

Exam Preparation Notes

Your Name

26 июля 2025 г.

Содержание

1	What is the difference between supervised and unsupervised learning?	3
2	Explain the concept of overfitting in machine learning.	3

1 What is the difference between supervised and unsupervised learning?

Answer

Supervised learning and unsupervised learning are two main categories of machine learning algorithms.

Supervised Learning

In supervised learning, the algorithm learns from a labeled dataset. This means that each data point is tagged with a correct output or label. The goal is to learn a mapping function that can predict the output for new, unseen data.

Example: Predicting house prices based on features like area, number of bedrooms, and location. The model is trained on a dataset of houses with known prices.

Unsupervised Learning

The algorithm learns from an unlabeled dataset. The goal is to find hidden patterns, structures, or relationships within the data without any predefined labels.

Example: Clustering customers into different segments based on their purchasing behavior. The algorithm identifies the clusters without any prior knowledge of customer groups.

2 Explain the concept of overfitting in machine learning.

Answer

Overfitting is a common problem in machine learning where a model learns the training data too well, to the point that it also learns the noise and random fluctuations in the data. This negatively impacts the model's performance on new, unseen data.

Causes of Overfitting

- **Model Complexity:** A model that is too complex for the given data can easily overfit.
- **Insufficient Data:** With a small training dataset, the model may not be able to generalize well.

How to Prevent Overfitting

- **Cross-Validation:** Use techniques like k-fold cross-validation to ensure the model performs well on unseen data.
- **Regularization:** Add a penalty term to the loss function to penalize large coefficients.
- **Simpler Model:** Choose a less complex model.