MUHAMMED ADEEM SHAIK

📍 1448 W Fillmore St, Chicago, IL - 60607   in linkedin.com/in/adeemshaik

✉ mshaik35@uic.edu 📱 609-509-9876   ⓖ github.com/adeemshaik

EDUCATION

Masters in Computer Science - University Of Illinois Chicago, Illinois   📅 Aug 2021 - May 2023
Courses: Alogrithms, Networking, Parallel & Distributed Systems, Cloud Computing, Information Retrieval

PG Diploma in Data Science - IIIT Bangalore, India   📅 Jan 2020 - Jan 2021
Courses: Machine Learning, NLP, Advanced SQL, Principles for Big Data.

Bachelors in Computer Applications - Bangalore University, India   📅 May 2016 - June 2019
Courses: OOPS Java, Data Structures, Software Engineering, Unix Programming, Operating Systems.

PROFESSIONAL EXPERIENCE

Software Engineer, Deloitte, Bengaluru, India   📅 August 2019 - July 2021

- **Tech Stack** - **Azure Databricks, Spark, Scala, Airflow, Python**
- Worked on an application to assist packaged consumer manufacturers in identifying components that impact sales, such as pricing, distribution, promotion, and advertising, and to compare their performance with competitors.
- Created a data-based framework to read 30 TB of data from multiple sales acquisition channels and analyzed sales activity for 5 years in various regions across the globe.
- Ordering UI (ideation, planning and coordination): Worked with the UI team to create a platform where the product leaders can set up jobs with required parameters and jobs run cadence.
- Actively worked on the orchestration of the pipeline where the application trigger points were reduced from two to a zero-touch application with the help of Apache Airflow, that reduced the cost of operations by 30 %.

- **Tech Stack** - **AWS S3, Redshift, Spark, Kafka, Airflow, Tableau, Python**
- Built ETL pipelines to extract, transform, and load large volumes of data from various sources into data warehouse enabling near-real-time data analysis and decision-making.
- Worked on implementing data models and schemas for optimal storage, retrieval, and analysis of data.
- Developed and maintained data pipelines for real-time data processing, using technologies like Kafka and Spark.
- Built automated data quality and validation checks to ensure data accuracy and consistency, reducing errors by 40%.
- Maintained and monitored production data systems to ensure high availability and security.

PROJECTS

Log Generator: Map Reduce   *AWS EMR, Apache Hadoop*
- Implemented distributed parallel processing in Hadoop Map Reduce to classify logs based on time interval and frequency, ensuring accurate and up-to-date log analysis.
- Source code is deployed as JAR file onto AWS Elastic Map Reduce Cluster and Python lambda function is used to analyze log files.

Log File Monitoring and Alert System   *AWS(EKS, EMR, SNS, ECR), Apache(Spark, Kafka)*
- Designed and Developed an Cloud Architecture which uses a log file generator application deployed onto an AWS Elastic Kubernetes cluster to write logs into AWS Elasticache Redis Database.
- Utilized the AKKA actor system's ability to continuously monitor the AWS Elasticache Redis Database and forward errors to a Kafka Topic.
- Implemented an AWS EMR Spark Cluster to stream logs from the Kafka Topic to alert stakeholders.

Aviation Data Pipeline   *GCP(GCS, Big Query, Data Studio, Wrangler), Python, Docker, Airflow*
- Built an end-end data engineering pipeline to analyze the aviation data of individual airline, airline sectors and industry as a whole. Created an ETL pipeline using Data Fusion(GCP)to visualize the data from airline API using Kafka and batch processing.
- Visualized the processed data in big query using Data studio to compare the performances of various airlines.

SKILLS

Tools&Frameworks:	GIT, SAP, Apache(Hadoop, Kafka, Spark, Airflow), Docker, Power BI
Databases&Cloud:	MySQL, Mongo DB, Redis, AWS (EMR, EKS, EC2), GCP
Programming:	Java, Python, JavaScript, SQL, HTML/CSS, PHP, Scala
Web Technologies:	HTML5, CSS, Bootstrap