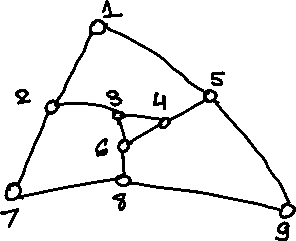
**1. (30 marks) Case Study: Traffic congestion modeling using GNN**

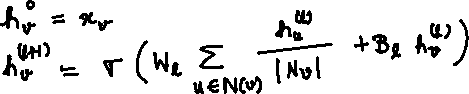
The city of Raipur is trying to reduce traffic congestion in the Jaistambh chowk during busy hours. A simple layout of 9 intersections in shown below. Intersections 1, 5 and 7 are congested. Intersections 2, 3, 4 are NOT congested. You have to classify intersections 6, 8, 9 using a 1-layer GNN followed by sigmoid function.



|  |  |
| --- | --- |
| **1** | **10, 9, 9, 8** |
| **2** | **3,4,5,3** |
| **3** | **3,4,3,2** |
| **4** | **3,5,6,4** |
| **5** | **9,8,9,10** |
| **6** | **8,7,7,3** |
| **7** | **8,9,10,9** |
| **8** | **5,5,5,5** |
| **9** | **4,4,6,8** |



**Update Equations:**



(15 points) Assume W1 = 0.2\*I (Identity Matrix) and B1 = 0.2\*Ones Matrix of appropriate dimensions. Sigma () is the ReLU function. Using the update equations given above, and feature vectors (x) in the above table, calculate the embedding for nodes 3, 5, 6, 8 and 9.

(15 points) The Sigmoid function in the last layer for node classification takes the embedding of each node as input and outputs “congested” if its value is more than 0.5.

The weight vector, w, used is [0.1, 0.1, 0.1, 0.1] and bias is b.

Choose a value of b that classifies 3 and 5 correctly.



Classify nodes 6,8 and 9 based on the above rule.



**2. (30 marks) Objective type questions**

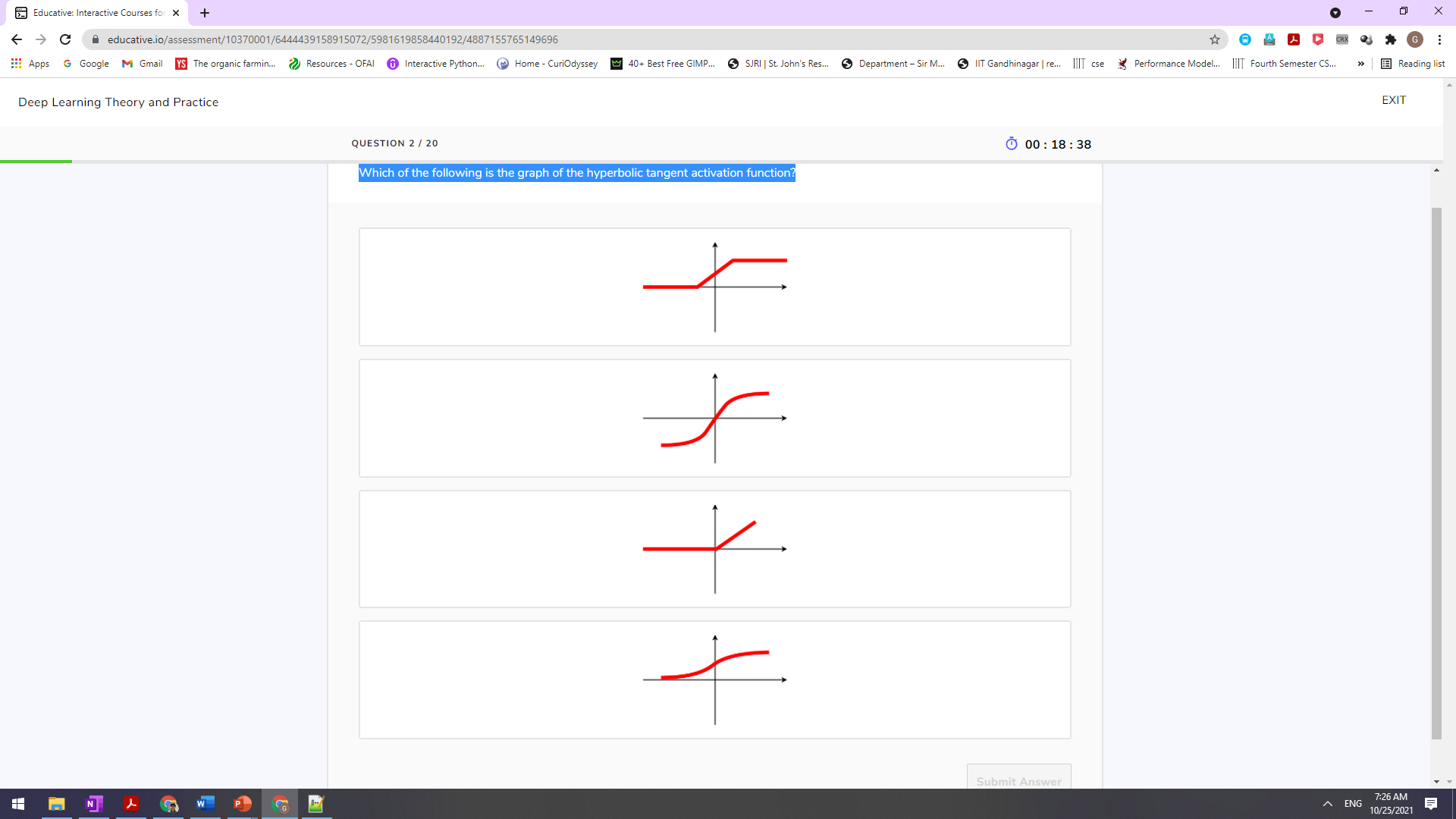
**i. In machine learning, what is the purpose of splitting a dataset into training and test sets?**

