100 XP

What is machine learning?

5 minutes

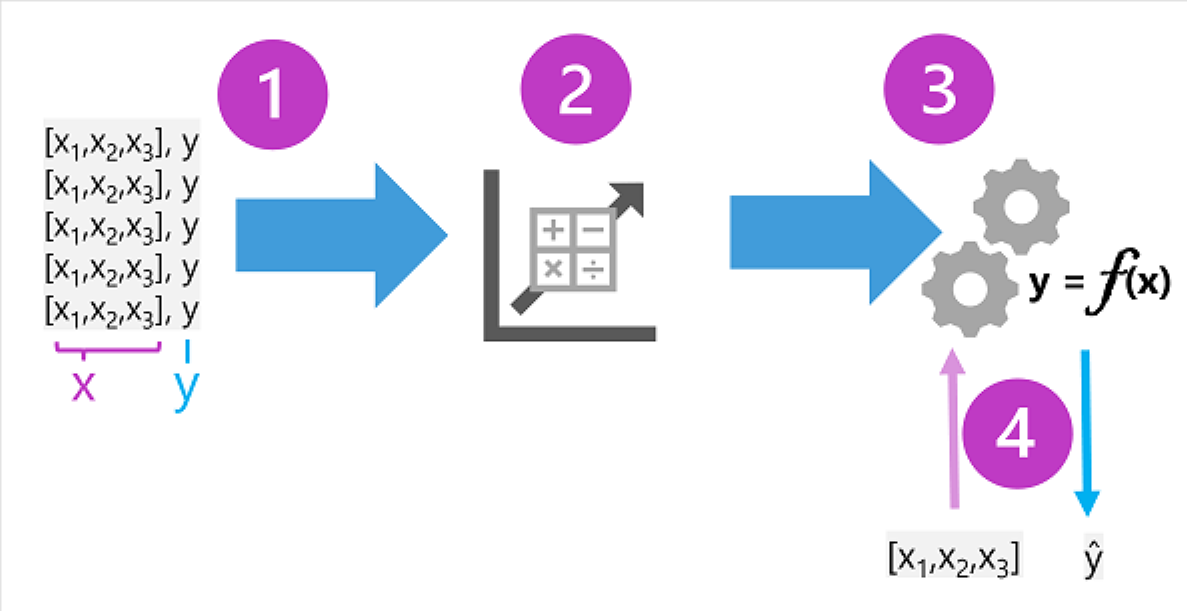
Machine learning has its origins in statistics and mathematical modeling of data. The fundamental idea of machine learning is to use data from past observations to predict unknown outcomes or values. For example:

- The proprietor of an ice cream store might use an app that combines historical sales and weather records to predict how many ice creams they're likely to sell on a given day, based on the weather forecast.
- A doctor might use clinical data from past patients to run automated tests that predict whether a new patient is at risk from diabetes based on factors like weight, blood glucose level, and other measurements.
- A researcher in the Antarctic might use past observations automate the identification of different penguin species (such as *Adelie*, *Gentoo*, or *Chinstrap*) based on measurements of a bird's flippers, bill, and other physical attributes.

Machine learning as a *function*

Because machine learning is based on mathematics and statistics, it's common to think about machine learning models in mathematical terms. Fundamentally, a machine learning model is a software application that encapsulates a *function* to calculate an output value based on one or more input values. The process of defining that function is known as *training*. After the function has been defined, you can use it to predict new values in a process called *inferencing*.

Let's explore the steps involved in training and inferencing.



1. The training data consists of past observations. In most cases, the observations include the observed attributes or *features* of the thing being observed, and the known value of the thing you want to train a model to predict (known as the *label*).

   In mathematical terms, you'll often see the features referred to using the shorthand variable name ***x***, and the label referred to as ***y***. Usually, an observation consists of multiple feature values, so ***x*** is actually a *vector* (an array with multiple values), like this: $[x_1,x_2,x_3,...]$.

   To make this clearer, let's consider the examples described previously:

   - In the ice cream sales scenario, our goal is to train a model that can predict the number of ice cream sales based on the weather. The weather measurements for the day (temperature, rainfall, windspeed, and so on) would be the *features* (***x***), and the number of ice creams sold on each day would be the *label* (***y***).
   - In the medical scenario, the goal is to predict whether or not a patient is at risk of diabetes based on their clinical measurements. The patient's measurements (weight, blood glucose level, and so on) are the *features* (***x***), and the likelihood of diabetes (for example, *1* for at risk, *0* for not at risk) is the *label* (***y***).
   - In the Antarctic research scenario, we want to predict the species of a penguin based on its physical attributes. The key measurements of the penguin (length of its flippers, width of its bill, and so on) are the *features* (***x***), and the species (for example, *0* for Adelie, *1* for Gentoo, or *2* for Chinstrap) is the *label* (***y***).

2. An *algorithm* is applied to the data to try to determine a relationship between the features and the label, and generalize that relationship as a calculation that can be performed on ***x*** to calculate ***y***. The specific algorithm used depends on the kind of predictive problem you're

trying to solve (more about this later), but the basic principle is to try to *fit* the data to a function in which the values of the features can be used to calculate the label.

3. The result of the algorithm is a *model* that encapsulates the calculation derived by the algorithm as a *function* - let's call it ***f***. In mathematical notation:

   $y = f(x)$

4. Now that the *training* phase is complete, the trained model can be used for *inferencing*. The model is essentially a software program that encapsulates the function produced by the training process. You can input a set of feature values, and receive as an output a prediction of the corresponding label. Because the output from the model is a prediction that was calculated by the function, and not an observed value, you'll often see the output from the function shown as $\hat{y}$ (which is rather delightfully verbalized as "y-hat").

Next unit: Types of machine learning

Continue >