

Given 3 users with ratings...

U1: 1 3

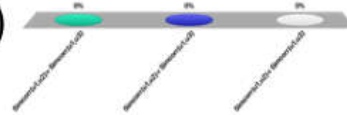
U2: 2 4

U3: 1 4

A. $\text{Sim}_{\text{corr}}(u1, u2) > \text{Sim}_{\text{corr}}(u1, u3)$

B. $\text{Sim}_{\text{corr}}(u1, u2) = \text{Sim}_{\text{corr}}(u1, u3)$

C. $\text{Sim}_{\text{corr}}(u1, u2) < \text{Sim}_{\text{corr}}(u1, u3)$

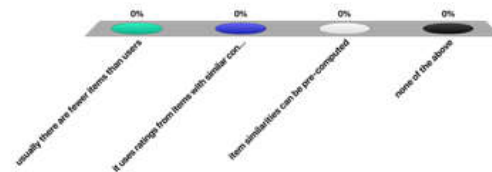


- we take the relative behaviour w.r.t. the mean

Item-based collaborative filtering addresses better the cold-start problem because ...

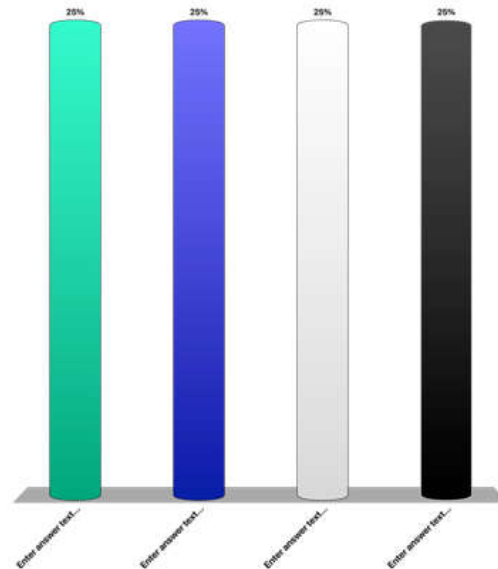
- A. usually there are fewer items than users
- B. it uses ratings from items with similar content
- C. item similarities can be pre-computed

D. none of the above



For a user that has not done any ratings,
which method can make a prediction?

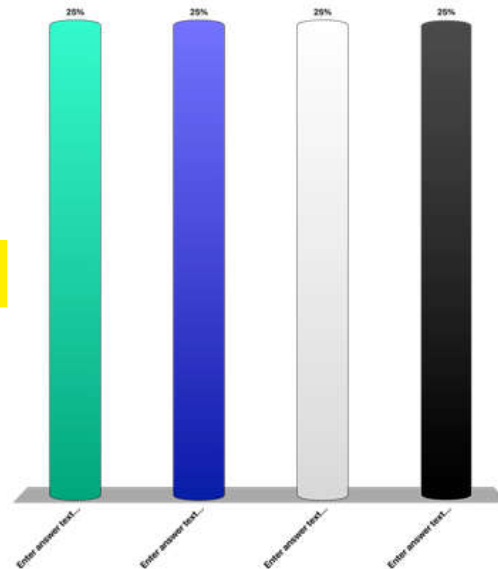
- A. User-based collaborative RS
- B. Item-based collaborative RS
- C. Content-based RS
- D. None of the above**



- even without any ratings, we can compare user profiles based on some external information available for them

For an item that has not received any ratings,
which method can make a prediction?

- A. User-based collaborative RS
- B. Item-based collaborative RS
- C. Content-based RS**
- D. None of the above



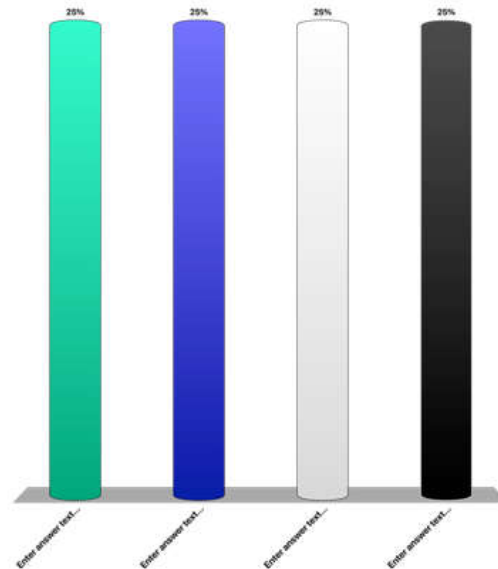
Which of the following techniques do you believe would be most appropriate to identify communities on a social graph?

