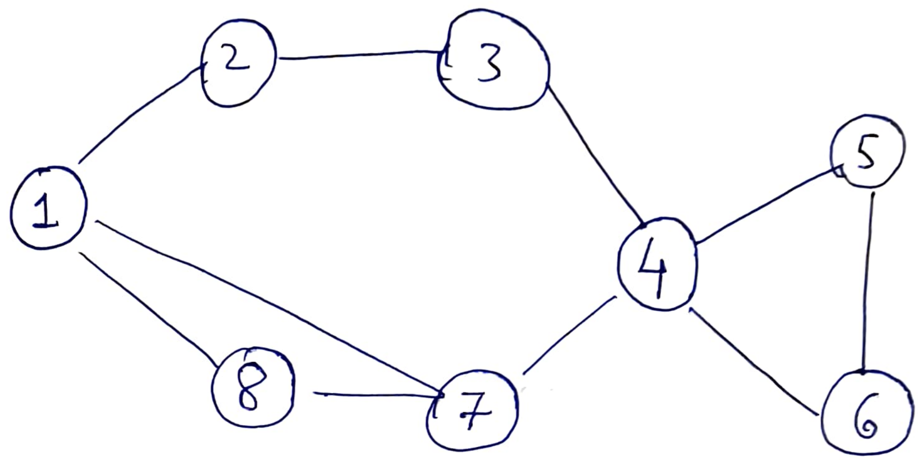
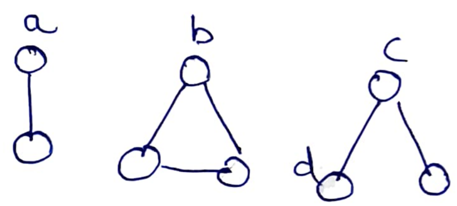
Name: P.V. Sriram
Roll No: 1801CS37

Q1) Given the graph,



We are required to find the Graphlet Degree Vector of node 1.

i.e) $GDV(1)_3$ = Count vector of graphlets rooted at node 1



a	b	c	d
1-2	1-7-8	8-1-2	1-2-3
1-7		7-1-2	1-7-4
1-8	(1)	(2)	(2)
(3)			

$\Rightarrow$ $\boxed{3, 1, 2, 2}$

Q2) Using logistic regression, we could classify nodes into Red / Green based on their GDVs

Given, for such a logistic regression model
$P(\text{Color}(X) = \text{Green} \mid \text{GDV}(X))$

The parameter vector is $\vec{\theta} = [\underbrace{-5}_{\text{bias}} \ 1 \ 2 \ 1 \ 0.5]^T$

$$\therefore \quad P(X) = \frac{1}{1 + e^{-\theta^T X}}$$

here $X = [\underbrace{1}_{\text{bias}}, 3, 1, 2, 2]$

$$\theta^T X = (-5 + 3*1 + 1*2 + 2*1 + 2*0.5)$$

$$= (\cancel{3} + 2 + \cancel{2} + 1 - \cancel{5}) = \underline{\underline{3}}$$

$$\Rightarrow P(X) = \frac{1}{1 + e^{-3}} = \frac{1}{1 + 0.0498} = \frac{1}{1.0498}$$

$$\approx 0.95256 \approx \boxed{0.953}$$

$\therefore$ Its probability of being green is 0.953

Q3) From previous question, we get

$$P(X = Green \mid GDV(X)) = 0.953$$

$$\Rightarrow P(X = Red \mid GDV(X)) = 1 - \cancel{P(X} \; 0.953 = 0.047$$

$$\therefore P(X = Green) > P(X = Red)$$

$$\boxed{\therefore \text{ Green Class}}$$

The Node 1 is green in color.

Q4) We can see that the predictive model takes following function.

$$f(X) = x_1 + 2x_2 + x_3 + 0.5x_4 - 5$$

~~Also,~~ where $X = [x_1 \; x_2 \; x_3 \; x_4]$

Now, for this hyperplane, vector normal to it is

$$\vec{P} = [1, 2, 1, 0.5] \Rightarrow |\vec{P}| = \sqrt{1+4+1+0.25}$$

$$= \sqrt{6.25} = 2.5$$

$$\hat{P} = \frac{\vec{P}}{|\vec{P}|} = \frac{[1, 2, 1, 0.5]}{2.5} = \boxed{[0.4, 0.8, 0.4, 0.2]}$$

