Problem 1

Sunday, February 6, 2022 13:14

[USING 1 GRACE DAY]

Q.1 Use variable elimination to compute $P(TM|L)$

C	P(C)
0	0.8
1	0.2

L	P(L)
0	0.4
1	0.6

C	RN	P(RN|C)
0	0	0.95
0	1	0.05
1	0	0.35
1	1	0.65

C	L	S	P(S|C,L)
0	0	0	0.99
0	0	1	0.01
0	1	0	0.7
0	1	1	0.3
1	0	0	0.5
1	0	1	0.5
1	1	0	0.4
1	1	1	0.6

S	RN	TM	P(TM|S,RN)
0	0	0	0.99
0	0	1	0.01
0	1	0	0.75
0	1	1	0.25
1	0	0	0.9
1	0	1	0.1
1	1	0	0.45
1	1	1	0.55

$$P(C, L, S, RN, TM) = P(C)\,P(L)\,P(RN|S)\,P(S|C,L)\,P(TM|S,RN)$$

$$P(TM|L) = \frac{P(TM, L)}{P(L)} = \frac{\sum_C \sum_S \sum_{RN} P(C)\,P(L)\,P(RN|S)\,P(S|C,L)\,P(TM|S,RN)}{P(L)}$$

$$P(TM, L) = \sum_C \sum_S \sum_{RN} P(C, L, S, RN, TM)$$

Step 1: Eliminate C

$$P(TM, L) = \sum_S \sum_{RN} \sum_C P(C, L, S, RN, TM)$$

$$= \sum_S \sum_{RN} \sum_C P(C)\,P(L)\,P(RN|C)\,P(S|C,L)\,P(TM|S,RN)$$

$$= \sum_S \sum_{RN} P(L)\,P(TM|S,RN) \underbrace{\sum_C P(C)\,P(RN|C)\,P(S|C,L)}_{\text{Let } \tau_C(S, L, RN)}$$

$$\tau_C(S, L, RN) = \sum_C P(C)\,P(RN|C)\,P(S|C,L)$$

$$\tau_C(S, L, RN) = P(C=0)\,P(RN|C=0)\,P(S|C=0, L) + P(C=1)\,P(RN|C=1)\,P(S|C=1, L)$$

S	L	RN	$\tau_C(S, L, RN)$
0	0	0	$0.8 \times 0.95 \times 0.99 + 0.2 \times 0.35 \times 0.5 =$ 0.7874
0	0	1	$0.8 \times 0.05 \times 0.99 + 0.2 \times 0.65 \times 0.5 =$ 0.1046
0	1	0	$0.8 \times 0.95 \times 0.7 + 0.2 \times 0.35 \times 0.4 =$ 0.56
0	1	1	$0.8 \times 0.05 \times 0.7 + 0.2 \times 0.65 \times 0.4 =$ 0.08
1	0	0	$0.8 \times 0.95 \times 0.01 + 0.2 \times 0.35 \times 0.5 =$ 0.0426
1	0	1	$0.8 \times 0.05 \times 0.01 + 0.2 \times 0.65 \times 0.5 =$ 0.0654
1	1	0	$0.8 \times 0.95 \times 0.3 + 0.2 \times 0.35 \times 0.6 =$ 0.27
1	1	1	$0.8 \times 0.05 \times 0.3 + 0.2 \times 0.65 \times 0.6 =$ 0.09

now, $P(TM, L) = \sum_S \sum_{RN} P(L)\,P(TM|S,RN)\,\tau_C(S, L, RN)$

Step 2: Eliminate S

$$P(TM, L) = \sum_{RN} P(L) \underbrace{\sum_S P(TM|S,RN)\,\tau_C(S, L, RN)}_{\text{Let } \tau_S(L, RN, TM)}$$

$$\tau_S(L, RN, TM) = \sum_S P(TM|S,RN)\,\tau_C(S, L, RN)$$

$$= P(TM|S=0, RN)\,\tau_C(S=0, L, RN) + P(TM|S=1, RN)\,\tau_C(S=1, L, RN)$$

L	RN	TM	$\tau_S(L, RN, TM)$
0	0	0	$0.99 \times 0.7874 + 0.99 \times 0.0426 =$ 0.817866
0	0	1	$0.01 \times 0.7874 + 0.1 \times 0.0426 =$ 0.012134
0	1	0	$0.75 \times 0.1046 + 0.45 \times 0.0654 =$ 0.107888
0	1	1	$0.25 \times 0.1046 + 0.55 \times 0.0654 =$ 0.061212
1	0	0	$0.99 \times 0.56 + 0.9 \times 0.27 =$ 0.7974
1	0	1	$0.01 \times 0.56 + 0.1 \times 0.27 =$ 0.0326
1	1	0	$0.75 \times 0.08 + 0.45 \times 0.09 =$ 0.1005
1	1	1	$0.25 \times 0.08 + 0.55 \times 0.09 =$ 0.0695

