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

Answer :- Sure, let's use a simple example to illustrate these concepts, often used in Bayesian statistics:

Imagine you have a medical test to detect a rare disease, and the test has a certain accuracy rate. Here's how the concepts of Prior, Posterior, and Likelihood apply:

1. **Prior Probability**: This is your initial belief about the likelihood of having the disease before any test results. Let's say based on known statistics, the prior probability of having the disease in the general population is 1%.
2. **Likelihood**: This is the probability of observing the test result given that you have (or do not have) the disease. For example, if the test is 95% accurate in detecting the disease (sensitivity), the likelihood of testing positive given that you have the disease is 95%. Conversely, if the test has a 90% accuracy in ruling out the disease (specificity), the likelihood of testing negative given that you do not have the disease is 90%.
3. **Posterior Probability**: This is your updated belief about the probability of having the disease after taking into account the test results. It combines your prior belief and the likelihood of the test results. Using Bayesian inference, you calculate this by multiplying the prior probability by the likelihood and normalizing to sum to 1 (since probabilities must sum to 1).

For example, if your prior probability of having the disease was 1% (0.01) and the test result came back positive with a likelihood of 95% (0.95), the posterior probability would be:

Posterior probability=Prior probability×LikelihoodNormalization constant\text{Posterior probability} = \frac{\text{Prior probability} \times \text{Likelihood}}{\text{Normalization constant}}Posterior probability=Normalization constantPrior probability×Likelihood​

Posterior probability=0.01×0.95Normalization constant\text{Posterior probability} = \frac{0.01 \times 0.95}{\text{Normalization constant}}Posterior probability=Normalization constant0.01×0.95​

The normalization constant ensures that the posterior probability sums to 1, accounting for all possible outcomes (positive and negative test results).

This example shows how Bayesian reasoning updates your beliefs (prior probability) with new evidence (likelihood from the test result) to form a new belief (posterior probability) about the likelihood of having the disease.

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

Answer :- Bayes' theorem plays a fundamental role in the concept learning principle, particularly in Bayesian inference, which is essential for updating beliefs about hypotheses or concepts based on new evidence. Here's how Bayes' theorem connects to concept learning:

1. **Updating Probabilities**: In concept learning, you start with a prior probability distribution over different hypotheses or concepts. These priors represent your initial beliefs about the likelihood of each hypothesis being true before any evidence is considered.
2. **Incorporating Evidence**: As new evidence (data) becomes available, Bayes' theorem allows you to update your beliefs (posterior probabilities) about the hypotheses. It combines your prior beliefs with the likelihood of observing the data given each hypothesis.
3. **Iterative Learning Process**: Concept learning often involves an iterative process where new evidence continuously updates your beliefs. Bayes' theorem provides a formal mechanism to adjust these beliefs after each new piece of evidence, refining your understanding of which hypotheses or concepts are more likely.
4. **Optimal Decision Making**: By using Bayes' theorem, you can make decisions that are optimal in the sense of maximizing the expected utility or minimizing expected loss, taking into account both prior beliefs and the probabilistic implications of new evidence.

In summary, Bayes' theorem serves as the mathematical foundation for updating beliefs in concept learning, enabling a principled approach to incorporating evidence and refining hypotheses or concepts based on observed data. It underpins the Bayesian approach to learning and inference, which is widely used in various fields including machine learning, statistics, and cognitive science.

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

Answer :- Sure! One practical application of the Naive Bayes classifier is in email spam filtering. Here's how it works:

Example: Email Spam Filtering

1. Problem Statement:
   * You receive a lot of emails daily, and you want to automatically filter out spam emails to keep your inbox clean.
   * You have historical data of emails categorized as either spam or not spam (ham).
2. Data Preparation:
   * Each email is represented as a set of features, such as the presence or absence of certain words (e.g., "free", "win", "offer", etc.).
   * These features are extracted from the content and possibly other metadata of each email.
3. Training:
   * Using the historical data, the Naive Bayes classifier learns the probabilities of each feature (word) occurring in spam emails and non-spam emails (ham).
   * It calculates:
     + Prior probabilities: The overall probabilities of an email being spam or ham.
     + Likelihoods: The probabilities of each word occurring in spam or ham emails.
   * The "Naive" assumption here is that the presence of each word is independent of the others, simplifying the calculation but not always holding true in reality.
4. Classification:
   * When a new email arrives, the classifier calculates the probability that it belongs to each class (spam or ham) based on the features (words) present in the email.
   * It computes these probabilities using Bayes' theorem and selects the class with the highest posterior probability (most likely).
5. Filtering:
   * If the probability that the email is spam exceeds a certain threshold (often set based on performance metrics like precision and recall), the email is flagged as spam and can be directed to a spam folder or deleted.
   * If not, it is classified as ham and delivered to your inbox.

