Effat University

SPRING 20223

COLLEGE OF ENGINEERING – DEPARTMENT OF COMPUTER SCIENCE

Lab 5 Bayesian Network May 12, 2023

Name 1: Aicha Sidiya – S20106146 Section: 1

Name 2: Razan Almahdi – S20106649

1- Solve the following problems.
2- Find a BN code, use any BN and infer a query, then give a conclusion of the output.

Problem 1:

In this example, let us imagine that we are given the task of modeling a student's marks (***m***) for an exam he has just given. From the given Bayesian Network Graph below, we see that the marks depend upon two other variables. They are,

- Exam Level (***e***)– This discrete variable denotes the difficulty of the exam and has two values (0 for easy and 1 for difficult)
- IQ Level (***i***) – This represents the Intelligence Quotient level of the student and is also discrete in nature having two values (0 for low and 1 for high)

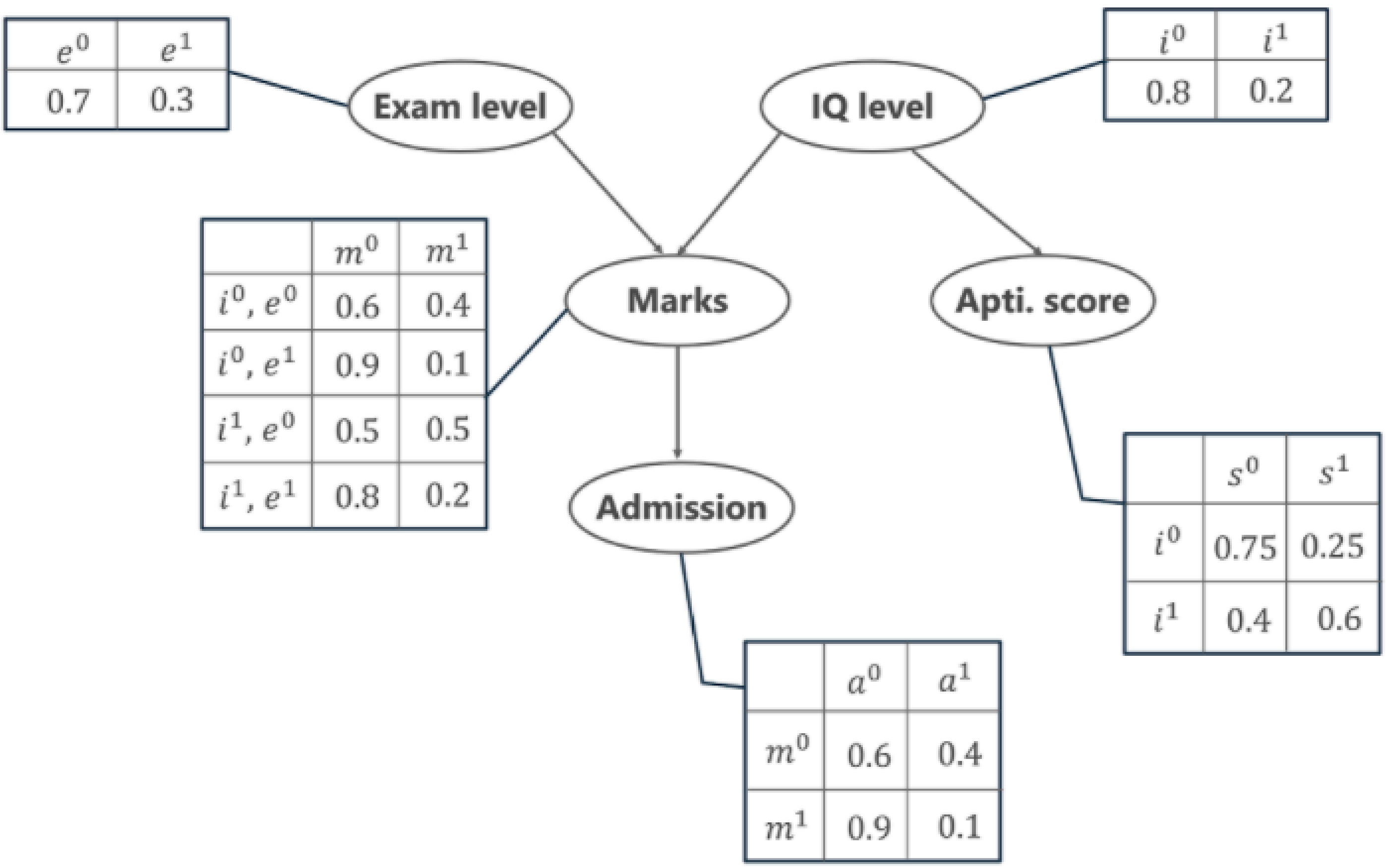
Additionally, the IQ level of the student also leads us to another variable, which is the Aptitude Score of the student (***s***). Now, with marks the student has scored, he can secure admission to a particular university. The probability distribution for getting admitted (***a***) to a university is also given below.

