Smit Kumbhani

New York | smitkumbhani101@gmail.com | +1(631)948-4041

| linkedin.com/in/smit-kumbhani-44b07615a/ | Google Scholar | github.com/smit8800

EDUCATION

Stony Brook University New York, USA
Master of Science - Computer Science (GPA: 3.78/4) *Jan. 2023 - Present*
Courses: Computer Vision(Ph.D. level), Machine Learning(Ph.D. level), Databases, Probability and Statistics

PROFESSIONAL EXPERIENCE

Research Project Assistant, Stony Brook University: Aug. 2023 – Present
- Collaborated with Prof. Klaus Mueller on the image synthesis project and Developing denoising diffusion models for 3D CT image generation.
- Devised and executed Paired and Unpaired Training via CycleGAN to generate CT image segmentation maps for visual Turing tests and conducted comprehensive performance analyses compared with other generative models, including StyleGAN and diffusion models.
- Implemented the state-of-the-art research paper titled "Improved Denoising Diffusion Probabilistic Models" from OpenAI, significantly enhancing CT Segmentation Map Generation.

Machine Learning Engineer, SigmaRed Technologies: Jun. 2021 – Mar. 2022
- Developed an end-to-end ML monitoring platform in collaboration with the MLOps team to monitor and detect real-time data drift and anomalies in containerised machine learning pipelines.
- Deployed an improved version of the probabilistic and statistical divergence-based drift detection algorithm to detect data-distribution shifts in image and text-embedding data that may deteriorate the model's performance in run-time.
- Designed and implemented a feature store database using MongoDB and built a data pipeline using Apache Kafka to accommodate outlier and drift detection mechanisms in run-time and achieved 50% improvement in real-time data retrieval for the model performance analytics dashboard.
- Collaborated with an explainable AI research team to build RPC(remote procedure call framework) based microservice to integrate real-time model explainability functionality.

Machine Learning Engineering Intern, SigmaRed Technologies: Mar. 2021 – May 2021
- Worked with the core team of ML monitoring and contributed to developing adversarial and drift detection mechanisms for tabular data.
- 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
- The Department of Science and Technology, India-funded project teamed with IIT Bombay to leverage the research and development of artificial intelligence and IOT for the Metal Casting Industry.
- Collaborated with the backend engineering team to develop a deep neural network training and prediction platform for sensor data.
- Reduced 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.

SELECTED TECHNICAL PROJECTS and RESEARCH PUBLICATIONS

- **Explainable Pneumothorax Detection Framework, [Computer Vision]**: Designed and developed bi-model architecture to add CAM-based interpretability in Pneumothorax detection and segmentation. **[Link, article]**
- **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. **[Link]**
- **GCP-based Malware Detection System, [Python, JavaScript]**: 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. **[Link]**
- **Multi-Modal Fusion of Speech Recognition, [CNN/LSTM, Transformers]**: Developed Multi-Model architecture by combining CNN and seq-seq-based deep learning network to detect spoken digits. **[Link]**
- 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.**
- K. Mridha, S. Kumbhani, S. Jha, D. Joshi, A. Ghosh and R. N. Shaw, "Deep Learning Algorithms are used to Automatically Detection Invasive Ducal Carcinoma in Whole Slide Images," 2021 IEEE 6th International Conference on Computing, Communication and Automation (ICCCA), 2021, pp. 123-129, **doi: 10.1109/ICCCA52192.2021.9666302.**

TECHNICAL SKILLS

- **Languages**: Python, C/C++, Go, SQL — **Library and Frameworks:** TensorFlow, TFLite, PyTorch, Flask, Django, Keras, Scikit-learn, NLTK, Pandas, Numpy, OpenCV — **Databases:** MongoDB, MySQL — **Cloud and Tools:** AWS(EC2, S3), GCP(Compute Engine, Storage), Linux, Docker, GIT, Kubernetes — **Big-Data:** Spark, Hadoop, Kafka.