Benefits of Naive Bayes Classifier in Email Spam Filtering:

* Efficiency: Naive Bayes classifiers are computationally efficient and can handle large amounts of data quickly.
* Effectiveness: Despite their simple assumptions (naivety), Naive Bayes classifiers often perform well in practice for tasks like spam filtering, especially when trained on a diverse and representative dataset.
* Scalability: They scale well with the number of features (words) and can be updated easily with new data, making them suitable for dynamic environments like email filtering where new spam tactics emerge frequently.

This example illustrates how the Naive Bayes classifier leverages Bayes' theorem to classify emails based on the likelihood of features (words) occurring in spam or non-spam contexts, providing a practical and effective solution to the problem of email spam detection

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

Answer :- Yes, the Naive Bayes classifier can be adapted to handle continuous numeric data, although it's more commonly used for categorical data. When dealing with continuous data, you typically need to make some assumptions or use techniques that allow you to estimate probabilities from the data distribution. Here are a couple of approaches to handle continuous numeric data with Naive Bayes:

1. Gaussian Naive Bayes

Assumption: Each feature (numeric variable) follows a Gaussian (normal) distribution.

Steps to Implement:

* Data Preparation: Calculate the mean and standard deviation of each feature for each class (e.g., spam or ham).
* Training:
  + For each class, compute the mean (μ\muμ) and standard deviation (σ\sigmaσ) of each feature.
  + Use these statistics to estimate the probability density function (PDF) of each feature for each class, assuming it follows a Gaussian distribution: P(xi∣Ck)=12πσk,i2exp⁡(−(xi−μk,i)22σk,i2)P(x\_i \mid C\_k) = \frac{1}{\sqrt{2\pi\sigma\_{k,i}^2}} \exp\left(-\frac{(x\_i - \mu\_{k,i})^2}{2\sigma\_{k,i}^2}\right)P(xi​∣Ck​)=2πσk,i2​​1​exp(−2σk,i2​(xi​−μk,i​)2​)
  + Where xix\_ixi​ is the value of the iii-th feature, μk,i\mu\_{k,i}μk,i​ and σk,i\sigma\_{k,i}σk,i​ are the mean and standard deviation of the iii-th feature in class CkC\_kCk​.
* Classification:
  + Given a new data point with continuous features x=(x1,x2,...,xn)x = (x\_1, x\_2, ..., x\_n)x=(x1​,x2​,...,xn​):
    - Compute the posterior probability for each class CkC\_kCk​ using Bayes' theorem: P(Ck∣x)∝P(Ck)∏i=1nP(xi∣Ck)P(C\_k \mid x) \propto P(C\_k) \prod\_{i=1}^{n} P(x\_i \mid C\_k)P(Ck​∣x)∝P(Ck​)∏i=1n​P(xi​∣Ck​)
    - Predict the class with the highest posterior probability.

2. Kernel Density Estimation (KDE) with Naive Bayes

Approach: Use KDE to estimate the probability density function for each feature, allowing for a more flexible estimation of the underlying distribution.

Steps to Implement:

* Data Preparation: Estimate the probability density function (PDF) for each feature in each class using KDE.
* Training:
  + For each class, estimate the PDF of each feature using KDE, which provides a non-parametric estimate of the distribution of the feature in that class.
  + KDE computes an estimate of the probability density function based on the observed data points, smoothing out the distribution to provide a continuous estimate of the probability.
* Classification:
  + Given a new data point with continuous features x=(x1,x2,...,xn)x = (x\_1, x\_2, ..., x\_n)x=(x1​,x2​,...,xn​):
    - Calculate the likelihood of xxx under each class using the estimated KDEs.
    - Apply Bayes' theorem to compute the posterior probabilities and classify the data point accordingly.

Considerations:

* Feature Independence: Naive Bayes assumes that features are conditionally independent given the class. This assumption may not hold in practice, especially for continuous data where correlations between features might exist.
* Data Scaling: It's often beneficial to scale or normalize the continuous features to ensure they contribute equally to the classifier's decision-making process.
* Performance: Despite its simplifying assumptions, Naive Bayes can perform well in practice, especially when the independence assumption is not severely violated and when the data distribution assumptions are reasonably met.

These adaptations allow the Naive Bayes classifier to handle continuous numeric data by estimating or assuming the underlying probability distributions of the features within each class, facilitating classification based on these probabilistic models.

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?

