

## **MACHINE LEARNING CONCEPTS**

### **1. Machine learning concepts**

00:00 - 00:06

Nice job so far! Now, we'll cover important concepts in machine learning.

### **2. Three types of machine learning**

00:06 - 00:35

There are three types of machine learning. The first listed, reinforcement learning, is used for deciding sequential actions, like a robot deciding its path or its next move in a chess game. Reinforcement learning is not as common as the others and uses complex mathematics, like game theory. It won't be covered any further. The most common types are supervised and unsupervised learning. Their main difference lies in their training data.

### **3. Training data**

00:35 - 00:55

Remember, how we said machine learning "learns" patterns from existing data and applies it to new data? We call this existing data "training data". When a model is being built and learning from training data, we call this "training a model". This can take nanoseconds to weeks depending on the size of the data.

### **4. Supervised learning training data**

00:55 - 01:10

Let's look at training data for a supervised learning model. We'd like to train a model to predict whether a patient has heart disease. We have existing records from patients who've experienced chest pains and been tested for heart disease.

### **5. Supervised learning training data**

01:10 - 01:15

Our target variable is "heart disease", because this is what we want to predict.

### **6. Supervised learning training data**

01:15 - 01:29

The values "True" and "False" are labels for the target variable, meaning whether it's true or false that a patient has heart disease. Labels don't have to come in this form - they can be numbers or categories.

### **7. Supervised learning training data**

01:29 - 01:39

These rows are the observations or examples that our model will learn from. You should get as many of these as possible.

## **8. Supervised learning training data**

01:39 - 02:07

And, these columns are features. Features are different pieces of information that might help predict the target. Age, cholesterol, and smoking habits are known factors of heart disease. The magic of machine learning is that we can analyze many features at once, even the ones we're unsure about, and find relationships between different features. We input labels and features as data to train the model.

## **9. After training (supervised learning)**

02:07 - 02:15

Once training is done, we can give the model new input. In our case, a new patient.

## **10. After training (supervised learning)**

02:15 - 02:17

The features are inputted

## **11. After training (supervised learning)**

02:17 - 02:19

and, the model outputs its prediction.

## **12. Supervised vs unsupervised learning**

02:19 - 02:52

In supervised learning, the training data is "labeled", meaning the values of our target are known. For instance, we knew if previous patients had heart disease based on the labels "true" and "false". In unsupervised learning, we don't have labels, only features. What can we do with this? Usually tasks like anomaly detection and clustering, which divides data into groups based on similarity. Let's explore this with our dataset.

## **13. Unsupervised learning training data**

02:52 - 03:12

There are different treatments for heart disease. Different types of patients respond better or worse to certain treatments. We can use unsupervised learning to understand the different types of patients we have. Let's filter our dataset to only include patients with heart disease. We can pass it into a clustering model

## **14. Unsupervised learning training data**

03:12 - 03:36

and get categories of patients based on feature similarity. For example, one category could be patients with high cholesterol and blood sugar level of a certain age range. Note, we didn't know these categories and, even, the number of categories before running this. With this output, we can group patients and research better treatments for each group.

### **15. After training (unsupervised learning)**

03:36 - 03:45

Now, with a new patient, we can input the features into the model and get which patient type they best fit into.

### **16. Unsupervised Learning**

03:45 - 04:18

In reality, data doesn't always come with labels. Either it's too much manual work to label or we don't even know what the labels are. Think of the effort it would take to label millions of road images for self-driving cars. This is when unsupervised learning shines. In this sense, the model is unsupervised and finds it's own patterns. In chapter 2, we'll dig more into how unsupervised and supervised learning work and their use-cases.