A. Cliques

B. Random Walks

C. Shortest Paths

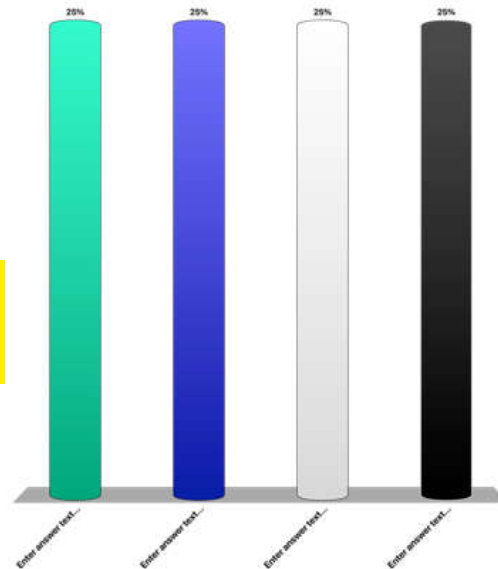
D. Association rules



* clique = subset of vertices of an undirected graph such that every two distinct vertices in the clique are adjacent; that is, its induced subgraph is complete.

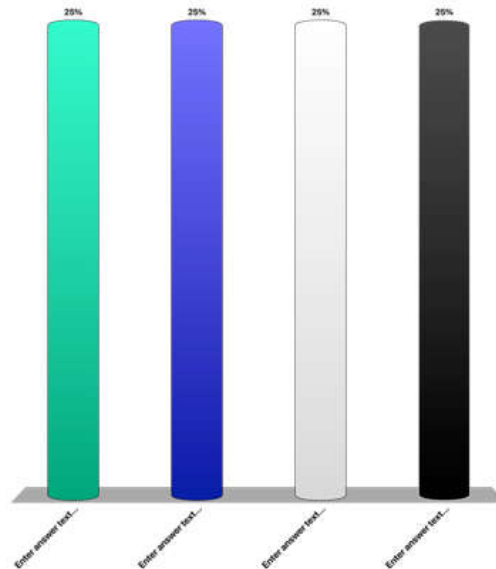
Modularity clustering will end up always with a single community at the top level?

- A. true
- B. Only for dense graphs
- C. Only for connected graphs
- D. never



Modularity clustering will end up always with the same community structure?

- A. true
- B. Only for connected graphs
- C. Only for cliques
- D. false**



- the order of the processing of the nodes might result in different communities
- there might be ties, so a node can choose to be in any of the multiple possible communities
- for cliques it will always produce the same result, maybe for some other structures as well, but in general it is not a stable clustering technique

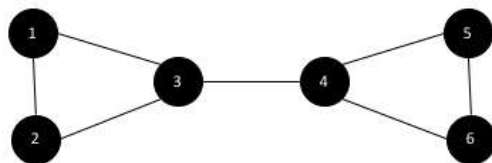
Betweenness of edge 3-4 is ...

A. 16

B. 12

C. 9

D. 4



© 2018, Karl Aberer, EPFL-IC, Laboratoire de systèmes d'informations répartis

Applied Classification - 64

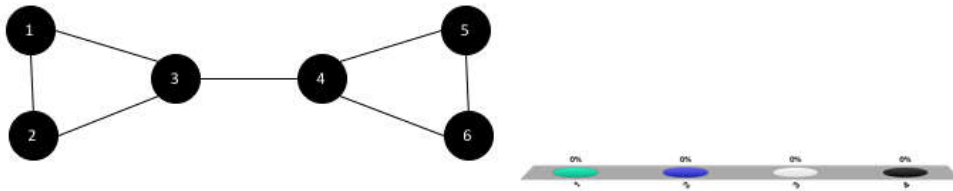
When computing path counts for node 1 with BFS, the count at 6 is ...

A. 1

B. 2

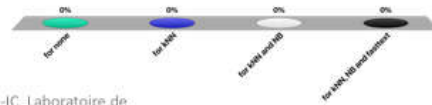
C. 3

D. 4



The dimensionality of the feature space
depends on the vocabulary size ...

- A. for none
- B. for kNN
- C. for kNN and NB
- D. for kNN, NB and
fasttext



© 2018, Karl Aberer, EPFL-IC, Laboratoire de
systèmes d'informations répartis