



Daily Machine Learning Interview Questions





11. What Is ‘naive’ in the Naive Bayes Classifier?





The classifier is called 'naive' because it makes assumptions that may or may not turn out to be correct. The algorithm assumes that the presence of one feature of a class is not related to the presence of any other feature (absolute independence of features), given the class variable.





For instance, a fruit may be considered to be a cherry if it is red in color and round in shape, regardless of other features. **This assumption may or may not be right (as an apple also matches the description).**





12. How Will You Know Which Machine Learning Algorithm to Choose for Your Classification Problem?





While there is no fixed rule to choose an algorithm for a classification problem, you can follow these guidelines:

- **If accuracy is a concern, test different algorithms and cross-validate them .**
- **If the training dataset is small, use models that have low variance and high bias .**
- **If the training dataset is large, use models that have high variance and little bias.**





13. When Will You Use Classification over Regression?





Classification is used when your target is categorical, while regression is used when your target variable is continuous.

Both classification and regression belong to the category of supervised machine learning algorithms.

Examples of classification problems include:



- **Predicting yes or no**
- **Estimating gender**
- **Breed of an animal**
- **Type of color**

Examples of regression problems include:

- **Estimating sales and price of a product**
- **Predicting the score of a team.**
- **Predicting the amount of rainfall.**



14. What is a Random Forest?





A ‘random forest’ is a supervised machine learning algorithm that is generally used for classification problems. It operates by constructing multiple decision trees during the training phase. The random forest chooses the decision of the majority of the trees as the final decision.





15. What is the Trade-off Between Bias and Variance?



Bias :

Bias in a machine learning model occurs when the predicted values are further from the actual values. Low bias indicates a model where the prediction values are very close to the actual ones.

Underfitting: High bias can cause an algorithm to miss the relevant relations between features and target outputs.

Variance:

Variance refers to the amount the target model will change when trained with different training data. For a good model, the variance should be minimized.

Overfitting:

High variance can cause an algorithm to model the random noise in the training data rather than the intended outputs.



Thank You

