

# RISHIKA GURRAM

gurram.ri@northeastern.edu • 857.313.5483 • [LinkedIn](#) • [GitHub](#) • [Portfolio](#)

## EDUCATION

### Northeastern University, Boston, MA

*Expected Dec 2024*

Master of Science in Information Systems (Computer and Software Engineering), **CGPA: 3.7**

*Relevant Coursework:* Network Structures and Cloud Computing, High Parallel Machine Learning and AI, Program Structures and Algorithms, Web Design and User Experience, Data Science Engineering Methods and Tools

### Jawaharlal Nehru Technological University, Hyderabad, India

*May 2020*

Bachelor of Technology in Electronics and Communications Engineering, **CGPA: 8.6**

## TECHNICAL SKILLS

**Languages:** Python, C++, SQL, JavaScript, Java, Unix Shell Scripting

**Cloud Technologies:** AWS (EC2, RDS, S3, CloudWatch, Route53, IAM), Terraform, GitHub Actions, Azure

**Web Technologies:** React JS, Node JS, Express, HTML/CSS

**Databases:** Oracle, PostgreSQL, MySQL, MongoDB

**Tools/IDE:** Git, Docker, HPC Clusters (OpenMP, DASK), RESTful APIs, Postman, Packer, Oracle Hyperion

## WORK EXPERIENCE

### Graduate Teaching Assistant

*Aug 2024 - Dec 2024*

*Northeastern University, Boston, USA*

- Mentored 50 students in CSYE7015 course on high-performance computing and parallel machine learning, guiding them through parallel architectures, HPC clusters, and multi-CPU/GPU data processing and machine learning implementations. Supported hands-on labs on Northeastern's Discovery cluster

### Software Engineering Intern

*Jul 2023 - Dec 2023*

*12-15 Molecular Diagnostics, CT, USA*

- Spearheaded a code refactoring initiative on a healthcare device, leading to a 45% decrease in error rates and the development of a highly efficient desktop application to streamline healthcare operations
- Adapted and optimized algorithms for client-specific use cases, implementing custom configurations and refining computational logic to address diverse healthcare
- Developed and deployed an ETL pipeline to automate complex data analysis, achieving a 50% reduction in processing time and significantly improving data accuracy and reliability
- Architected and implemented a scalable microservices architecture using Docker and AWS, integrating desktop applications with cloud-based systems for real-time user experiences and enhanced flexibility

### Associate Software Engineer

*Dec 2020 - Aug 2022*

*Accenture, India*

- Architected and implemented a comprehensive access control framework within the Oracle Hyperion workspace, resulting in a 45% reduction in system implementation time and improved data security
- Streamlined Dev, UAT and Prod environment consoles during business phases to complete tasks including migration, patching and hierarchy changes, yielding up to 60% reduced downtime and errors

## ACADEMIC PROJECTS

### Leveraging Parallel Processing for AI image detection (PyTorch, CUDA, Discovery Cluster, NVIDIA GPU) *Apr 2024*

- Orchestrated parallel processing on 6 GB dataset utilizing multiprocessing, DDP, and DDP AMP across diverse CPU and GPU resources on OOD, optimizing computational efficiency to amplify performance
- Implemented efficient utilization of Tesla NVIDIA GPU resources, cutting processing time by 93% and achieving image detection in 32 minutes, with an accuracy of 75%, marking an advance from serial processing
- Employed a 4-GPU AMP configuration, to establish a performance benchmark, demonstrating a 13% speedup over standard GPU setups and showcasing expertise in optimizing AI workloads for maximum efficiency

### Restful Services on Cloud (AWS, Python FastAPI, PostgreSQL)

*Apr 2023*

- Architected a three-tier RESTful API infrastructure on AWS optimizing infrastructure provisioning using Terraform templates for automation of EC2/RDS instances, VPC, Subnets, and Security Groups
- Engineered a CI/CD pipeline using GitHub Actions, Packer AMI builder, and CodeDeploy, reducing deployment time from several hours to 20 minutes
- Utilized CloudWatch to monitor EC2 and CPU RAM to create triggers for scale-in and scale-out rules for the Application Load Balancer, guaranteeing it could control up to 10,000 requests concurrently