1. What is prior probability? Give an example.

Prior probability, also known as prior belief or prior distribution, refers to the initial probability assigned to an event or hypothesis before considering any evidence or data. It represents our knowledge or belief about the event or hypothesis before observing any new information.

Example: Let's say we want to determine the probability of a person having a certain disease. Before conducting any tests or gathering specific information about the person, we may have a prior probability estimate based on factors such as their age, gender, and family medical history. This initial estimate represents our prior belief about the person's likelihood of having the disease.

2. What is posterior probability? Give an example.

Posterior probability, also known as the updated probability or posterior distribution, refers to the probability of an event or hypothesis after considering new evidence or data. It is calculated by combining the prior probability with the likelihood of the observed data.

Example: Continuing from the previous example, suppose we conduct medical tests on the person and collect additional data. By using Bayes' theorem, we can update our prior probability of the person having the disease to the posterior probability. The posterior probability takes into account the new test results, incorporating the prior belief along with the observed evidence.

3. What is likelihood probability? Give an example.

Likelihood probability represents the probability of observing a specific outcome given a certain hypothesis or model. It measures how well the hypothesis explains the observed data.

Example: Suppose we have a coin and want to determine if it is fair or biased. We can flip the coin multiple times and count the number of heads and tails observed. The likelihood probability would indicate how likely we would obtain the observed sequence of heads and tails if the coin were fair or biased. It quantifies the goodness of fit between the data and the hypothesis of a fair or biased coin.

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

The Naïve Bayes classifier is a probabilistic machine learning algorithm that uses Bayes' theorem to classify data into different classes or categories. It is called "naïve" because it makes a strong assumption of feature independence, assuming that the features of the data are conditionally independent given the class variable.

Despite its simplistic assumption, the Naïve Bayes classifier is widely used and can perform well in many real-world applications. The assumption of feature independence allows for fast and efficient training and classification, making it suitable for large datasets. However, in reality, features are often dependent to some extent, and this assumption may not hold true.

5. What is optimal Bayes classifier?

The optimal Bayes classifier, also known as the Bayes optimal classifier, is a theoretical concept that represents the best achievable classification performance given complete knowledge of the true underlying probability distributions of the data. It serves as a benchmark against which other classifiers are compared.

The optimal Bayes classifier assigns each data sample to the class with the highest posterior probability. It considers the prior probabilities, likelihoods, and the costs or utilities associated with different classification decisions. It provides the highest accuracy that can be achieved when the true probability distributions are known.

In practice, however, the true probability distributions are usually unknown, and we rely on estimating them or using other approximate classifiers such as Naïve Bayes, logistic regression, or decision trees.

6. Write any two features of Bayesian learning methods.

Two features of Bayesian learning methods are:

1. Probabilistic Framework: Bayesian learning methods are based on a probabilistic framework that allows us to model and reason about uncertainty. Instead of providing a single point estimate, Bayesian methods provide a distribution over the possible values of parameters or predictions. This uncertainty representation is valuable when making decisions or assessing the reliability of predictions.

2. Incorporation of Prior Knowledge: Bayesian learning methods allow the incorporation of prior knowledge or beliefs into the learning process. By specifying prior distributions, we can express our prior beliefs about the parameters or structure of the model. This prior knowledge is combined with

observed data through Bayes' theorem to obtain the posterior distribution, which represents the updated beliefs after considering the data.

7. Define the concept of consistent learners.

In machine learning, a learner is considered consistent if, given sufficient data, it can converge to the true underlying model or hypothesis. In other words, as the amount of training data approaches infinity, a consistent learner will produce a model that matches the true data-generating distribution.

A consistent learner minimizes the gap between the model it learns and the true model, ensuring that the learner can generalize well to unseen data. Consistency is an important property in machine learning algorithms as it guarantees convergence to the true model when the learner has access to enough data.

8. Write any two strengths of Bayes classifier.

Two strengths of the Bayes classifier are:

1. Efficient and Fast: The Bayes classifier can be very efficient and computationally fast, especially the Naïve Bayes variant. It only requires calculating the necessary probabilities and combining them using Bayes' theorem. The assumption of feature independence in Naïve Bayes simplifies the calculations, allowing for rapid training and classification even with large datasets.

2. Handles Irrelevant Features: The Bayes classifier can handle irrelevant features gracefully. Due to the assumption of feature independence in Naïve Bayes, irrelevant features have little impact on the classification decision. They do not affect the probabilities of the features that contribute to the classification, making the classifier robust to irrelevant or noisy features.

9. Write any two weaknesses of Bayes classifier.

Two weaknesses of the Bayes classifier are:

1. Strong Feature Independence Assumption: The Naïve Bayes classifier assumes that features are conditionally independent given the class variable. In reality, this assumption may not hold true, as features often exhibit dependencies. This assumption can lead to suboptimal performance if the features are not truly independent. However,

in practice, Naïve Bayes can still perform well despite this simplification.

2. Limited Representation Power: The Naïve Bayes classifier has limited representation power due to its simplistic assumption and linear decision boundaries. It cannot model complex relationships between features. If the true underlying distribution of the data is highly non-linear or contains complex interactions between features, Naïve Bayes may not be the most suitable classifier. Other more flexible classifiers, such as neural networks or decision trees, may be more appropriate.

10. Explain how Naïve Bayes classifier is used for:

1. Text classification:

Naïve Bayes is commonly used for text classification tasks, such as email spam detection or sentiment analysis. In text classification, each document or text sample is represented as a set of words or features. The Naïve Bayes classifier assumes that the occurrence of each word is conditionally independent given the class label. By estimating the probabilities of words belonging to each class based on training data, the classifier can assign a class label to new, unseen documents based on the likelihoods calculated from the observed words.

2. Spam filtering:

Naïve Bayes is widely used in spam filtering applications. In this context, the classifier is trained on a labeled dataset consisting of both spam and non-spam emails. The Naïve Bayes classifier learns the conditional probabilities of different words or features appearing in spam or non-spam emails. When applied to new incoming emails, the classifier calculates the probability that the email is spam or non-spam based on the observed words and assigns the appropriate label.

3. Market sentiment analysis:

Naïve Bayes can be used for market sentiment analysis, which involves determining the sentiment or opinion expressed in textual data, such as social media posts or customer reviews, about a particular company, product, or market. The classifier is trained on labeled data where the sentiment (positive, negative, or neutral) is associated with each text sample. By estimating the conditional probabilities of different words or features given each sentiment class, the classifier can predict the sentiment of new, unseen text samples based on the likelihoods calculated from the observed words.