Mohammad Rafi Shaik

52 Guild st, Boston, MA 02119 • +16176859445 • raafy.cv@gmail.com • https://www.linkedin.com/in/raafy

Education

MS, Data Analytics Engineering

Northeastern University, Boston, MA, USA   Expected Apr 2023

Relevant Courses: Machine Learning, Data Mining, Database Design and Visualizations, Probability & Statistics

B.Tech, Computer Science and Engineering

Sri Venkateswara University, Tirupati, AP, India   May 2018

Relevant Courses: Data Structures, Algorithms, Operating Systems, Artificial Intelligence, Cloud Computing

Technical Skills

Analytical Tools: Apache Airflow, SSIS, Tableau, SSRS

Languages/Scripting: Python, SQL, R, UNIX

Technologies & Techniques: Machine Learning, Text Classification, Text Summarization, Time Series Analysis

Databases: Snowflake, MS SQL Server

Big Data: Kafka, Ni-Fi, HDFS, Hive

Certifications: Apache Airflow x2, Snowflake, Microsoft Azure DP-200

Professional Experience

Data Engineer   May 2018 - Jul 2021

ACS Solutions, Hyderabad, India

- Built integrations for different Ad Platforms such as Facebook, Google AdWords, Pinterest, Impact Radius, Smartly, Snapchat and Email partners with Snowflake using Python and Apache Airflow
- Structured more than 130 data pipelines hosted on Apache Airflow dealing with 30+ different API's
- Developed and deployed python integrations in AWS using Docker, AWS Batch, EC2, AWS ECR and Lambda
- Expertise in Snowflake cloud data warehousing and its unique architecture and features like zero-copy cloning, snowpipe, data sharing, time travel, streams, tasks and virtual warehouses
- Created complex SQL procedures and SSIS packages using proper control and data flow elements with error handling
- Engaged in all phases of Data practice tasks Data Integration, Data warehousing, Reporting and ETL

Projects

Student Job Tracking Database application, Northeastern University   Sep 2021 - Nov 2021

- Designed and developed normalized application database for university student job tracking platform
- Implemented various database features triggers, encryption, data masking, data validation, row level security, indexes, stored procedures for complex functionality and PowerBI dashboards & reports

Accident Data Analysis in Chicago, Northeastern University   Sep 2021 - Nov 2021

- Exhibited various steps of data analysis including Data exploration, Data Cleaning, Data Wrangling, Data Clustering, Time Series Analysis, and Trend Analysis
- Implemented feature engineering to identify important features using statistic techniques and PCA

Streaming application - Kafka & NiFi, ACS Solutions   Jun 2019 - Nov 2019

- Developed a streaming application for pushing the CDC (Change Data Capture) data from SQL Server to HDFS and then to Hive leveraging Kafka topics and NiFi flows

"Sociopinion" - Twitter Data Analysis, SV University   Sep 2017 - Apr 2018

- Led team of two in design and development of Sociopinion app, is dynamic voice of society today
- Visualized current trend and gain statistical overview of products by comparing and contrasting social opinion today in twitter