In the below graph, we see several tables representing the probability distribution values of the given 5 variables. These tables are called the Conditional Probabilities Table or CPT. There are a few properties of the CPT given below –

- The sum of the CPT values in each row must be equal to 1 because all the possible cases for a particular variable are exhaustive (representing all possibilities).
- If a variable that is Boolean in nature has k Boolean parents, then in the CPT it has 2K probability values.

Coming back to our problem, let us first list all the possible events that are occurring in the above-given table.

1. Exam Level (e)
2. IQ Level (i)
3. Aptitude Score (s)
4. Marks (m)
5. Admission (a)



e^0	e^1
0.7	0.3

i^0	i^1
0.8	0.2

	m^0	m^1
i^0, e^0	0.6	0.4
i^0, e^1	0.9	0.1
i^1, e^0	0.5	0.5
i^1, e^1	0.8	0.2

	s^0	s^1
i^0	0.75	0.25
i^1	0.4	0.6

	a^0	a^1
m^0	0.6	0.4
m^1	0.9	0.1

CALCULATION OF JOINT PROBABILITY DISTRIBUTION

P[a, m, i, e, s]= P(a | m) . P(m | i, e) . P(i) . P(e) . P(s | i)

Case 1: Calculate the probability that in spite of the exam level being difficult, the student having a low IQ level and a low Aptitude Score, manages to pass the exam and secure admission to the university.
Solution:

P(a = 1, m = 1, i = 0, e = 1, s = 0)
= P(a = 1|m = 1) . P(m = 1|i = 0, e = 1) . P(i = 0) . P(e = 1) . P(s = 0|i = 0)
$= 0.1 \times 0.1 \times 0.8 \times 0.3 \times 0.75 = 0.0018$

There is 0.18% chance that the student with a low IQ passed a difficult exam.

Case 2: In another case, calculate the probability that the student has a High IQ level and Aptitude Score, the exam being easy yet fails to pass and does not secure admission to the university.
Solution:

P(a = 0, m = 0, i = 1, e = 0, s = 1)
= P(a = 0|m = 0) . P(m = 0|i = 1, e = 0) . P(i = 1) . P(e = 0) . P(s = 1|i = 1)
$= 0.6 \times 0.5 \times 0.2 \times 0.7 \times 0.6 = 0.0252$

There is 2.52% chance that the student with a high IQ failed an easy exam.

Problem 2:

When IRS receives tax forms, it puts them through a computer to flag forms that need to be investigated further. The computer looks for mistakes in the forms, for example addition mistakes or incorrect deduction amounts. Suppose the computer correctly flags 80% of all returns that have mistaken, and it incorrectly flags 5% of error-free returns. Further, suppose that 15% of all tax returns have errors.

A tax return is flagged by the computer. What is the chance that it actually contains mistakes, given that the computer flagged it?

Solution:
P(F|E) = 0.8
P(F|$\bar{E}$) = 0.05
P(E) = 0.15
P($\bar{E}$) $= 0.85$
P(F) = P(F|E).P(E) + P(F|$\bar{E}$).P($\bar{E}$)
P(E|F) = $\frac{P(F|E).P(E)}{P(F)} = \frac{0.8 \times 0.15}{0.8\times0.15 + 0.05\times0.85} = 0.74$

There is 74% chance that the return has an error.

Problem 3:

Legal cases of disputed paternity in many countries are resolved using blood tests. Laboratories make genetic determinations concerning the mother, child, and alleged father.

You are on a jury considering a paternity suit. The mother has blood type O, and the alleged father has blood type AB.

A blood test shows that the child has blood type B. What is the chance that the alleged father is in fact the real father, given that the child has blood type B?

Here's some information we need to solve the problem. According to genetics, there is a 50% chance that this child will have blood type B if this alleged father is the real father. Furthermore, based on incidence rates of B genes in the population, there is a 9% chance that this child would have blood type B if this alleged father is not the real father.

Based on other evidence (e.g., testimonials, physical evidence, records) presented before the DNA test, you believe there is a 75% chance that the alleged father is the real father. This assessment is your prior belief. Now, we need to use Bayes Rule to update it for the results of the child's blood test.

Solution:

$P(B|R) = 0.5$

$P(B|\bar{R}) = 0.09$

$P(R) = 0.75$

$P(\bar{R}) = 0.25$

$P(B) = P(B|R).P(R) + P(B|\bar{R}).P(\bar{R})$

$$P(R|B) = \frac{P(B|R) * P(R)}{P(B)} = \frac{0.5 \times 0.75}{0.5 \times 0.75 + 0.09 \times 0.25} = 0.94$$

There is 94% chance that the alleged father is the real father.

