**1. What is prior probability? Give an example.**

**Ans:** Prior probability, also known as prior belief or simply prior, represents our initial belief or probability assignment about an event or outcome before considering any new evidence or data. It reflects what we believe about the event's likelihood based on existing knowledge or experience. Mathematically, it is denoted as P(A), where A is an event.

Example: In a medical diagnosis context, the prior probability of a patient having a particular disease before any test results might be based on the historical prevalence of the disease in the population. If 5% of people in the population have the disease, then P(Disease) = 0.05 could be the prior probability.

**2. What is posterior probability? Give an example.**

**Ans:** Posterior probability represents the updated probability of an event or outcome after taking into account new evidence or data. It is calculated using Bayes' theorem and combines the prior probability with the likelihood of observing the data given the event. Mathematically, it is denoted as P(A | B), where A is the event of interest, and B is the observed evidence or data.

Example: Continuing with the medical diagnosis example, after conducting tests on a patient, the posterior probability of the patient having the disease (P(Disease | Test Result)) is calculated based on both the prior probability (P(Disease)) and the test's accuracy (likelihood) in detecting the disease.

**3. What is likelihood probability? Give an example.**

**Ans:** Likelihood probability, often denoted as P(B | A), represents the probability of observing evidence or data B given that a particular event A has occurred. It quantifies how well the data supports or is consistent with the event of interest.

Example: In a coin-flipping experiment, the likelihood probability P(Heads | Coin is fair) represents the probability of getting a heads (evidence) when the coin is known to be fair (event A). If the coin is indeed fair, this likelihood might be 0.5, indicating a 50% chance of getting heads.

**4. What is Naïve Bayes classifier? Why is it named so?**

**Ans:** The Naïve Bayes classifier is a probabilistic machine learning algorithm used for classification tasks. It's named "naïve" because it makes a simplifying assumption of independence among the features used for classification. This assumption implies that the presence or absence of a particular feature is unrelated to the presence or absence of any other feature. Despite this simplification, Naïve Bayes often performs well in practice, especially for text classification and spam filtering.

**5. What is optimal Bayes classifier?**

**Ans:** The optimal Bayes classifier, also known as the Bayes optimal classifier, is a theoretical concept in machine learning. It represents the best possible classification performance achievable for a given problem when we have complete knowledge of the underlying probability distributions. It's the classifier that minimizes the classification error rate.

**6. Write any two features of Bayesian learning methods.**

**Ans:** Bayesian learning methods provide a principled way to handle uncertainty and incorporate prior knowledge into modeling.

These methods produce probability distributions over possible outcomes, allowing for more nuanced decision-making.

**7. Define the concept of consistent learners.**

**Ans:** Consistent learners are machine learning algorithms that, as the amount of training data increases, converge to a model that approximates the true underlying data distribution. In other words, they become more accurate and approach the best possible performance with an infinite amount of data.

**8. Write any two strengths of Bayes classifier.**

**Ans:** It handles well in situations with small datasets or when there's limited training data.

It's interpretable and provides probabilistic outputs, which can be valuable for decision-making and uncertainty estimation.

**9. Write any two weaknesses of Bayes classifier.**

**Ans:** It relies on the "naïve" independence assumption, which may not hold in real-world datasets, leading to suboptimal performance.

It can struggle with high-dimensional data because the number of parameters grows exponentially with the number of features.

**10. Explain how Naïve Bayes classifier is used for**

**1. Text classification**

**2. Spam filtering**

**3. Market sentiment analysis**

**Ans:** Text classification: Naïve Bayes is commonly used for text classification tasks such as spam detection and sentiment analysis. It represents documents as vectors of word frequencies or presence/absence of words and calculates the probability of each class (e.g., spam or not spam) given the document's features (words). The class with the highest probability is the predicted class.

Spam filtering: In spam filtering, Naïve Bayes can be trained on a dataset of emails labeled as spam or not spam. It calculates the likelihood of observing certain words or patterns in spam and non-spam emails. When a new email arrives, it calculates the probability that the email belongs to each class (spam or not spam) based on its content and assigns the label with the highest probability.

Market sentiment analysis: In sentiment analysis, Naïve Bayes can be used to determine the sentiment (positive, negative, neutral) expressed in a piece of text. It's trained on labeled data with text and corresponding sentiment labels. It then estimates the probability of a given sentiment class based on the words and phrases in the text, allowing businesses to gauge public sentiment towards products, brands, or topics.