Answer :- Bayesian Belief Networks (BBNs), also known as Bayesian Networks (BNs) or probabilistic graphical models, are powerful tools for representing and reasoning about uncertainty and probabilistic dependencies among a set of random variables. Here’s an overview of what they are, how they work, their applications, and their capabilities:

What are Bayesian Belief Networks?

* Representation: BBNs represent a set of random variables and their conditional dependencies via a directed acyclic graph (DAG). Each node in the graph represents a random variable, and edges between nodes indicate probabilistic dependencies (i.e., causal relationships or conditional influences).
* Probabilistic Inference: BBNs use Bayes' theorem and conditional probability to quantify the uncertainty about variables given evidence and to update beliefs as new evidence becomes available.
* Structure: The structure of a BBN encodes causal relationships and conditional dependencies among variables, while the associated probabilities (conditional probability tables, CPTs) quantify these relationships probabilistically.

How do Bayesian Belief Networks Work?

1. Modeling:
   * Graph Structure: Define a DAG where each node corresponds to a random variable and edges indicate direct dependencies.
   * Conditional Probability Tables (CPTs): Assign probabilities to each node based on the states of its parent nodes in the graph.
2. Inference:
   * Updating Beliefs: Given observed evidence (data or knowledge about some variables), BBNs calculate posterior probabilities of other variables using Bayes' theorem and propagate this information through the network.
   * Variable Elimination: Techniques like variable elimination or message passing algorithms (e.g., belief propagation) are often used to efficiently compute posterior probabilities.
3. Applications of Bayesian Belief Networks:
   * Medical Diagnosis: BBNs can model relationships among symptoms, diseases, and test results to assist in medical diagnosis by calculating the likelihood of different diseases given observed symptoms.
   * Risk Assessment: They are used in risk assessment and management to model dependencies among various risk factors and assess the likelihood and consequences of different risks.
   * Fault Diagnosis: BBNs help in diagnosing faults in complex systems by modeling dependencies among components and observed symptoms.
   * Finance and Business: Applications include fraud detection, credit scoring, market analysis, and decision support systems where uncertainty and probabilistic reasoning are critical.

Capabilities of Bayesian Belief Networks

* Handling Uncertainty: BBNs explicitly model and reason with uncertainty, making them suitable for domains where data is incomplete, noisy, or uncertain.
* Complex Dependencies: They can model complex dependencies among variables, including non-linear and feedback relationships, by structuring the graph appropriately.
* Inference Efficiency: Various algorithms and techniques (e.g., exact and approximate inference methods) allow BBNs to efficiently compute probabilities and make predictions even with large and complex networks.

Conclusion

Bayesian Belief Networks are versatile tools that leverage probabilistic reasoning to model and solve a wide range of problems involving uncertainty and complex relationships among variables. Their ability to handle uncertainty and infer probabilistic dependencies makes them applicable in diverse fields such as healthcare, finance, engineering, and more, where decision-making under uncertainty is crucial. While they excel in many scenarios, their effectiveness depends on the accuracy of the model structure and the quality of available data.

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?

Answer :- To find the probability that an alarm is triggered given that an individual is actually an intruder (P(A = 1 | I = 1)), we can use Bayes' theorem. Here's how we calculate it:

Given:

* Probability of an alarm being triggered when there is an intruder: P(A=1∣I=1)=0.98P(A = 1 | I = 1) = 0.98P(A=1∣I=1)=0.98
* Probability of an alarm being triggered when there is no intruder: P(A=1∣I=0)=0.001P(A = 1 | I = 0) = 0.001P(A=1∣I=0)=0.001
* Probability of there being an intruder in the population: P(I=1)=0.00001P(I = 1) = 0.00001P(I=1)=0.00001
* Probability of no intruder: P(I=0)=1−P(I=1)=0.99999P(I = 0) = 1 - P(I = 1) = 0.99999P(I=0)=1−P(I=1)=0.99999

We need to find P(A=1∣I=1)P(A = 1 | I = 1)P(A=1∣I=1), which is the probability that the alarm is triggered given that there is an intruder.

Using Bayes' theorem: P(A=1∣I=1)=P(I=1∣A=1)⋅P(A=1)P(I=1)P(A = 1 | I = 1) = \frac{P(I = 1 | A = 1) \cdot P(A = 1)}{P(I = 1)}P(A=1∣I=1)=P(I=1)P(I=1∣A=1)⋅P(A=1)​

Where:

* P(I=1∣A=1)P(I = 1 | A = 1)P(I=1∣A=1) is the probability that there is an intruder given that the alarm is triggered.
* P(A=1)P(A = 1)P(A=1) is the total probability that the alarm is triggered.

