Assignment 13 Solutions

1. Provide an example of the concepts of Prior, Posterior, and Likelihood ?

```
Prior: The prior probability refers to the initial belief or probability of
an event before any new evidence is observed.
In this case, let's say we want to determine the probability of a patient
having a certain disease (D) before conducting any tests.
The prior probability of the patient having the disease could be based on
factors such as prevalence in the general population or
the patient's medical history.

Likelihood: The likelihood is the probability of observing a particular
evidence or test result given a specific hypothesis or condition.
In our example, the likelihood would be the probability of obtaining a
positive test result (T+) given that the patient actually has the
disease (D). The likelihood is typically determined through previous data or
knowledge about the accuracy of the test.

Posterior: The posterior probability is the updated probability of a
hypothesis or condition after taking into account the
observed evidence or test result. Using Bayesian inference, we can calculate
the posterior probability by combining the prior
probability and the likelihood. In our example, the posterior probability
would be the probability of the patient having the disease (D)
given the positive test result (T+).
```

2. What role does Bayes' theorem play in the concept learning principle ?

```
Bayes' theorem is a fundamental principle in the concept learning process as
it provides a framework for updating beliefs or
probabilities based on new evidence. It allows for the calculation of the
posterior probability, which represents the updated
probability of a hypothesis or concept given the observed evidence. By
incorporating prior knowledge and the likelihood of the
evidence, Bayes' theorem helps in making informed decisions and updating
hypotheses in the concept learning process.
```

3. Offer an example of how the Nave Bayes classifier is used in real life ?

```
A common example of the Naive Bayes classifier in real life is spam email
filtering. It is used to classify incoming emails as either
spam or non-spam based on the presence of certain features or words. The
classifier learns from a training dataset containing labeled
emails (spam or non-spam) and calculates the probabilities of different
features occurring in each class. When a new email arrives,
the Naive Bayes classifier uses these probabilities to calculate the
likelihood of it being spam or non-spam. This helps in effectively
```

```
filtering out unwanted spam emails and improving the email user experience.
```

4. Can the Nave Bayes classifier be used on continuous numeric data? If so, how can you go about doing it ?

```
Yes, the Naive Bayes classifier can be used on continuous numeric data by
discretizing it into categories or bins. Discretization can
be done using techniques like equal width or equal frequency binning. Once
the data is discretized, the Naive Bayes classifier can be
applied to calculate the probabilities of each category and make predictions
based on the likelihood of the observed features.
```

5. What are Bayesian Belief Networks, and how do they work? What are their applications? Are they capable of resolving a wide range of issues ?

```
Bayesian Belief Networks (BBNs) are graphical models that represent and
reason about uncertain relationships between variables using
Bayesian probability theory. They consist of nodes representing variables
and directed edges representing probabilistic dependencies
between variables. BBNs use conditional probability distributions to model
the relationships and perform probabilistic inference.

BBNs find applications in various fields, including healthcare, finance,
risk analysis, and decision support systems.
They can be used for risk assessment, diagnosis, prediction, decision-
making, and modeling complex systems. BBNs excel in handling
uncertainty, incorporating prior knowledge, and updating beliefs based on
observed evidence.
```

6. Passengers are checked in an airport screening system to see if there is an intruder. Let I be the random variable that indicates whether someone is an intruder I = 1) or not I = 0), and A be the variable that indicates alarm I = 0). If an intruder is detected with probability P(A = 1|I = 1) = 0.98 and a non-intruder is detected with probability P(A = 1|I = 0) = 0.001, an alarm will be triggered, implying the error factor. The likelihood of an intruder in the passenger population is P(I = 1) = 0.00001. What are the chances that an alarm would be triggered when an individual is actually an intruder ?

```
To find the chances that an alarm would be triggered when an individual is
actually an intruder, we need to calculate the conditional
probability P(I = 1|A = 1), which represents the probability of being an
intruder given that an alarm has been triggered.

We can use Bayes' theorem to calculate this probability:

P(I = 1|A = 1) = (P(A = 1|I = 1) * P(I = 1)) / P(A = 1)

Given:
P(A = 1|I = 1) = 0.98
P(A = 1|I = 0) = 0.001
P(I = 1) = 0.00001

We need to calculate P(A = 1), the probability of an alarm being triggered.
This can be calculated using the law of total probability:
```

```
P(A = 1) = P(A = 1|I = 1) * P(I = 1) + P(A = 1|I = 0) * P(I = 0)

Since P(I = 0) = 1 - P(I = 1), we can substitute it into the equation:

P(A = 1) = P(A = 1|I = 1) * P(I = 1) + P(A = 1|I = 0) * (1 - P(I = 1))

Substituting the given values:

P(A = 1) = 0.98 * 0.00001 + 0.001 * (1 - 0.00001)

Now, we can calculate P(I = 1|A = 1) using Bayes' theorem:

P(I = 1|A = 1) = (P(A = 1|I = 1) * P(I = 1)) / P(A = 1)

Substituting the values:

P(I = 1|A = 1) = (0.98 * 0.00001) / P(A = 1)

By substituting the previously calculated value of P(A = 1), we can
determine the chances that an alarm would be triggered when an individual is
actually an intruder.
```

