SHUBHODEEP MITRA

(602) 849-3716 | *smitra27@asu.edu* | *linkedin.com/in/shubhodeep-mitra* | shubhodeepmitra.github.io

EDUCATION

Master of Science in Computer Science October 2024
Arizona State University, Tempe, AZ **GPA**: 3.96/4

- Foundations of Algorithms, Distributed Database Systems, Privacy and Machine Learning, Cloud Computing.

Bachelor of Engineering in Computer Science and Engineering May 2018
The National Institute of Engineering, Mysuru, India **GPA**: 8.74/10

TECHNICAL SKILLS

Programming Languages: Java, C/C++, Bash, Python (NumPy, pandas), Android, Kotlin, Go, Linux/Unix Programming.
Skills: SpringBoot, ReactJS, Redux, MapReduce, LLM, RAG, Kafka, Docker, REST, Ansible, AWS, GraphQL, PyTorch, Terraform, Kubernetes, SQL, PostgreSQL, Redis, MongoDB, Firebase, Git, GitHub, DataDog, gdb, ElasticSearch.

EXPERIENCE

Software Developer 2 - Marqeta, Bay Area, USA November 2024 – Present

- Building scalable and available Infrastructure for network systems to support mission-critical financial systems.

Graduate Research Assistant (Distributed Systems) - EMITLab ASU, Tempe, USA December 2022 – July 2024

- **Led developments of scalable ML infrastructure** to support scaling of large machine learning models, optimizing data pipelines and concurrent workflows on AWS, which increased model diversity and training efficiency by **4x**.
- **Improved parallel processing** and automated ML workload deployments using Ansible on AWS EC2 clusters, reducing experiment runtime by **27%** through improved parallelism and resource utilization.
- **Revamped in-memory graph data structures** and integrated local caching mechanisms, reducing MongoDB calls by over **50,000** and speeding up generation of provenance graphs.
- **Built an interactive data visualization platform** with React.js and Next.js to summarize and analyze multi-variate time series, aiding data-driven decision-making for large-scale datasets.

Software Engineer 2 - Hewlett Packard Enterprise (HPE), Bangalore, India July 2018 – July 2022

- **Led development efforts in C/C++ for critical networking components**, including internal-VLAN, L3 counters, Netdev, and Ofproto, pivotal for IP-Subinterface, facilitating seamless traffic flow across **17 protocols**.
- Redesigned multicast protocols, IGMP and MLD to integrate real-time packet flow monitoring, boosting reliability.
- **Developed CLI infrastructure for L2 protocols**, VLAN Translation and Multi-Zone User-Based Tunneling, integrating SDN features that enabled network management and simplified configuration for large-scale deployments.
- **Engineered test automation suite** for all development work, achieving **98%** code coverage and empowering DevOps team to perform continuous system health checks.

RELEVANT PROJECTS

ColumnarDB March 2024

- **Designed a Columnar Database system** in Java with features including BitMap and BTree indexing, compressed BitMap, Columnar Joins, Columnar Sort, Scan, and Delete optimizing data operations efficiency for **50k entries**.

Live Video Streaming Application December 2023

- Architected and implemented a high-performance **live video streaming application** in **Go** and **WebRTC**, optimizing for low latency and scalability, employing advanced techniques to enhance video delivery quality.

Image Retrieval and Recommendation October 2023

- Designed and built an **end-to-end image retrieval and recommendation system** with 82% accuracy, using neural network feature extraction, dimensionality reduction, Personalized PageRank, LSH, and a vector database.

Elastic Face Recognition Service March 2023

- **Developed a scalable face recognition service** using AWS EC2, S3, DynamoDB, and SQS, achieving 98% accuracy in facial recognition from video streams while dynamically scaling to handle 10,000+ concurrent requests.

Scalable Aesthetic-Preserving Face De-Identification November 2022

- Created a Kotlin Android app employing an **ML-kit** and openCV foundation to detect individuals in photograph backgrounds through face recognition, applying an aesthetic-preserving filter to safeguard bystanders' privacy.

Real-Time Parking Spot Notification, Hewlett Packard Enterprise January 2020

- Engineered a scalable system to monitor over 500 vehicles entering and exiting campus. Leveraged surveillance footage and employed OpenCV and TensorFlow, providing Android users with live parking availability updates.

PUBLICATIONS

Novel TLS Signature Extraction for Malware Detection, IEEE CONECCT July 2020

- Researched a solution at Hewlett Packard Enterprise to identify presence of malware in a network flow from an initial unencrypted Client Hello packet of TLS with 92.4 percent accuracy.