Exercise H – 5

AIM

Product Recommender System using Amazon Review dataset

PROCEDURE

Recommender systems help customers by suggesting probable list of products from which they can easily select the right one. They make customers aware of new and/or similar products available for purchase by providing comparable costs, features, delivery times etc.
Recommender systems can be built with two different methods: Content Based Filtering, and Collaborative Filtering.

Content Based Filtering

In content-based filtering, the similarity between different products is calculated on the basis of the attributes of the products. For instance, in a content-based book recommender system, the similarity between the books is calculated on the basis of genres, the author of the book, the publisher of the book, title of the book etc.

Collaborative filtering

Collaborative filtering is commonly used for recommender systems. These techniques aim to fill in the missing entries of a user-item association matrix. Collaborative Filtering (CF) approach is used.
CF is based on the idea that the best recommendations come from people who have similar tastes. In other words, it uses historical item ratings of like-minded people to predict how someone would rate an item. Collaborative filtering has two sub-categories that are generally called memory based and model-based approaches.

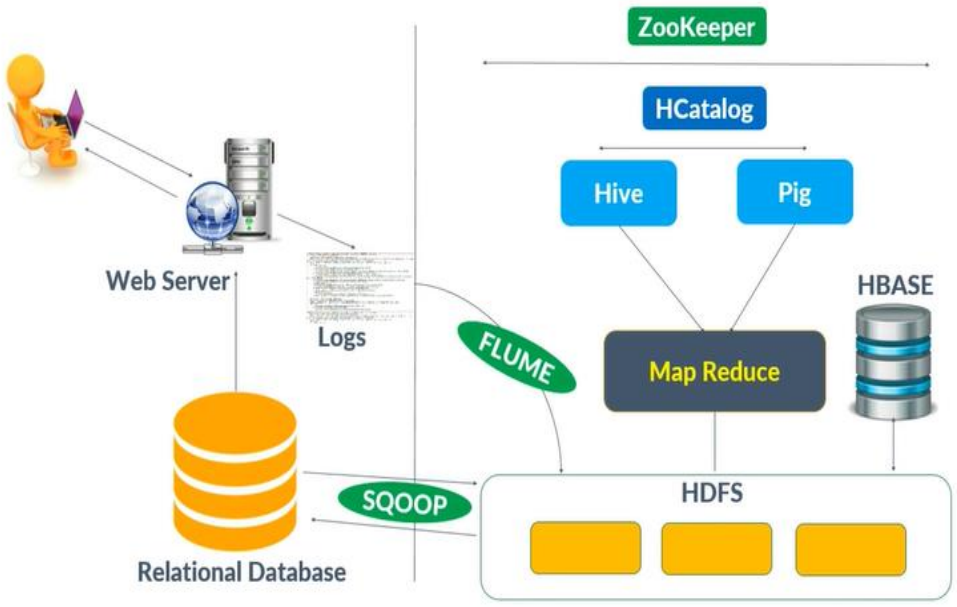
Memory-based

There are two approaches: the first one identifies clusters of users and utilizes the interactions of one specific user to predict the interactions of other similar users. The second approach identifies clusters of items that have been rated by user A and utilizes them to predict the interaction of user A with a different but similar item B. These methods usually encounter major problems with large sparse matrices, since the number of user-item interactions can be too low for generating high quality clusters.

Model-based

These methods are based on machine learning and data mining techniques. The goal is to train models to be able to make predictions. For example, we could use existing user-item interactions to train a model to predict the top-5 items that a user might like the most. One advantage of these methods is that they are able to recommend a larger number of items to a larger number of users, compared to other methods like memory based approach. They have large coverage, even when working with large sparse matrices.

ALGORITHM



RESULT

The above methods are effectively used in Product Recommender using Amazon Review dataset.

OUTPUT