Assignment 10: Implementation of Recommendation System

Aim:
Implement any machine learning techniques using available dataset to develop a recommendation system.

LO4:- To learn various machine learning techniques to solve complex realworld problems

Theory:
A recommendation system is a type of information filtering system that predicts the preferences or ratings a user would give to a particular item. It suggests items to users based on their past behaviors, preferences, or similarities to other users. There are mainly three types of recommendation systems:

-Content-Based Filtering: This type of recommendation system suggests items similar to those the user has liked in the past. It analyzes the content of the items (e.g., movie descriptions, product features) and recommends items with similar content.

- Collaborative Filtering: Collaborative filtering recommends items by leveraging the preferences and behaviors of other users. It identifies users with similar preferences and recommends items liked by those users but not yet seen by the target user.

-Hybrid Recommendation Systems: Hybrid recommendation systems combine multiple recommendation techniques to provide more accurate and diverse recommendations. They can incorporate content-based filtering, collaborative filtering, and other approaches to overcome the limitations of individual methods.

How it works:
- Content-Based Filtering: This method analyzes the attributes of items and the user's past interactions to recommend similar items. For example, in a movie recommendation system, if a user has liked action movies in the past, the system will recommend other action movies based on their similarity in terms of genre, actors, directors, etc.

- Collaborative Filtering: Collaborative filtering works by identifying patterns in user behavior. It recommends items that similar users have liked in the past but the target user hasn't interacted with yet. It can be user-based or item-based. User-based collaborative filtering identifies similar users based on their past interactions and recommends items liked by those similar users. Item-based collaborative filtering recommends items similar to those the user has liked in the past.

Advantages and disadvantages of each type of recommendation system:

- Content-Based Filtering:
- Advantages: Can recommend new and niche items, does not require a large user base, transparent reasoning for recommendations.
- Disadvantages: Limited to recommending items with similar attributes, struggles with recommending diverse items, relies heavily on the quality of item descriptions.

- Collaborative Filtering:
- Advantages: Can recommend unexpected items based on user behavior, does not rely on item attributes, effective for large datasets.
- Disadvantages: Cold start problem for new users or items, requires a large user base to generate accurate recommendations, susceptible to popularity bias.

Cosine similarity is a measure used to determine how similar two vectors are in a multi-dimensional space. It calculates the cosine of the angle between two vectors, providing a measure of similarity irrespective of their magnitude. In the context of recommendation systems and natural language processing, cosine similarity is often used to measure the similarity between documents, texts, or items based on their features or attributes.

Mathematically, the cosine similarity between two vectors A and B is calculated using the following formula:

$$\text{cosine similarity} = \frac{A \cdot B}{\|A\| \|B\|}$$

Where:
- $A \cdot B$ represents the dot product of vectors A and B,
- $\|A\|$ represents the Euclidean norm (magnitude) of vector A, - $\|B\|$ represents the Euclidean norm (magnitude) of vector B.

Cosine similarity values range from -1 to 1:
- 1 indicates that the two vectors are perfectly similar,
- 0 indicates no similarity (orthogonal vectors), - -1 indicates perfect dissimilarity (opposite directions).

Conclusion:
We have used the concept of Cosine Similarity to build a content based recommendation system, which suggests movies related to what the user searches for.