

# Zaid Chaudhary

716 547 1188 | [mohdzaid@buffalo.edu](mailto:mohdzaid@buffalo.edu) | [www.linkedin.com/in/zaidchaudhary/](http://www.linkedin.com/in/zaidchaudhary/) | [Github](#)

## Education

### University at Buffalo, The State University of New York

**Master of Science, Management Information Systems** (STEM Designated), GPA: 4.0/4.0

Jun. 2024

**Relevant Courses:** Health Care Analytics, Data Warehousing, Database Management Systems, Predictive Analytics, Big Data & Distributed Computing, Technology Consulting for Social Impact.

### University at Buffalo, The State University of New York

**Master of Science: Data Science (STEM Designated)**, GPA: 3.5/4.0

Feb. 2023

## Skills

**Applications:** ,Microsoft .Net Integration,SAAS, ,SSIS, SAP Cloud Platform Integration, Salesforce,AWS,Azure,Excel

**Language:** SAP ABAP, SQL, C, Java, Python, HTML, XML, JSON, Groovy, R,Pyspark,C#,ASP, Visual Studio

**ETL:** Apave Kafka, Apache Spark, Redshift JDBC/ODBC drivers, Azure Synapse,Azure Data Factory,Hadoop,MS SQL server

**Tools:** Tableau, Power BI, Talend, Tableau Prep, Eclipse, NWDS, Postman, SoapUI, ServiceNow, ADO, Figma,MATLAB,Scala

**Data Pipelines:** Airflow,AWS Step Functions,Amazon Redshift,Apache NiFi,Snowflake,AWS Lambda, AWS Event Bridge,GCP

## Experience

### University at Buffalo

Buffalo, New York

#### Graduate Teaching Assistant, Data Intensive Computing

Aug. 2023-Dec. 2023

- **Led a team of 4** Teaching Assistants and personally engaged with around **200+ students** to address their queries by taking weekly office hours.
- Graded assignments, exams and projects providing constructive feedback for improvement.

## Academic Projects

### Realtime Socket Streaming using Apache Spark

Aug. 2023-Oct. 2023

- Developed a real-time sentiment analysis system, processing 7M+ records via Apache Spark and LLM, with data streaming through Kafka, indexed in Elasticsearch.
- Automated sentiment analysis of 1,000 daily Yelp reviews using Spark and developed a Kibana dashboard for real-time sentiment visualization, categorizing insights and ranking top restaurants to inform and enhance retail strategies.

### Azure-Based ETL Pipeline for OTT Platform Data

Mar. 2023-Apr. 2023

- Developed an ETL pipeline for OTT data on Azure, processing Kaggle datasets with automated updates, achieving data ingestion and security improvements.
- Optimized data transformation with Azure Synapse and Spark, enhancing processing speed by 20% and storage efficiency with Parquet format.
- Developed a PowerBI dashboard for OTT analytics with four KPIs and streamlined code deployment via Azure DevOps CI/CD, boosting deployment efficiency and stakeholder engagement by 20%.

### AWS-Based ETL Pipeline for Reddit Data

Sep. 2021-Dec. 2021

- Developed an automated ETL pipeline using Apache Airflow, PRAW, and AWS services, processing 100+ top posts daily from Reddit's "MachineLearning" subreddit for analytics.
- Implemented data validation and de-duplication workflows with AWS Glue, enhancing data integrity for real-time Tableau visualizations from Amazon Redshift.
- Managed version control of ETL scripts and workflows on GitHub, facilitating efficient collaboration and updates within the data engineering team.

### ETL pipeline with Talend Open Studio and Oracle Cloud Platform

Sep. 2023-Oct. 2023

- Orchestrated the development of an efficient ETL pipeline via Talend and Oracle Cloud, streamlining and securing data flows.
- Spearheaded the resolution of Slowly Changing Dimensions in Talend, safeguarding the precision of data transformations and the preservation of historical integrity.
- Crafted a dynamic Tableau dashboard to dissect key business indicators, bolstering strategic decisions in marketing and production planning.

## Certifications

**Microsoft Certified Azure Data Engineer Associate**, Certification by Microsoft

Dec. 2023

**Tableau Desktop Specialist**, Certification by Tableau

Feb. 2023

**AWS Certified Cloud Practitioner**, Certification by AWS

Jun. 2023

**Certified Scrum Master** Certification by Scrum.org

Apr. 2023

**Product Management Basics Certification**, Certification by Pendo.io

Sep. 2023

**Engineer Data in Google Cloud**, Certification by Google

Sep. 2023