Smit Kumbhani

smitkumbhani100@gmail.com | portfolio
| linkedin.com/in/smit-kumbhani-44b07615a/ | Google Scholar | github.com/smit8800

EXPERIENCE

Research Project Assistant, VAI Lab, Stony Brook University: Aug. 2023 – Apr. 2024 (expected)

- Collaborated with Prof. Klaus Mueller on the image synthesis project specializing in improving conditional denoising diffusion models for CT/MRI image generation.
- Developed a cutting-edge Multi-Conditioned Iterative Latent Variable Refinement-based Diffusion Model and introduced an innovative Image-to-Image Multi-Guided Diffusion Mode for CT Image Synthesis.
- Implementing novel training-free Style Injection method within a Multi-Condition Latent Diffusion Model to significantly improve image-guided Latent Refinement in mammogram synthesis
- Conducted comprehensive performance analyses compared with other generative models, including StyleGAN and diffusion models

Machine Learning Engineer, SigmaRed Technologies: Jun. 2021 – Mar. 2022

- Developed an end-to-end machine learning monitoring platform in collaboration with the MLOps team to monitor and detect real-time data drift and anomalies in containerised machine learning pipelines.
- Designed and integrated an inference data pipeline for sales forecasting, seamlessly connecting it with post-model container-based micro-services for model monitoring and AI bias detection teams.
- Deployed a 13% improved version of the probabilistic and statistical divergence-based drift detection algorithm to detect data-distribution shifts in data embeddings that may deteriorate the model's performance in run-time.
- Designed and implemented a feature store database and built highly scalable data ingestion infrastructure using Apache Kafka and PySpark to accommodate outlier and drift detection mechanisms in run-time and achieved a 50% improvement in real-time data retrieval for the model performance analytics dashboard.

Machine Learning Engineering Intern, SigmaRed Technologies: Mar. 2021 – May 2021

- Contributed to developing model training workflows using Kubeflow for training data pipelines.
- Performed well-detailed research on various drift and outliers detection algorithms for the tabular dataset.

Research Intern, DST-SERB SMART Foundry Lab, Marwadi University: Jan. 2020 – Nov. 2020

- Collaborated with the backend engineering team to develop a deep neural network training and prediction platform for sensor data. optimized 20% training time by implementing a self-feature selection pipeline for deep neural networks.
- Studied and integrated genetic algorithm to further optimise the time complexity for auto hyper-parameter tuning of the deep neural network architecture.

EDUCATION

- **Stony Brook University** New York, USA
  Master of Science - Computer Science *Jan. 2023 - Dec. 2024 (expected)*
  Courses: Natural language processing, Computer vision, Machine learning, Databases, Operating systems

SELECTED TECHNICAL PROJECTS and RESEARCH PUBLICATIONS

- **Multi-Conditioned Denoising Diffusion Probabilistic Model (mDDPM) for Medical Image Synthesis (MICCAI) 2024**: Proposed a controlled generation framework for synthetic images with annotations, utilizing multiple conditional specifications as input to enable precise control over the output. **menuscript submitted**
- **Explainable Pneumothorax Detection Framework, [Computer Vision]**: Designed and developed bi-model architecture employing UNet++ and ResNet for precise pneumothorax segmentation in CT images. Implemented Grad-CAM to enhance the inference interpretability of pneumothorax detection and segmentation. **[Code, article]**
- **GCP-based Malware Detection System, [Python, Google Cloud]**: Implemented a Malware detection system capable of identifying nine distinct categories of malware using source code from .byte and .asm files. Implemented a feature transformation of a 200 GB dataset on GCP, converting byte and assembly program files into Image data input, resulting in an outstanding micro-f1 score of 0.999. **[Code]**
- **Luong-Style-Attention-based-Languange-Translation, [Python, NLP]**: Proposed optimised translation model using the Luong-style attention mechanism and achieved the final BLEU score of 0.6667. **[Code]**
- Smit Kumbhani, Vishesh Dharaiya, "A Custom Stacking-based Ensemble Learning approach to predict failure of stripper well", Proceedings of International Conference on Communication and Artificial Intelligence (ICCAI) 2021, Springer Publication, **doi: 10.1007/978-981-19-0976-4-28.**

TECHNICAL SKILLS

- **Languages**: Python, Golang, Java, C++ — **Library and Frameworks:** SLURM for GPU-accelerated training, TensorFlow, CUDA, TFLite, PyTorch, Flask, Django, Keras, Scikit-learn, NLTK, OpenCV, GraphQL, React.js, Node.JS — **Databases:** MongoDB, MySQL — **Cloud and Tools:** AWS, GCP, Linux, Docker, GIT, Kubernetes — **Big-Data:** Spark, Kafka.