

[Dashboard](#)/ [My courses](#)/ [SD212](#)/ [Exam](#)/ [Exam](#)

Started on	Tuesday, 27 June 2023, 1:30 PM
State	Finished
Completed on	Tuesday, 27 June 2023, 3:29 PM
Time taken	1 hour 59 mins
Marks	23.00/40.00
Grade	5.75 out of 10.00 (57.5%)

Question 1

Correct

Mark 1.00 out of 1.00

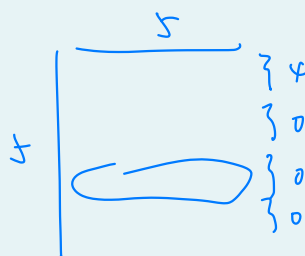
What is the sum of the elements in the third row of the following matrix, coded in the CSR format of scipy?

```
shape = (5, 5)
```

```
indices = [0, 1, 3, 4, 0, 3]
```

```
indptr = [0, 4, 4, 4, 4, 6]
```

```
data = [-1, 2, 2, 1, 2, 1]
```



Answer:



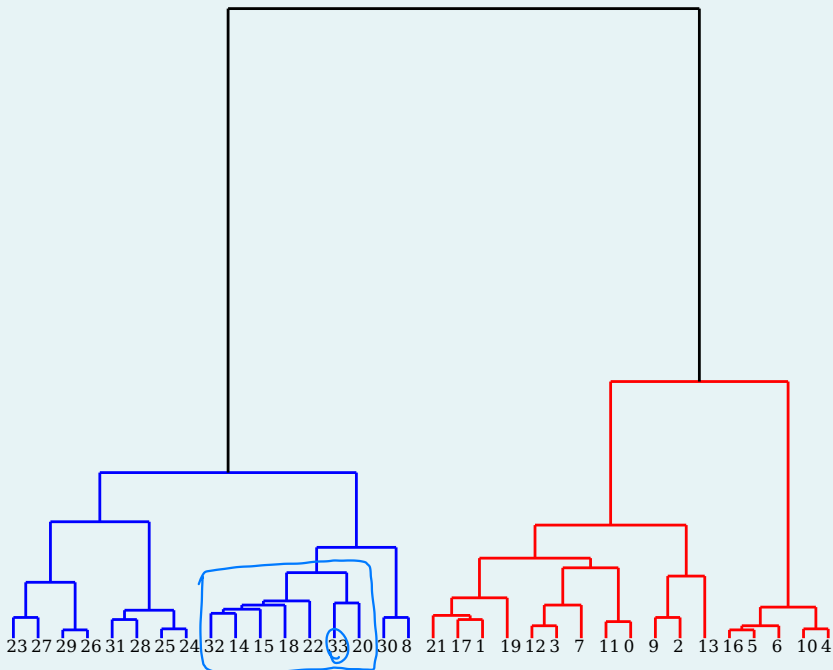
0

Question 2

Incorrect

Mark 0.00 out of 1.00

Consider the following dendrogram:



You apply the following Python function to this dendrogram:

```
def cut(dendrogram, node, level):
    n = dendrogram.shape[0] + 1

    cluster = {i: [i] for i in range(n)}
    cluster_index = node

    k = 0
    for t in range(n - 1):
        i = int(dendrogram[t][0])
        j = int(dendrogram[t][1])
        cluster[n + t] = cluster.pop(i) + cluster.pop(j)
        if cluster_index in {i, j}:
            cluster_index = n + t
            k += 1
        if k == level:
            return cluster[cluster_index]
```

with the following parameters:

cut(dendrogram, 33, 2)

What is the length of the returned list?

Type your answer as an integer.

Example:

3

Answer:

4

7

✖

Question 3

Correct

Mark 1.00 out of 1.00

You want to learn the representation of a graph with a GNN.

The graph has 2000 nodes and 20,000 edges. Each node is represented by a vector of features of dimension 10 and has a label in the set 1, 2, 3, 4.

