Machine Learning #13

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

Let's say we want to determine the probability that a person has diabetes given that they have a positive test result. Here, the concept of Prior, Posterior, and Likelihood can be demonstrated as follows:

- ***Prior probability:*** ***The probability of having diabetes before we perform the test.*** Let's assume that 10% of the population has diabetes, so the prior probability is 0.1.
- ***Likelihood probability:*** ***The probability of getting a positive test result given that the person has diabetes.*** Let's assume that the sensitivity of the test is 80%, so the likelihood probability is 0.8.
- ***Posterior probability:*** ***The probability of having diabetes given that the person has a positive test result.*** We can calculate the posterior probability using Bayes' theorem:

```
Posterior probability = (Likelihood probability * Prior probability) / Probability of a positive test res
ult
```

Let's assume that the specificity of the test is also 80%. This means that there is a 20% chance of getting a false positive result. Therefore, the probability of a positive test result is:

```
Probability of a positive test result = (Likelihood probability * Prior probability) + (Probability of a
 false positive * Probability of not having diabetes)
= (0.8 * 0.1) + (0.2 * 0.9) = 0.26
```

Now, we can calculate the posterior probability:

```
Posterior probability = (0.8 * 0.1) / 0.26 = 0.308
```

This means that if a person has a positive test result, there is a 30.8% chance that they actually have diabetes.

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

Bayes' theorem plays a fundamental role in the concept learning principle as it provides a mathematical framework to update the probabilities of hypotheses in the light of new evidence. In concept learning, the goal is to infer a hypothesis or concept from a set of examples or observations. ***Bayes' theorem is used to calculate the probability of a hypothesis given the evidence or observations.***

The concept learning principle follows a Bayesian approach by starting with a prior probability distribution over the space of hypotheses and updating it based on the observed data. The posterior probability distribution obtained after applying Bayes' theorem represents the updated belief about the hypotheses. The most probable hypothesis is then selected as the final concept or hypothesis.

Bayes' theorem provides a principled way to incorporate new evidence or data into the learning process and adjust the probabilities of hypotheses accordingly. This enables the learning system to adapt to new situations and make more accurate predictions based on the available evidence.

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

One example of how the Naïve Bayes classifier is used in real life is in spam filtering. The classifier can be trained on a dataset of emails, where each email is labeled as either "spam" or "not spam." The classifier then uses the probabilities of certain words or features in the email (such as the presence of certain keywords) to predict whether a new email is likely to be spam or not. This allows the email provider to automatically filter out unwanted messages and protect the user's inbox from being filled with irrelevant or harmful content.

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

Yes, the Naïve Bayes classifier can be used on continuous numeric data. One common way to handle continuous data is to assume a Gaussian (normal) distribution and use the mean and variance of the data to calculate the probability of a given class. This approach is called Gaussian Naïve Bayes. Another approach is to discretize the continuous data into a set of intervals or bins and treat the data as categorical. This approach is called Discretized Naïve Bayes. The choice between these approaches depends on the nature of the data and the problem at hand.

▼ 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 probabilistic graphical models that can represent and reason about uncertain knowledge in a domain. BBNs are composed of nodes and directed edges, with each node representing a random variable and each directed edge indicating a conditional dependence between the connected nodes. The directed acyclic graph (DAG) structure of BBNs makes them efficient and powerful tools for probabilistic reasoning and decision making under uncertainty.

BBNs work by using Bayes' theorem to update the probabilities of a node given new evidence. The probability of a node is initially set based on prior knowledge, which is then updated as new evidence is observed, and the probabilities of connected nodes are also updated accordingly. The final probability of a node is called its posterior probability, which represents the updated belief about that node given the observed evidence.

BBNs have a wide range of applications, including decision making, risk analysis, diagnosis, prediction, and classification. They can be used in various fields such as healthcare, finance, engineering, and environmental science. For example, BBNs can be used to diagnose medical conditions based on patient symptoms, predict equipment failure in manufacturing, and evaluate the impact of climate change on ecosystems.

BBNs are capable of resolving a wide range of issues because of their ability to handle complex probabilistic relationships between variables and to reason under uncertainty. However, the accuracy and effectiveness of BBNs depend on the quality of the prior knowledge and the quality and quantity of available data. Moreover, building a BBN can be a challenging task that requires expert knowledge and skills in probability theory, statistics, and computer science.

▼ 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?

We can use Bayes' theorem to calculate the probability that an individual is actually an intruder given that an alarm has been triggered.

Let B be the event that an alarm is triggered. Then, we want to calculate P(I=1|B), the probability that an individual is an intruder given that an alarm has been triggered.

By Bayes' theorem, we have:

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

where P(B|I=1) is the probability of an alarm being triggered when an individual is an intruder, P(I=1) is the prior probability of an individual being an intruder, and P(B) is the total probability of an alarm being triggered.

