# United International University (UIU)

## Dept. of Computer Science & Engineering (CSE)

## Final Assignment: Spring 2020

**Course Code: CSI 341, Course Title: Artificial Intelligence**

Total Marks: 40 Duration: 2 hours

**Answer all questions**. Marks are indicated in the right side of each question

---

1. a. Two production lines produce the same part. Line 1 produces 1,000 parts per week of which 100 are defective. Line 2 produces 2,000 parts per week of which 150 are defective. If you choose a part randomly from the stock what is the probability it is defective? If it is defective what is the probability it was produced by line 1? **[2]**

   b. A survey has been done by a kids play zone to assess the interest of kids in different play activities. The data obtained is as follows:

   200 children were observed for the survey, half of them boys. Among the boys, 50 were 4 year old, 30 were 5 year old and the rest were 6 year old. Among the 4 year old boys, 40% liked the outdoor rides, 40% liked the sand box and the rest liked the puzzles and games. Among the 5 year old boys, the ratio was 4:2:4. In case of the 6 year old boys, it was 3:3:4. For the girls the statistics was a bit different. 40 were 4 year old, 20 were 5 year old and the rest were 6 year old. Among the 4 year old girls, 40% liked the outdoor rides, 30% liked the sand box and the rest liked the puzzles and games. Among the 5 year old girls, the ratio was 3:3:4. In case of the 6 year old girls, it was 3:2:5.

   i. Based on this data, construct a full joint probability distribution among the three random variables Gender(G), Age(A) and Play type(P). **[2]**

   ii. Calculate the following probabilities from your table:

   - Probability of a child being a 4 year old. **[1]**
   - Probability of a girl child not being interested in sand box. **[1]**
   - Probability of a child being a 6 year old given that he/she is interested in puzzles and games. **[1]**

2. a. Draw a bayes net to represent the following scenario: **[1]**

   Lots of factors affect the harvesting season in Bangladesh. If there is a cyclone or a flood, harvest is poor. When harvest is bad the economy is affected directly. The economy is also declined by low industrial production and also lower rate of tourism. Cyclone affects tourism in Bangladesh since most tourist destinations are near the sea.

   Now answer the following questions:

   i. How many probability entries are required for your bayes net? **[1]**

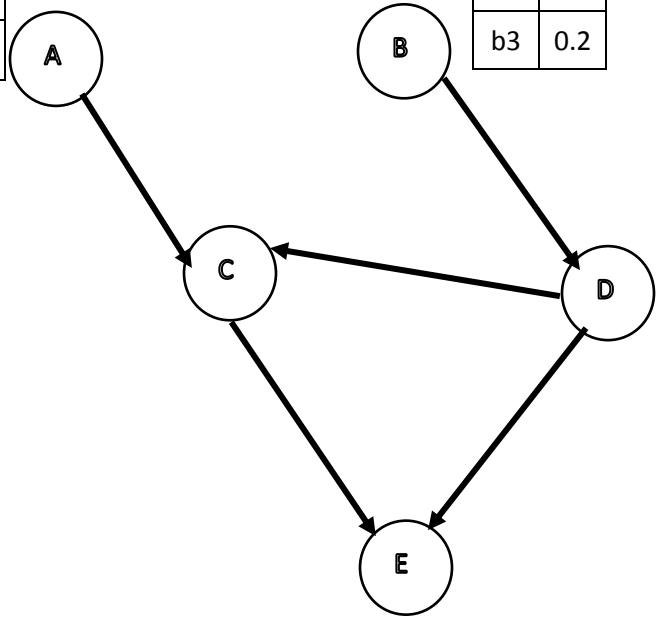
   ii. How many probability entries are required to represent your random variables in a full joint probability distribution? **[1]**

b. Draw the Bayesian Network that corresponds to the following conditional probability: **[1]**
P(H) P(G | H) P(F | G) P(E | C) P(D| C, E) P(C | F, G) P(B | F) P(A | B, C, D)

c. Consider the following Bayes Net with five random variables: A={a1, a2, a3}, B={b1, b2, b3}, C={c1, c2}, D={d1, d2}, E={e1, e2}. The conditional probability tables are given below:

| A | P(A) |
|---|---|
| a1 | 0.2 |
| a2 | 0.1 |
| a3 | 0.7 |

| B | P(B) |
|---|---|
| b1 | 0.5 |
| b2 | 0.3 |
| b3 | 0.2 |




| B | D | P(D\|B) |
|---|---|---|
| b1 | d1 | 0.75 |
| b1 | d2 | 0.25 |
| b2 | d1 | 0.55 |
| b2 | d2 | 0.45 |
| b3 | d1 | 0.85 |
| b3 | d2 | 0.15 |