You decide to train a GNN with a single hidden layer of dimension 20 to learn the labels.

How many parameters must be learned (including the bias terms)?

Answer:

304

304



$$10 \times 20 + 20 + 20 \times 4 + 4$$

$$220 + 84$$

Question 4

Correct

Mark 1.00 out of 1.00

Which format would you recommend if you must update frequently a large sparse matrix?

- ☐ Compressed sparse column
- ☐ Compressed sparse row
- ☐ None of them
- ☐ Coordinate
- ☒ List of list ✓

Question 5

Incorrect

Mark 0.00 out of 1.00

What is the time complexity to get the degree of each node of a graph?

There are n nodes and m edges.

The graph is undirected.

The adjacency matrix is stored in CSR format.

sum of row

- ☐ $O(n^2)$
- ☐ $O(nm)$
- ☒ $O(m)$ ✗
- ☒ $O(n)$ ✓

Question 6

Correct

Mark 1.00 out of 1.00

In a connected, undirected graph, the frequency of visit of each node by a random walk is:

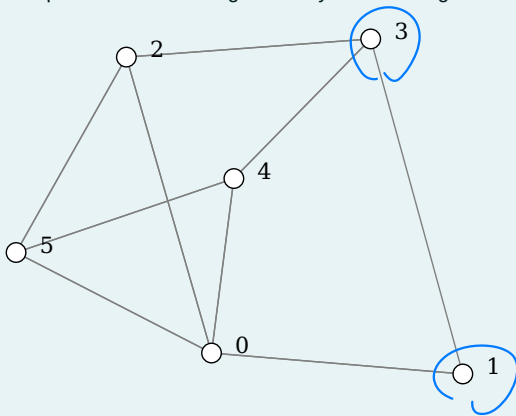
- ☒ proportional to its degree ✓
- ☐ inversely proportional to its degree
- ☐ none of them

Question 7

Correct

Mark 1.00 out of 1.00

Which pair of nodes is merged first by the Paris algorithm?



Type your answer as a list of space-separated integers in increasing order (e.g., 2 3).

Answer: 1 3

1 3



Question 8

Correct

Mark 1.00 out of 1.00

A graph of 6 nodes has a clustering coefficient of $9/20$.

The nodes have respective degrees 2, 2, 3, 3, 4, 4.

How many triangles are there in this graph?

Type your answer as an integer.

$$\binom{2}{2} \times 2 + \binom{3}{2} \times 2 + \binom{4}{2} \times 2 = 2 + 6 + 12 = 20$$

$$\# \text{ triangles} = \frac{20}{3} = 3$$

Answer: 3

3



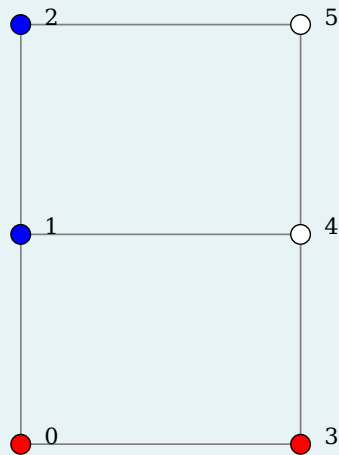
Question 9

Correct

Mark 1.00 out of 1.00

You apply heat diffusion with temperature centering for node classification.

What is the label of node 4 predicted by this classifier in this graph?



Blue:

$$T_5 = \frac{1}{2} + \frac{1}{2}T_4$$

$$T_4 = \frac{1}{3}T_5 + \frac{1}{3}$$

$$\Rightarrow \begin{cases} T_4 = \frac{3}{5} \\ T_5 = \frac{4}{5} \end{cases} \quad \bar{T} = \frac{2 + \frac{2}{5}}{6} = \frac{12}{30}$$

$$T_4 - \bar{T} = \frac{1}{30}$$

Red:

$$T_5 = \frac{1}{2}T_4$$

$$T_4 = \frac{1}{3}T_5 + \frac{1}{3}$$

$$\begin{cases} T_4 = \frac{2}{5} \\ T_5 = \frac{1}{5} \end{cases} \quad \bar{T} = \frac{2 + \frac{3}{5}}{6} = \frac{13}{30}$$

$$T_4 - \bar{T} = -\frac{1}{30}$$

Select one:

