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

Ans : The Bayes rule can be expressed in many forms. The simplest one is in terms of odds. The idea is to take the odds for something happening (against it not happening), which we'll write as prior odds.

The word prior refers to our assessment of the odds before obtaining some new information that may be relevant. The purpose of the formula is to update the prior odds when new information becomes available, to obtain the posterior odds, or the odds after obtaining the information (the dictionary meaning of posterior is "something that comes after, later.)"

The likelihood ratio is the probability of the observation in case the event of interest, divided by the probability of the observation in case of no event.

Example: the odds that it will rain later today. Imagine getting up in the morning in Finland. The chances of rain are 206 in 365, so prior odds = 206:159.

Suppose the chances of having a cloudy morning on a rainy day are 9 out of 10 - that means that only one out of 10 rainy days start out with blue skies. The chances of clouds are nine times higher on a rainy day than on a rainless day.

Likelihood ratio = (9/10) / (1/10) = 9

BAYES RULE

The Bayes rule for converting prior odds into posterior odds as follows:

posterior odds = likelihood ratio × prior odds

Ex.:

Prior odds = 206:159

Likelihood odds = 9

Posterior odds = 206:159 × 9 = 1854:159

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

Ans : Bayes Theorem is also used widely in machine learning, where it is a simple, effective way to predict classes with precision and accuracy. The Bayesian method of calculating conditional probabilities is used in machine learning applications that involve classification tasks.Bayes theorem helps us to calculate the single term P(B|A) in terms of P(A|B), P(B), and P(A). This rule is very helpful in such scenarios where we have a good probability of P(A|B), P(B), and P(A) and need to determine the fourth term.Bayes' Theorem allows you to update the predicted probabilities of an event by incorporating new information. Bayes' Theorem was named after 18th-century mathematician Thomas Bayes. It is often employed in finance in calculating or updating risk evaluation.

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

Ans : Real time Prediction: Naive Bayes is an eager learning classifier and it is sure fast. Thus, it could be used for making predictions in real time. Multi class Prediction: This algorithm is also well known for multi class prediction feature. Here we can predict the probability of multiple classes of target variable.

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

Ans : , how do we perform classification using Naive Bayes when the data we have is continuous in nature. If an instance in test data set has a category that was not present during training then it will assign it “Zero” probability and won't be able to make prediction. This is known as Zero frequency problem.Gaussian Naive Bayes

This type of Naive Bayes is used when variables are continuous in nature. It assumes that all the variables have a normal distribution.

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 ?

Ans : it can also be used in various tasks including prediction, anomaly detection, diagnostics, automated insight, reasoning, time series prediction, and decision making under uncertainty. Bayesian Network can be used for building models from data and experts opinions, and it consists of two parts: Directed Acyclic Graph.Bayesian networks are a widely-used class of probabilistic graphical models. They consist of two parts: a structure and parameters. The structure is a directed acyclic graph (DAG) that expresses conditional independencies and dependencies among ran- dom variables associated with nodes.

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 ?

Ans : This can be solved directly with the Bayesian theorem.

P(T = 1|A = 1) = P(A = 1|T = 1)P(T = 1)

P(A = 1)

=

P(A = 1|T = 1)P(T = 1)

P(A = 1|T = 1)P(T = 1) + P(A = 1|T = 0)P(T = 0)

=

0.98 × 0.00001

0.98 × 0.00001 + 0.001 × (1 − 0.00001) = 0.0097

≈

0.00001

0.001

= 0.01

It is interesting that even though for any passenger it can be decided with high reliability (98% and 99.9%)

whether (s)he is a terrorist or not, if somebody gets arrested as a terrorist, (s)he is still most likely not a

terrorist (with a probability of 99%).

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) ?