Now, $P(TM, L) = \sum_{RN} P(L)\,\tau_S(L, RN, TM)$

Step 3: Eliminate RN

$$P(TM, L) = P(L) \underbrace{\sum_{RN} \tau_S(L, RN, TM)}_{\text{Let } \tau_{RN}(TM, L)}$$

$$\tau_{RN}(TM, L) = \sum_{RN} \tau_S(L, RN, TM)$$

$$= \tau_S(L, RN=0, TM) + \tau_S(L, RN=1, TM)$$

L	TM	$\tau_{RN}(L, TM)$
0	0	$0.817866 + 0.107888 =$ 0.925746
0	1	$0.012134 + 0.061212 =$ 0.074254
1	0	$0.7974 + 0.1005 =$ 0.8979
1	1	$0.0326 + 0.0695 =$ 0.1021

$$P(TM, L) = P(L)\,\tau_{RN}(TM, L)$$

L	TM	P(TM, L)
0	0	$0.925746 \times 0.4 =$ 0.3702984
0	1	$0.074254 \times 0.4 =$ 0.0297016
1	0	$0.8979 \times 0.6 =$ 0.53874
1	1	$0.1021 \times 0.6 =$ 0.06126

$$P(TM|L) = \frac{P(TM, L)}{P(L)} = \tau_{RN}(TM, L)$$

TM	L	P(TM|L)
0	0	0.925746
0	1	0.8979
1	0	0.074254
1	1	0.1021

Problem 2

Sunday, February 6, 2022 14:46

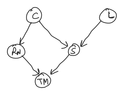
Q.2) Use graph elimination to compute $P(TM|L)$

Joint probability distribution

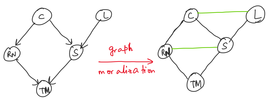
$P(C,L,S,RN,TM) =$

$P(C)\,P(L)\,P(S|C,L)\,P(RN|C)$

$P(TM|RN,S)$



→ Convert bayesian network graph to moralized graph.

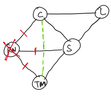


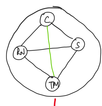
Graph Elimination Ordering: $RN \rightarrow S \rightarrow C$

$$P(TM|L) = \frac{P(TM,L)}{P(L)}$$

$P(TM,L) = P(C)\,P(L)\,P(RN|C)\,P(S|C,L)\,P(TM|S,RN)$

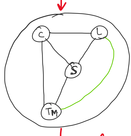
Eliminate RN





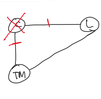
$m_{RN}(C,S,TM)$

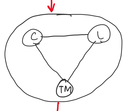
Eliminate S



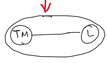
$m_S(C,L,TM)$

Eliminate C





$m_C(L,TM)$



$$m_{RN}(C,S,TM) = \sum_{RN} P(RN|C)\,P(TM|S,RN)$$

$$m_S(C,L,TM) = \sum_{S} m_{RN}(C,S,TM)\,P(S|C,L)$$

$$m_C(L,TM) = \sum_{C} m_S(C,L,TM)\,P(C)$$

Now, we need to calculate these messages

$$m_{RN}(C,S,TM) = P(RN=0|C)\,P(TM|S,RN=0) + P(RN=1|C)\,P(TM|S,RN=1)$$

C	S	TM	$m_{RN}(C,S,TM)$
0	0	0	$0.95\times0.99 + 0.05\times0.75 = 0.978$
0	0	1	$0.95\times0.01 + 0.05\times0.25 = 0.022$
0	1	0	$0.95\times0.9 + 0.05\times0.45 = 0.8775$
0	1	1	$0.95\times0.1 + 0.05\times0.55 = 0.1225$
1	0	0	$0.35\times0.99 + 0.65\times0.75 = 0.834$
1	0	1	$0.35\times0.01 + 0.65\times0.25 = 0.166$
1	1	0	$0.35\times0.9 + 0.65\times0.45 = 0.6075$
1	1	1	$0.35\times0.1 + 0.65\times0.55 = 0.3925$

