Types of Machine Learning ⌄

Quiz

1 **What is the primary goal of machine learning?**

- ● The primary goal of machine learning is to enable computers to learn from data and improve their performance on specific tasks over time.
- ○ The primary goal of machine learning is to create machines that can perform only a single task.
- ○ The primary goal of machine learning is to replace human decision-making entirely.

Submitted   Correct!

🐞 Report issue

2 **What is the difference between supervised and unsupervised learning?**

- ☐ Supervised learning only uses unlabeled data and does not require any correct answers during training.
- ☑ Supervised learning involves training models on labeled data, where the correct answers are provided, and the model learns to map input data to output labels.
- ☐ Supervised learning and unsupervised learning are interchangeable terms with no distinct differences.
- ☑ Unsupervised learning analyzes unlabeled data to discover patterns or relationships without explicit output labels.

Submitted   Correct!

🐞 Report issue

Quiz ⌄

3 **Select two examples of applications for machine learning in real-world scenarios.**

- ☐ Machine learning is exclusively used in robotics
- ☐ Machine learning is only used in theoretical research and does not have practical applications in real-world scenarios.
- ☑ Machine learning is applied in healthcare for disease prediction and diagnosis, using patient data to identify potential health issues.
- ☑ In marketing, machine learning is utilized for customer segmentation, allowing businesses to target specific groups with personalized advertisements.

Submitted   Correct!

1 **True/False: The primary goal of the case study is to understand the practical aspects of applying machine learning techniques to predict and mitigate customer churn.** ◎

- ○ True
- ● False

Submitted   Correct!

🐞 Report issue

2 **True/False: The dataset for this case study includes customer demographics, usage patterns, contract details, and customer service interactions.** ◎

- ● True
- ○ False

Submitted   Correct!

1 **What does the hypothesis (h (x)) represent in machine learning?** ◎

Activity description...

- ○ The target variable to be predicted.
- ○ The ground truth or actual output of the training data.
- ○ The features of the dataset.
- ● The predicted output based on the input features and model parameters.

Submitted   Correct!

🐞 Report issue

3 **Scenario Question**

You are a data scientist working on a binary classification problem. You have tried two different models for the task. Model A uses a simple hypothesis space with a linear model, while Model B employs a more complex hypothesis space with a high-degree polynomial. After evaluating both models, you notice that Model B fits the training data almost perfectly, but its performance on new, unseen data is not as good. On the other hand, Model A generalizes better to unseen data.

Based on this scenario, which model is likely suffering from overfitting?

- ● Model B, because it fits the training data almost perfectly.
- ○ Model A, because it uses a linear hypothesis space.
- ○ Neither model is suffering from overfitting.
- ○ Both models are suffering from overfitting.

Submitted   Correct!

🐞 Report issue

3 Why is model evaluation crucial in the machine learning workflow?

○ Model evaluation measures the performance of the model and ensures its effectiveness on new, unseen data.

○ Model evaluation is unnecessary as machine learning models always perform perfectly.

○ Model evaluation provides a way to quickly complete the machine learning project.

Submitted   Correct!

⚐ Report issue

---

3 Select the correct machine learning type for Supervised learning

○ Uses labeled data to make predictions based on previously known outcomes.

○ Involves agents interacting with an environment to maximize rewards through learned actions.

○ Trains models on unlabeled data to discover hidden patterns and groupings.

Submitted   Correct!

⚐ Report issue

Quiz

1 What is the key distinction between supervised and unsupervised learning?

○ Supervised learning uses labeled data, while unsupervised learning works with unlabeled data to find patterns

○ Unsupervised learning uses labeled data, while supervised learning works with unlabeled data to find patterns

Submitted   Correct!

⚐ Report issue

2 Select the correct example of a real-world application for each type of machine learning.

○ Supervised learning in email spam filtering, reinforcement learning in customer segmentation, and unsupervised learning in training self-driving cars.

○ Supervised learning in email spam filtering, unsupervised learning in customer segmentation, and reinforcement learning in training self-driving cars.

Submitted   Correct!

Completed activities 3/3 100%

---

1 What is the difference between supervised and unsupervised machine learning?. Which of the following statements are true?