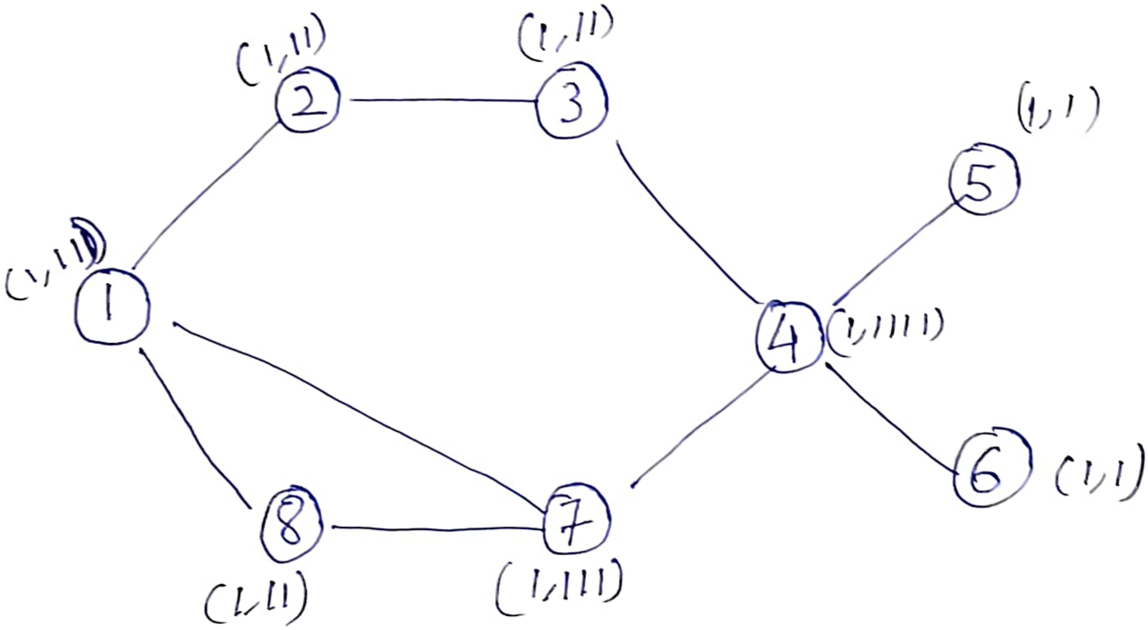
5) Required to represent the given graph in Weisfeiler Lehman graph Kernel.

Given,

HASH function = $x \bmod 13$ where x is sum of colors at node.

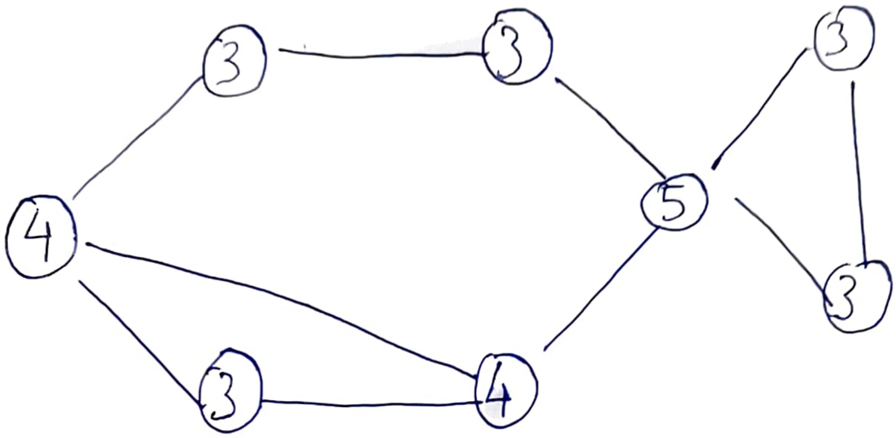
i.e) $c^{(K+1)}(v) = \left(c^{(K)}(v) + \sum_{u \in N(v)} c^{K}(w) \right) \bmod 13$

<u>Graph</u> :