$$m_S(C,L,TM) = m_{RN}(C,S=0,TM)\,P(S=0|C,L) + m_{RN}(C,S=1,TM)\,P(S=1|C,L)$$

C	L	TM	$m_S(C,L,TM)$
0	0	0	$0.978\times0.99 + 0.8775\times0.01 = 0.976995$
0	0	1	$0.022\times0.99 + 0.1225\times0.01 = 0.023005$
0	1	0	$0.978\times0.7 + 0.8775\times0.3 = 0.94785$
0	1	1	$0.022\times0.7 + 0.1225\times0.3 = 0.05215$
1	0	0	$0.834\times0.5 + 0.6075\times0.5 = 0.72075$
1	0	1	$0.166\times0.5 + 0.3925\times0.5 = 0.27925$
1	1	0	$0.834\times0.4 + 0.6075\times0.6 = 0.6981$
1	1	1	$0.166\times0.4 + 0.3925\times0.6 = 0.3019$

$$m_C(L,TM) = m_S(C=0,L,TM)\,P(C=0) + m_S(C=1,L,TM)\,P(C=1)$$

L	TM	$m_C(L,TM)$
0	0	$0.976995\times0.8 + 0.72075\times0.2 = 0.925946$
0	1	$0.023005\times0.8 + 0.27925\times0.2 = 0.074254$
1	0	$0.94785\times0.8 + 0.6981\times0.2 = 0.8979$
1	1	$0.05215\times0.8 + 0.3019\times0.2 = 0.1021$

Now,

$$P(TM,L) = P(L)\,m_C(TM,L)$$

$$P(TM|L) = \frac{P(TM,L)}{P(L)} = m_C(TM,L)$$

Q) For finding P(TM|L), find out the optimal elimination sequence with lowest complexity.

(a) For elimination order C → S → RN (used in problem 1 for variable elimination)

# of multiplications (N_M) = 32 + 16
# of additions (N_A) = 8 + 8 + 4
total = 48 N_M + 20 N_A

(b) For elimination order RN → S → C (used in problem 2 for graph elimination)

N_M = 16 + 16 + 8 = 40

N_A = 8 + 8 + 4 = 20

(c) For elimination order C → RN → S

N_M = 32 + 16 = 48

N_A = 8 + 8 + 4 = 20

(d) For elimination order RN → C → S

N_M = 16 + 32 = 48

N_A = 8 + 8 + 4 = 20

(e) For elimination order S → C → RN

N_M = 32 + 32 = 64

N_A = 16 + 8 + 4 = 28

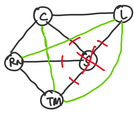
(f) For elimination order S → RN → C

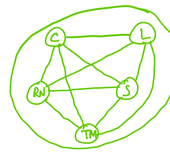
N_M = 32 + 16 + 8 = 56

N_A = 16 + 8 + 4 = 28

∴ Elimination order requiring lowest number of multiplications and additions operations: RN → S → C

→ For elimination ordering starting with S, it creates a elimination clique of tree width = 5



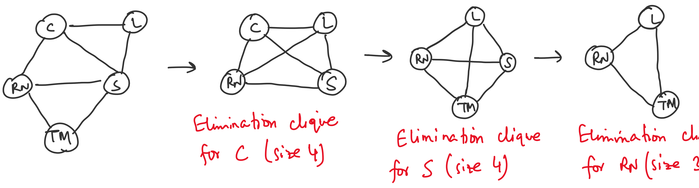


Elimination clique for (S)

Tree width = 5 ⇒ Complexity = $O(e^5)$

Hence, elimination sequence starting with S has the highest complexity as evident from (e) & (f).

→ For elimination order (a) C → S → RN

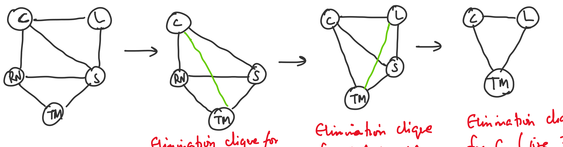


Elimination clique for C (size 4)

Elimination clique for S (size 4)

Elimination clique for RN (size 3)

→ For elimination order (b) RN → S → C

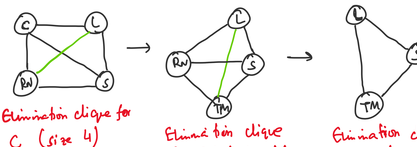


Elimination clique for RN (size 4)

Elimination clique for S (size 4)

Elimination clique for C (size 3)

→ Elimination order (c) C → RN → S

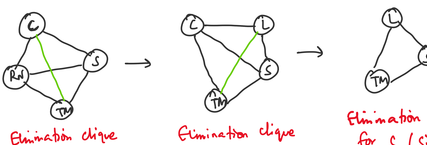


Elimination clique for C (size 4)

Elimination clique for RN (size 4)

Elimination clique for S (size 3)

→ Elimination order (d) RN → C → S



Elimination clique for RN (size 4)

Elimination clique for C (size 4)

Elimination clique for S (size 3)

→ Overall complexity is determined by the number of largest elimination clique when each node is eliminated.

