SOHRAB AHMED

SENIOR SYSTEM ANALYST

SUMMARY

Over 5 years of profound experience in Big Data ecosystem components with a passion to identify competitive advantages by transforming raw data into powerful information with analytics-driven business insights.

EDUCATION

MS in Machine Learning & AI
        Liverpool John Moores University | 2022
B.Tech in Computer Science & Engineering
        Uttar Pradesh Technical University | 2014

CERTIFICATION

Microsoft Certified Trainer | Microsoft |2020
Data Engineer Associate | Microsoft | 2020
Developer Associate | Microsoft | 2020
Azure Data Fundatamentals | Microsoft | 2020
Azure Fundamentals | Microsoft | 2020

CONTACT

- ✉ sohrabahmed@outlook.com
- ☎ +60 1732679457
- in linkedin.com/in/sohrabahmed
- ⚲ Subang Jaya, Malaysia

CORE SKILLS

- Databricks, Scala, Spark, SQL, Logic App
- Azure, HDFS, Amazon S3
- File Format: Delta, ORC, Avro, Parquet, Sequence
- Azure Data Factory, Talend, Azkaban, Airflow
- Shell Script, Azure CLI, Sqoop
- Hive, SQL-Server, MySQL, Amazon Redshift
- Azure DevOps, Git, AAS

EXPERIENCE

Senior System Analyst

PPG Industries | Subang Jaya | Nov 2019 - present

- Currently Designing Delta Lake using Databricks
- Creating modern data architecture with Azure Data Factory, Azure Databricks, Azure Synapse Analytics, and Power BI
- Implemented Azure Databricks CI CD with Azure DevOps
- Productionized the pipeline stack in Talend, Logic App and Azure Data Factory v2
- Developed Logic App for AAS Model and Power BI Dataset refresh

Big Data Analyst

Neustar Inc | Hyderabad | Dec 2018 - Oct 2019

- Automated the data flow and fill rate generation
- Moved business logic from on-premise to AWS
- Designed Data availability dashboard
- Moved repo from SVN to GIT
- Enhancement of on-premise Hadoop and SQL Server systems
- Implemented compression and data formats for optimal data storage and fast access

Data Engineer

Sphota biz | Bengaluru | Apr 2018 - Dec 2018

- Worked on designing and implementing an automated data ingestion framework for seamless flow of data from source to data lake with change data
- Productionized the entire pipeline stack including job scheduling in Azkaban, monitoring in Power-BI and alerting in Slack
- Implemented low latency data query platform for ad-hoc access of clickstream data
- Worked on designing storage structure for data lake
- Implemented GDPR
- Built the business solution in Cloud (AWS, Qubole), enhancement of on-premise Hadoop and Teradata systems
- Worked on analyzing different compression and data formats for optimal data storage and fast access in the data lake

Programmer Analyst

Cognizant | Gurugram | Jun 2015 - Mar 2018

- Collected and analyzed large-scale data sets using various frameworks
- Moved data between HDFS and RDBMS using Sqoop
- Developed Java UDFs for Hive