We can calculate these probabilities as follows:

P(B|I=1) = 0.98 (given in the problem statement)
P(B|I=0) = 0.001 (given in the problem statement)
P(I=1) = 0.00001 (given in the problem statement)
P(I=0) = 1 - P(I=1) = 0.99999

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

P(B) = P(B|I=1) * P(I=1) + P(B|I=0) * P(I=0)

= 0.98 * 0.00001 + 0.001 * 0.99999

= 0.0010098

Finally, we can substitute these values into Bayes' theorem:

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

= 0.98 * 0.00001 / 0.0010098

= 0.0097

Therefore, the probability that an individual is actually an intruder given that an alarm has been triggered is 0.0097 or approximately 0.97%.

▼ 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).

Let's first define the events and the given probabilities:

- D: person is immune to antibiotic (what we want to find)
- T: test result is positive
- P(T=1 | D=0) = 0.01: probability of a false positive (test result is positive, but person is not immune)
- P(T=0 | D=1) = 0.05: probability of a false negative (test result is negative, but person is immune)

- P(D=1) = 0.02: probability of a person being immune to antibiotic (prior probability)

We want to calculate P(D=1 | T=1), the probability of a person being immune given that the test result is positive. Using Bayes' theorem, we can write:

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

The denominator P(T=1) is the total probability of getting a positive test result, which can be calculated using the law of total probability:

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

Substituting the values:

P(T=1) = 0.95 * 0.02 + 0.01 * 0.98 = 0.0291

Now we can calculate P(D=1 | T=1):

P(D=1 | T=1) = 0.95 * 0.02 / 0.0291 = 0.651

Therefore, the likelihood that a person who tests positive is actually immune is 0.651 or about 65.1%.

▼ 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?

1. The likelihood that the student can solve the exam problem can be calculated using Bayes' theorem. Let A be the event that the problem is of form A, B be the event that the problem is of form B, and C be the event that the problem is of form C. Let S be the event that the student can solve the problem. Then, we have:

   P(A) = 0.3
   P(B) = 0.2
   P(C) = 0.5
   P(S|A) = 9/10 = 0.9
   P(S|B) = 2/10 = 0.2
   P(S|C) = 6/10 = 0.6

   The probability of the student being able to solve the problem can be calculated as:

   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
   = 0.51

   Therefore, the likelihood that the student can solve the exam problem is 0.51.

2. Given that the student can solve the problem, the likelihood that the problem was of form A can be calculated using Bayes' theorem as well. Let's use the same notation as before. Then, we have:

   P(A|S) = P(S|A)P(A)/P(S)
   = 0.9*0.3/0.51
   = 0.529

Therefore, the likelihood that the problem was of form A given the student's solution is 0.529.

▼ 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?

1. The bank is open for 10 hours, which is 600 minutes. Since there is a 5% chance of a customer coming in each 5-minute time period, the expected number of customers in a 10-hour period is:

Expected number of customers = 600/5 * 0.05 = 60

Therefore, we can expect an average of 60 customers to come into the bank during a 10-hour period.

1. For each 5-minute time period, there are four possible outcomes: (1) a customer comes in and is detected, (2) a customer comes in and is not detected, (3) no customer comes in, but a false photograph is taken, and (4) no customer comes in and no photograph is taken.

Using the probabilities given, we can calculate the expected number of photographs in each category per day:

- Photographs of customers: 60 * 0.99 = 59.4 (since the detection probability is 99%)
- Missed photographs: 60 * 0.01 = 0.6 (since the probability of not detecting a customer is 1%)
- False photographs: 540 * 0.1 = 54 (since there are 540 5-minute time periods without a customer)

Therefore, we can expect an average of 59.4 photographs of customers, 0.6 missed photographs, and 54 false photographs per day.

1. If there is a photograph, there are two possible scenarios: either a customer came in and was detected, or there was no customer, but a false photograph was taken. Let P(C) denote the probability that a customer came in and P(F) denote the probability that a false photograph was taken. Then, the probability of a photograph given that there is a customer is:

P(Photograph | C) = P(Detection | C) = 0.99

Similarly, the probability of a photograph given that there is no customer is:

P(Photograph | F) = P(False photograph | F) = 0.1

Using Bayes' theorem, we can calculate the probability of a customer given that there is a photograph:

P(C | Photograph) = P(Photograph | C) * P(C) / [P(Photograph | C) * P(C) + P(Photograph | F) * P(F)]

Substituting the probabilities, we get:

P(C | Photograph) = 0.99 * 0.05 / [0.99 * 0.05 + 0.1 * (1 - 0.05)] = 0.33

Therefore, if there is a photograph, the probability that there is a customer is about 33%.

▼ 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.

Here is an example conditional probability table for the "Won Toss" node in the Bayesian Belief Network for the match winning prediction problem:

Won Toss	P(Won Toss)
True	0.5
False	0.5

In this table, "Won Toss" is the node, and the probabilities are listed for each of the two possible outcomes: "True" (i.e., the team won the toss) and "False" (i.e., the team lost the toss). The probabilities for each outcome should sum to 1.0, and in this example, they are both 0.5, which reflects the assumption of no prior knowledge or bias about the outcome of the toss.