First, calculate P(A=1)P(A = 1)P(A=1) 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)P(A = 1) = P(A = 1 | I = 1) \cdot P(I = 1) + P(A = 1 | I = 0) \cdot P(I = 0)P(A=1)=P(A=1∣I=1)⋅P(I=1)+P(A=1∣I=0)⋅P(I=0)

Substitute the given values: P(A=1)=0.98⋅0.00001+0.001⋅0.99999P(A = 1) = 0.98 \cdot 0.00001 + 0.001 \cdot 0.99999P(A=1)=0.98⋅0.00001+0.001⋅0.99999 P(A=1)=0.0000098+0.00099999P(A = 1) = 0.0000098 + 0.00099999P(A=1)=0.0000098+0.00099999 P(A=1)=0.00100979P(A = 1) = 0.00100979P(A=1)=0.00100979

Now, calculate P(I=1∣A=1)P(I = 1 | A = 1)P(I=1∣A=1) using Bayes' theorem: P(I=1∣A=1)=P(A=1∣I=1)⋅P(I=1)P(A=1)P(I = 1 | A = 1) = \frac{P(A = 1 | I = 1) \cdot P(I = 1)}{P(A = 1)}P(I=1∣A=1)=P(A=1)P(A=1∣I=1)⋅P(I=1)​ P(I=1∣A=1)=0.98⋅0.000010.00100979P(I = 1 | A = 1) = \frac{0.98 \cdot 0.00001}{0.00100979}P(I=1∣A=1)=0.001009790.98⋅0.00001​ P(I=1∣A=1)=0.000000980.00100979P(I = 1 | A = 1) = \frac{0.00000098}{0.00100979}P(I=1∣A=1)=0.001009790.00000098​ P(I=1∣A=1)≈0.000971P(I = 1 | A = 1) \approx 0.000971P(I=1∣A=1)≈0.000971

Therefore, the probability that an alarm would be triggered when an individual is actually an intruder is approximately 0.000971\boxed{0.000971}0.000971​, or about 0.0971%0.0971\%0.0971%.

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).

Answer :- To find the likelihood that a person who tests positive is actually immune to the antibiotic (random variable D = 1), we need to calculate the conditional probability P(D=1∣T=1)P(D = 1 | T = 1)P(D=1∣T=1) using Bayes' theorem.

Given:

* Probability of a false positive (testing positive when not immune): P(T=1∣D=0)=0.01P(T = 1 | D = 0) = 0.01P(T=1∣D=0)=0.01
* Probability of a false negative (testing negative when immune): P(T=0∣D=1)=0.05P(T = 0 | D = 1) = 0.05P(T=0∣D=1)=0.05
* Probability that a person is immune (antibiotic-resistant): P(D=1)=0.02P(D = 1) = 0.02P(D=1)=0.02
* Probability that a person is not immune: P(D=0)=1−P(D=1)=0.98P(D = 0) = 1 - P(D = 1) = 0.98P(D=0)=1−P(D=1)=0.98

First, we calculate P(T=1)P(T = 1)P(T=1), the total probability of testing positive: P(T=1)=P(T=1∣D=1)⋅P(D=1)+P(T=1∣D=0)⋅P(D=0)P(T = 1) = P(T = 1 | D = 1) \cdot P(D = 1) + P(T = 1 | D = 0) \cdot P(D = 0)P(T=1)=P(T=1∣D=1)⋅P(D=1)+P(T=1∣D=0)⋅P(D=0)

Calculate P(T=1∣D=1)P(T = 1 | D = 1)P(T=1∣D=1): Since the false negative rate is P(T=0∣D=1)=0.05P(T = 0 | D = 1) = 0.05P(T=0∣D=1)=0.05, P(T=1∣D=1)=1−P(T=0∣D=1)=1−0.05=0.95P(T = 1 | D = 1) = 1 - P(T = 0 | D = 1) = 1 - 0.05 = 0.95P(T=1∣D=1)=1−P(T=0∣D=1)=1−0.05=0.95

Now, calculate P(T=1)P(T = 1)P(T=1): P(T=1)=0.95⋅0.02+0.01⋅0.98P(T = 1) = 0.95 \cdot 0.02 + 0.01 \cdot 0.98P(T=1)=0.95⋅0.02+0.01⋅0.98 P(T=1)=0.019+0.0098P(T = 1) = 0.019 + 0.0098P(T=1)=0.019+0.0098 P(T=1)=0.0288P(T = 1) = 0.0288P(T=1)=0.0288

