1. What is prior probability? Give an example.

Answers: Prior probability, in Bayesian statistical inference, is the probability of an event before new data is collected. This is the best rational assessment of the probability of an outcome based on the current knowledge before an experiment is performed.

For example, historical data suggests that around 60% of students who start college will graduate within 6 years. This is the prior probability.

2. What is posterior probability? Give an example.

Answers: Posterior probability is a revised probability that takes into account new available information. Posterior probability is an adjustment on prior probability: Posterior probability = prior probability + new evidence (called likelihood).

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. This is the a priori probability. If we are given an additional piece of information that a ball was drawn at random from the selected urn, and that ball was black, what is the probability that the chosen urn is urn A? Posterior probability takes into account this additional information and revises the probability downward from 0.5 to 0.333 according to Bayes´ theorem, because a black ball is more probable from urn B than urn A.

3. What is likelihood probability? Give an example.

Answers: Probability is used to finding the chance of occurrence of a particular situation, whereas Likelihood is used to generally maximizing the chances of a particular situation to occur. Probabilities attach to results; likelihoods attach to hypotheses.

Example: Two samples, same distribution: two samples drawn from the same distribution have the same likelihood. This means both samples would lead us to the same conclusion.

Example: Same sample, two distributions: The likelihood functions of both distributions are proportional implying that they should reach the same conclusion but potentially leading to different conclusions.

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

Answers: 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?

Answers: 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.

Answers:

Each observed training example can incrementally decrease or increase the estimated

probability that a hypothesis is correct.

– This provides a more flexible approach to learning than algorithms that completely eliminate a hypothesis if it is found to be inconsistent with any single example.

• Prior knowledge can be combined with observed data to determine the final probability of a hypothesis. In Bayesian learning, prior knowledge is provided by asserting

– a prior probability for each candidate hypothesis, and

– a probability distribution over observed data for each possible hypothesis.

• Bayesian methods can accommodate hypotheses that make probabilistic predictions

• New instances can be classified by combining the predictions of multiple hypotheses, weighted by their probabilities.

• Even in cases where Bayesian methods prove computationally intractable, they can provide a standard of optimal decision making against which other practical methods can be measured

7. Define the concept of consistent learners.

Answers: A consistent learning algorithm is simply required to output a hypothesis that is consistent with all the training data provided to it. This notion of consistency is closely related to the empirical risk minimisation principle in the machine learning literature, where the risk is defined using the zero-one loss.

8. Write any two strengths of Bayes classifier.

Answers:

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.

Naive Bayes is better suited for categorical input variables than numerical variables.

9. Write any two weaknesses of Bayes classifier.

Answers:

Naive Bayes assumes that all predictors (or features) are independent, rarely happening in real life. This limits the applicability of this algorithm in real-world use cases.

This algorithm faces the ‘zero-frequency problem’ where it assigns zero probability to a categorical variable whose category in the test data set wasn’t available in the training dataset. It would be best if you used a smoothing technique to overcome this issue.

Its estimations can be wrong in some cases, so you shouldn’t take its probability outputs very seriously.

10. Explain how Naïve Bayes classifier is used for

1. Text classification

2. Spam filtering

3. Market sentiment analysis

Answers:

1. The Naive Bayes classifier is a simple classifier that classifies based on probabilities of events. It is the applied commonly to text classification. Though it is a simple algorithm, it performs well in many text classification problems. Other Pros include less training time and less training data. That is, less CPU and Memory consumption. As with any machine learning model, we need to have an existing set of examples (training set) for each category (class).
2. 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.
3. 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.