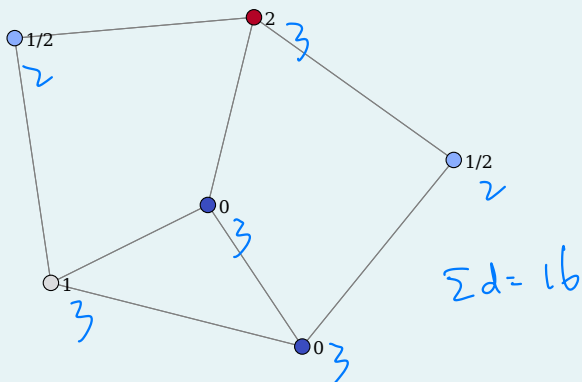
- ☒ Blue ✓
- ☐ Red

Question 10

Incorrect

Mark 0.00 out of 1.00

What is the common temperature of all nodes at equilibrium after heat diffusion in discrete time in this graph?



$$2 \times 3 + 2 \times \frac{1}{2} + 2 \times \frac{1}{2} + (1 \times 3)$$

$$6 + 1 + 1 + 3$$

$$11$$

$$\Sigma d = 16$$

The numbers correspond to the temperatures in the initial state.

Type your answer as an integer or an irreducible fraction (e.g., 2/3).

Answer: 2/3

$$\frac{11}{16}$$

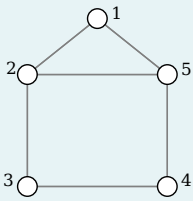
✗

Question 11

Incorrect

Mark 0.00 out of 1.00

Consider a graph neural network trained on the following graph:



The features of the nodes 1, 2, 3, 4, 5 are the rows of the following matrix:

$$\begin{pmatrix} 0 & 1 & -2 \\ 0 & 1 & 0 \\ 2 & 3 & -1 \\ -2 & -1 & 0 \\ 2 & 0 & -1 \end{pmatrix}$$

Each layer consists of the sum of the embedding of the node and the average embedding of the neighbors, followed by a ReLu activation function.

Consider a neuron of the first layer, with weights $w = (1, 0, -\frac{1}{2})$ and bias $b = 1$.

What is the output of this neuron for node 4?

Type your answer as an integer or an irreducible fraction (e.g., 2/3).

Answer: 3/2

Handwritten: 3/2

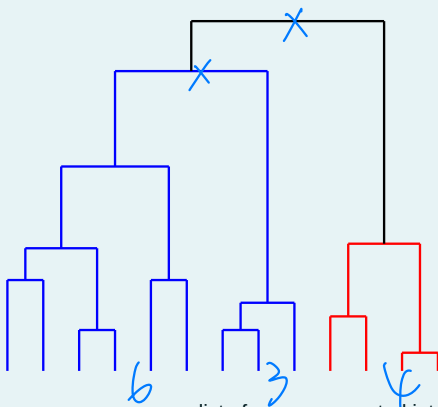
✗

Question 12

Correct

Mark 1.00 out of 1.00

What are the sizes of the top-3 clusters associated with this dendrogram?



Type your answer as a list of space-separated integers in increasing order (e.g., 2 4 6).

Answer: 3 4 6

Handwritten: 3 4 6

✓

Question 13

Correct

Mark 1.00 out of 1.00

What is the **indice pointer (indptr)** vector of the following matrix in the CSR format of scipy?

$$\begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 2 & 1 \end{bmatrix}$$

Type your answer as a list of space-separated integers.

