ML Assignment 13

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

 Prior could be the probability distribution representing the relative proportions of voters who will vote for a particular politician in a future election. The unknown quantity may be a parameter of the model or a latent variable rather than an observable variable.

Posterior probability is a revised probability that takes into account new available information. For example, let there be two urns, urn A having 5 black balls and 10 red balls and urn B having 10 black balls and 5 red balls. Now if an urn is selected at random, the probability that urn A is chosen is 0.5.

Likelihood the probability of getting heads when flipping a fair coin is 0.5 because there are two possible outcomes (heads or tails), and each outcome has an equal likelihood of occurring.

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

Bayes theorem plays an important role in the concept learning principle. It is used to calculate the probability of a concept given evidence. Specifically, it helps to calculate the probability of a concept belonging to a certain class given the evidence that it has certain properties. The probability is then used to help determine which concept is most likely to be the correct one. This helps to improve the accuracy of concept learning by helping to reduce the chances of incorrect decisions being made.

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

The Naïve Bayes classifier is a popular supervised machine learning algorithm used for classification tasks such as text classification. It belongs to the family of generative learning algorithms, which means that it models the distribution of inputs for a given class or category.

The Naive Bayes classifier is used in real life in many applications, such as email spam filtering, document classification, and medical diagnosis. For example, a Naive Bayes classifier could be used to classify emails as spam or legitimate. The classifier would use the presence of certain words or phrases in the email to determine whether it is spam or not.

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

the Naive Bayes classifier can be used on continuous numeric data. To do this, you have to bin the data into discrete intervals or categories. This can be done by defining a function that takes the numeric data and assigns it to different bins. Once the data is binned, the Naive Bayes classifier can be used as usual.

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 are used to represent uncertain knowledge and make decisions based on that knowledge. They are a type of Bayesian network, which is a graphical model that represents probabilistic relationships between variables.

Applications of Bayesian Network

Gene Regulatory Network. GRN is Gene Regulatory Network or Genetic Regulatory Network.

Medicine. It is the science or practice of diagnosis.

Biomonitoring.

Document Classification.

Information Retrieval.

Semantic Search.

Image Processing.

Spam Filter

BBNs are capable of resolving a wide range of issues, but it is important to note that their accuracy is dependent on the quality of the data used to create the model. BBNs can be used to solve problems in many different fields, including finance, economics, engineering, and medicine. However, it is important to remember that they are not a silver bullet, and they require careful analysis and interpretation to be effective.

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?

This can be solved directly with the Bayesian theorem.

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

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

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

= 0.0097

≈ 0.00001/ 0.001 = 0.01

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) ?

P(D = p|T = p)

= P(T = p|D = p)P(D = p) P(T = p) (7)

= 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

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?

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

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

= (9/10 \* 30/100)/(61/100 )

= (27/100)/( 61/100)

= 27 /61 = 0.442

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)?

P(Customer coming into the bank in each 5-minute time period) = 0.05 Since each hour has 12 5-minute time period, then there will be a total of 10\*12 = 120 ,5-min time periods per day.

Therefore, 120\*0.05 = 6 customers can come into the bank on a daily basis.

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?

P(If there is no customer, the camera takes a false photograph) = 0.1

P(If there is a client, the CCTV will detect them) = 0.99

For fake photographs,

(120 - 6)\*0.1 = 11.4 or 11 fake photographs.

For missed photographs,

6\*0.1 = 0.6 or 60% of the times.

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

P(If there is a client, the CCTV will detect them) = 0.99

P(Customer) = 0.05

For, P(Customer is there| Photo is taken)

(0.99\*0.05)/((0.99\*0.5)+0.1(1-0.05))

which is 0.34

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.