Now, use Bayes' theorem to find P(D=1∣T=1)P(D = 1 | T = 1)P(D=1∣T=1): P(D=1∣T=1)=P(T=1∣D=1)⋅P(D=1)P(T=1)P(D = 1 | T = 1) = \frac{P(T = 1 | D = 1) \cdot P(D = 1)}{P(T = 1)}P(D=1∣T=1)=P(T=1)P(T=1∣D=1)⋅P(D=1)​ P(D=1∣T=1)=0.95⋅0.020.0288P(D = 1 | T = 1) = \frac{0.95 \cdot 0.02}{0.0288}P(D=1∣T=1)=0.02880.95⋅0.02​ P(D=1∣T=1)=0.0190.0288P(D = 1 | T = 1) = \frac{0.019}{0.0288}P(D=1∣T=1)=0.02880.019​ P(D=1∣T=1)≈0.6597P(D = 1 | T = 1) \approx 0.6597P(D=1∣T=1)≈0.6597

Therefore, the likelihood that a person who tests positive is actually immune to the antibiotic (antibiotic-resistant) is approximately 0.6597\boxed{0.6597}0.6597​, or about 65.97%65.97\%65.97%.

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?

Answer :- To find the likelihood that the student can solve the exam problem, we need to calculate the probability that the chosen question is one that the student has practiced for. Here’s how we can approach it:

Given:

* Probability of question being from form A: P(A)=0.30P(A) = 0.30P(A)=0.30
* Probability of question being from form B: P(B)=0.20P(B) = 0.20P(B)=0.20
* Probability of question being from form C: P(C)=0.50P(C) = 0.50P(C)=0.50
* Number of type A problems practiced: 9 out of 10
* Number of type B problems practiced: 2 out of 10
* Number of type C problems practiced: 6 out of 10

First, calculate the total probability of each form being chosen: P(A)=0.30P(A) = 0.30P(A)=0.30 P(B)=0.20P(B) = 0.20P(B)=0.20 P(C)=0.50P(C) = 0.50P(C)=0.50

Now, calculate the probability that the chosen question is one that the student has practiced (let's denote this event as SSS): P(S)=P(A)⋅910+P(B)⋅210+P(C)⋅610P(S) = P(A) \cdot \frac{9}{10} + P(B) \cdot \frac{2}{10} + P(C) \cdot \frac{6}{10}P(S)=P(A)⋅109​+P(B)⋅102​+P(C)⋅106​

Calculate each term: P(S)=0.30⋅0.9+0.20⋅0.2+0.50⋅0.6P(S) = 0.30 \cdot 0.9 + 0.20 \cdot 0.2 + 0.50 \cdot 0.6P(S)=0.30⋅0.9+0.20⋅0.2+0.50⋅0.6 P(S)=0.27+0.04+0.30P(S) = 0.27 + 0.04 + 0.30P(S)=0.27+0.04+0.30 P(S)=0.61P(S) = 0.61P(S)=0.61

Therefore, the likelihood that the student can solve the exam problem, which means the probability that the chosen question is one that the student has practiced for, is 0.61\boxed{0.61}0.61​, or 61%. This reflects the combined probabilities of each form of the question being chosen and the student's preparation for each type.

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

Answer :- To determine the likelihood that the problem on the exam was of form A given that the student solved it, we can use Bayes' theorem. Here are the steps and calculations based on the given probabilities:

Given:

* Probability of the problem being form A: P(A)=0.30P(A) = 0.30P(A)=0.30
* Probability of the problem being form B: P(B)=0.20P(B) = 0.20P(B)=0.20
* Probability of the problem being form C: P(C)=0.50P(C) = 0.50P(C)=0.50
* Number of problems solved:
  + Type A: 9 out of 10
  + Type B: 2 out of 10
  + Type C: 6 out of 10

First, calculate the total number of problems solved: Total problems solved=9+2+6=17\text{Total problems solved} = 9 + 2 + 6 = 17Total problems solved=9+2+6=17

Now, calculate the probability of the student solving a problem given its form (conditional probabilities):

* P(solve∣A)=910=0.90P(\text{solve} \mid A) = \frac{9}{10} = 0.90P(solve∣A)=109​=0.90
* P(solve∣B)=210=0.20P(\text{solve} \mid B) = \frac{2}{10} = 0.20P(solve∣B)=102​=0.20
* P(solve∣C)=610=0.60P(\text{solve} \mid C) = \frac{6}{10} = 0.60P(solve∣C)=106​=0.60