Example:

0 1 3

Answer: 0 2 2 5

0 2 2 5



Question 14

Incorrect

Mark 0.00 out of 1.00

After clustering, you get an aggregate graph with the following adjacency matrix:

$$\begin{bmatrix} 6 & 3 & 3 \\ 3 & 12 & 1 \\ 3 & 1 & 8 \end{bmatrix}$$

$$\sum_k \frac{m_k}{m} - \sum_k \left(\frac{V_k}{V} \right)^2$$

What is the modularity of this clustering?

Type your answer as an integer or an irreducible fraction (e.g., 1/3).

Answer: 29/100

31/100

$$\frac{26}{40} - \left(\frac{12}{40} \right)^2 - \left(\frac{16}{40} \right)^2 - \left(\frac{12}{40} \right)^2$$

$$\frac{13}{20} - \frac{9}{100} - \frac{16}{100} - \frac{9}{100}$$

$$\frac{65-34}{100}$$



Question 15

Correct

Mark 1.00 out of 1.00

You want to select nodes that are strongly connected to some nodes and weakly connected to some other nodes.

Which approach would you recommend?

- ☐ Spectral method
- ☐ PageRank
- ☐ Graph neural networks
- ☒ Heat diffusion ✓

Question 16

Incorrect

Mark 0.00 out of 1.00

Consider the spectral embedding of Wikivitals in dimension 20, based on the transition matrix.

We say that a category is topical if its average pairwise cosine similarity is high (close to 1).

What is the most topical category?

Select one:

- ☒ Mathematics
- ☐ History
- ☐ Biological and health sciences
- ☒ People ✖
- ☐ Technology
- ☐ Arts
- ☐ Philosophy and religion
- ☐ Everyday life
- ☐ Physical sciences
- ☐ Geography
- ☐ Society and social sciences

Question 17

Correct

Mark 1.00 out of 1.00

What is the maximum modularity of a clustering with 9 clusters?

Type your answer as an irreducible fraction (e.g., 2/3).

Answer: 8/9

8/9



Question 18

Correct

Mark 1.00 out of 1.00

You cluster some undirected graph of n nodes with adjacency matrix A using Louvain. You obtain k labels and form the corresponding membership matrix M , a binary matrix of size $n \times k$.

Consider the aggregate graph, with adjacency matrix $M^T A M$ of size $k \times k$.

What represent the sum of each row of this matrix?

V_k

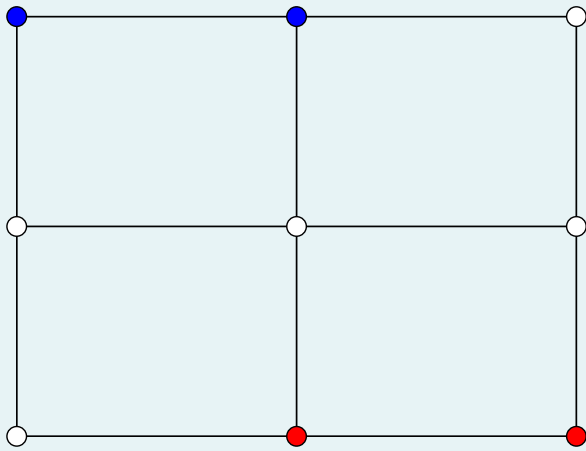
- ☒ the volume of each cluster ✓
- ☐ the strength of each cluster
- ☐ the number of edges of each cluster
- ☐ the number of nodes of each cluster

Question 19

Correct

Mark 1.00 out of 1.00

What are the temperatures of nodes in the center row at equilibrium?



$$\begin{aligned} T_3 &= \frac{1}{2}T_6 \\ T_4 &= \frac{1}{3}T_5 + \frac{1}{3}T_7 \\ T_5 &= \frac{1}{4}T_4 + \frac{1}{4}T_6 + \frac{1}{4} \\ T_6 &= \frac{1}{3}T_3 + \frac{1}{3}T_5 + \frac{1}{3} \\ T_7 &= \frac{1}{2}T_4 + \frac{1}{2} \end{aligned}$$

$$\begin{aligned} T_3 &= \frac{3}{10} \\ T_4 &= \frac{2}{5} \\ T_5 &= \frac{1}{2} \\ T_6 &= \frac{3}{5} \\ T_7 &= \frac{7}{10} \end{aligned}$$

The 2 blue nodes have fixed temperature 0, the 2 red nodes have fixed temperature 1.

