Movie Recommendation System Using Hadoop and Mahout

The Movie Recommendation System is a Hadoop-based tool that leverages Apache Mahout to provide personalized movie suggestions based on user preferences and viewing history. It allows streaming platforms, entertainment services, or academic demos to implement collaborative filtering at scale — enabling better user engagement and viewing experience.

Key Features:

- Recommend movies to users based on viewing history and ratings
- Use collaborative filtering techniques (user-based or item-based)
- Input: User ID, Movie ID, Rating, Timestamp
- Process massive datasets using Hadoop for scalability
- Generate top-N recommendations for each user
- Train and evaluate the model using Mahout's built-in algorithms
- Export results for integration with dashboards or apps

Libraries & Tools Used:

- **Apache Hadoop** – for distributed data storage and processing
- **Apache Mahout** – for building machine learning recommendation models
- **HDFS** – to store large datasets like movie ratings and metadata
- **MapReduce** – for pre-processing and filtering large-scale data
- **Java or Scala** – to implement and customize Mahout pipelines
- **MySQL / Hive** (optional) – for storing and querying recommendation results
- **Visualization tools** (optional) – for result presentation (e.g., Power BI, Tableau)

24M11MC278
Tula Sunil
Aditya University