

Question and Answers Document

1. What is the trade-off between bias and variance?

Bias is the error that a model makes when it is systematically too high or too low. Variance is the error that a model makes when it is too sensitive to small changes in the data. The goal of machine learning is to find a model that has low bias and low variance. However, it is often difficult to find a model that perfectly balances these two errors. In some cases, it may be necessary to sacrifice some bias in order to reduce variance, or vice versa.

2. What are the different types of classification and regression tasks?

Classification tasks are those in which the goal is to predict a class label for each data point. Regression tasks are those in which the goal is to predict a continuous value for each data point.

3. What is an imbalanced dataset? How would you handle it?

An imbalanced dataset is one in which the classes are not evenly represented. This can make it difficult to train a model that can accurately predict the minority class. One way to handle an imbalanced dataset is to use oversampling or undersampling. Oversampling involves duplicating the minority class data points, while undersampling involves removing some of the majority class data points.

4. Why is the XOR problem cannot be solved by a single perceptron?

The XOR problem is a Boolean function that takes two inputs and outputs 1 if the inputs are different and 0 if the inputs are the same. A single perceptron can only learn linear functions, which cannot represent the XOR function.

5. What is a Markov decision process?

A Markov decision process (MDP) is a mathematical model that describes a system that evolves over time. The state of the system at any given time is determined by the previous state and the actions that have been taken. The goal of an MDP is to find a policy that maximizes the expected reward over time.

6. Discuss the importance of pooling layer in convolutional neural networks.

Pooling layers are used in convolutional neural networks to reduce the size of the feature maps while preserving the most important features. This is important because it allows the network to learn more complex features without overfitting the data.

7. What is F1 score? Explain its significance.

The F1 score is a measure of the performance of a classification model. It is calculated as the harmonic mean of the precision and recall. The precision is the fraction of predicted positive instances that are actually positive, and the recall is the fraction of actual positive instances that are predicted positive. The F1 score is a good measure of overall model performance because it takes into account both precision and recall.

8. Explain overfitting and underfitting using a simple example.

Overfitting is a problem that occurs when a model learns the training data too well and does not generalize well to new data. This can happen when the model is too complex or when the training data is not representative of the data that the model will be used on.

Underfitting is a problem that occurs when a model does not learn the training data well enough and does not generalize well to new data. This can happen when the model is too simple or when the training data is not large enough.

9. Explain the various methods to perform dimensionality reduction.

Dimensionality reduction is a technique that is used to reduce the number of features in a dataset. This can be useful when the dataset is too large or when the features are correlated. There are many different methods for dimensionality reduction, including principal component analysis (PCA), linear discriminant analysis (LDA), and t-distributed stochastic neighbor embedding (t-SNE).

10. What are the different types of AI algorithms?

There are many different types of AI algorithms, each with its own strengths and weaknesses. Some common types of AI algorithms include supervised learning, unsupervised learning, reinforcement learning, and natural language processing.