Type your answer as a list of 3 space-separated integers or irreducible fractions, corresponding to the temperatures of the nodes of the center row listed from left to right.

Example:

0 1/2 1

Answer: 2/5 1/2 3/5

$$\frac{2}{5} \quad \frac{1}{2} \quad \frac{3}{5}$$

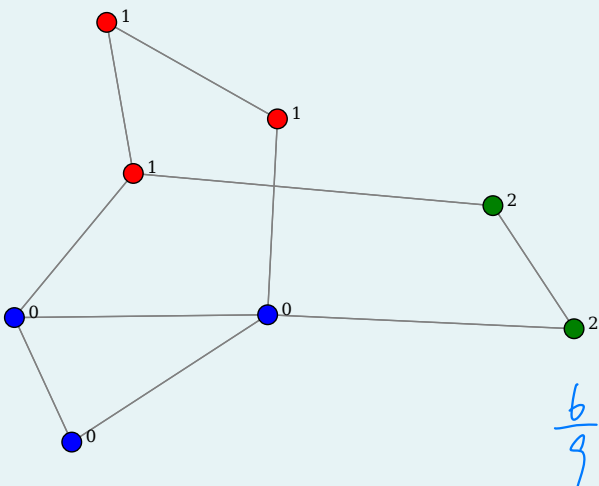


Question 20

Correct

Mark 1.00 out of 1.00

What is the strength of the blue cluster (label 0)?



The numbers correspond to the labels.

Type your answer as an irreducible fraction (e.g., 2/3).

Answer: 2/3

$$\frac{2}{3}$$



Question 21

Correct

Mark 1.00 out of 1.00

What is the Bacon number of Elvis Presley (using the Cinema graph)?

Answer:

3

3

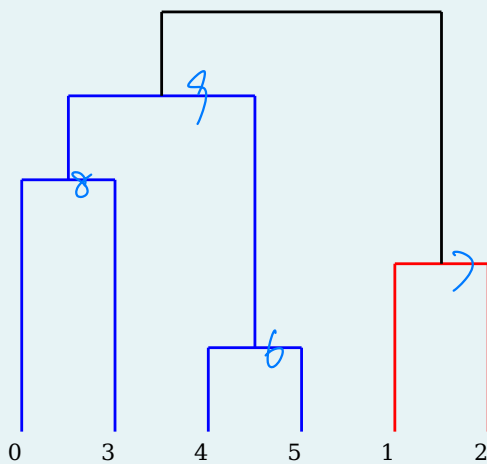


Question 22

Correct

Mark 1.00 out of 1.00

Give the matrix representation of the following dendrogram, including heights and cluster sizes. Heights are integers from 1 to 5.



Type your answer, reading the elements of the matrix as a text (left to right, top-down), as a sequence of space-separated integers (e.g., 0 1 2 1 0 2 4).

Merged nodes must be listed in increasing order of indices (e.g., 1 2 and not 2 1).

4 5 1 2 1 2 2 2 0 3 3 2 6 8 4 4 7 9 5 6

Answer:

4 5 1 2 1 2 2 2 0 3 3 2 6 8 4 4 7 9 5 6



Question 23

Correct

Mark 1.00 out of 1.00

The clustering of a graph has modularity 0.68.

Select one:

- ☐ The number of clusters cannot exceed 3 \times $0.68 > \frac{2}{3}$
- ☒ The number of clusters cannot be lower than 4 \checkmark
- ☐ The number of clusters cannot be lower than 3 \times
- ☐ The number of clusters cannot exceed 4 \times

Question 24

Correct

Mark 1.00 out of 1.00

In the Openflights graph, what is the closest airport from Charles de Gaulle International Airport and Beijing Capital International Airport in terms of Personalized PageRank?

