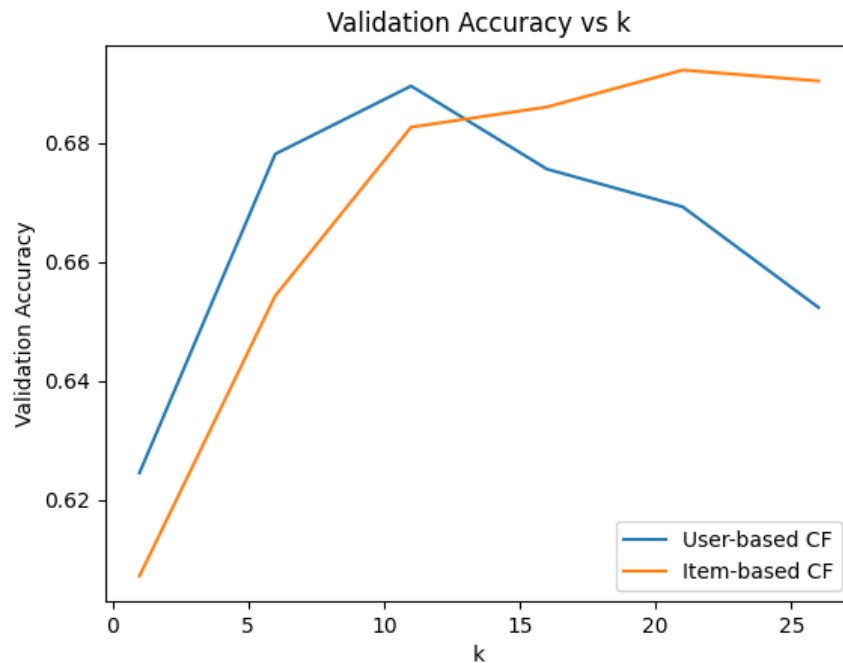


Question 1

- (a) (b) (c) The accuracy on the validation data with $k \in \{1, 6, 11, 16, 21, 26\}$ on user-based and item-based collaborative filtering is as follows:



Test Accuracy on user-based CF with $k^* = 11$: 0.6841659610499576

Test Accuracy on item-based CF with $k^* = 21$: 0.6816257408975445

- (d) The test on user-based CF is slightly better than item-based CF.
 Additionally, the test accuracy on user-based CF cost less time than item-based CF.
 Therefore, user-based CF is better than item-based CF in this case.
- (e)
- The KNN algorithm is computational expensive for large datasets.
 - The Curse of Dimensionality: In high dimensions, “most” points are approximately the same distance and the nearest neighbors are not very useful.