☑ Supervised learning is used when we want to predict a certain outcome from a given input.

☐ Supervised learning problems can be grouped into clustering and association.

☑ The goal for unsupervised learning is to model the underlying structure or distribution in the data.

☐ There are two major types of supervised learning problems, called clustering and regression.

Submitted   Correct!

⚐ Report issue

2 Select which of the following scenarios are regression problems.

☑ Predict the score that a student will achieve in an exam whose grade can be 0.1, 2,. . . , 10

☐ Predict whether a user will churn from the service.

☑ Predict how much a company will spend on electricity the next semester.

☐ Given a tweet, determine whether or not it contains text against or on favor for a presidential candidate.

Submitted   Correct!

Completed activities 4/4 100%

⚐ Report issue

---

3 Select which of the following scenarios are classification problems.

☑ An algorithm is trained to recognize spam email by learning the characteristics of what constitutes spam vs non-spam email.

☐ Predict the prices of a house in Boston based on zipcode, neighbourhood, the per capita crime rate by town, etc

☑ Determine whether a customer is likely to purchase more items or not

☐ Impact of blood alcohol content on coordination

Submitted   Correct!

⚐ Report issue

4 Suppose you want to develop a supervised machine learning model to predict whether a superhero will fly or not. Which of the following statements are true?

○ We'll use unlabeled examples to train the model.

○ A regression model is the best way to predict the probability to fly.

○ This is not a machine learning problem

○ A classification model provide the best approach.

Submitted   Correct!

⚐ Report issue

Completed activities 4/4 100%

---

1 True or False: Clustering is a common task in unsupervised learning, where data points are grouped together based on similarity

○ True

○ False

Submitted   Correct!

⚐ Report issue

2 Imagine you work for an e-commerce company. The company has a large database of customer transactions, and your task is to segment customers into different groups for targeted marketing. Which unsupervised learning technique would you use?

○ For customer segmentation in e-commerce, you can use a dimensional reduction algorithms like PCA.

○ For customer segmentation in e-commerce, you can use clustering algorithms like K-means or hierarchical clustering.

Submitted   Correct!

⚐ Report issue

1 **Customer Segmentation for Marketing Strategy**

Given a dataset containing customer demographics and purchase history, how can we group customers based on their similarities to tailor marketing strategies?

○ Supervised Learning
◉ Unsupervised Learning

Submitted   Correct!

⚐ Report issue

3 **You are a financial analyst at a bank and need to detect potentially fraudulent transactions in a credit card dataset. How could unsupervised learning be applied to identify unusual transaction patterns indicative of fraud?**

○ Unsupervised learning can be applied to detect fraudulent transactions by training a supervised machine learning model on a labeled dataset of fraudulent and non-fraudulent transactions.

◉ Unsupervised learning can be applied to detect fraudulent transactions by clustering normal and abnormal transaction patterns or using outlier detection methods to identify unusual transaction behavior.

○ Unsupervised learning can NOT be applied to detect fraudulent transactions.

Submitted   Correct!

⚐ Report issue

2 **Anomaly Detection in Network Traffic**

How can we identify unusual patterns or anomalies in network traffic that may indicate a security breach, without having prior labeled examples of such incidents?

○ Both options are possible
○ Supervised Learning
◉ Unsupervised Learning

Submitted   Correct!

⚐ Report issue

Completed activities 6/6   100%

3 **Predicting Housing Prices based on Features**

Given historical data on housing prices and features such as location, size, and amenities, can we build a model to predict the prices of new houses?

○ Unsupervised Learning
○ Both options are possible
◉ Supervised Learning

Submitted   Correct!

⚐ Report issue

4 **Predicting Customer Satisfaction for an E-commerce Website**

Given customer feedback data on an e-commerce website where customers can rate their satisfaction on a scale from 1 to 5, should we model this problem as a classification or a regression task?

○ Regression problem
◉ Both options are possible
○ Classification problem

Submitted   Correct!

⚐ Report issue

Completed activities 6/6   100%

5 **You are a healthcare researcher aiming to identify potential subgroups of patients based on their medical records to personalize treatment plans. How could unsupervised learning be utilized to uncover distinct patient clusters with similar medical profiles, allowing for more targeted and effective healthcare interventions?**

○ Unsupervised learning is not applicable in this scenario, and supervised learning should be used instead.