Calculate the total probability of the student solving a problem: P(solve)=P(solve∣A)⋅P(A)+P(solve∣B)⋅P(B)+P(solve∣C)⋅P(C)P(\text{solve}) = P(\text{solve} \mid A) \cdot P(A) + P(\text{solve} \mid B) \cdot P(B) + P(\text{solve} \mid C) \cdot P(C)P(solve)=P(solve∣A)⋅P(A)+P(solve∣B)⋅P(B)+P(solve∣C)⋅P(C) P(solve)=0.90⋅0.30+0.20⋅0.20+0.60⋅0.50P(\text{solve}) = 0.90 \cdot 0.30 + 0.20 \cdot 0.20 + 0.60 \cdot 0.50P(solve)=0.90⋅0.30+0.20⋅0.20+0.60⋅0.50 P(solve)=0.27+0.04+0.30P(\text{solve}) = 0.27 + 0.04 + 0.30P(solve)=0.27+0.04+0.30 P(solve)=0.61P(\text{solve}) = 0.61P(solve)=0.61

Now, apply Bayes' theorem to find P(A∣solve)P(A \mid \text{solve})P(A∣solve), the probability that the problem was form A given that the student solved it: P(A∣solve)=P(solve∣A)⋅P(A)P(solve)P(A \mid \text{solve}) = \frac{P(\text{solve} \mid A) \cdot P(A)}{P(\text{solve})}P(A∣solve)=P(solve)P(solve∣A)⋅P(A)​ P(A∣solve)=0.90⋅0.300.61P(A \mid \text{solve}) = \frac{0.90 \cdot 0.30}{0.61}P(A∣solve)=0.610.90⋅0.30​ P(A∣solve)=0.270.61P(A \mid \text{solve}) = \frac{0.27}{0.61}P(A∣solve)=0.610.27​ P(A∣solve)≈0.4426P(A \mid \text{solve}) \approx 0.4426P(A∣solve)≈0.4426

Therefore, the likelihood that the problem was of form A given that the student solved it is approximately 0.4426\boxed{0.4426}0.4426​, or about 44.26%44.26\%44.26%.

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)?

Answer :- To determine the expected number of customers entering the bank on a daily basis, we need to calculate the average number of customers per 5-minute bin and then scale it up to a 10-hour (600 minutes) period.

Given:

* Probability of a customer entering the bank in a 5-minute bin: p=0.05p = 0.05p=0.05
* Probability of no customer entering the bank in a 5-minute bin: 1−p=0.951 - p = 0.951−p=0.95
* Probability of detecting a customer if they are present: P(detect∣customer)=0.99P(\text{detect} \mid \text{customer}) = 0.99P(detect∣customer)=0.99
* Probability of a false detection if no customer is present: P(detect∣no customer)=0.10P(\text{detect} \mid \text{no customer}) = 0.10P(detect∣no customer)=0.10

Step-by-Step Calculation:

1. Expected number of customers per 5-minute bin:

E[customers per bin]=p⋅1+(1−p)⋅0=0.05E[\text{customers per bin}] = p \cdot 1 + (1 - p) \cdot 0 = 0.05E[customers per bin]=p⋅1+(1−p)⋅0=0.05

1. Expected number of detections per 5-minute bin:

E[detections per bin]=p⋅P(detect∣customer)+(1−p)⋅P(detect∣no customer)E[\text{detections per bin}] = p \cdot P(\text{detect} \mid \text{customer}) + (1 - p) \cdot P(\text{detect} \mid \text{no customer})E[detections per bin]=p⋅P(detect∣customer)+(1−p)⋅P(detect∣no customer) E[detections per bin]=0.05⋅0.99+0.95⋅0.10E[\text{detections per bin}] = 0.05 \cdot 0.99 + 0.95 \cdot 0.10E[detections per bin]=0.05⋅0.99+0.95⋅0.10 E[detections per bin]=0.0495+0.095E[\text{detections per bin}] = 0.0495 + 0.095E[detections per bin]=0.0495+0.095 E[detections per bin]=0.1445E[\text{detections per bin}] = 0.1445E[detections per bin]=0.1445

1. Expected number of detections per day (10 hours = 600 minutes):

E[detections per day]=E[detections per bin]×6005E[\text{detections per day}] = E[\text{detections per bin}] \times \frac{600}{5}E[detections per day]=E[detections per bin]×5600​ E[detections per day]=0.1445×120E[\text{detections per day}] = 0.1445 \times 120E[detections per day]=0.1445×120 E[detections per day]=17.34E[\text{detections per day}] = 17.34E[detections per day]=17.34

Therefore, the expected number of customers coming into the bank on a daily basis, based on the CCTV detections, is approximately 17.34\boxed{17.34}17.34​. This means we expect about 17 customers per day on average, considering the probabilities and detection rates provided.

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?

