higher education |*Bachelor of Technology in Data Science and Artificial Intelligence* **Aug 2019- June 2023**

SKILLS

- **Technologies**: Computer Vision, Machine Learning, Neural Networks, Deep Learning, Transformers, GANs, LLMs, Generative AI, NLP, LSTM, RAG, Object Detection.
- **Programming Languages**: Python, JavaScript, SQL, Node.js.
- **Frameworks & Tools**: TensorFlow, PyTorch, Keras, Hugging Face, Scikit-learn, OpenCV, Pandas, Numpy, Seaborn, Apache Spark.
- **Cloud Platforms:** AWS (S3, Lambda, SageMaker, EC2, RDS, Redshift , Google Cloud Platform (GCP), Firebase.
- **MLOps/DevOps:** Docker, Kubernetes, CI/CD Pipelines, MLflow.
- **Version Control:** Git, GitHub/GitLab
- **Data Visualization:** Tableau, Matplotlib, Seaborn, Power BI.

PROFESSIONAL EXPERIENCE

Stanford Research Institute International(SRI) | *AI Research intern* **June 2024 - Dec 2024**

- Developed a **multimodal system** for content recognition in videos, targeting specific content not recognizable by standard LLMs.
- Developed pipelines for processing **audio, visual, and text data** using **ASR** for audio, **contrastive models** for caption generation, and advanced **filtering techniques** for extracting **semantic embeddings** from video.
- **Fine-tuned multiclass Hugging Face transformers**, achieving 90% classification accuracy.
- **Fine-tuned LLMs** and **Vision-Language Models (VLMs)** for **content understanding** and enable **advanced question answering**.
- Built and optimized AI-powered chatbots using **Retrieval-Augmented Generation (RAG**), **vector bases,** and **embeddings**.
- Optimized chatbot performance with a **reward model** using human feedback (**RHLF**), advanced prompt engineering (**chain-of-thought, prompt chaining**), and **reranking methods** for improved accuracy.
- Designed and implemented **MLOps** pipelines using **Docker** and **MLflow** for scalable deployment.

Intellect Design Arena | *Machine Learning intern* **Jan 2023 - April 2023**

- Built **predictive time-series** models for error forecasting in logs , improving operational efficiency.
- Developed and deployed **data preprocessing** tools to structure **large datasets** for effective training.
- Designed **RESTful APIs** and automated data visualization systems, enhancing performance monitoring.
- Applied **A/B testing** and **clustering techniques** to **validate models** and identify operational patterns.
- **Optimized model training pipelines** for increased inference efficiency, leading to a 15% productivity gain.
- Deployed models using **AWS SageMaker** and used **AWS RDS** for scalable data storage.

Bridgera | *Software developer intern* **April 2023 - July 2023**

- Enhanced system reliability through data preprocessing and dynamic field auto-population using Node.js.
- Collaborated with teams to integrate **data science insights** into software applications for **improved analytics**.
- Designed scalable production-ready machine learning pipelines, increasing overall system efficiency.
- Implemented **search and ranking mechanisms** to improve **data retrieval accuracy** within internal systems.

PROJECTS

Music recommendation System *(**Tech**: Scikit-Learn,Pandas, NumPy, Matplotlib, LightGBM, CatBoost.)*

- Built **scalable recommendation pipelines** with **content and collaborative filtering** and **gradient boosting** achieving 70% accuracy.
- Managed updates with **PySpark** for **data preprocessing** and **feature engineering**.
- **Fine-tuned Transformers** to analyze **contextual relationships** in song metadata, lyrics, and user interactions, improving recommendation relevance and **personalization**.

Car insurance automation using Object detection and Blockchain *(**Tech**:Python, Pytorch, Detectron2, React,web3.py,Firebase, AWS)*

- Automated insurance processes using **image segmentation models** with 96% accuracy and **blockchain** for secure claims.
- Designed scalable architecture for **fraud detection** and real-time data processing.
- Leveraged **AWS SageMaker** for training and deploying object detection models to ensure scalability and efficiency.

Automated Video Highlight Detection Using LLM *(**Tech**:Python, OpenCV, LLM, Computer Vision, NLP, BLIP)*

- Developed a **multi-modal LLM pipeline** for generating video highlights, leveraging **BLIP** for **vision-to-language transformation** and **GPT-2** for summarization. Integrated **scene understanding**, and motion detection for video analysis.

Medical Notes Generator and Question-Answering System *(Tech: Python, PyTorch, RAG Pipelines, FAISS, LLaMA-2,* Gemma-2, Streamlit*)*

- Built a system for automating SOAP note generation and interactive QA from medical transcripts.
- Utilized a **RAG pipeline** with **FAISS** for **retrieval** and models like **Gemma-2** and **LLaMA-2** for response generation.
- Designed prompts for **structured data extraction** and ensured fast, accurate **information retrieval**.
- Developed a web-based system for automating SOAP note generation and **interactive question answering** from medical transcripts.