→ (C → S → RN), (C → RN → S), (RN → C → S), (RN → S → C) have same number of largest elimination clique. (=4)

Hence, these 4 elimination orders are optimal.

→ (S → C → RN) and (S → RN → C) have more number of largest elimination clique (=5). Hence these elimination orders are not optimal.

Problem 4

Tuesday, February 8, 2022 9:42

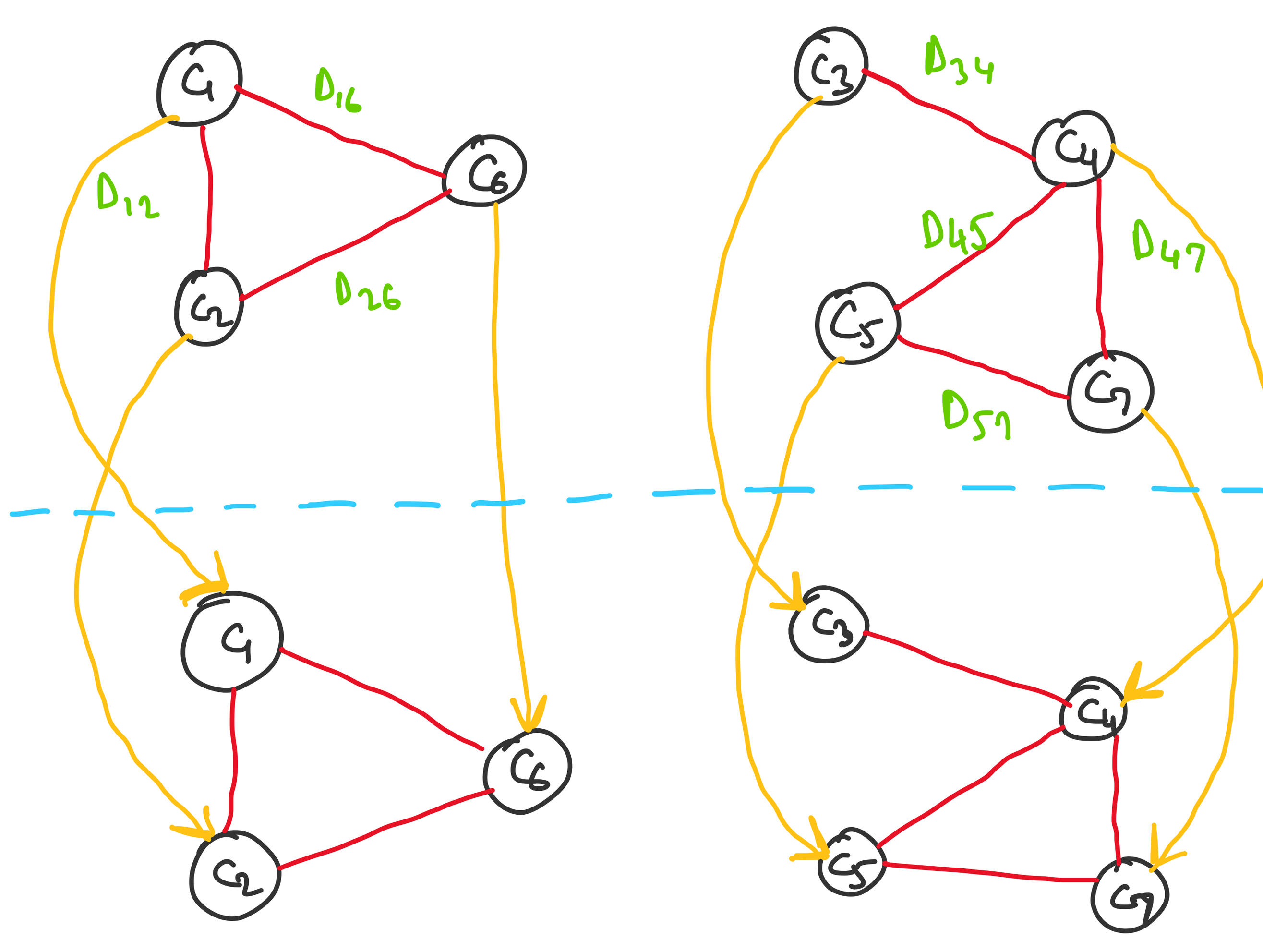
Q1 for COVID-19 prediction task (for cities in SD area), design a graph neural network to perform this task.