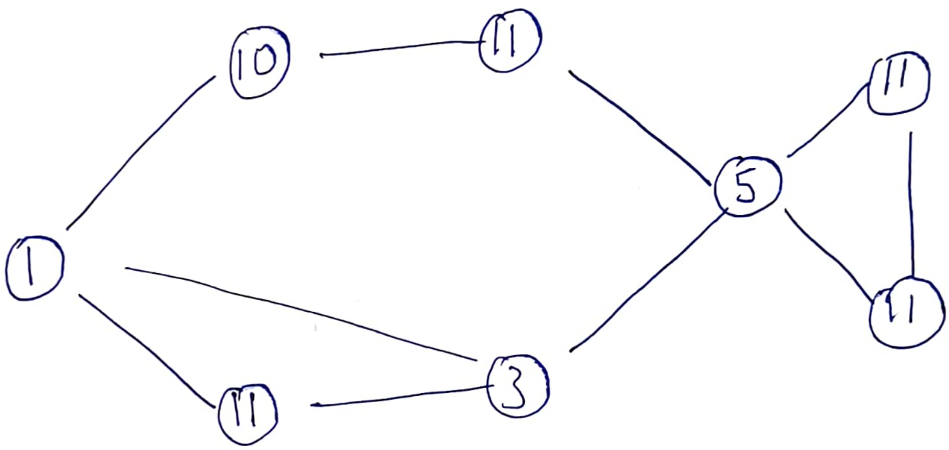


Now, we perform color refinement for 4 steps

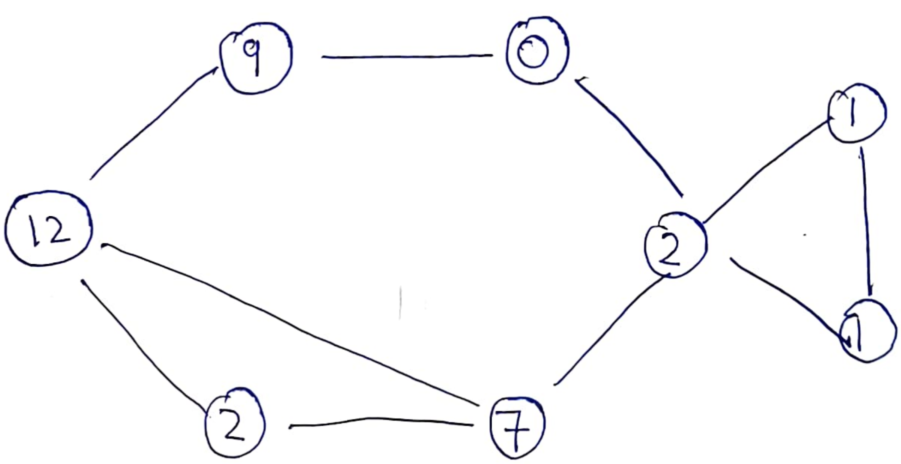
Step-1



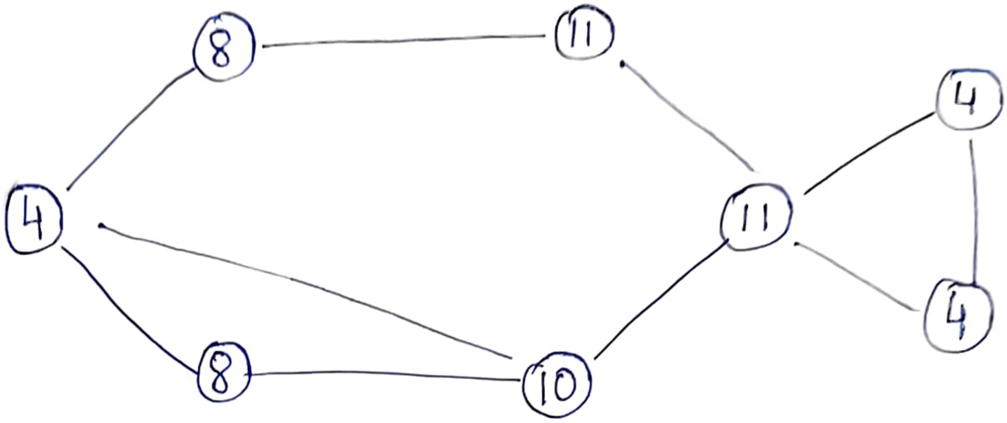
Step-2



Step-3



Step-4



Therefore, color of node 4 after first step is $\boxed{5}$

6) From above steps, we can see color of node 4 after second step is $\boxed{5}$

7) We can also see color of node 4 after fourth step is $\boxed{11}$

8) From the above color refinement, we can make the following color count vector

$$\boxed{CCV(G) = \begin{array}{ccccccccccccc} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 \\ 1, & 13, & 2, & 6, & 5, & 2, & 0, & 1, & 2, & 1, & 2, & 6, & 1 \end{array}}$$

Q9) We consider a message to be data transferred from one node to another

And in a color refinement algorithm the nodes on a node interact with each other

i.e) In a single step there are 2 × no. of edges interactions.

~~&~~ and there are 4 steps

∴ Total no. of messages

$$= 2 \times \underbrace{10}_{\text{no. of edges}} \times \overbrace{4}^{\text{no. of steps}} = \boxed{80}$$

10) For the given graph, required to find Assortativity coefficient based on degree.

i.e)
$$\gamma = \frac{\sum_{ij}\left(A_{ij} - \frac{k_i k_j}{2m}\right) k_i k_j}{\sum_{ij}\left(k_i \delta_{ij} - \frac{k_i k_j}{2m}\right) k_i k_j}$$

where A is adjacency matrix
k_i is degree of node i
m is number of edges $= 10$
$\delta_{ij} = 0$ if $k_i \neq k_j$; 1 if $k_i = k_j$

Degrees

1 → 3
2 → 2
3 → 2
4 → 4
5 → 2
6 → 2
7 → 3
8 → 2

Adjacency Matrix

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

Numerator

$$\sum_{ij} A_{ij} k_i k_j = \overset{21}{(6+9+6)} + \overset{10}{(6+4)} + \overset{12}{(4+8)}$$
$$+ \overset{36}{(8+8+8+12)} + \overset{12}{(8+4)} + \overset{12}{(8+4)}$$
$$+ \overset{27}{(9+12+6)} + \overset{12}{(6+6)} = 142$$

$$\sum_{ij} \frac{(k_i k_j)^2}{2m} = 145.8$$

Denominator

$$\sum_{ij} k_i \delta_{ij} k_i k_j = 27 \times 2 + 8 \times 5 + 64 \times 1$$
$$+ 8 \times 5 + 8 \times 5 + 8 \times 5$$
$$+ 27 \times 2 + 8 \times 5$$
$$= 372$$

$$\therefore \quad r = \frac{142 - 145.8}{372 - 145.8} = \frac{-3.8}{226.2} = -0.016799$$

$$\boxed{\approx -0.017}$$