Answer :- Let's break down the problem based on the probabilities given:

Given:

* Probability of a customer arriving in a 5-minute interval: P(customer)=0.05P(\text{customer}) = 0.05P(customer)=0.05
* Probability of no customer arriving: P(no customer)=1−P(customer)=0.95P(\text{no customer}) = 1 - P(\text{customer}) = 0.95P(no customer)=1−P(customer)=0.95
* Probability of detecting a customer (true positive rate): P(detect∣customer)=0.99P(\text{detect} \mid \text{customer}) = 0.99P(detect∣customer)=0.99
* Probability of a false detection (false positive rate): P(detect∣no customer)=0.10P(\text{detect} \mid \text{no customer}) = 0.10P(detect∣no customer)=0.10

Now, let's calculate the expected number of photographs:

1. Fake Photographs (when there is no customer):

This is the number of photographs taken when there actually isn't a customer.

Expected number of false photographs per 5-minute interval: Expected false photographs=P(detect∣no customer)⋅P(no customer)\text{Expected false photographs} = P(\text{detect} \mid \text{no customer}) \cdot P(\text{no customer})Expected false photographs=P(detect∣no customer)⋅P(no customer)

Calculate: Expected false photographs=0.10⋅0.95\text{Expected false photographs} = 0.10 \cdot 0.95Expected false photographs=0.10⋅0.95 Expected false photographs=0.095\text{Expected false photographs} = 0.095Expected false photographs=0.095

To find the daily number of false photographs, multiply by the number of 5-minute intervals in a day. Assuming a typical business day of 8 hours (480 minutes): Daily false photographs=0.095⋅24×605\text{Daily false photographs} = 0.095 \cdot \frac{24 \times 60}{5}Daily false photographs=0.095⋅524×60​ Daily false photographs=0.095⋅288\text{Daily false photographs} = 0.095 \cdot 288Daily false photographs=0.095⋅288 Daily false photographs≈27.36\text{Daily false photographs} \approx 27.36Daily false photographs≈27.36

So, on average, there are approximately 27\boxed{27}27​ fake photographs taken daily.

2. Missed Photographs (when there is a customer):

This is the number of photographs missed when there actually is a customer.

Expected number of missed photographs per 5-minute interval: Expected missed photographs=1−P(detect∣customer)⋅P(customer)\text{Expected missed photographs} = 1 - P(\text{detect} \mid \text{customer}) \cdot P(\text{customer})Expected missed photographs=1−P(detect∣customer)⋅P(customer)

Calculate: Expected missed photographs=(1−0.99)⋅0.05\text{Expected missed photographs} = (1 - 0.99) \cdot 0.05Expected missed photographs=(1−0.99)⋅0.05 Expected missed photographs=0.01⋅0.05\text{Expected missed photographs} = 0.01 \cdot 0.05Expected missed photographs=0.01⋅0.05 Expected missed photographs=0.0005\text{Expected missed photographs} = 0.0005Expected missed photographs=0.0005

To find the daily number of missed photographs: Daily missed photographs=0.0005⋅24×605\text{Daily missed photographs} = 0.0005 \cdot \frac{24 \times 60}{5}Daily missed photographs=0.0005⋅524×60​ Daily missed photographs=0.0005⋅288\text{Daily missed photographs} = 0.0005 \cdot 288Daily missed photographs=0.0005⋅288 Daily missed photographs=0.144\text{Daily missed photographs} = 0.144Daily missed photographs=0.144

So, on average, there are approximately 0\boxed{0}0​ missed photographs taken daily. This is because the probability of missing a customer detection is very low (0.05%). Therefore, missed photographs are negligible in this scenario.

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

Answer :- To find the likelihood that there is a customer if there is a photograph taken by the CCTV system, we need to calculate the conditional probability P(customer∣photograph)P(\text{customer} \mid \text{photograph})P(customer∣photograph) using Bayes' theorem. Here are the given probabilities:

Given:

* Probability of a customer arriving in a 5-minute period: P(customer)=0.05P(\text{customer}) = 0.05P(customer)=0.05
* Probability of no customer arriving in a 5-minute period: P(no customer)=1−P(customer)=0.95P(\text{no customer}) = 1 - P(\text{customer}) = 0.95P(no customer)=1−P(customer)=0.95
* Probability of the CCTV detecting a customer: P(photograph∣customer)=0.99P(\text{photograph} \mid \text{customer}) = 0.99P(photograph∣customer)=0.99
* Probability of the CCTV taking a false photograph (false positive): P(photograph∣no customer)=0.10P(\text{photograph} \mid \text{no customer}) = 0.10P(photograph∣no customer)=0.10