7. An antibiotic resistance test (random variable T) has 1% false positives (i.e., 1% of those who are not immune to an antibiotic display a positive result in the test) and 5% false negatives (i.e., 1% of those who are not resistant to an antibiotic show a positive result in the test) (i.e. 5 percent of those actually resistant to an antibiotic test negative). Assume that 2% of those who were screened were antibiotic-resistant. Calculate the likelihood that a person who tests positive is actually immune (random variable D) ?

```
To calculate the likelihood that a person who tests positive is actually
immune (D), we need to find the conditional probability
P(D = 1|T = 1), which represents the probability of being immune given a
positive test result.

We can use Bayes' theorem to calculate this probability:

P(D = 1|T = 1) = (P(T = 1|D = 1) * P(D = 1)) / P(T = 1)

Given:
P(T = 1|D = 0) = 0.01 (false positive rate)
P(T = 1|D = 1) = 0.95 (true positive rate)
P(D = 1) = 0.02 (proportion of antibiotic-resistant individuals)

We need to calculate P(T = 1), the probability of a positive test result.
This can be calculated using the law of total probability:

P(T = 1) = P(T = 1|D = 0) * P(D = 0) + P(T = 1|D = 1) * P(D = 1)

Since P(D = 0) = 1 - P(D = 1), we can substitute it into the equation:

P(T = 1) = P(T = 1|D = 0) * (1 - P(D = 1)) + P(T = 1|D = 1) * P(D = 1)

Substituting the given values:
```

```
P(T = 1) = 0.01 * (1 - 0.02) + 0.95 * 0.02

Now, we can calculate P(D = 1|T = 1) using Bayes' theorem:

P(D = 1|T = 1) = (P(T = 1|D = 1) * P(D = 1)) / P(T = 1)

Substituting the values:

P(D = 1|T = 1) = (0.95 * 0.02) / P(T = 1)

By substituting the previously calculated value of P(T = 1), we can
determine the likelihood that a person who tests positive is
actually immune.
```

8. In order to prepare for the test, a student knows that there will be one question in the exam that is either form A, B, or C. The chances of getting an A, B, or C on the exam are 30 percent, 20%, and 50 percent, respectively. During the planning, the student solved 9 of 10 type A problems, 2 of 10 type B problems, and 6 of 10 type C problems.

1. What is the likelihood that the student can solve the exam problem?
2. Given the student's solution, what is the likelihood that the problem was of form A?

```
To calculate the likelihood that the student can solve the exam problem, we
need to consider the student's performance on each type
of problem.
The probability of the student solving a problem of form A is 9 out of 10,
which is 0.9.
The probability of the student solving a problem of form B is 2 out of 10,
which is 0.2.
The probability of the student solving a problem of form C is 6 out of 10,
which is 0.6.

Since there is only one question on the exam and it can be of form A, B, or
C, we can calculate the overall likelihood as the weighted average of the
probabilities for each form:

Likelihood = (Probability of form A * Likelihood of solving form A) +
(Probability of form B * Likelihood of solving form B) + (Probability of
form C * Likelihood of solving form C)
= (0.3 * 0.9) + (0.2 * 0.2) + (0.5 * 0.6)
= 0.27 + 0.04 + 0.3
= 0.61

Therefore, the likelihood that the student can solve the exam problem is
0.61, or 61%.

Given the student's solution, we can calculate the likelihood that the
problem was of form A using Bayes' theorem.
Let P(A) be the probability of the problem being of form A, which is 0.3.
Let P(S|A) be the probability of the student solving a form A problem, which
is 0.9.

We need to find P(A|S), the probability of the problem being of form A given
that the student solved it.
```

```
Using Bayes' theorem:
P(A|S) = (P(S|A) * P(A)) / P(S)

To calculate P(S), we need to consider the probabilities of the student
solving problems of all forms:

P(S) = (P(S|A) * P(A)) + (P(S|B) * P(B)) + (P(S|C) * P(C))
= (0.9 * 0.3) + (0.2 * 0.2) + (0.6 * 0.5)
= 0.27 + 0.04 + 0.3
= 0.61

Now, we can substitute the values into Bayes' theorem:

P(A|S) = (0.9 * 0.3) / 0.61
= 0.27 / 0.61
≈ 0.443

Therefore, the likelihood that the problem was of form A given the student's
solution is approximately 0.443, or 44.3%.
```

