

YASWANTH ADARI

☎ +1 214 5384308 ✉ yaswanthadari.utdallas@gmail.com [in linkedin.com/in/yaswanth-adari](https://www.linkedin.com/in/yaswanth-adari) github.com/YaswanthAd

EDUCATION

University of Texas at Dallas, Dallas, TX Expected May 2024
Master of Science in Computer Science | CGPA: 3.94 / 4.0

National Institute of Technology Tiruchirappalli, India Graduated May 2019
Bachelors of Technology | CGPA: 8.40 / 10.0

WORK EXPERIENCE

CGI Aug 2019 - Jun 2022
Software Engineer Bangalore, India

- Engineered ETL pipeline using Apache Spark, transforming data from IBM DOORS, JAMA, POLARIN to Amazon Redshift on AWS with PySpark, AWS Glue, and integrated Apache Kafka for real-time streaming, achieving 60% automation and ensuring scalability.
- Enhanced data workflow management with Apache Airflow, streamlining scheduling and monitoring in ETL processes, resulting in a 15% increase in overall task efficiency and a 20% reduction in execution errors.
- Configured middle-wares like Prometheus, Grafana, Kafka and MySQL for monitoring, logging, and event streaming, achieving a 25% decrease in system downtime and boosting performance by 30%.
- Orchestrated performance optimizations for databases, fine-tuning MySQL and Snowflake configurations to achieve 12% reductions in query response time.
- Implemented Git-integrated CI/CD pipelines with Jenkins, reducing release cycle times by 20% and lowering post-release issues by 10%.

Indian Institute of Management, Bangalore May 2018 - Jul 2018
Software Engineer Intern Bangalore, India

- Extracted data from 100+ articles & research blogs on food aggregators Swiggy, Zomato with using web scraping tools such as BeautifulSoup in Python, and stored the processed data in AWS S3.
- Enabled University-wide access to data by deploying AWS-based APIs via API Gateway and AWS Lambda, ensuring scalability and efficiency

ACADEMIC PROJECTS

- Social Network Analysis with Spark GraphX/GraphFrame | Utilized Spark GraphX/GraphFrame to analyze social network data from the SNAP repository, extracting valuable insights. Developed graph queries and algorithms to identify top nodes by outdegree, indegree, PageRank, connected components, and triangle counts. [Github](#)
- E-commerce Application | Engineered a streamlined database architecture for Refurb's e-commerce platform, incorporating normal forms, entity mapping, and constraints. Developed an optimized SQL schema to ensure data integrity and efficiency, seamlessly integrating it with the front end for a cohesive user experience. [Github](#)
- Tutor Challenge | Innovatively designed a resource allocation solution for the "Tutor Assignment Challenge" using threads, locks, and semaphores, resulting in a 30% reduction in allocation time. [Github](#)

PERSONAL PROJECTS

- Real-Time Named Entity Analysis with Spark Streaming and Kafka | Implemented a real-time data analysis pipeline using Spark Streaming, Kafka, Elasticsearch, and Kibana to extract and visualize named entity trends from dynamic text sources. [Github](#)
- Text Classification Using Naive Bayes Classifier | Developed a PySpark-driven Naive Bayes classifier, ensuring high accuracy in text categorization and revealing insights into class prior probabilities. Executed comprehensive optimization throughout the process, including data preprocessing, model development, and result evaluation for diverse text corpora. [Github](#)

TECHNICAL SKILLS

Languages: C/C++, Python, Java, JavaScript, HTML, CSS, SQL

Developer Tools: Spring, Docker, AWS S3, AWS Redshift, Spark, Hive, Hadoop, Git, Kafka, Helm, GraphX, Hbase

Databases: SQL, PostgreSQL, MongoDB, Cassandra, Redis.

Frameworks & Libraries: React, Node.js, Express.js, Redux, Firebase, PySpark, RESTful web services, Linux, PyTorch, TensorFlow.

Relevant Coursework: Operating Systems, Database Design, Data Structures and Algorithms, Design and Analysis of Algorithms, Machine Learning, Big Data Management and Analytics, Natural Language Processing, Computer Vision.