Let cities in SD area be denoted by C_i, $i = 1, 2, \ldots 7$

Let distance between cities C_i and C_j be denoted by D_{ij}

$D_{ii} = 0$ and $D_{ij} = D_{ji}$ $\forall$ $i, j = 1, 2, \ldots 7$

Let X_i denote the attribute / feature vector for city C_i

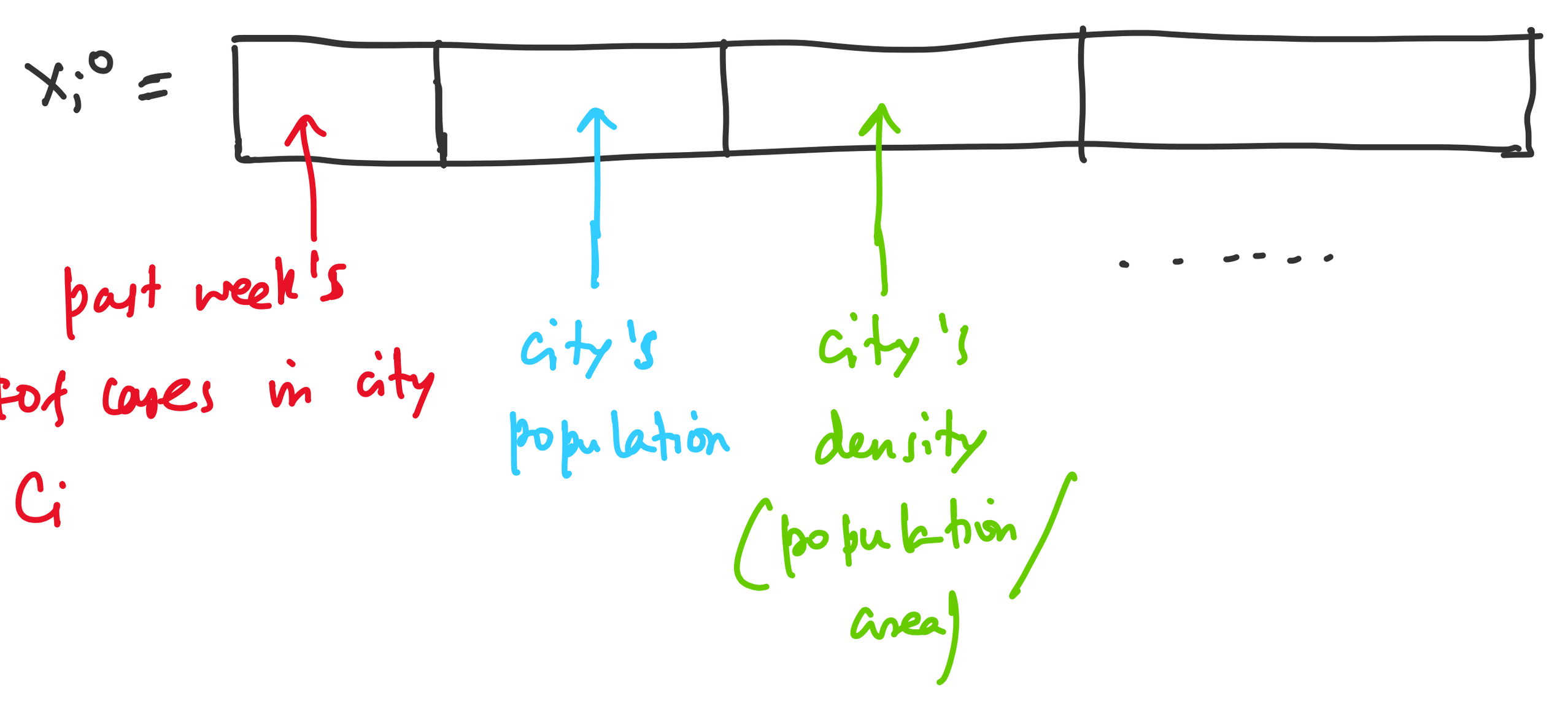


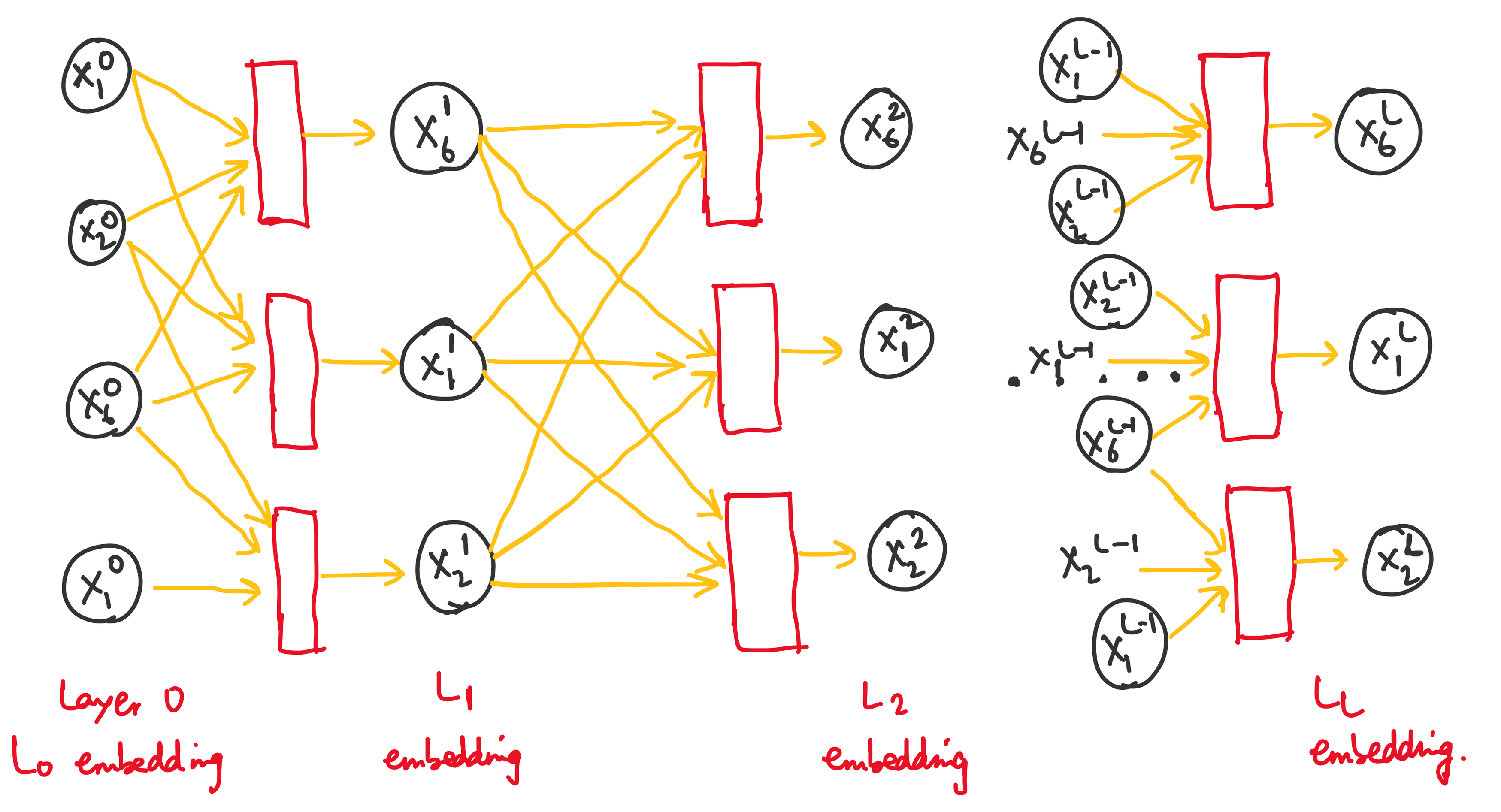
G_p : city graph with nodes representing past week cases

G_c : city graph with nodes representing current week's cases

Graph Neural Network

X_i^0 : initial feature vector for city C_i, $i = 1, 2, \ldots, 7$





X_i^0 : initial embedding / feature vector $\in R^d$.

$$X_i^k = \text{RELU}\left[W_k \sum_{j \in N(C_i)} \frac{X_j^{k-1}}{|N(C_i)|} + B_k X_i^{k-1} \right] \quad \forall\ k = 1, 2, \ldots L \text{ and } i = 1, 2, \ldots, 7$$

RELU: non-linear activation function.

