Sumukha Manjunath

sumukha1996@gmail.com | Ontario, CA | linkedin.com/in/sumukha-manjunath/ | 9846833259 | sumukha21.github.io

Machine Learning Engineer with a Master's degree specializing in Machine Learning and over 5 years of experience in research, designing, prototyping, and deploying machine learning algorithms across healthcare, automotive, and agriculture sectors.

Experience

Raven Industries*, Machine Learning Engineer (Internship and Co-op)* May 2023 – April 2024

- Primary inventor of the filed patent "Attention Based Feature for object-oriented granular neighbor search".
- Improved YOLO-NAS model performance for obstacle detection by 18% through increased data diversity by training an image2image diffusion model and engineering a Generative AI-driven data augmentation pipeline.
- Reduced model unit testing and evaluation pipeline design time by 90% and streamlined CI/CD of models to production by developing a model agnostic pipeline design and orchestration library using DAG data structure and PyTorch.
- Accelerated data curation time by 80% by automating it through containerized deployment of Python-based image recommendation system using AWS S3, Lambda, DynamoDB, Mlflow and Facebook AI Similarity Search index (FAISS).
- Achieved a 99% reduction in memory consumption for image storage by training a Vision Transformer for image representation as 1D NumPy array using self-supervised distributed training on DGX cluster with self-distillation and masked-image-modelling.

Carl Zeiss, *Senior Data Scientist* (Consultant) May 2021 – June 2022

- Led the improvement of Multimodal CNN algorithms by 16% for Diabetic Retinopathy, Age related Macular Degeneration and Diabetic Macular Edema detection using medical images and patient metadata.
- Collaborated with retina specialists and external data science team for experiment design, data collection, algorithm development and clinical validation of machine learning models.
- Expedited patient diagnosis by 60% by leading the engineering of deep learning algorithms based diagnostic pipelines using Python, TensorFlow, Azure Services, FastAPI and Docker for Fundus and OCT systems.

Robert Bosch, ***Senior Software Engineer-Machine Learning*** August 2018 – May 2021

- Led a cross functional team in the end-to-end agile development of a blueberry farm's real-time crop count regression and localization tool, achieving 50% accuracy enhancement in harvest forecasting through a custom Unet and LSTM models.
- Improved obstacle and free space detection by 12% by integrating a Fully Convolutional Network (FCN) trained using TensorFlow to an Advanced Driver Assistance System (ADAS) component for automated lane changes.
- Increased driver safety by improving road sign detection by 42% through Faster-RCNN integration to ADAS feature.
- Reduced agriculture supplier customer identification cost by 26% by developing a tool for farmland analysis using rasterio, TensorFlow based segmentation algorithm and Flask on sentinel satellite images.

Technical Skills

Languages: Python, C++, MATLAB, SQL, HTML

Database, Cloud Services, Tools: Git, Docker, Kubernetes, Terraform, Jenkins, AWS, Azure, Weaviate, Streamlit

Machine Learning: TensorFlow, PyTorch, Keras, Spacy, NLTK, Huggingface, LlamaIndex, Mlflow, Scikit-learn, Databricks

Other Libraries: OpenCV, Pandas, Pyspark, FastAPI, Matplotlib, Plotly, NumPy, Pillow, Flask

Projects

ResearchSurveyLLM (Project repository link)**:**

- Developed an automated summary paper generator from research papers using RAG framework with LlamaIndex, leveraging open source large language models LLama-2-7b for natural language generation and Mistral-7B for vector embedding.
- Created an ETL pipeline to scrape and filter academic papers from arXiv, storing the refined dataset in Weaviate DB.

FashionXChange (Project repository link)**:**

- Developed an application with Streamlit for text-based modification of outfits of people in images using Grounding DINO, Segment Anything Model and Stable Diffusion.
- Achieved superior pose preservation and reduction in distortion of results by finetuning Stable Diffusion inpainting model with DeepFashion dataset, leveraging Low Rank Adaptation (LoRA) and custom data masking.

Education

Master of Science, Electrical Engineering August 2022 – May 2024

North Carolina State University, Raleigh, NC GPA: 4.0/4.0