CSE427s - 6 Lab6 (RS2)   78% (7/9)

✓ 1. Select all that applies to a **Top-N list** recommendation approach. From the point of view of the user that gets the recommendataion, Top-N-list recommendations are

(A) based on the user's preferences

(B) the same for every user

(C) similar to the item the user is currenbtly interested in

(D) niche items

✓ 2. Select all that applies to a **Frequently Bought Together** recommendation approach. From the point of view of the user that gets the recommendataion, FBT recommendations are

(A) based on the user's preferences

(B) the same for every user

(C) similar to the item the user is currenbtly interested in

(D) niche items

⊘ 3. Job 1: Provide the **final output** of job 1 for user2 for the provided example data.

Use ( ) for key-value pairs and [ ] for lists.

(user2,[item2,item3])

✓ 4. Job 1: The Reducer of the first job...

(A) ...is the IdentityReducer

(B) ...is not required (we can do a map-only job)

(C) ...passes through the key and list-of-values

✓ 5. Job 2 (*pairs*): What is the **Mapper output** for the following Mapper input (user2, [item3,item2]).

(A) ([item3, item2], 1)

(B) ([item2, item3], 1)

(C) ([item3, item2], 1)
      ([item2, item3], 1)

(D) Both A and B are correct.

(E) None of the above.

✗ 6. **Job 2 (pairs): From the list below, identify key-value paris that appear in the entire job's output.**

(A) ([item1 ,item2], 1)

(B) ([item1 ,item2], 3)

(C) ([item2 ,item4], 1)

(D) ([item4 ,item2], 2)

(E) ([item5 ,item6], 3)

✓ 7. **Job 2 (*stripes*): Identify correct Mapper output key-value pairs for all items bought by user1 in the provided example data using the stripes approach.**

(A) (item1, {(item1, 1), (item2, 1), (item3 ,1)})

(B) (item1, {(item2, 1), (item3 ,1)})

(C) (item2, {(item1 ,1),(item3 ,1)})

(D) (item2, {(item3 ,1)})

(E) (item3, _)

✗ 8. **Job 2 (*stripes*): How many integers need to be buffered in the reduce() function that receives the input for item1 as key? We do not use a Combiner.**

(A) 3

(B) 5

(C) 7

(D) 9

(E) 11

(F) none of the above

✓ 9. **Which approach (pairs or stripes) is more efficient w.r.t. communication cost?**

(A) Pairs

(B) Stripes

(C) Same efficieny for both.

✓ 10. **Which approach (pairs or stripes) scales better to Big data w.r.t. memory usage in the compute nodes? Consider the edge case that your job will actullay fail due to an out of memory error.**

(A) Pairs

(B) Stripes

(C) Same efficiency for both.