1. The training set is for model training and the test set is for model validation.



1. The test set is for additional training after training with the training set.
2. The training and test sets are for training different models in parallel.
3. The test set is for pre-training the network before training with the training set.

**ii. Match the activation functions**





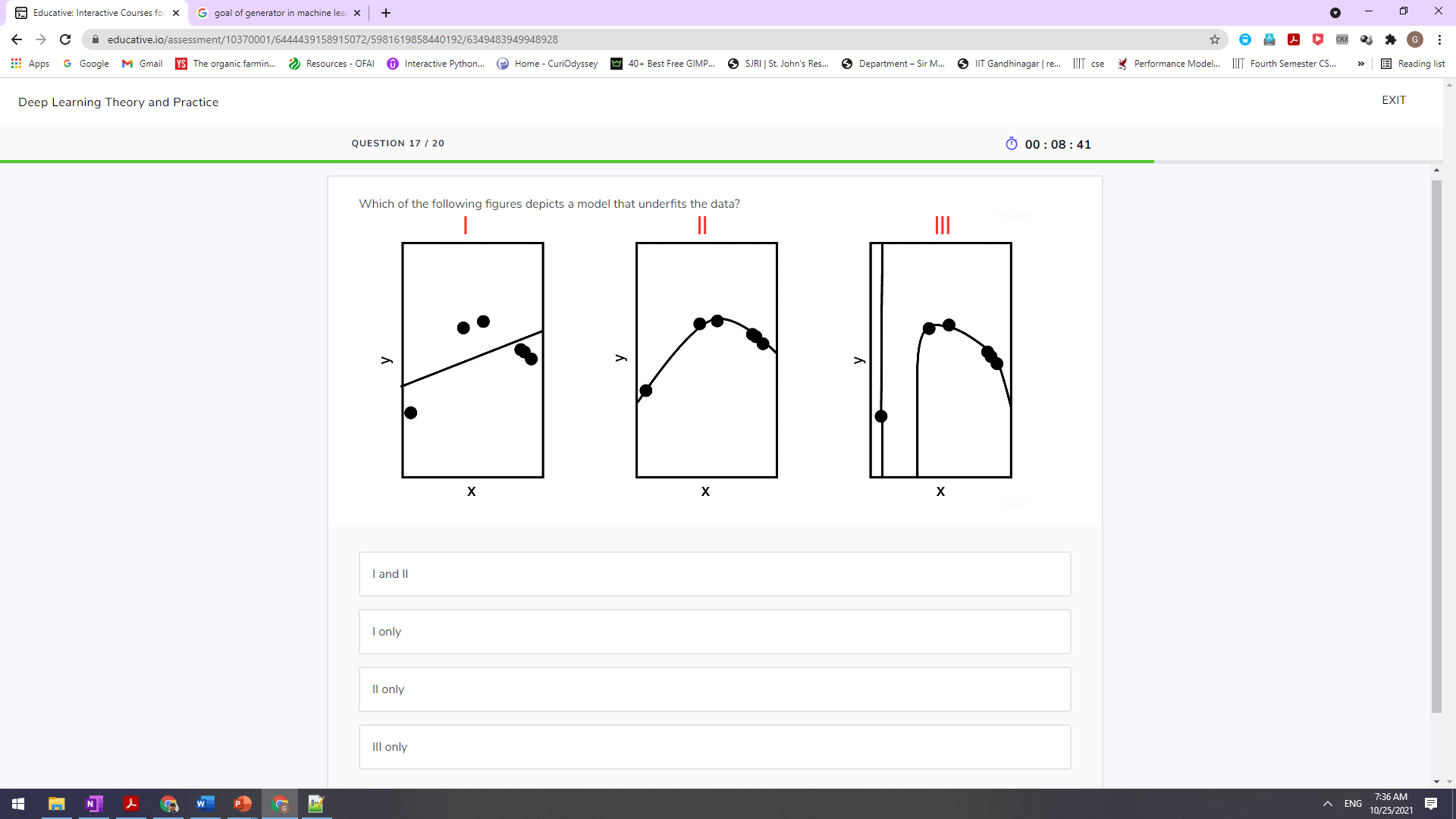
**iii. Given that n training samples are available to train a neural network. How many training samples are processed by stochastic gradient descent at each iteration of the optimization?**

1. one training sample
2. n training samples
3. k training samples, chosen at random, where 1<k<n/2
4. k training samples, selected in order, where 1<k<n/2

**iv. Which component of the artificial neural networks (ANNs) enables the modeling of nonlinear functions?**

1. bias
2. weights
3. depth of the network
4. activation functions

**v. Which of the following depicts a model that underfits the data?**

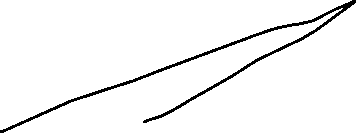
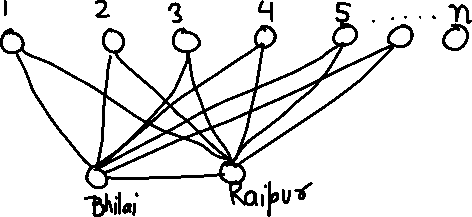


**vi. A neuron has input vector X = weight vector w= and bias = 0.12. What is its output?**



**3. (20 marks) Case Study: Taxi Company Graph**

There are two transport hubs for a taxi company: Bhilai and Raipur. Their cabs either go from Bhilai to Raipur OR make a trip to a city and come back to either Bhilai or Raipur. Their connectivity graph looks like the following:



n is the number of cities, served by them. The weight of each edge is 1. Note that there is no edge between other cities, apart from what is shown in the graph above. This is called the double-star graph. For example, there is no edge between 3 and 5.

a. (5 marks) Calculate the degree distribution of this graph. What is the mean degree of the graph?

b. (5 marks) Show the form of the adjacency matrix representation of this graph?

c. (5 marks) What is the clustering coefficient of Bhilai and city 1?

Hint: Clustering coefficient of node i: where is its degree and is the number of edges between its neighbors.

d. (5 marks) What is the betweenness of Bhilai? Is it the same as that of Raipur?

Hint:

