MUHAMMED ADEEM SHAIK

📍620 Comstock St, Seattle, WA - 98109

✉shaikmuhammedadeem@gmail.com 📱 609-509-9876

in linkedin.com/in/adeemshaik

○ github.com/adeemshaik

PROFESSIONAL EXPERIENCE

Data Infra Engineer, Apple, Seattle 📅 May 2023 - Present
- **Tech Stack** - **AWS, Kubernetes, Spark, Ranger, Airflow, Solr, Hive, Flink, JupyterHub, Python**
- Contributed to the development of a hybrid-cloud, multi-tenant data platform, aiming to optimize Data Engineering, Machine Learning, and Analytical processes for enhanced scalability and efficiency across diverse teams.
- Enhanced core products such as Apple Maps, Search, and Siri by integrating our cutting-edge data platform, which utilizes the latest big data technologies, boosting their AI/ML functionalities.
- Assisted in the design and deployment of a high-capacity streaming ingestion platform, handling up to 1 trillion events daily, leveraging Apache Kafka and Spark Streaming within a Kubernetes environment.
- Contributed to the strategic implementation of security protocols for data handling, collaborating with tools like Spark, Ranger, and Solr on Kubernetes to fortify data integrity and optimize accessibility.

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 %.
- Got Two Spot Awards for the contribution to the team.

- **Tech Stack** - **AWS, 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%.
- Got One Spot Award for the contribution to the team.

PROJECTS

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 database and forward errors to a Kafka Topic and used an AWS EMR Spark Cluster to stream logs from the 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.

AzureFlow *Azure(DevOps), Terraform, Jenkins, Python, Bash*
- Developed an efficient Azure cloud infrastructure for large-scale research data management. Also, automated a CI/CD pipeline with Azure DevOps, Terraform, and Jenkins for optimized software builds, tests, and deployments.

SKILLS

Data Processing & Management Apache (Hadoop, Kafka, Spark, Airflow, Flink, Ranger), Trino, Docker, Kubernetes

Monitoring, Infrastructure & CI/CD Datadog, Jenkins, Splunk, Pulumi

Databases & Cloud MySQL, MongoDB, Redis, AWS, Azure, GCP

Programming Java, Python, JavaScript, SQL, HTML/CSS, PHP, Scala

Web Technologies HTML5, CSS, Bootstrap

EDUCATION

Masters in Computer Science - University Of Illinois Chicago, Illinois 📅 Aug 2021 - May 2023
PG Diploma in Data Science - IIIT Bangalore, India 📅 Jan 2020 - Jan 2021
Bachelors in Computer Applications - Bangalore University, India 📅 May 2016 - June 2019