X_i^k: layer L_k embedding

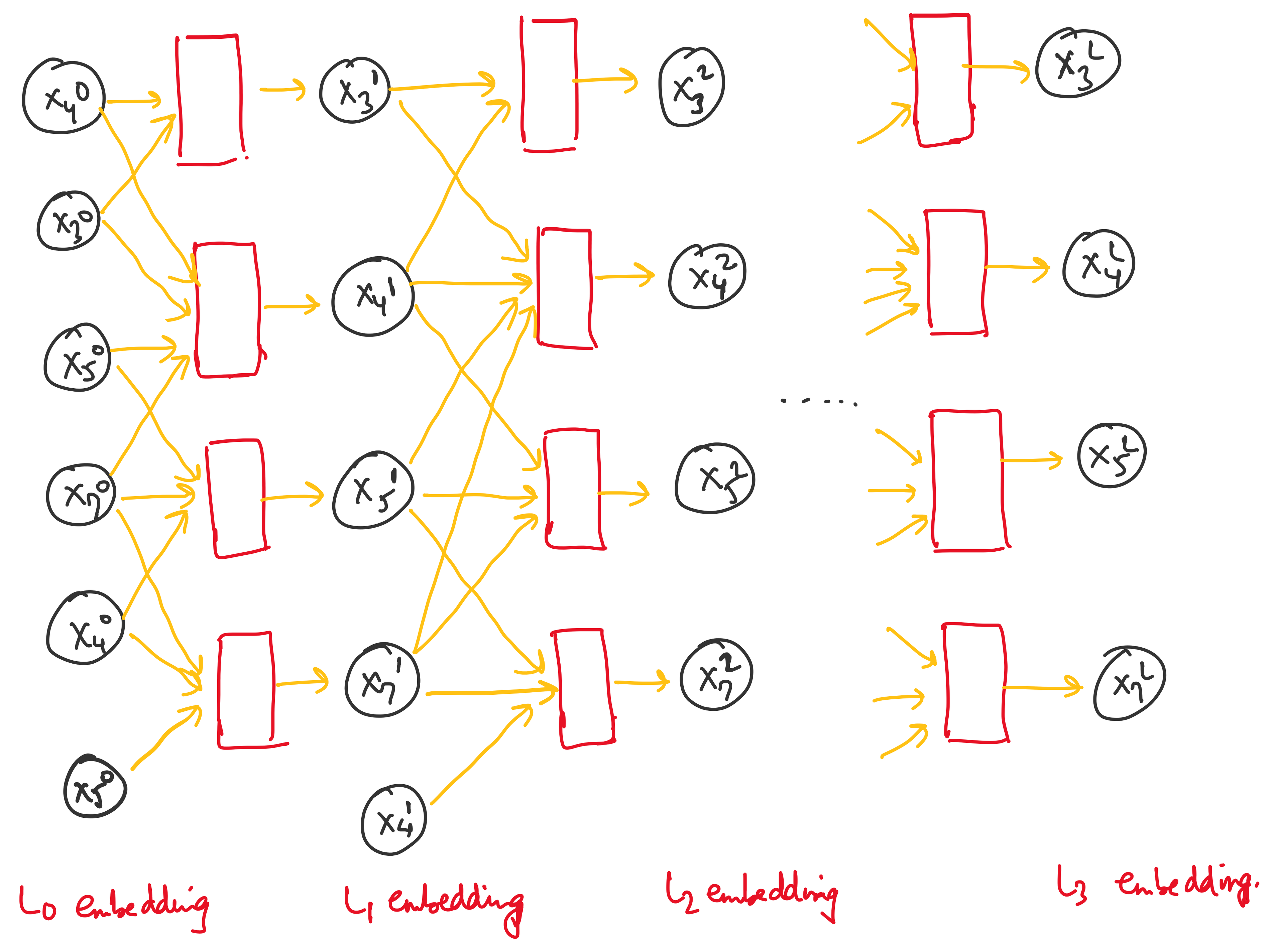
W_k: neighbor aggregation weight matrix

$\sum_{j \in N(C_i)} \frac{X_j^{k-1}}{|N(C_i)|}$: average features of neighbouring nodes from previous layer.

B_k: self-information weight matrix

$W_k, B_k \in R^{d \times d}$ are trainable weight matrices.

