ASSIGNMENT - 12

1. What is prior probability? Give an example.

Ans: Prior probability refers to the initial belief or chance of an event happening before we consider any evidence. It reflects our existing knowledge or assumptions.

Example: You're flipping a fair coin. The prior probability of getting heads (or tails) is 50% since there's no reason to believe one side is more likely to land than the other.

2. What is posterior probability? Give an example.

Ans: Posterior probability is the updated probability of an event occurring after we have considered evidence. It combines the prior probability with the new information.

Example: You flip the coin and see heads. Now, the posterior probability of the next flip being heads is still 50% (assuming the coin is fair). However, if you flip the coin 10 times and get heads 8 times, the posterior probability of the next flip being heads increases (though it won't be exactly 80% due to randomness).

3. What is likelihood probability? Give an example.

Ans: Likelihood probability is the probability of observing a particular piece of evidence given that a specific hypothesis is true. It focuses on how likely the evidence is under a certain assumption.

Example: In the coin flip example, the likelihood of seeing heads given the coin is fair is 50%. However, if the coin is biased towards heads, the likelihood of seeing heads would be higher (e.g., 70%).

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

Ans: A Naive Bayes classifier is a classification algorithm based on Bayes' theorem. It predicts the class (category) of a new data point by considering the probability of each class given the features of the data point.

It's called "naive" because it assumes independence between features, which may not always hold true in real-world data. This simplification allows for efficient calculations but can affect accuracy.

5. What is optimal Bayes classifier?

Ans: The optimal Bayes classifier is a theoretical concept that represents the best possible classification method based on Bayes' theorem. It would require knowing the true probabilities of all features and classes, which is often impractical in real-world scenarios with limited data.

6. Write any two features of Bayesian learning methods.

Ans: Incorporation of Prior Knowledge: Unlike some algorithms that only learn from data, Bayesian methods allow you to integrate existing knowledge through prior probabilities.

Probabilistic Predictions: They provide probability estimates for class membership, offering a sense of confidence in the predictions.

7. Define the concept of consistent learners.

Ans: Consistent learners are algorithms that, with increasing amounts of data, converge towards the true underlying relationship between features and classes. In simpler terms, their predictions become more accurate as they learn from more data.

8. Write any two strengths of Bayes classifier.

Ans: Strengths of Bayes Classifiers:

* Simple and Efficient: Naive Bayes classifiers are relatively easy to understand and implement, making them a good choice for quick classification tasks.
* Effective for High-Dimensional Data: They can handle datasets with many features without significant performance degradation, unlike some other algorithms.

9. Write any two weaknesses of Bayes classifier.

Ans: Weaknesses of Bayes Classifiers:

* Naive Independence Assumption: The assumption of independence between features can lead to inaccurate predictions if features are actually correlated.
* Sensitivity to Data Quality: The performance of Naive Bayes classifiers can be affected by noisy or irrelevant data in the training set.

10. Explain how Naïve Bayes classifier is used for

* Text classification
  + Feature Extraction: Text is converted into features, often using a Bag-of-Words (BoW) approach. Each word becomes a feature, and its frequency in the document is its value.
  + Training: The classifier is trained on labeled data. Each document has a category (e.g., sports news, politics) associated with it. The Naive Bayes model calculates the probability of each word appearing in each category.
  + Classification: For a new document, the model calculates the probability of the document belonging to each category based on the word frequencies. It uses Bayes' theorem to consider both the prior probability of each category and the likelihood of the words appearing in that category.
    - The document is assigned to the category with the highest posterior probability.
* Spam filtering
  + Similar to text classification, Naive Bayes excels in spam filtering:
  + Features: Words, presence of URLs, sender information, etc., become features.
  + Training: The model learns the probabilities of these features appearing in spam and non-spam emails from labeled data sets.
  + Classification: New emails are analyzed for feature presence. The model calculates the probability of the email being spam based on the likelihood of these features in spam emails.
    - Emails with a high probability of being spam are flagged.
* Market sentiment analysis
  + Features: Words and phrases that express positive, negative, or neutral sentiment become features.
  + Training: The model is trained on labeled financial news articles or social media posts where the sentiment (positive, negative, or neutral) towards the market is identified.
  + Classification: For new text, the model calculates the probability of it belonging to each sentiment class based on the presence of sentiment-indicating words.
    - The text is classified as expressing positive, negative, or neutral sentiment based on the highest probability.