Take respective weights 2 and 1 for these two airports.

Type either the name of the city or the name of the airport.

Example:

Madrid

Answer: Amsterdam Airport Schiphol

Amsterdam Airport Schiphol



Question 25

Incorrect

Mark 0.00 out of 1.00

On a 9×9 grid network, Louvain always finds the clustering with maximum modularity.

$O(n)$ per iter

MP-hand

Select one:

☐ True ✗

☒ False

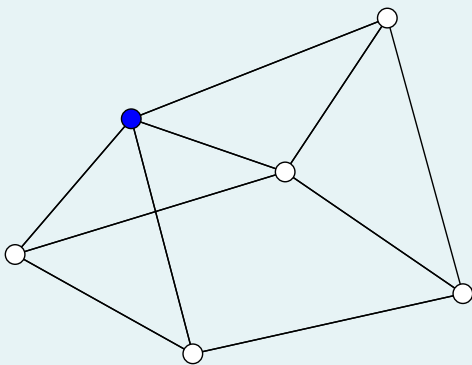
Question 26

Incorrect

Mark 0.00 out of 1.00

Consider a random walk in this graph, without restarts.

What is the frequency of visits to the blue node?



Type your answer as an irreducible fraction (e.g., $2/3$).

The graph has 10 edges.

Answer: $2/10$

$1/5$

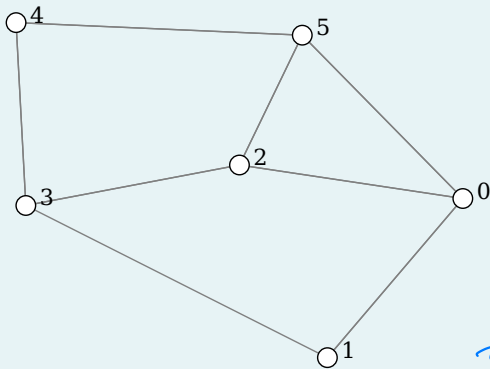


Question 27

Correct

Mark 1.00 out of 1.00

Consider PageRank with damping factor $\alpha = 0.01$ in this graph.



$$\pi' = \pi P(1-\alpha) + \frac{\alpha}{n} \mathbf{1}^T$$

What is the best ranked node?

Type an integer.

Answer:

3

3

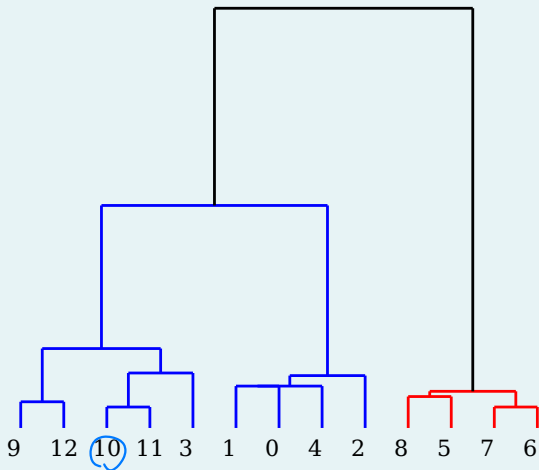


Question 28

Correct

Mark 1.00 out of 1.00

What is the set of nested clusters associated with this dendrogram, starting from node 10?



Type your answer as a list of 13 space-separated integers, starting from 10 (e.g., 10 0 1 2 3 8 9 11 12 4 5 6 7).

Each node must appear exactly once in this list. Within each cluster, nodes must be listed in increasing order.

Answer:

10 11 3 9 12 0 1 2 4 5 6 7 8



10 11 3 9 12 0 1 2 4 5 6 7 8

Question 29

Incorrect

Mark 0.00 out of 1.00

How many articles of Wikivitals have links to each of the following articles: France, Japan and Egypt?