Ans : T = p means Test positive,

T = n means Test negative,

D = p means person takes drug,

D = n means person does not take drugs

We know:

P(T = p|D = n) = 0.01 (false positives)

(false negatives) P(T = n|D = p) = 0.05 =⇒ P(T = p|D = p) = 0.95 (true positives)

P(D = p) = 0.02 =⇒ P(D = n) = 0.98

We want to know the probability that somebody who tests positive is actually taking drugs:

P(D = p|T = p) = P(T = p|D = p)P(D = p)

P(T = p)

(Bayes theorem)

We do not know P(T = p):

P(T = p) = P(T = p|D = p)P(D = p) + P(T = p|D = n)P(D = n)

We get:

P(D = p|T = p) = P(T = p|D = p)P(D = p)

P(T = p)

=

P(T = p|D = p)P(D = p)

P(T = p|D = p)P(D = p) + P(T = p|D = n)P(D = n)

=

0.95 · 0.02

0.95 · 0.02 + 0.01 · 0.98

= 0.019/0.0288 ≈ 0.66

There is a chance of only two thirds that someone with a positive test is actually taking drugs.

An alternative way to solve this exercise is using decision trees. Let’s assume there are 1000 people tested.

Now we can put this together in a contingency table:

D = p D = n sum

T = p 19 9.8 28.8

T = n 1 970.2 971.2

sum 20 980 1000

To determine the probability that somebody who tests positive is actually taking drugs we have to calculate:

taking drugs and positive test

all positive test =

19

28.8

≈ 0.66

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.

What is the likelihood that the student can solve the exam problem?

Given the student's solution, what is the likelihood that the problem was of form A?

Ans : What is the probability that you will solve the problem of the exam?

Solution: The probability to solve the problem of the exam is the probability of getting a problem of

a certain type times the probability of solving such a problem, summed over all types. This is known

as the total probability.

P(solved) = P(solved|A)P(A) + P(solved|B)P(B) + P(solved|C)P(C)

= 9/10 · 30% + 2/10 · 20% + 6/10 · 50%

= 27/100 + 4/100 + 30/100 = 61/100 = 0.61 .

(b) Given you have solved the problem, what is the probability that it was of type A?

Solution: For this to answer we need Bayes theorem.

P(A|solved) = P(solved|A)P(A)

P(solved)

=

9/10 · 30%

61/100

=

27/100

61/100

=

27

61

= 0.442... .

So we see that given you have solved the problem, the a posteriori probability that the problem was of

type A is greater than its a priori probability of 30%, because problems of type A are relatively easy to solve.

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.

How many customers come into the bank on a daily basis (10 hours)?

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?

Explain likelihood that there is a customer if there is a photograph?

Ans : How many airplanes fly over the radar station on average per day (24 hours)?

Solution: There are 24×12 = 288 five-minute periods per day. In each period there is a probability of

5% for an airplane being present. Thus the average number of airplanes is 288×5% = 288×0.05 = 14.4.

2. How many false alarms (there is an alarm even though there is no airplane) and how many false

no-alarms (there is no alarm even though there is an airplane) are there on average per day.

Solution: On average there is no airplane in 288 − 14.4 of the five-minute periods. This times the

probability of 10% per period for a false alarm yields (288 − 14.4) × 10% = 273.6 × 0.1 = 27.36 false

alarms.

On average there are 14.4 airplanes, each of which has a probability of 1% of getting missed. Thus the

number of false no-alarms is 14.4 × 1% = 14.4 × 0.01 = 0.144.

there is an alarm, what is the probability that there is indeed an airplane?

Solution: For this question we need Bayes theorem.

P(airplane|alarm)

=

P(alarm|airplane)P(airplane)

P(alarm)

=

P(alarm|airplane)P(airplane)

P(alarm|airplane)P(airplane) + P(alarm|no airplane)P(no airplane)

=

0.99 · 0.05

0.99 · 0.05 + 0.1 · (1 − 0.05) = 0.342...

≈

0.05

0.05 + 0.1

= 0.333... .

It might be somewhat surprising that the probability of an airplane being present given an alarm is

only 34% even though the detection of an airplane is so reliable (99%). The reason is that airplanes are

not so frequent (only 5%) and the probability for an alarm given no airplane is relatively high (10%)

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.?