

Applications of data science

1. Applications of data science

00:00 - 00:10

Previously, you learned the definition of data science and the steps in a data science workflow. In this lesson, you'll learn how data science can be applied to real-world problems.

2. More case studies

00:10 - 00:20

Let's take a deep dive into three exciting areas of data science: traditional machine learning, the Internet of Things, and deep learning.

3. Case study: fraud detection

00:20 - 00:30

Suppose you work in fraud detection at a large bank. You'd like to use data to determine the probability that the transaction is fake.

4. Case study: fraud detection

00:30 - 01:11

To answer this question, you might start by gathering information about each purchase, such as the amount, date, location, purchase type, and card-holder address. You'll need many examples of transactions, including this information, as well as a label that tells you whether each transaction is valid or fraudulent. Luckily, you probably have this information in a database. These records are called "training data", and are used to build an algorithm. Each time a new transaction occurs, you'll give your algorithm information, like amount and date, and it will answer the original question: What is the probability that this transaction is fraudulent?

5. What do we need for machine learning?

01:11 - 01:53

Before we can answer that question, let's walk through our example and highlight what we need for machine learning to work its magic. First, a data science problem begins with a well-defined question. Our question was "What is the probability that this transaction is fraudulent?" Next, we need some data to analyze. We have months of old credit card transactions and associated metadata, like date and location, that have already been identified as either fraudulent or valid. Finally, we need additional data every time we want to make a new prediction. We need to have the same type of information on every new purchase so that we could label it as either "fraudulent" or "valid".

6. Case study: smart watch

01:53 - 02:29

Now, suppose you're trying to build a smart watch to monitor physical activity. You want to be able to auto-detect different activities, such as walking or running. Your smart watch is equipped with a special sensor, called an "accelerometer", that monitors motion in three dimensions. The data generated by this sensor is the basis of your machine learning problem. You could ask several volunteers to wear your watch and record when they are running or walking. You could then develop an algorithm that recognizes accelerometer data as representing one of those two states: walking or running.

7. Internet of Things (IoT)

02:29 - 03:04

Your smart watch is part of a fast growing field called "the Internet of Things", also known as IoT, which is often combined with Data Science. IoT refers to gadgets that are not standard computers, but still have the ability to transmit data. This includes smart watches, internet-connected home security systems, electronic toll collection systems, building energy management systems, and much, much more. IoT data is a great resource for data science projects!

8. Case study: image recognition

03:04 - 03:16

Let's tackle another example. A key task for self-driving cars is identifying when an image contains a human. What would be the dataset for this problem?

9. Case study: image recognition

03:16 - 03:32

We could express the picture as a matrix of numbers where each number represents a pixel. However, this approach would probably fail if we fed the matrix into a traditional machine learning model. There's simply too much input data!

10. Deep learning

03:32 - 04:03

We need more advanced algorithms from a subfield of machine learning called deep learning. In deep learning, multiple layers of mini-algorithms, called "neurons", work together to draw complex conclusions. Deep learning takes much, much more training data than a traditional machine learning model, but is also able to learn relationships that traditional models cannot. Deep learning is used to solve data-intensive problems, such as image classification or language understanding.

11. Let's practice!

04:03 - 04:10

You're now familiar with several different applications of Data Science. Let's practice!

