|  |  |
| --- | --- |
| 1 | What is prior probability? Give an example. |
| Ans. | Prior probability shows the likelihood of an outcome in a given dataset. For example, in the mortgage case, P(Y) is the default rate on a home mortgage, which is 2%. P(Y|X) is called the conditional probability, which provides the probability of an outcome given the evidence, that is, when the value of X is known. |
| 2 | What is posterior probability? Give an example. |
| Ans. | 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. |
| 3 | What is likelihood probability? Give an example. |
| Ans. | Likelihood Function in Machine Learning and Data Science is the joint probability distribution (jpd) of the dataset given as a function of the parameter. Think of it as the probability of obtaining the observed data given the parameter values. |
| 4 | What is Naïve Bayes classifier? Why is it named so? |
| Ans. | Naive Bayes is a simple and powerful algorithm for predictive modeling. Naive Bayes is called naive because it assumes that each input variable is independent. This is a strong assumption and unrealistic for real data; however, the technique is very effective on a large range of complex problems. |
| 5 | What is optimal Bayes classifier? |
| Ans. | The Bayes Optimal Classifier is a probabilistic model that makes the most probable prediction for a new example. Bayes Optimal Classifier is a probabilistic model that finds the most probable prediction using the training data and space of hypotheses to make a prediction for a new data instance. |
| 6 | Write any two features of Bayesian learning methods. |
| Ans. | A probability distribution over observed data for each possible hypothesis. New instances can be classified by combining the predictions of multiple hypotheses, weighted by their probabilities. |
| 7 | Define the concept of consistent learners. |
| Ans. | Consistent Learners: A learner L using a hypothesis H and training data D is said to be a consistent learner if it always outputs a hypothesis with zero error on D whenever H contains such a hypothesis. • By definition, a consistent learner must produce a hypothesis in the version space for H given D. |
| 8 | Write any two strengths of Bayes classifier. |
| Ans. | This algorithm works quickly and can save a lot of time. Naive Bayes is suitable for solving multi-class prediction problems. If its assumption of the independence of features holds true, it can perform better than other models and requires much less training data. |
| 9 | Write any two weaknesses of Bayes classifier. |
| Ans. | The greatest weakness of the naïve Bayes classifier is that it relies on an often-faulty assumption of equally important and independent features which results in biased posterior probabilities. |
| 10 | Explain how Naïve Bayes classifier is used for |
| Ans. | * **Text classification**   The Naive Bayes classifier is a simple classifier that classifies based on probabilities of events. It is applied commonly to text classification. With the training set, we can train a Naive Bayes classifier which we can use to automatically categorize a new sentence.   * **Spam filtering**   Naive Bayes classifiers work by correlating the use of tokens (typically words, or sometimes other things), with spam and non-spam e-mails and then using Bayes' theorem to calculate a probability that an email is or is not spam. It is one of the oldest ways of doing spam filtering, with roots in the 1990s.   * **Market sentiment analysis**   Market Sentiment analysis is a field dedicated to extracting subjective emotions and feelings from text. One common use of sentiment analysis is to figure out if a text expresses negative or positive feelings. Naive Bayes is a popular algorithm for classifying text. |