RAKTIMJYOTI PARASHAR

P: +1-(267)-221-1000| aarjayparashar@gmail.com | https://github.com/raktimparashar-upenn
https://www.linkedin.com/in/raktim-parashar-upenn/

SUMMARY

Accomplished Mechanical Engineer and Solution Data Scientist with 3+ years of experience in research and industry applications involving multi-cloud solutions in AWS and Snowflake. Expert in designing complex robotic systems, deep learning, and end-to-end machine learning pipelines, complemented with a Master's degree from University of Pennsylvania, specializing in Robotics/AI and Mechatronics. Significant experience in large language models (LLM), Computer Vision and Contemporary AI research.

WORK EXPERIENCE

GRASP LAB, University of Pennsylvania
Research Engineer (June 2022 – Present)

- Collaboratively led the design and implementation of a cutting-edge 6-DOF, occlusion-aware **object tracking pipeline** using state-of-the-art **Computer Vision** models. This system utilizes segmented point cloud data and a **deep neural network** to compute optimal grasp poses for robotic handling of novel objects, significantly enhancing automation capabilities in unstructured environments.
- Engineered a robust pipeline for reconstructing detailed 3D models from raw point cloud data using Log-GPIS and **Open3D** frameworks. This initiative has improved the accuracy of robot perception systems, enabling finer control and interaction with complex environments, and has been instrumental in refining the lab's autonomous robotic systems.
- Collaborated with a cross-functional team to create a **Neural Radiance Field (NeRF)** representation of indoor spaces to accurately model and navigate around obstacles using vision-based systems only. This work has paved the way for developing more intuitive and safer autonomous navigation solutions in environments where traditional sensors might fail.
- Worked on **LiDAR-Depth Camera Sensor Fusion** onboard AgileX SCOUT using **ROS**.

PENN MEDICINE, University of Pennsylvania
Data Scientist (March 2022 – August 2022)

- Architected and implemented a scalable multi-cloud **Snowflake-AWS** solution to analyze patient length of stay data across multiple hospital systems. This initiative supported advanced analytics to forecast patient flows and optimize hospital resource allocation, contributing significantly to the efficiency of healthcare delivery.
- Developed data-driven strategies to enhance the hospital supply chain management by integrating patient stay insights with supply chain needs. Enabled precise inventory control and reduced waste by aligning supply levels with anticipated patient care requirements, leveraging cloud-based analytics platforms.
- Employed **MongoDB** as a core technology for data management and analysis in biotech research, enabling sophisticated data aggregation, storage, and retrieval processes. Utilized MongoDB to facilitate efficient data pulling and integration within Snowflake and AWS platforms, supporting complex scientific computations and insights.
- Led the design and deployment of bioinformatics workflows that utilize patient data to improve hospital logistics. Created robust data pipelines that support the continuous analysis of large datasets, enhancing the decision-making processes in hospital management and care provision.
- Established strict data governance frameworks to ensure compliance with healthcare regulations, including HIPAA. Implemented secure data storage and transfer protocols within multi-cloud environments, safeguarding sensitive patient information against unauthorized access.

SKILLS

Technical: Windows, Linux, C++, Python, R, Tableau, PowerBI, Pandas, Matplotlib, Seaborn, Scikit-learn, Logistic Regression, SVM, Decision Trees, Random Forests, XGBoost, LightGBM, Deep Learning, TensorFlow, PyTorch, Flask, FastAPI, SQL, MySQL, PostgreSQL, Spark, MongoDB, Git, GitHub, HuggingFace, BERT, GPT, ViT, Transformers, RAG, Docker, CI/CD pipelines, HTML, CSS, Agile, Databricks, AWS, Snowflake, OpenCV, Open3D.

Cloud: AWS SageMaker, Rekognition, IAM, EC2, Lambda, S3, AWS Serverless, Snowflake, SnowSpark.

PROJECTS

- Applied projective transformations and homographies in Python and OpenCV to project logos onto video sequence.
- **Augmented Reality (AR)** with AprilTags: Developed an AR application by solving the PnP, P3P and Procrustes problem.
- **Structure from Motion (SfM)**: Implemented SfM pipeline to recover 3D transformation between two views using SIFT feature matching, 8-point algorithm with RANSAC filtering, camera pose estimation and point triangulation.
- **3D Reconstruction** using Stereo Vision: Performed stereo rectification, block matching with SAD, SSD and ZNCC kernels, and left-right consistency check to build disparity map and reconstruct point cloud.
- **Cancer Prediction:** Pioneered the implementation of advanced machine learning and deep learning models for the accurate detection and segmentation of polyps from colonoscopic images, using **PyTorch, OpenCV, Scikit-learn, pandas, numpy, albumentations**. Enhanced diagnostic precision by developing algorithms that surpass traditional methods, aiding medical professionals in identifying subtle abnormalities that might not be visible to the human eye, thereby significantly improving patient outcomes in colorectal cancer screenings.
- **Search Engine:** Engineered a sophisticated machine learning application to analyze linguistic relationships and patterns in medical terminology using NLTK, Scikit-Learn, Pandas, Numpy, Streamlit. Developed a smart search engine capable of querying complex medical records for specific terms, enhancing the retrieval of critical data. Successfully deployed and scaled the solution using Azure Machine Learning pipelines (**Azure Data Factory, Azure Blob Storage, Azure Databricks**), facilitating real-time data processing and significantly improving research and clinical decision-making efficiency.
- **Enhancing Business Engagement: Advanced AI and LLM for Detoxifying and Moderating Hate Speech in Online Communities:** Leveraged advanced AI and **Large Language Models (LLM)** to combat toxicity in online communities. Our findings revealed how sophisticated models like FLAN-T5, enhanced through Proximal Policy Optimization (PPO), can significantly reduce hate speech and improve digital interactions.
- **End to end deep learning project** to classify disease in potato plant as either early blight or late blight. Technology stack of this project consists of **TensorFlow, tf dataset, tf serving, FastAPI, data augmentation, react js, react native, Google cloud platform (for deployment), quantization, tf lite.**
- **Chicken Disease Classification:** End to end deep learning project using **MLOps** tools such as **DVC** along with deployment using **GitHub Actions** in AWS Cloud using **EC2** instance.
- **Text Summarizer:** End to end **NLP** project along with deployment using **GitHub Actions** in **AWS Cloud** using EC2 instance.
- People Counter using **YOLO-NAS** and **DeepSORT Object Tracking**.
- Push Up Counter using **YOLOv7 Pose Estimation**.
- Built a **Speech Recognition and Summarization** System using vosk, pydub and transformers packages.
- Stock Price Predictor using yfinance package.
- mRNA Degradation Prediction using deep neural networks.

EDUCATION

UNIVERSITY OF PENNSYLVANIA
Master of Science in Mechanical Engineering and Applied Mechanics
Concentration: Robotics and Mechatronics (GPA: 3.68/4.0)
MANIPAL INSTITUTE OF TECHNOLOGY
Bachelor of Technology in Mechanical Engineering and Manufacturing Technology (8.5/10)