9. A bank installs a CCTV system to track and photograph incoming customers. Despite the constant influx of customers, we divide the timeline into 5 minute bins. There may be a customer coming into the bank with a 5% chance in each 5-minute time period, or there may be no customer (again, for simplicity, we assume that either there is 1 customer or none, not the case of multiple customers). If there is a client, the CCTV will detect them with a 99 percent probability. If there is no customer, the camera can take a false photograph with a 10% chance of detecting movement from other objects.

1. How many customers come into the bank on a daily basis (10 hours)?
2. On a daily basis, how many fake photographs (photographs taken when there is no customer) and how many missed photographs (photographs taken when there is a customer) are there?
3. Explain likelihood that there is a customer if there is a photograph?

```
To calculate the number of customers coming into the bank on a daily basis
(10 hours), we need to determine the number of 5-minute time periods in 10
hours and multiply it by the probability of a customer coming in each time
period.
There are 10 hours * 60 minutes / 5 minutes = 120 time periods in 10 hours.

The probability of a customer coming in each time period is 5%.

Therefore, the number of customers coming into the bank on a daily basis is
120 time periods * 5% = 6 customers.

On a daily basis, the number of fake photographs (photographs taken when
there is no customer) can be calculated by multiplying the number of time
periods without a customer by the probability of a false photograph being
taken.
The number of time periods without a customer is 120 time periods - 6
customers = 114 time periods.

The probability of a false photograph being taken when there is no customer
is 10%.
```

```
Therefore, the number of fake photographs on a daily basis is 114 time
periods * 10% = 11.4 photographs, which we can round to 11 photographs.

The number of missed photographs (photographs taken when there is a
customer) on a daily basis is equal to the number of customers coming into
the bank, which is 6 customers.

The likelihood that there is a customer if there is a photograph can be
calculated using Bayes' theorem.
Let P(C) be the probability of a customer being present, which is 5%.
Let P(P) be the probability of a photograph being taken, which is equal to
the number of time periods (120) divided by the total number of time periods
without a customer (114 + 6 = 120).

We need to find P(C|P), the probability of a customer being present given
that a photograph is taken.

Using Bayes' theorem:
P(C|P) = (P(P|C) * P(C)) / P(P)

The probability of a photograph being taken when there is a customer is 99%.

Substituting the values into Bayes' theorem:
P(C|P) = (0.99 * 0.05) / P(P)

Since P(P) is equal to 1 (as there will always be a photograph taken in each
time period), we can simplify the equation:

P(C|P) = 0.99 * 0.05
= 0.0495

Therefore, the likelihood that there is a customer if there is a photograph
is 0.0495, or approximately 4.95%.
```

10. Create the conditional probability table associated with the node Won Toss in the Bayesian Belief network to represent the conditional independence assumptions of the Nave Bayes classifier for the match winning prediction problem in Section 6.4.4.?

```
To create the conditional probability table (CPT) associated with the node
"Won Toss" in the Naive Bayes classifier for the match
winning prediction problem, we need to define the conditional probabilities
of "Won Toss" given each feature in the classifier.
However, without the specific details of the features and their values
mentioned in Section 6.4.4 of the context you provided,
it is not possible to create the exact CPT.

The CPT for the "Won Toss" node would typically include the conditional
probabilities of "Won Toss" being true or false given different
combinations of feature values. Each row in the table represents a specific
combination of feature values, and the columns represent
the probabilities of "Won Toss" being true or false.

Here's a general example of how the CPT for "Won Toss" could look:

Feature 1   Feature 2   Feature 3   ... P(Won Toss = True)  P(Won Toss =
False)
```

```
Value 1 Value 1 Value 1 ... 0.7 0.3
Value 1 Value 1 Value 2 ... 0.6 0.4
Value 1 Value 2 Value 1 ... 0.8 0.2
Value 2 Value 1 Value 1 ... 0.9 0.1
Value 2 Value 2 Value 2 ... 0.5 0.5
 ...    ... ... ... ... ...
In this example, "Feature 1," "Feature 2," "Feature 3," etc., represent the
different features used in the Naive Bayes classifier.
Each combination of feature values has an associated probability of "Won
Toss" being true or false.

Please note that this is a simplified example, and the actual CPT for the
"Won Toss" node would depend on the specific features and
their values used in the match winning prediction problem mentioned in
Section 6.4.4.
```