Answer: 565 225 356

✗

1390 1079 960

Question 30

Incorrect

Mark 0.00 out of 1.00

In the aggregation step of the Louvain algorithm, the modularity...

- ☒ increases ✗
- ☐ decreases
- ☒ remains the same
- ☐ may increase or decrease depending on the graph

Question 31

Correct

Mark 1.00 out of 1.00

The modularity of a clustering can be negative.

Select one:

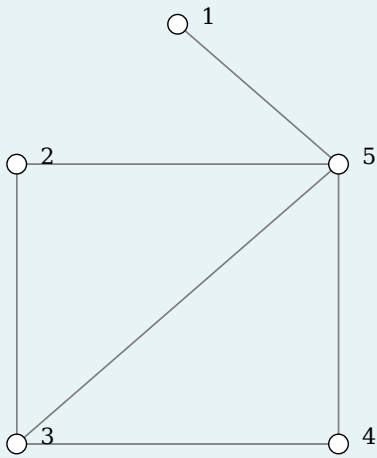
- ☒ True ✓
- ☐ False

Question 32

Incorrect

Mark 0.00 out of 1.00

Consider heat diffusion in continuous time in this graph:



$$L = D - A = \begin{bmatrix} 1 & 0 & 0 & 0 & -1 \\ 0 & 2 & -1 & 0 & -1 \\ 0 & -1 & 3 & -1 & -1 \\ 0 & 0 & -1 & 2 & -1 \\ -1 & -1 & -1 & -1 & 4 \end{bmatrix}$$

$$T = T(0) - L t T(0) + o(t)$$

$$L T(0) = [0, 1, 2, -1, -2]$$

What is the ranking of nodes after heat diffusion at time $t = 0^+$, starting from the following vector of temperatures?

$$T(0) = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

Type your answer as a list of 5 space-separated integers, listing nodes in decreasing order of temperature.

Example:

1 3 2 4 5

Answer:

2 4 5 3 1

2 3 5 4 1

✗

Question 33

Incorrect

Mark 0.00 out of 1.00

Consider the transition matrix of an undirected graph.

The number of zero eigenvalues is equal to the number of connected components of the graph.

Select one:

☒ True ✗

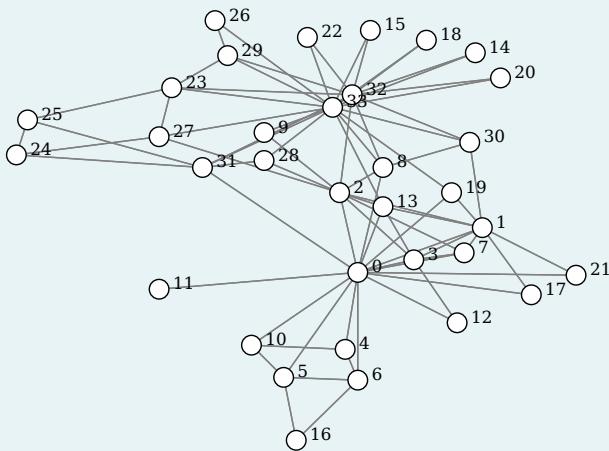
☒ False ✓

Question 34

Incorrect

Mark 0.00 out of 1.00

Consider the spectral embedding in dimension 3 of the Karate-Club graph, based on the transition matrix.



Which fraction of node pairs have a negative cosine similarity in the embedding space?

Consider all $\binom{n}{2}$ node pairs of the graph.

Type your answer as a number between 0 and 1 (precision +/- 0.01).

