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

* **User-Based Collaborative Filtering (UBCF):**
  + **Idea:** Finds similar users based on their past interactions (e.g., ratings, purchases) and recommends items liked by similar users.
  + **How it works:** If two users have similar preferences or behaviour patterns, items liked by one user can be recommended to the other.
  + **Example:** If User A and User B both like Movies X and Y, and User A also likes Movie Z, then Movie Z will be recommended to User B.
  + **Challenge:** This method may struggle with scaling when there are many users or sparse data.
* **Item-Based Collaborative Filtering (IBCF):**
  + **Idea:** Finds similar items based on user interactions and recommends items that are similar to those the user has already interacted with.
  + **How it works:** If a user has liked or rated certain items, the system identifies other items that are most similar to those. This is based on how users generally perceive those items.
  + **Example:** If a user likes Movie X and Movie Y, and Movie Z is highly rated by people who liked both X and Y, then Movie Z is recommended to the user.
  + **Challenge:** It requires identifying item similarities, which may not be straightforward for new items.
* **Key Differences:**
  + UBCF focuses on **user similarity**, while IBCF focuses on **item similarity**.
  + UBCF is better for scenarios where there is ample data on user behaviour, whereas IBCF is more robust when there are many items but fewer user interactions.

**2. What is Collaborative Filtering, and How Does it Work?**

* **Collaborative Filtering:** It’s a recommendation algorithm that predicts a user’s interest by collecting preferences from many users. The underlying principle is that if two users have similar preferences in the past, they will likely have similar tastes in the future.
* **How it Works:**
  + Collaborative filtering creates recommendations based on historical user-item interactions such as purchases, ratings, or clicks.
  + The system leverages patterns or correlations between users and/or items to make predictions about what a user might like.

There are two main types:

* + **User-based collaborative filtering** (described above)
  + **Item-based collaborative filtering** (described above)
* **Common Techniques:**
  + **Matrix Factorization:** Decomposes the user-item interaction matrix into two smaller matrices to learn latent factors.
  + **K-Nearest Neighbors (KNN):** Measures the similarity between users or items using methods like cosine similarity or Pearson correlation.
* **Challenges:**
  + **Cold Start Problem:** Difficulty making recommendations for new users or new items without prior data.
  + **Sparsity:** When user-item interaction data is sparse (many users haven’t interacted with many items), it’s harder to find patterns.