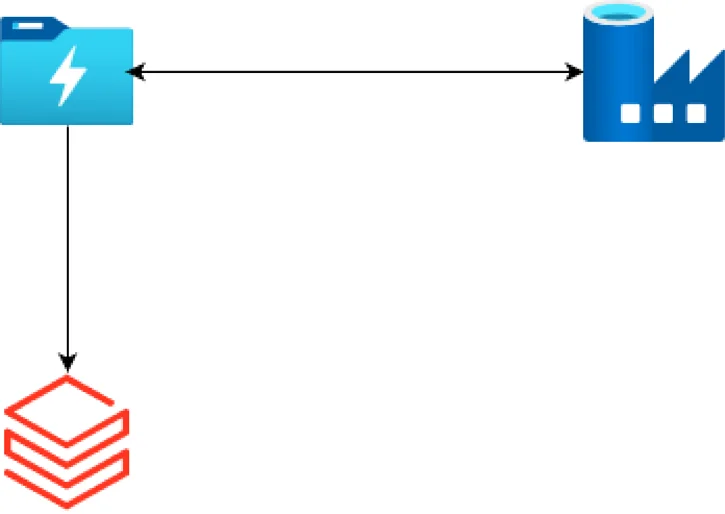
1.Movielens Data Analysis for Recommendations

Recommender System is a system that seeks to predict or filter preferences according to the user's choices. Recommender systems are utilized in various areas, including movies, music, news, books, research articles, search queries, social tags, and products in general. This use case focuses on the movie recommendation system used by top streaming services like Netflix, Amazon Prime, Hulu, Hotstar, etc, to recommend movies to their users based on historical viewing patterns. Before the final recommendation is made, a complex data pipeline brings data from many sources to the recommendation engin



in this project, you will explore the usage of Databricks Spark on Azure with Spark SQL and build this data pipeline.

* Download the dataset from GroupLens Research, which is a research group in the Department of Computer Science and Engineering at the University of Minnesota.
* Upload it to Azure Data lake storage manually .
* Create a Data Factory pipeline to ingest files. Then you use databricks to analyze the dataset for user recommendation.

2.Analyse Yelp Dataset with Spark & Parquet Format on Azure Databricks

In this Databricks Azure project, you will use Spark & Parquet file formats to analyse the Yelp reviews dataset. As part of this you will deploy Azure data factory, data pipelines and visualise the analysis.

https://www.yelp.com/dataset