| A | D | C | P(C\|A,D) |
|---|---|---|---|
| a1 | d1 | c1 | 0.5 |
| a1 | d1 | c2 | 0.5 |
| a1 | d2 | c1 | 0.4 |
| a1 | d2 | c2 | 0.6 |
| a2 | d1 | c1 | 0.2 |
| a2 | d1 | c2 | 0.8 |
| a2 | d2 | c1 | 0.7 |
| a2 | d2 | c2 | 0.3 |
| a3 | d1 | c1 | 0.1 |
| a3 | d1 | c2 | 0.9 |
| a3 | d2 | c1 | 0.8 |
| a3 | d2 | c2 | 0.2 |

| C | D | E | P(E\|C,D) |
|---|---|---|---|
| c1 | d1 | e1 | 035 |
| c1 | d1 | e2 | 0.65 |
| c1 | d2 | e1 | 0.15 |
| c1 | d2 | e2 | 0.85 |
| c2 | d1 | e1 | 0.8 |
| c2 | d1 | e2 | 0.2 |
| c2 | d2 | e1 | 0.7 |
| c2 | d2 | e2 | 0.3 |

Now find out the following probabilities:

i. P(d2) **[3]**
ii. P(e2|a3, c1, d2) **[1.5]**
iii. P(b1|c2, d1) **[2.5]**

3. In the Dark Ages, *Harvard, Dartmouth, and Yale* admitted only male students. Assume that, at that time, *80 percent* of the sons of Harvard men went to Harvard and the rest went to Yale, *40 percent* of the sons of Yale men went to Yale, and the rest split evenly between Harvard and Dartmouth; and of the sons of Dartmouth men, *70 percent* went to Dartmouth, *20 percent* to Harvard, and *10 percent* to Yale.

   Now model the scenario as a Markov model and answer the following questions:

   a. Show the transition probability matrix. **[1]**
   b. Calculate the probability that the grandson of a man from Yale went to Harvard. **[2]**
   c. Calculate the probability that the next generation (i.e. child) of Harvard man went to Yale and the next generation after that (i.e. grandchild) went to Harvard. **[2]**
   d. Find the stationary probability distribution for this Markov model. **[3]**

4. Consider a set of documents, each of which either categorized as *Sports* or *Informatics* category. These documents are categorized based on the presence of 8 words within that document: **{ goal, tutor, variance, speed, drink, defence, performance, field }**

**Training data set:**

| goal | Tutor | variance | speed | drink | defence | performance | field | category |
|---|---|---|---|---|---|---|---|---|
| Yes | No | No | No | Yes | Yes | Yes | Yes | *Sports* |
| No | No | Yes | No | Yes | Yes | No | No | *Sports* |
| No | Yes | No | Yes | No | Yes | Yes | No | *Sports* |
| No | No | Yes | No | Yes | No | Yes | No | *Informatics* |
| No | No | No | No | No | No | No | No | *Informatics* |
| No | No | Yes | Yes | No | No | Yes | Yes | *Sports* |
| No | Yes | Yes | No | No | Yes | No | No | *Informatics* |
| Yes | No | No | No | Yes | No | Yes | Yes | *Sports* |
| Yes | Yes | No | Yes | No | No | Yes | Yes | *Informatics* |
| No | Yes | Yes | No | No | No | Yes | No | *Informatics* |
| Yes | No | No | Yes | No | yes | No | Yes | *Sports* |

a. Now predict the category of a new document containing the following words: [goal=Yes, tutor=No, variance=No, speed=Yes, drink=Yes, defence=No, performance=Yes, field=Yes] **[5]**

b. Briefly explain why we need Laplacian Smoothing during Naïve Bayes Classification. **[2]**

5. a. Suppose for the next trimester 5 AI theory sections (A, B, C, D, E) have been offered. But only 4 faculties are interested to take the AI course. Also they have some time slot preferences. So your task is to assign the interested faculties to these 5 sections by satisfying their preferences.

   The preferences are given below:

   - Faculty 2 doesn't want to take section A and E
   - Faculty 3 doesn't want to take section D
   - Faculty 4 doesn't want to take section C, E
   - Same faculty can't be assigned to section D, E

   The class timings are:

   - Section A, 8:30 AM – 10:00 AM
   - Section B, 9:00 AM – 10:30 AM
   - Section C, 10:15 AM – 11:45 AM
   - Section D, 9:30 AM – 11:00 AM
   - Section E, 11:30 AM – 01:00 PM

   [Hint: same faculty can't be assigned to time conflicted sections]

   i. Formulate this problem as a CSP problem that is show the set of variables, set of domains and set of constraints. **[1+1+1]**

   ii. Show the constraint graph for this CSP problem. **[1]**

   b. Consider the following CSP problem: **[3]**

   Set of variables = {A, B, C, D, E}
   Set of domains = { {3,4}, {2,3,4}, {1,3,4}, {1,2,4}, {1,2,3,4} }
   Set of Constraints = { C < D, B < C, A ≠ E, E ≠ D, B < D }

   Now show the steps followed by the backtracking algorithm with the minimum remaining values heuristic and the degree heuristic.