◉ Unsupervised learning can segment patients into clusters based on their medical records, revealing distinct subgroups without using any predefined labels.

○ Unsupervised learning can only be used if the dataset contains labeled patient groups.

○ Unsupervised learning can classify patients into predefined categories based on their medical records.

Submitted   Correct!

⚐ Report issue

6 **You are an e-commerce manager and want to predict whether customers are likely to make a purchase during their website visit. How can supervised learning be applied to develop a predictive model that helps classify customers into 'potential buyers' and 'non-buyers' based on historical data and labeled purchase information?**

Supervised learning can only be used if the purchase data is anonymized and doesn't contain

Completed activities 6/6   100%

1 True or False: Incorporating irrelevant features in a machine learning model can improve its predictive performance

○ True
● False

Submitted   Correct!

⚐ Report issue

2 True or False: Feature selection in machine learning helps in reducing the dimensionality of the dataset.

○ False
● True

Submitted   Correct!

A. Sender's Email Address B. Word Frequency in the Email C. Font Style and Size D. Number of Links in the Email E. Email Attachment Size

Select the correct options:

○ A, C, E
○ B, D, A
● B, D, E
○ A, D, E

Submitted   Correct!

⚐ Report issue

4 Scenario Question

You are working on a machine learning project to predict customer churn for a subscription-based service. The dataset you have includes customer information such as age, usage patterns, subscription type, and a unique customer ID for each record.

Why would using the customer ID as a feature for predicting churn be ineffective?

○ Using the customer ID as a feature may raise privacy concerns and ethical considerations, as it directly identifies individuals.

○ The customer ID is a unique identifier and does not contain predictive information regarding the likelihood of churn.

○ The customer ID is not related to the underlying reasons for churn and including it may introduce

Completed activities 4/4   100%

---

1 True or False: Having a large dataset guarantees high data quality

● False
○ True

Submitted   Correct!

⚐ Report issue

2 True or False: Data quality is a subjective term and does not have a universally agreed-upon definition

● True
○ False

Submitted   Correct!

⚐ Report issue

3 Scenario Question

You are working on a project to predict student performance based on various factors such as study hours, attendance, and extracurricular activities. The dataset you have includes information about these features for each student.

In this scenario, what would be an example of a high-quality dataset ?

A. A dataset with accurate and up-to-date information about study hours, attendance, and extracurricular activities for a diverse set of students.

B. A dataset with study hours as the only feature for a small subset of students.

C. A dataset with random and inconsistent entries for study hours, attendance, and extracurricular activities.

D. A dataset with missing values for most of the entries related to study hours, attendance, and extracurricular activities.

Please select the most appropriate option and provide reasoning for your choice.

○ B
○ C
○ D
● A

Did you like this project?

---

1 True or False: A larger hypothesis space always leads to better model performance.

● False
○ True

Submitted   Correct!

⚐ Report issue

2 True or False: Overfitting occurs when the hypothesis space is too complex for the given data.

○ False
● True

Submitted   Correct!

⚐ Report issue

---

3 Scenario Question

You are a data scientist working on a binary classification problem. You have tried two different models for the task. Model A uses a simple hypothesis space with a linear model, while Model B employs a more complex hypothesis space with a high-degree polynomial. After evaluating both models, you notice that Model B fits the training data almost perfectly, but its performance on new, unseen data is not as good. On the other hand, Model A generalizes better to unseen data.

Based on this scenario, which model is likely suffering from overfitting?

● Model B, because it fits the training data almost perfectly.

○ Model A, because it uses a linear hypothesis space.

○ Neither model is suffering from overfitting.

○ Both models are suffering from overfitting.

Submitted   Correct!

⚐ Report issue

1 True/False: The primary goal of the case study is to understand the practical aspects of applying machine learning techniques to predict and mitigate customer churn.

○ True

◉ False

Submitted   Correct!

⊙ Report issue

2 True/False: The dataset for this case study includes customer demographics, usage patterns, contract details, and customer service interactions.

◉ True

○ False

Submitted   Correct!