HOMEWORK 7

Devansh Goenka
908 335 1354

Instructions: There is no need to submit the latex source or any code. You can choose any programming language, as long as you implement the algorithm from scratch. Use this latex file as a template to develop your homework. Submit your homework on time as a single pdf file to Canvas. Please check Piazza for updates about the homework.

1 Chow-Liu Algorithm [100 pts]

Suppose we wish to construct a directed graphical model for 3 features X, Y, and Z using the Chow-Liu algorithm. We are given data from 100 independent experiments where each feature is binary and takes value T or F. Below is a table summarizing the observations of the experiment:

X	Y	Z	Count
T	T	T	36
T	T	F	4
T	F	T	2
T	F	F	8
F	T	T	9
F	T	F	1
F	F	T	8
F	F	F	32

1. Compute the mutual information $I(X,Y)$ based on the frequencies observed in the data. (20 pts)
2. Compute the mutual information $I(X,Z)$ based on the frequencies observed in the data. (20 pts)
3. Compute the mutual information $I(Z,Y)$ based on the frequencies observed in the data. (20 pts)
4. Which undirected edges will be selected by the Chow-Liu algorithm as the maximum spanning tree? (20 pts)
5. Root your tree at node X, assign directions to the selected edges. (20 pts)

2 Answers

2.1 1.1

We want to compute the mutual information $I(X;Y)$.
We know that:

$$I(X;Y) = H(Y) - H(Y|X)$$

And also,

$$H(Y|X) = H(X,Y) - H(X)$$

Thus, we get:

$$I(X;Y) = H(Y) + H(X) - H(X,Y)$$

Using the observations above, we get: $H(X) = -\frac{1}{2}\log_2\frac{1}{2} - \frac{1}{2}\log_2\frac{1}{2} = 1.00$, $H(Y) = -\frac{1}{2}\log_2\frac{1}{2} - \frac{1}{2}\log_2\frac{1}{2} = 1.00$

And we can create the joint distribution table for $P(X,Y)$ as follows:

X/Y	T	F
T	40/100	10/100
F	10/100	40/100

From which, we get: $H(X,Y) = -\frac{4}{10}\log_2\frac{4}{10} - \frac{4}{10}\log_2\frac{4}{10} - \frac{1}{10}\log_2\frac{1}{10} - \frac{1}{10}\log_2\frac{1}{10} = 1.722$

Thus, $I(X;Y) = 1.00 + 1.00 - 1.722 = 0.278$

2.2 1.2

Similarly, we can compute $I(X;Z)$ as:

$$I(X;Z) = H(Z) - H(Z|X)$$

And using the same steps as above, we get:

$$I(X;Z) = H(X) + H(Z) - H(X,Z)$$

We know $H(X) = 1.00$.

We can also compute $H(Z) = -\frac{55}{100}\log_2\frac{55}{100} - \frac{45}{100}\log_2\frac{45}{100} = 0.990$

And we can create the joint distribution table for $P(X,Z)$ as follows:

X/Z	T	F
T	38/100	12/100
F	17/100	33/100

From which, we get: $H(X,Z) = -\frac{38}{100}\log_2\frac{38}{100} - \frac{12}{100}\log_2\frac{12}{100} - \frac{17}{100}\log_2\frac{17}{100} - \frac{33}{100}\log_2\frac{33}{100} = 1.852$

Thus, $I(X;Z) = 1 + 0.990 - 1.852 = 0.138$

2.3 1.3

Similarly, we can compute $I(Y;Z)$ as:

$$I(Y;Z) = H(Y) - H(Y|Z)$$

And using the same steps as above, we get:

$$I(Y;Z) = H(Y) + H(Z) - H(Y,Z)$$

We know $H(Y) = 1.00$.

We also know $H(Z) = 0.990$.

And we can create the joint distribution table for $P(Y,Z)$ as follows:

Z/Y	T	F
T	45/100	10/100
F	5/100	40/100

From which, we get: $H(Y,Z) = -\frac{45}{100}\log_2\frac{45}{100} - \frac{10}{100}\log_2\frac{10}{100} - \frac{5}{100}\log_2\frac{5}{100} - \frac{40}{100}\log_2\frac{40}{100} = 1.593$

Thus, $I(Y;Z) = 1.00 + 0.990 - 1.593 = 0.398$

2.4 1.4

Here, we have 3 random variables and hence, the minimum spanning tree will have 3 vertices and 2 edges. The Chow-Liu algorithm greedily picks the set of vertices having the maximum information gain and creates an edge between them.

Here, we have the information gain between all edges as follows:

$$I(Y;Z) = 0.398$$
$$I(X;Y) = 0.278$$
$$I(X;Z) = 0.138$$

Therefore, the Chow-Liu algorithm picks the highest information gain, i.e, $I(Y;Z)$ and creates and edge between them and then picks the next highest information gain, i.e., $I(X;Y)$ and creates an edge between them and stops as the minimum spanning tree is complete.

The resulting undirected tree is as follows:

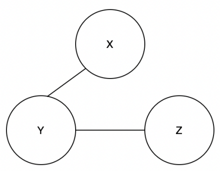

Figure 1: Undirected tree created by Chow-Liu algorithm

2.5 1.4

If we select X to be the root, we need to assign all the directed edges going away from the root. Hence, the tree becomes as:

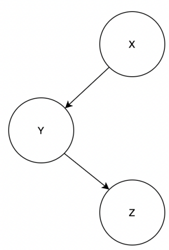

Figure 2: Directed tree after selecting X as the root