Answer: 0.6311

0.6613



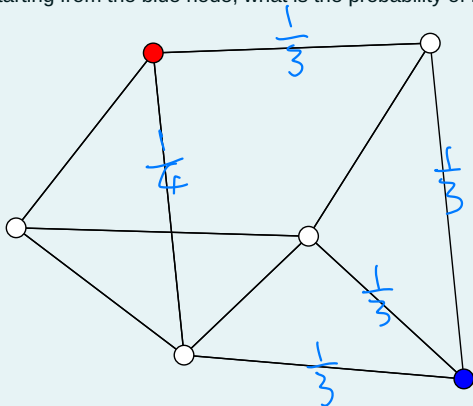
Question 35

Incorrect

Mark 0.00 out of 1.00

Consider a random walk in this graph.

Starting from the blue node, what is the probability of moving in 2 hops to the red node?



$$\begin{aligned} & \frac{1}{3} \times \frac{1}{4} + \frac{1}{3} \times \frac{1}{3} \\ &= \frac{1}{12} + \frac{1}{9} \\ &= \frac{7}{36} \end{aligned}$$

Type your answer as an irreducible fraction (e.g., 2/3).

Answer: 1/72

7/36

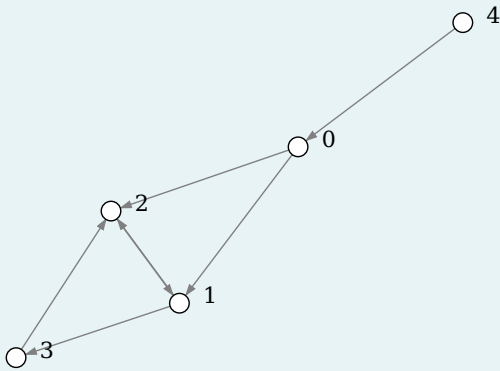


Question 36

Incorrect

Mark 0.00 out of 1.00

What is the indice pointer (indptr) vector of the adjacency matrix of this graph, in CSR format?



$$\begin{bmatrix} 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Type your answer as a list of space-separated integers (e.g., 0 1 2).

Answer:

0 2 4 4 5 6

0 2 4 5 6 7

✗

Question 37

Incorrect

Mark 0.00 out of 1.00

Consider a 3D grid of $3 \times 5 \times 7$ nodes.

bipartite

The corresponding transition matrix P has eigenvalue 0.

Select one:

☒ True

☐ False ✗

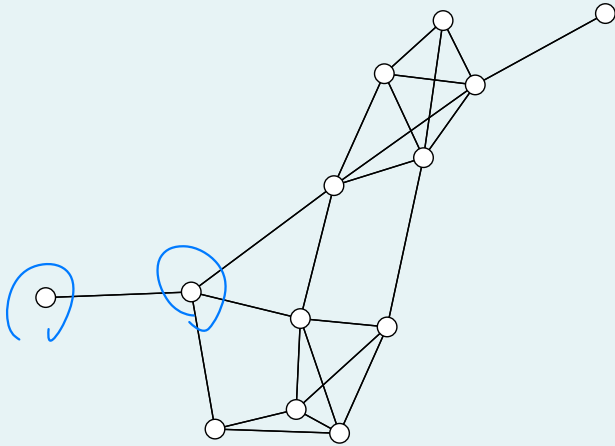
Question 38

Incorrect

Mark 0.00 out of 1.00

What is the resolution limit of this graph (above which Louvain gives one cluster per node)?

There are 24 edges.



$$\frac{48 \times 1}{1 \times 4}$$

Type the answer as an integer or an irreducible fraction (e.g., 2/3).

Answer: 1/12

12



Question 39

Correct

Mark 1.00 out of 1.00

Consider the eigenvectors of the Laplacian matrix associated with an undirected, connected graph.

All eigenvectors except one sum to 0.

Select one:

- ☒ True ✓
☐ False

L symmetric, eigenvectors orthogonal
 1^T is eigenvector $\Rightarrow 1^T v = 0$ for all other eigenvectors

Question 40

Correct

Mark 1.00 out of 1.00

In heat diffusion with boundary constraints (Dirichlet problem), the average temperature weighted by the degrees remains unchanged.

Select one:

- ☐ True
☒ False ✓

◀ Sign

Jump to...



Your questions during the exam ▶