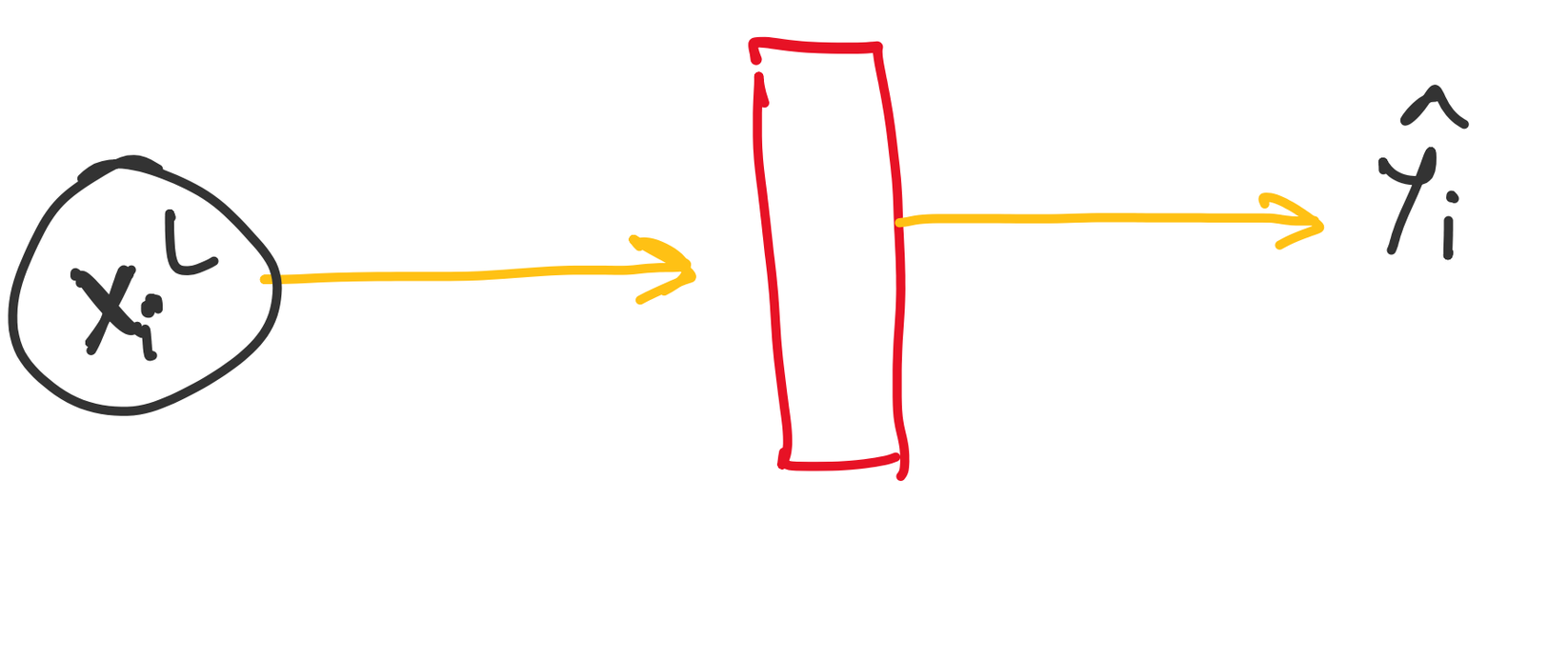
$X_i^k \in R^d$ are features/embeddings for city C_i at k^{th} layer.



→ After getting the embedding vectors $X_i^L \in R^d$, these can be used to train a neural network to predict the number of infected cases in each city.

→ Let $D = \{(x_1, y_1), (x_2, y_2) \ldots (x_n, y_n)\}$ represent the data where $x_i \in R^d$ and $y_i \in Z$ (infected number of cases) which has been collected from previous weeks.

→ Regression Neural Network for prediction.



$\hat{y}_i$ be the prediction for this week's infected # of cases in each city

Prediction $\hat{y}_i = \text{ReLU}(W X_i^L + b)$, $X_i^L \in R^d$, $W \in R^{d \times d}$, $b \in R^d$

W, b: trainable parameters.

→ The graph neural network can be trained such that cities that are close by and neighbors should have similar number of infected cases.

$$W_k^*, B_k^* = \underset{\substack{W_k, B_k \\ 1 \le k \le L}}{\text{argmax}} \sum_{i=1}^{N} \sum_{j \in N(C_i)} (X_i^L)^T (X_j^L) \qquad 1 \le k \le L$$

→ The above optimization problem can be optimized using stochastic gradient descent.

→ For the regression NN, the loss function can be written as

$$W^*, b^* = \underset{W, b}{\text{argmin}} \sum_{i=1}^{N} (y_i^G - \hat{y}_i)^2$$

N: # of training data.

y_i^G: ground truth data

→ The regression loss function can also be optimized using stochastic gradient descent algorithm.