1)Bayes Theorem is a method to determine conditional probabilities – that is, the probability of one event occurring given that another event has already occurred.

Naïve Bayes theorem is also a supervised algorithm, which is based on Bayes theorem and used to solve classification problems. It is one of the most simple and effective classification algorithms in Machine Learning which enables us to build various ML models for quick predictions. It is a probabilistic classifier that means it predicts on the basis of probability of an object.

Naïve Bayes classifier is one of the simplest applications of Bayes theorem which is used in classification algorithms to isolate data as per accuracy, speed and classes.

2) Conditional probability is defined as the likelihood of an event or outcome occurring, based on the occurrence of a previous event or outcome.

Bayes theorem calculates the conditional probability of the occurrence of an event based on prior knowledge of conditions that might be related to the event.

3) Naïve Bayes Imputation (NBI) is used to fill in missing values by replacing the attribute information according to the probability estimate. The NBI process divides the whole data into two sub-sets is the complete data and data containing missing data. Complete data is used for the imputation process at the lost value.

4) Naive Bayes works best in two cases: completely independent features and functionally dependent features.

Real world examples are:

sentimental analysis- Check a piece of text expressing positive emotions, or negative emotions?

classifying new articles- Classify a news article about technology, politics, or sports ?

spam filtration- To mark an email as spam, or not spam ?

5) some assumptions that the Naive Bayers algorithm makes: The main assumption is that it assumes that the features are conditionally independent of each other. Each of the features is equal in terms of weightage and importance. The algorithm assumes that the features follow a normal distribution.

6)Laplace smoothing is used to address the issue of zero probabilities for certain events in the dataset. It prevents the model from assigning zero probability to an event that it has not seen during training.

Laplacian smoothing addresses this issue by adding a small constant to the count of each possible feature value for each class. This effectively "smoothes" the probability distribution and prevents the occurrence of zero probabilities. By doing so, Laplacian smoothing helps to improve the generalization of the Naive Bayes classifier and makes it less sensitive to rare or unseen feature values.

7)The independence assumption in Naive Bayes states that the features are conditionally independent given the class. This assumption simplifies the calculation of the likelihood and allows the model to estimate the probability of a class based on the presence of individual features.

However, in real-world datasets, this assumption is often violated due to the presence of correlated, irrelevant, and uncertain variables.

8)A posterior probability is the revised or updated probability of an event occurring after taking into consideration new information. The posterior probability is calculated by updating the prior probability using Bayes' theorem.

The prior probability in Naive Bayes classification refers to the probability of each class occurring in the dataset, without considering any features. It represents the initial belief about the distribution of the classes in the absence of any evidence from the features.

The formula for calculating the posterior probability in Naive Bayes classification is P(c|x) = P(c) \* P(x|c) / P(x), where P(c|x) is the posterior probability of class c given feature x, P(c) is the prior probability of class c, P(x|c) is the likelihood of feature x given class c, and P(x) is the probability of feature x.

As a result, the naive Bayes classifier is a powerful tool in machine learning, particularly in text classification, spam filtering, and sentiment analysis, among others.

9) Overfitting occurs when the model fits more data than required, and it tries to capture each and every datapoint fed to it. Hence it starts capturing noise and inaccurate data from the dataset, which degrades the performance of the model. An overfitted model is said to have low bias and high variance.

Ways to prevent overfitting:

1. Early Stopping
2. Train with more data
3. Feature Selection
4. Cross-Validation
5. Data Augmentation
6. Regularization

10) Increasing the number of features in a naive Bayes classifier does not always guarantee an improvement in performance. While more features can potentially capture more information, they can also lead to overfitting, increased computational complexity, and the inclusion of irrelevant or redundant information. It's important to carefully consider the quality and relevance of the features being added to ensure that they contribute positively to the classifier's performance.