**1- Recommender Systems (RS) are based on one of two strategies – content filtering and collaborative filtering. Write down key features and properties of each of them.**

The content filtering approach is based on creating a profile for each user or product to characterize its nature. For example, for a movie profile, it consists of some attributes like its genre, the participating actors, its box office popularity and etc. for user profiles, there is a proper questionnaire in which there are demographic information or answers. These profiles allow programs to associate users with matching products. Unfortunately, content-based strategies require gathering external information that might not be available or easy to collect.

An alternative to content filtering relies only on past user behavior, for instance, previous transactions or product ratings. It doesn’t need the creation of explicit profiles. This approach is called collaborative filtering. In fact, collaborative filtering analyzes relationships between users and interdependencies among products to identify new user-item associations.

The attractive thing related to collaborative filtering is being domain free. Generally collaborative filtering is more accurate than content-based techniques, but it suffers from what is called the cold start problem. It’s because of its inability to address the system’s new products and users. In this aspect content filtering is superior.

**2- Two primary areas of collaborative filtering are neighborhood methods and latent factor models. Write down key features of each of them.**

Neighborhood methods are based on computing the relationships between items or, alternatively, between users. The item-oriented approach evaluates a user’s preference for an item based on ratings of “neighboring” items by the same user. A product’s neighbors are other products that tend to get similar ratings when rated by the same user. We need to know that a product’s neighbors are other products that tend to get similar ratings when rated by the same user.

In user-oriented approach, it identifies like-minded users who can complement each other’s ratings.

Latent factor models are an alternative approach that tries to explain the ratings by characterizing both items and users on, like 20 to 100 factors comprise a computerized alternative to the aforementioned human-created song genes.

**3- What do the “factors” in latent factor represent? How are these factors discovered?**

**Ans**) In the Latent factor models, ‘factors’ comprise a computerized alternative to the aforementioned human-created song genes. For movies, the factors that are discovered might measure dimensions such as comedy vs drama, amount of action, and dimensions like depth of character. For users, each discovered measures how much the user likes movies that score high on corresponding movie factor.

**4)What is the difference between explicit and implicit feedback for RS? What do you think are advantages and disadvantages of each?**

**Ans**) The ‘Explicit feedback’ includes explicit input by users that are made regarding their interests in products. For example, movie ratings by user. The matrix made from Explicit feedback is a sparse matrix, since single user will not rate all the possible items.

When explicit feedback is not available, ‘Implicit feedback’ come into play. Recommend systems can infer user preference using implicit feedback, which reflects opinion by considering the user behaviour including purchase history, browsing history and search patterns. This feedback is represented by dense matrix. For explicit feedback, It requires more involvement from the user which increases the cognitive load on the user. A user does not perceive some benefit from rating he might cease evaluating items or even leave the system and a numerical value may not be a good representation for the user’s perception of an item.

In implicit feedback even if it convenient for user, but more difficult to implement. For example in Amazon.com, When the entry of a book is displayed other book titles are shown which were purchased by customers who also bought the selected book. The time that a user spends on reading an article could also be considered as implicit feedback. The problem with this approach is that it is not possible to determine whether the user is actually reading the article or has taken a break.

**5- In the latent factor model for dimension f, the item i and user u are represented as vectors and . How is the interaction between item i and user u calculated?**

Matrix factorization models map both users and items to a joint latent factor space of dimensionality , such that user-item interactions are modeled as inner products in that space. For a given item , the elements of measure the extent to which the item possesses those factors, positive or negative. For a given user u, the elements of measure the extent of interest the user has in items that are high on the corresponding factors, again, positive or negative. The resulting dot product, , captures the interaction between user u and item I, which means the user’s overall interest in the item’s characteristics. This approximates user u’s rating of item I, which is denoted by , leading to this estimate.

The major challenge is computing the mapping of each item and user to factor vectors . After the recommender system completes this mapping, it can easily estimate the rating a user will give to any item by using equation .

**6- What does the learning system try to minimize? Understand the meaning of each term in the equation.**

To learn the factor vectors (), the system minimizes the regularized squared error on the set of known ratings. In the equation, is the set of the pairs for which is known the training set.

The system learns the model by fitting the previously observed ratings. However, the goal is to generalize those previous ratings in a way that predicts future, unknown ratings. Thus, the system should avoid overfitting the observed data by regularizing the learned parameters, whose magnitudes are penalized. The constant controls the extent of regularization and is usually determined by cross-validation.

**7- There are 2 learning algorithms for latent factorization – stochastic gradient descent (SGD) and alternating least squares (ALS). What are the advantages of ALS over SGD?**

Because both and are unknowns, Equation 2 is not convex. But, if we fix one of the unknowns, the optimization problem becomes quadratic and can be solved optimally. Thus, ALS techniques rotate between fixing the ’s and fixing the ’s. When all ’s are fixed, the system recomputes the ’s by solving a least-square problem, and vice versa. This ensures that each step decreases Equation 2 until convergence.

While in general stochastic gradient descent is easier and faster than ALS, ALS is favorable in at least two cases. The first one is when the system can use parallelization. In ALS, the system computes each independently of the other item factors and computes each independently of the other user factors. This gives rise to potentially massive parallelization of the algorithm. The second case is for systems centered on implicit data. Because the training set cannot be considered sparse, looping over each single training case, as gradient descent does, would not be practical. ALS can efficiently handle such cases.

**8- Read about the Netflix competition and the authors’ entry. What were the most descriptive dimensions (features) that their models discovered? Summarize briefly. Also mention what metric do they use to check the performance of their models.**

The metric that is used in the Netflix competition for finding out the performance of the model is Root Mean Square Error (RMSE). The model that improve the algorithm RSME performance by 10 percent is considered as a winner.

The most descriptive dimensions that the models discovered two factors from Netflix matrix factorization. The first factor vector has on one side lowbrow comedy and horror movies, aimed at male or adolescent, while other side contains drama or comedy with serious undertones and strong female leads. The second factorization axis has independent, critical acclaimed, quirky films on the top and on the bottom, mainstream formulaic films.