From the table in Figure 14-3, the **target variable denotes a class membership of heart disease or no heart disease**; hence, the target is categorical and can be termed as a **classification problem**.

How Do We Know that Learning Has Occurred?

This question is vital to determine if the learning algorithm can learn a useful pattern between the input features and the targets. Let's create a scenario that will give us better insights into appraising the question of determining when learning has occurred.

Assume a teacher takes a physics class for 3 months, and at the end of each session, the teacher administers a test to ascertain if the student has learned anything.

Let's consider two different scenarios the teacher might use in evaluating the students:

1. The teacher evaluates the student with the exact word-for-word questions that were used as sample problems while teaching.
2. The teacher evaluates the student with an entirely different but similar set of sample problems that are based on the principles taught in class.

In which of these subplots can the teacher ascertain that the student has learned? To figure this out, we must consider the two norms of learning:

1. Memorization: In the first subplot, it will be incorrect for the teacher to form a basis for learning because the student has seen and most likely memorized the examples during the class sessions. Memorization is when the exact snapshot of a sample is stored for future recollection. Therefore, it is inaccurate to use samples used in training to carry out learning evaluation. In machine learning, this is known as *data snooping*.
2. Generalization: In the second subplot, the teacher can be confident that the assessment serves as an accurate test to evaluate if the student has learned from the session. The ability to use the principles learned to solve previously unseen samples is known as *generalization*.

Hence, we can conclude that learning is the ability to generalize to previously unseen samples.

Training, Test, and Validation Datasets

The goal of supervised machine learning is to be able to predict or classify the targets on unseen examples correctly. We can misjudge the performance of our learning models if we evaluate the model performance with the same samples used to train the model as explained previously.

To properly evaluate the performance of a learning algorithm, we need to set aside some data for testing purposes. This held-out data is called a **test set**.

Another situation arises when we have trained the model on a dataset, and we now need to improve the performance of the model by adjusting some of the learning algorithm's parameters.

We cannot use the test set for model tuning because if we do that, the model's parameters are trained with the test dataset rendering it unusable as an unseen held-out sample for model evaluation. Hence, it is typical to divide the entire dataset into

- The training set (to train the model)
- The validation set (to tune the model)
- The test set (to evaluate the effectiveness of the model)

A common and straightforward strategy is to split 60% of the dataset for training, 20% for validation, and the final 20% for testing. This strategy is popularly known as the 60/20/20 rule. We will discuss more sophisticated methods for resampling (i.e., using subsets of available data) for machine learning later in this chapter. See Figure 14-4.