BOLLAPALLY MANIDEEPAK

Senior Machine Learning Engineer

9700506010 manideepakbollapally16@gmail.com Hyderabad, India manideepakb

INTRODUCTION

- I have accrued over **4.5** years of robust industry experience as an **AI** and **Machine Learning Engineer**, specializing in a diverse spectrum of domains including **Machine Learning, Computer Vision, Natural Language Processing (NLP), Generative AI,** and **MLOps**.
- My expertise includes extensive knowledge in Python, utilizing libraries such as **PyTorch, TensorFlow, Huggingface, OpenCV, NumPy, and Pandas.**
- Additionally, I have hands-on experience with **Nvidia AI** SDKs, including **Triton Inference server, Rapids, TAO Toolkit, Nvidia Nemo, and Nvidia Bot-Maker.**
- I have a proven track record of success in meeting tight deadlines and excelling in fast-paced environments.
- I am an enthusiastic team player and a ***Techno-Geek***.

EXPERIENCE

Senior Machine Learning Engineer

Quantiphi Analytics

Apr 2022 – Present Bangalore

- Extensive Involvement in Advancing **Generative AI** Capabilities.
- Successfully executed two Pilot projects, maintaining continuous communication with stakeholders and clients to iteratively enhance features based on their requirements. Implemented **ML Ops**, **CI/CD** pipelines to streamline project development.
- completed three End-to-end POC's, each including deployment, within 6-8 week cycles.
- Proficiently orchestrated two comprehensive on-boarding sessions covering **LLM'S** and **DGX Machines**, focusing on **Distributed Training** and **Inference**.
- As part of the **Nvidia R&D team**, I worked with various **NVIDIA Deep Learning SDK's**, including **Nvidia TAO Toolkit**, **Nvidia Nemo LLM's**, **Nvidia Rapids**, **Nvidia Riva** and **Nvidia Triton Inference Server**.

Lead Analyst (ML R&D)

Kantar

Aug 2020 – Apr 2022 Hyderabad

- Fast Proof of concepts for the Internal use cases.
- Played a key role in building the Image pipeline.
- Developed models using **Object detection**, **OCR extraction**, **One-Shot Learning** and **Image classification** techniques.
- End-to-end project management, communicating with stakeholders, understanding the business requirements and delivering projects accordingly.

Research Associate (Computer Vision)

Mirrorsize US Inc

Feb 2020 – Aug 2020 Noida

PROJECTS

Digital Avatar

- I have built a **LLM** based system that can understand customer queries, retrieve relevant information, and generate informative responses.
- In addition, I worked on Riva ASR (Automatic Speech Recognition) Models for tasks like **TTS** (Text To Speech) and **STT** (Speech To Text).
- I also enhanced the generator's(LLM's) accuracy through different fine-tuning techniques, such as **PEFT**, **LoRa** and **QLora.**
- Furthermore, Few Shot Learning and Prompt Tuning on LLMs for better responses.
- Deployed **NLP**, **Speech AI models**, LLMs on **Triton Inference Server.**
- Implemented the **Guardrail's** to restrict the LLM to the scope of the domain, and avoid the unwanted responses.
- Integration of multiple micro services such as **Speech AI**, **Dialog Manager**, **Databases** and **Fulfilment Module.**

Credit Engine

- Developed an end-to-end knowledge graph construction pipeline for unstructured text data.
- Extracted the Important Information from the Data provided Using the **Nemo Large Language** Models.
- Leveraged **Network-X** and **PyViz** to Construct a Network-Graph out of the Extracted and Preprocessed Information.
- Implemented Graph mining techniques such as **Page Rank** and **Community Detection** knowledge graphs such as node classification graph classification and link prediction.
- Used the Constructed graph to Inject into the LLM to Predict the next action to be performed.