First, calculate the total probability of a photograph being taken: P(photograph)=P(photograph∣customer)⋅P(customer)+P(photograph∣no customer)⋅P(no customer)P(\text{photograph}) = P(\text{photograph} \mid \text{customer}) \cdot P(\text{customer}) + P(\text{photograph} \mid \text{no customer}) \cdot P(\text{no customer})P(photograph)=P(photograph∣customer)⋅P(customer)+P(photograph∣no customer)⋅P(no customer) P(photograph)=0.99⋅0.05+0.10⋅0.95P(\text{photograph}) = 0.99 \cdot 0.05 + 0.10 \cdot 0.95P(photograph)=0.99⋅0.05+0.10⋅0.95 P(photograph)=0.0495+0.095P(\text{photograph}) = 0.0495 + 0.095P(photograph)=0.0495+0.095 P(photograph)=0.1445P(\text{photograph}) = 0.1445P(photograph)=0.1445

Now, apply Bayes' theorem to find P(customer∣photograph)P(\text{customer} \mid \text{photograph})P(customer∣photograph): P(customer∣photograph)=P(photograph∣customer)⋅P(customer)P(photograph)P(\text{customer} \mid \text{photograph}) = \frac{P(\text{photograph} \mid \text{customer}) \cdot P(\text{customer})}{P(\text{photograph})}P(customer∣photograph)=P(photograph)P(photograph∣customer)⋅P(customer)​ P(customer∣photograph)=0.99⋅0.050.1445P(\text{customer} \mid \text{photograph}) = \frac{0.99 \cdot 0.05}{0.1445}P(customer∣photograph)=0.14450.99⋅0.05​ P(customer∣photograph)=0.04950.1445P(\text{customer} \mid \text{photograph}) = \frac{0.0495}{0.1445}P(customer∣photograph)=0.14450.0495​ P(customer∣photograph)≈0.3429P(\text{customer} \mid \text{photograph}) \approx 0.3429P(customer∣photograph)≈0.3429

Therefore, the likelihood that there is a customer if there is a photograph taken by the CCTV system is approximately 0.3429\boxed{0.3429}0.3429​, or about 34.29%34.29\%34.29%. This means that given a photograph, there is about a 34.29% chance that it was taken because a customer entered the bank.

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.

Answer :- In a Naive Bayes classifier for match winning prediction, the variable "Won Toss" can be represented as a node in a Bayesian Belief Network (BBN). This node would typically have conditional probabilities based on the outcome of whether a team wins the match given whether they won the toss or not. Here's how you might construct a conditional probability table (CPT) for the node "Won Toss":

Conditional Probability Table for "Won Toss"

Let's denote:

* WWW: Event that the team wins the match.
* TTT: Event that the team won the toss.

Assume we have two possible outcomes for each variable:

* W=1W = 1W=1: Team wins the match.
* W=0W = 0W=0: Team does not win the match.
* T=1T = 1T=1: Team won the toss.
* T=0T = 0T=0: Team did not win the toss.

The conditional probabilities can be represented as follows:

| TTT | P(W=1∣T)P(W = 1 \mid T)P(W=1∣T) | P(W=0∣T)P(W = 0 \mid T)P(W=0∣T) |
| --- | --- | --- |
| T=1T = 1T=1 | p11p\_{11}p11​ | p10p\_{10}p10​ |
| T=0T = 0T=0 | p01p\_{01}p01​ | p00p\_{00}p00​ |

Where:

* p11=P(W=1∣T=1)p\_{11} = P(W = 1 \mid T = 1)p11​=P(W=1∣T=1): Probability that the team wins given they won the toss.
* p10=P(W=0∣T=1)p\_{10} = P(W = 0 \mid T = 1)p10​=P(W=0∣T=1): Probability that the team does not win given they won the toss.
* p01=P(W=1∣T=0)p\_{01} = P(W = 1 \mid T = 0)p01​=P(W=1∣T=0): Probability that the team wins given they did not win the toss.
* p00=P(W=0∣T=0)p\_{00} = P(W = 0 \mid T = 0)p00​=P(W=0∣T=0): Probability that the team does not win given they did not win the toss.

Example Values (Hypothetical):

Assume the following hypothetical values based on prior data or domain knowledge:

* p11=0.7p\_{11} = 0.7p11​=0.7: If a team wins the toss, there is a 70% chance they will win the match.
* p10=0.3p\_{10} = 0.3p10​=0.3: If a team wins the toss, there is a 30% chance they will not win the match.
* p01=0.5p\_{01} = 0.5p01​=0.5: If a team does not win the toss, there is a 50% chance they will win the match.
* p00=0.5p\_{00} = 0.5p00​=0.5: If a team does not win the toss, there is a 50% chance they will not win the match.

These values are illustrative and would be estimated or learned from historical match data in practice. The Naive Bayes classifier assumes conditional independence between the features (in this case, winning the toss) given the class (in this case, winning the match), which simplifies the modeling process while making a strong assumption about feature dependencies.