MOHAN DESHPANDE

+1(682) 360-5283 ◇ Dallas, TX
mohandeshpande@hotmail.com ◇ linkedin.com/in/mohandeshpande ◇ https://mohandeshpande.github.io

EXPERIENCE

Software Engineer July 2021 - Present
Capital One - Contract *Plano, TX*

- Modernized company infrastructure by leading a critical migration project to transition from a costly, outdated mainframe database
- Took ownership of migrating 3 billion records to a modern cloud-based solution (DynamoDB) and redesigning system architecture to improve efficiency, resiliency, scalability, and performance
- Executed the migration using Python, Apache Spark, and AWS EMR while upgrading Java Spring Boot components and integrating new Kafka streams and APIs
- Completed the migration 50% faster than the original estimate (15 days vs. 30 days), improved response times by 97% during peak hours, and eliminated data loss

Reseach Assistant July 2020 - June 2021
CSE Dept. At UTA *Arlington, TX*

- Developed core functionalities for an online learning platform using Angular, TypeScript, and RESTful APIs
- Designed database schema and developed essential CRUD operations for multiple modules, improving data management and platform functionality
- Improved site responsiveness by 20% through the implementation of lazy loading techniques
- Instructed students in Data Structures and Algorithms, enhancing their understanding and application of these critical concepts

SKILLS

Languages	Java, Python, SQL, JavaScript
Frameworks	Java Spring Boot, Apache Kafka, Apache Spark, Log4j, JDBC, REST, JUnit, Mockito
Cloud and Tools	AWS EMR, EC2, DynamoDB, S3, Snowflake, OneLake, Splunk, Git, CI/CD

EDUCATION

Master of Science in Computer Science Aug 2018 – May 2020
The University of Texas at Arlington *GPA 3.8*

Bachelor of Engineering in Information Technology June 2013 – May 2017
Pune University *GPA 3.4*

PROJECTS

Big Data Analysis. Analyzed over 4 million movie titles from the IMDB dataset using the Hadoop MapReduce framework and Java. Improved computation time by 15% through the use of in-mapper combiners and optimized mapper configurations. Deployed the project on a 2000-node cluster and conducted runtime analysis.

ETL Data Pipeline. Built an automated, scalable ETL data pipeline using Google Cloud Dataflow, Apache Beam, Pub/Sub, BigQuery, and Data Studio for big data analysis and visualization. Simulated real-time streaming and batch processing with Google Cloud templates, Apache Beam, and Pub/Sub. Automated data extraction and pipeline execution using scheduled cron jobs and Google Cloud Functions.

Graph processing. Accomplished efficient processing of over 100,000 graph vertices, achieving streamlined identification of connected components, by utilizing Spark framework in Scala and applying transformations like reduceByKey, flatMap, join, and map on AWS Elastic MapReduce (EMR), a cloud-native big data platform that simplifies running big data frameworks for processing and analyzing large datasets.