Recommendation System

Interview Questions:

1. Can you explain the difference between user-based and item-based collaborative filtering?

**Collaborative Filtering (CF)** is a technique used in recommendation systems to predict a user’s interests based on the preferences of similar users or items.

* **User-Based Collaborative Filtering (UBCF):**
  + **How It Works:** It recommends items to a user based on the preferences of similar users.
  + **Example:** If User A likes items X, Y, and Z, and User B has similar preferences, User B will be recommended items that User A likes but User B hasn’t tried yet.
  + **Strengths:**
    - Easy to understand and implement.
    - Effective when there are many active users with diverse preferences.
  + **Limitations:**
    - Struggles with the **cold start problem** for new users (since it needs historical data to find similarities).
    - Not scalable with large datasets because it requires comparing all users.
* **Item-Based Collaborative Filtering (IBCF):**
  + **How It Works:** It recommends items similar to what the user has already liked. It focuses on the similarity between items rather than users.
  + **Example:** If a user likes item X, and item Y is similar to X (based on user interactions), the user will be recommended item Y.
  + **Strengths:**
    - More scalable and performs better with a large number of users.
    - Handles the cold start problem for users better because items are often more stable over time.
  + **Limitations:**
    - May not capture the diversity of user preferences as effectively as user-based CF.

1. What is collaborative filtering, and how does it work?

**Collaborative Filtering** is a method used to make recommendations based on the past behaviors of users (or items) and the assumption that if two users (or items) agree in the past, they will agree in the future.

**How It Works:**

1. **Data Collection:** Gather data on user-item interactions (e.g., ratings, clicks, purchases).
2. **Similarity Calculation:**
   * In **user-based CF**, calculate similarity between users (e.g., using cosine similarity or Pearson correlation).
   * In **item-based CF**, calculate similarity between items based on how users have interacted with them.
3. **Prediction:**
   * For a given user, predict the rating or preference for an item based on the ratings from similar users (user-based) or based on the similarity of items the user has already interacted with (item-based).
4. **Recommendation:** Provide items with the highest predicted scores to the user.

**Types of Collaborative Filtering:**

* **Memory-Based CF:** Uses historical data to make recommendations (user-user or item-item).
* **Model-Based CF:** Uses machine learning models (like matrix factorization or neural networks) to predict user preferences.