Problem 4:

The doctor has prior information that 90% of sick children in that neighborhood have the flu, while the other 10% are sick with measles. Let F stand for an event of a child being sick with flu and M stand for an event of a child being sick with measles. Assume for simplicity that $F \cup M = \Omega$, i.e., that there no other maladies in that neighborhood.

A well-known symptom of measles is a rash (the event of having which we denote R). Assume that the probability of having a rash if one has measles is $P(R \mid M) = 0.95$. However, occasionally children with flu also develop rash, and the probability of having a rash if one has flu is $P(R \mid F) = 0.08$.

Upon examining the child, the doctor finds a rash. What is the probability that the child has measles?

Solution:

Substituting in these values and using the given information that 90% of sick children have the flu and 10% have measles, we get:

$P(F) = 0.90.$

$P(M) = 0.10.$

$P(R \mid M) = 0.95$

$P(R \mid F) = 0.08.$

$P(R) = P(R \mid M).P(M) + P(R \mid F).P(F) = (0.95)(0.10) + (0.08)(0.90) = 0.167$

Finally, we can substitute all of these values into Bayes' theorem to get:

$P(M \mid R) = P(R \mid M) * P(M) / P(R)$

$= 0.95 * 0.1 / 0.167$

$= 0.57$

Therefore, the probability that the child has measles given that the child has a rash is 0.625 or about 62.5%.

Problem 5:

It is estimated that 50% of emails are spam emails. Some software has been applied to filter these spam emails before they reach your inbox. A certain brand of software claims that it can detect 99% of spam emails, and the probability for a false positive (a non-spam email detected as spam) is 5%.

Now if an email is detected as spam, then what is the probability that it is in fact a non-spam email?

Solution:

To find P(B), we can use the law of total probability:

$P(B) = P(B \mid A) * P(A) + P(B \mid A') * P(A')$

$= 0.99 * 0.5 + 0.05 * 0.5$

$= 0.52$

Substituting in the values we have, we get:

$P(A' \mid B) = P(B \mid A') * P(A') / P(B)$

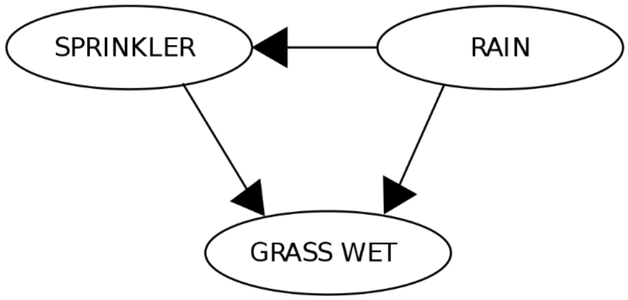
$= 0.05 * 0.5 / 0.52$

$= 0.048$

Therefore, the probability that an email is not spam given that the filtering software detects it as spam is about 4.8%.

Problem 6:

RAIN	SPRINKLER T	SPRINKLER F
F	0.4	0.6
T	0.01	0.99



RAIN T	RAIN F
0.2	0.8

SPRINKLER	RAIN	GRASS WET T	GRASS WET F
F	F	0.0	1.0
F	T	0.8	0.2
T	F	0.9	0.1
T	T	0.99	0.01

Two events can cause grass to be wet: an active sprinkler or rain. Rain has a direct effect on the use of the sprinkler (namely that when it rains, the sprinkler usually is not active). This situation can be modeled with a Bayesian network. Each variable has two possible values, T (for true) and F (for false).

The joint probability function is, by the chain rule of probability,

$$\Pr(G,S,R) = \Pr(G \mid S,R)\Pr(S \mid R)\Pr(R)$$

where G = "Grass wet (true/false)", S = "Sprinkler turned on (true/false)", and R = "Raining (true/false)".

"What is the probability that it is raining, given the grass is wet?"

Solution:

P(R = T | G = T)
= P(R = T, G = T) / P(G = T)
= (P(G = T, S = T, R = T) + P(G = T, S = F, R = T)) / P(G =T)

P(G = T)
= P(G = T, S = T, R = T) +
P(G = T, S = F, R = T) +
P(G = T, S = T, R = F) +
P(G = T, S = F, R = F)

= 0.99 * 0.01 * 0.2 + 0.8 * 0.99 * 0.2 + 0.9 * 0.4 * 0.8 + 0.0 * 0.6 * 0.8
= 0.00198 + 0.1584 + 0.288 + 0 = 0.44838

P(R = T | G = T) = $\frac{0.00198+0.1584}{0.44838} = 0.3577$

Therefore, the probability that it is raining, given the grass is wet is about 35.77%.