

GAUTHAM SHAJI

(945)-232-4345 • gauthamshajigs97@gmail.com • linkedin.com/in/gautham-shaji • github.com/gauthamshaji

EDUCATION

Master of Science in Computer Science

The University of Texas at Dallas, Richardson, TX

August 2022 - May 2024

3.82 GPA

Relevant coursework: Advanced Operating Systems(Distributed Systems), Database Design, Big Data.

Bachelor of Technology

National Institute of Technology Calicut, Kerala, India

August 2015 - May 2019

3.02 GPA

Relevant coursework: Data Structures, Computer Networks, Artificial Intelligence and Machine Learning.

TECHNICAL SKILLS

Programming: Java, Python, C/C++, Javascript, Go

Databases: MySQL, NoSQL, MongoDB, DynamoDB, PostgreSQL, Cassandra

Backend: Spring Boot, Flask, Hadoop, Docker, Apache Kafka, Node.js , OAuth2 , Apache Spark

Cloud Services: Amazon Web Service (Textract, Lambda, S3,EC2, DynamoDB), Google Cloud Platform (Kubernetes)

EXPERIENCE

The University of Texas at Dallas, : Graduate Software Engineer

June 2023 – Present

- Operationalized an automated URL testing framework, markedly enhancing quality assurance time by 40% (**Selenium, Java**)
- Integrated **OpenID Connect authentication**, improving system security and user verification processes within UTD SSO

Zycus Infotech, : Associate Technical Lead

April 2020 – June 2022

- Scaled monolithic python **autonomous header recognition system** pipeline by redesigning it with independent asynchronous Kubernetes CronJobs leading to 80% improvement in runtime performance.
- Implemented a **TensorFlow Serving**-based production system achieving over **10,000** queries per second throughput across a dual server setup. (**Python, LSTM, Tensorflow, AWS S3, EC2**)
- Engineered a **serverless** autonomous learning framework enabling dynamic, customized entity suggestions with real-time training updates (**Python, Flask, MongoDB, Universal Sentence Encoder, EC2**)

Quantiphi Analytics, : Software Developer Engineer

June 2019 – April 2020

- Architected an insurance assistant backend, facilitating scalable, high-throughput real-time data processing and leveraging advanced NLP LLM models (**Python, AWS Lex, BERT-based QnA, Word2Vec, AWS EC2**)
- Developed QDox, a document automation platform, achieving efficient **extraction and processing** from unstructured data
- Significantly reduced development time by 70% with pre-trained models across 100+ document types, integrating custom workflows and human-in-the-loop architecture for enhanced accuracy (**AWS Textract, AWS EC2, LSTM, OpenCV, Application Load Balancer, Classic Load Balancer, S3, AWS Lambda**)

ACADEMIC PROJECTS

Distributed Storage System with RAFT (Distributed Systems) [\[GitHub Link\]](#)

August 2023

- Engineered a robust distributed storage system integrating the **RAFT protocol** for consensus and replicated state machine technique for strong consistency
- Incorporated replicated state machines to synchronize and maintain a consistent application state across all nodes in the system (**Java, Git, Socket Programming**)

Peer-to-Peer System Implementation for File Sharing and Distributed Search(Advanced OS) [\[GitHub Link\]](#)

July 2023

- Forged a **Peer-to-Peer (P2P)** system enabling seamless file sharing and downloads among autonomous computers
- Orchestrated node joining and departure processes, hop-count, response times, networking protocols, scalability strategies, fault tolerance mechanisms, and synchronization techniques(mutex locks and semaphores) (**Java, Git**)

Authentication and Authorization in Spring Boot Applications Using JWT [\[GitHub Link\]](#)

September 2023

- Successfully integrated **JSON Web Tokens (JWT)** into a Spring Boot application for stateless authentication mechanism and authorization for secure claim representation between parties
- Streamlined JWT-based authentication, focusing on user validation, token management, and HTTP request filtering for security. Secured API endpoints are accessible solely by users with valid JWTs (**Spring Boot, Java, Spring Security , MySQL, JWT**)

Scalable Text Classification using Spark NLP and Apache Spark [\[GitHub Link\]](#)

November 2023

- Implemented an efficient text classification solution using **Apache Spark** and **Spark NLP**, capable of processing and analyzing large-scale text datasets.
- Developed two text classification models – a Logistic Regression model and a ClassifierDL deep learning model with f1 score of 80 and 87 respectively. (**Universal Sentence Encoder, Apache Spark, Spark NLP, Python, NLP, LLMs**)