IRFAN AHAMED MELEKKANDY PUTHALATH

East Newark, NJ — iam73@scarletmail.rutgers.edu — (862) 423-3940 — LinkedIn — GitHub — Portfolio

EDUCATION

RUTGERS UNIVERSITY *Sep 2024 – Dec 2025*

Master of Information Technology (Machine Learning)

Relevant coursework: Advanced Database Management (AWS, Snowflake), Data Structures And Algorithms, Machine Learning, Python for Analytics

RV COLLEGE OF ENGINEERING *Aug 2017 – Aug 2021*

Electronics And Instrumentation

EXPERIENCE

CGI INC. Bengaluru, India

Software Engineer Sep 2022 – Jul 2024

NLP Document Processing System

- Designed and deployed a hybrid CNN-BERT model in PyTorch 2.0 to automatically classify legal documents into 12 regulatory categories, achieving 94% F1-score on a dataset of 50,000 documents
- Optimized document processing by implementing dynamic batching on AWS Inferentia, reducing latency by 73% and enabling real-time processing of 15,000 documents per day
- Established a robust CI/CD pipeline using GitHub Actions and SageMaker Model Monitor, reducing deployment errors by 41% and ensuring reliable model updates

Microservices Optimization

- Developed a high-throughput fee calculation service using Vert.x, achieving 28,000 requests per second through lock-free algorithms and efficient resource management
- Led the migration of a legacy Java EE application to a Kubernetes-based microservices architecture (237 pods), cutting annual cloud infrastructure costs by $1.8M

RELEVANT PROJECTS

Multimodal Sentiment Analysis System *Jan 2025*

- Developed a sentiment analysis platform combining BERT and CNN models to process both text and image data, achieving 91% accuracy on benchmark datasets
- Improved model retraining efficiency by 15% through advanced data preprocessing techniques
- Deployed the solution as a FastAPI-based service in Docker containers, enabling automated retraining and seamless integration into production environments

React Notes Application Deployment with AWS Amplify *Nov 2024*

- Built and launched a scalable React-based notes app using AWS Amplify, maintaining 99.9% uptime and supporting over 5,000 active users
- Automated the deployment workflow with AWS Amplify, integrating continuous delivery and real-time monitoring, which resulted in a 40% reduction in manual deployment errors and accelerated feedback cycles
- Enhanced user experience by optimizing CDN edge delivery, reducing average image load times from 3 seconds to under 750 milliseconds and lowering bounce rates

Production-Grade Depth Regression Model *Oct 2024*

- Designed and implemented a neural network for monocular depth estimation, achieving 98% precision in object identification from single camera feeds
- Reduced inference time by 67% through model optimization, enabling real-time processing at 28 frames per second on edge devices
- Improved prediction robustness by 20% in challenging lighting conditions using ensemble techniques, and collaborated with a team to integrate global and local depth networks, boosting overall performance by 45%

SKILLS & CERTIFICATIONS

Certifications: AWS Certified Developer – Associate

Machine Learning Algorithms & Techniques: Supervised Learning, Deep Learning (CNNs, RNNs, Transformers), Unsupervised Learning (Clustering, PCA), Reinforcement Learning

ML Engineering & Deployment: MLOps (Kubeflow, MLflow), Model Serving (TensorFlow Serving, ONNX Runtime, TorchServe), Monitoring and Drift Detection

Programming Languages & Frameworks: Python, Java, SQL, C++, R, JavaScript, Angular, Vert.x, Bash, Linux

Tools & Platforms: Docker, Kubernetes, AWS SageMaker, MySQL, Redis, Tableau, VS Code, Eclipse, Git, Pandas, NumPy

Cloud & DevOps: AWS, Azure, CloudFormation, DynamoDB, ECS, CI/CD Pipelines, Containerization