Nvidia Launchpad Labs

1. **Automatic close Captioning**
   - Used the Nvidia Deep learning SDK's like **Nemo**,**TAO** (Train Adapt Optimize)Toolkit and **Riva** to perform **TTS.**
2. **Vehicle Damage Estimation**
   - Used the Nvidia **TAO** to train and fine-tune the Object detection model and Deployed on the **Triton Inference server** with the **Dynamic Batching.**
3. **Demand Forecasting**
   - Used Nvidia **Rapids** and **Dask** for the processing of Huge Data frames trained using the **Dask-XGboost** using parallel training ,deployed on the Nvidia **Triton Inference Server** and Used **Plotly Dash** for the Dash boarding.

Multi GPU Training

- Being a part of core Vison team performed fast prototyping on projects based on Computer Vision.
- Trained custom data using the State-of-the-art research papers like **Centermaskv2**, **Maskrcnn Benchmark**, **Yolact**, **Yolact ++**, **Detectron2**, **Pointrend** and **Unet.**
- Worked on **Open pose** and **MMpose** for **Pose Estimation.**
- Worked with the **3d computer vision** models like **HMR** and **SMPL** models for 3D reconstruction.

Machine Learning Intern

Gamenous Pvt Ltd

Jul 2019 - Feb 2020 Hyderabad

- Worked on the Road Crack Detection Problem using Various Object detection Algorithms Like **YOLO-(V2, V3)**, **Inception-V2**, **Mobile-Net-V2**, **EfficientNet.**
- Build our own web based Image Annotation tool with a database connectivity.

SKILLS

Python | Machine Learning | Computer Vision | Pytorch | TensorFlow | Deep Learning | OpenCV | Natural Language Processing | Generative AI | GANS | Large Language Models | Prompt Engineering | LangChain | LlamaIndex | Time Series Analysis

Docker | Kubernetes | Oracle Cloud | MLOps | GCP | AZURE | AWS | Google Vertex AI

Nvidia Triton Inference Server | Dask | Nvidia RAPIDS | Accelerated Machine Learning

MongoDB | Vector Databases | SQL | Flask | Streamlit | Plotly

CERTIFICATIONS

1. Oracle Cloud Infrastructure 2023 Certified Data Science Professional
2. Oracle Cloud Infrastructure 2023 AI Certified Foundations Associate
3. Certificate of Dean
4. TensorFlow Specialization: Advanced Techniques
5. Building Real-Time Video AI Applications

EDUCATION

Master of Computer Application

Chaitanya Deemed to be University

2020 – 2023

B.Sc Statistics

Chaitanya Degree and PG College

2016 – 2019

- The pre-training of the **GPT-20B** model has been effectively completed utilizing **DeepSpeed**'s **Pipeline parallelism** technique, executed on a Nvidia's **DGX Cluster.**

Kantar Vision

- Integrated multiple models into a Pipeline for the Text and Image extraction and Identification of Food items.
- Worked on the os2d for One shot learning, Used the **yolov7** for specific brand detection.
- Worked on merging the **OCR** and **NLP** techniques for extraction of the Brand information from images and videos.
- Worked on Face Analysis model which extracts the features like Age, Gender, Race, Region from the image.

MirrorSize

- Trained Fit detection which is an object detection model with **yolov5XL** in **AWS** using Distrubuted **Training Techniques.**
- Performed Custom Instance segmentation on **Yolact++** with **ResNet-101** Backbone on our data. With **mAP** Score of above 90%.
- Performed the segmentation on our custom data with SOTA papers like **Centermaskv2**, **PointRend**, **Detectron2**, **Maskrcnn-Benchmark** and **Unet.**

Highway Management System

- Object Detection for detecting different Assets on Highway (Signboards, Kilometer-stones, Bus-bays etc) and classification of respective condition.
- Collected Road Cracks Data and used Object detection Models for Identification of 5 different type of Cracks.

ACHIEVEMENTS

Super Ninja

Recipient of the "**Super Ninja**" award, recognized for exceptional multitasking prowess, unwavering attention to detail, and a proven track record of making substantial impacts while consistently delivering top-tier performance.

Selected as Dean for School Of AI

Conducted 5 Meetups on AI, Machine Learning and Deep Learning and a Global Hackathon on AI for Healthcare.

VOLUNTEERING

Dean

School of AI

Feb 2018 – Dec 2021

Served as Dean for Artificial Intelligence community in Hyderabad.