

Figure 15-1. *Batch learning*

In a situation where there is a need to train the model with data that is generated continuously from the source, batch learning becomes inappropriate to deal with that situation. In such a circumstance, we want to be able to update our learning model on the go, based on the new data samples that are available.

Online Learning

In online learning, data *streams* (either individually or in mini-batches) into the learning algorithm and updates the model. Online learning is ideal in situations where data is generated continuously in time, and we need to use real-time data samples to build a prediction model. A typical example of this case is in stock market prediction.

Online learning is illustrated in Figure 15-2.

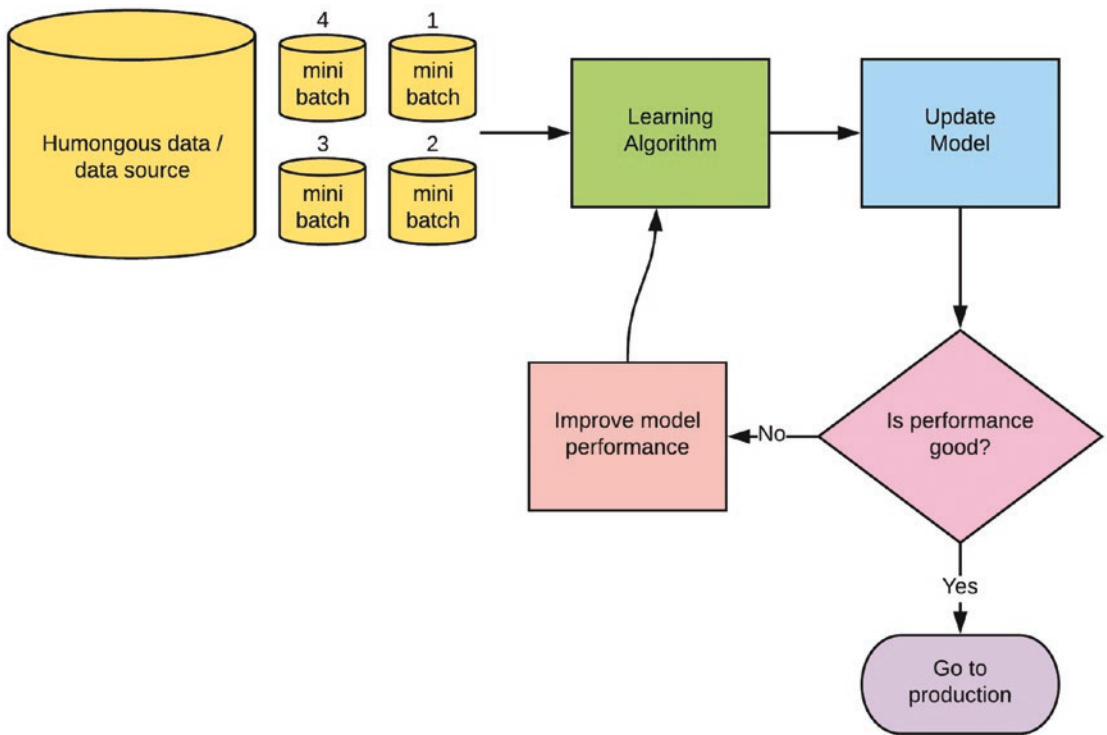


Figure 15-2. *Online learning*

This brief chapter explained the contrast between batch learning and online learning. In the next chapter, we will focus our attention on a vital optimization algorithm for machine learning, gradient descent.

CHAPTER 16

Optimization for Machine Learning: Gradient Descent

Gradient descent is an optimization algorithm that is used to minimize the cost function of a machine learning algorithm. Gradient descent is called an iterative optimization algorithm because, in a stepwise looping fashion, it tries to find an approximate solution by basing the next step off its present step until a terminating condition is reached that ends the loop.

Take the following convex function in Figure 16-1 as a visual of gradient descent finding the minimum point of a function space.

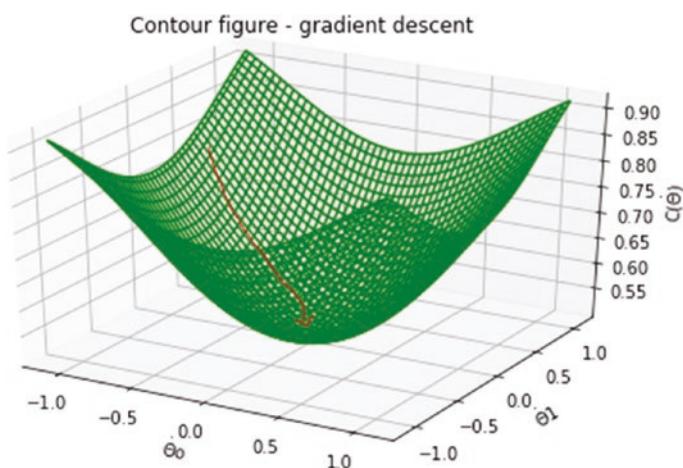


Figure 16-1. Contour figure – gradient descent