Shoumik Majumdar

LinkedIn: Shoumik Majumdar • Email: shoumikmajumdar@gmail.com

EDUCATION

Boston University, Boston, USA Sep 2019 - Jan 2021

- Master of Science, Computer Science

University of Mumbai, Mumbai, India Aug 2014 - May 2018

- Bachelor of Engineering, Computer Science

SKILLS

- **Languages**: Python, C, Java, R, Shell Script, SQL.
- **Databases**: mySQL, Big Table, Big Query.
- **Technologies/Platforms**: Linux, Bash, PyTorch, TensorFlow, LangChain, Keras, Pandas, Numpy, OpenCV, Scikit-learn, Flask, Git, Docker, MLFlow, SparkML, Google Cloud Platform, Databricks.

WORK EXPERIENCE

Quantiphi Inc: Machine Learning Engineer, Boston May 2021 - Present

- Machine Learning Engineer in the Applied AI team responsible for architecting, developing and deploying end-to-end Machine Learning solutions on Google Cloud Platform (GCP).
- As part of the Applied research team, responsible for designing research studies and documenting learning which led to publications and development of AI accelerators.
- Support GTM teams by conducting use case discoveries and workshops, prototyping and scoping statements of work (SOW).

Boston University: Research Associate (Machine Learning), Boston Sep 2019 - Jan 2021

- Collected, filtered and annotated the 1st human action recognition video dataset for **domain generalization**.
- Used the Inception I3D architecture to align shifts across both spatial and temporal domains. Extended **Adversarial Feature Augmentations** and **GradCAM saliency** mappings to incorporate model explainability for videos.

RELEVANT PROJECTS

Entity Extraction - Document AI Dec 2022 - Present

- Developing a **document classification** and **entity extraction** solution for a public sector organization.
- Training entity extraction models on Health Records to extract entities and identifying comorbidity. Enabling **continuous training** with **Human In the Loop** integration.
- The Solution will eliminate the need for entity extraction workflows currently in place, resulting in a substantial **reduction in the processing time** for insurance claims.

Knot Detection - Visual Inspection AI Jun 2023 - Aug 2023

- Created an **image segmentation** model utilizing GCP's **Visual Inspection AI** platform for real-time detection of knots in yarn as it passes through tufting machines.
- Developed and deployed ingestion and inference pipelines on the edge using **Kubernetes** to enable automatic shutdown of tufting machines using relay switches with minimal latency.
- The solution **improved manufacturing yield** by flagging knots accurately thus **reducing downtime and defects** and relieving the need for **manual inspectors**.

Composite Metrology/Hybrid Metrology Dec 2021 - Apr 2022

- Orchestrated an **end to end ML framework** on GCP for a metrology usecase for a semiconductor manufacturing company.
- Developed **Vertex AI** pipelines on GCP for data extraction, model training, batch predictions and model deployment with **automated triggers via Pub/Sub**.
- Migrated client's on-premise services to a **fully-managed, scalable** and **cost efficient** framework on GCP enabling easy transition for future usecases.
- The developed framework reduced average runtime of the workflow from **96 hours to 8 hours**.

Model Monitoring/Continuous Evaluation framework Aug 2021 - Nov 2021

- Developed a **model monitoring framework** on GCP for an object detection usecase for a semiconductor manufacturing company.
- Prototyped and analysed scalability of various drift detection algorithms like Maximum Mean Discrepancy, JS Divergence and KL Divergence.
- Developed a **continuous evaluation pipeline** to detect data and label drifts and trigger **model retraining** and deployment using **Cloud Composer**.

PROFESSIONAL CERTIFICATIONS

- Databricks Certified Machine Learning Professional. Jun 2023
- Databricks Certified Machine Learning Associate. May 2023
- Google Cloud Certified - Professional Machine Learning Engineer. Oct 2022
- Google Cloud Certified - Associate Cloud Engineer. Dec 2021