Machine Learning:

Understanding machine learning and the problems it can solve

Extract from Machine Learning Bookcamp by Alexey Grigorev, Manning Publications, 2020

Machine learning is part of applied mathematics and computer science. It uses tools from mathematical disciplines such as probability, statistics, and optimisation theory to extract patterns from data.

The main idea behind machine learning is learning from examples: we prepare a dataset with examples, and a machine learning system "learns" from this dataset. In other words, we give the system the input and the desired output, and the system tries to figure out how to do the conversion automatically, without asking a human.

We can collect a dataset with description of cars and their prices, for example. Then we provide a machine learning model with this dataset and "teach" it by showing cars and their prices. This process is called training or sometimes fitting.
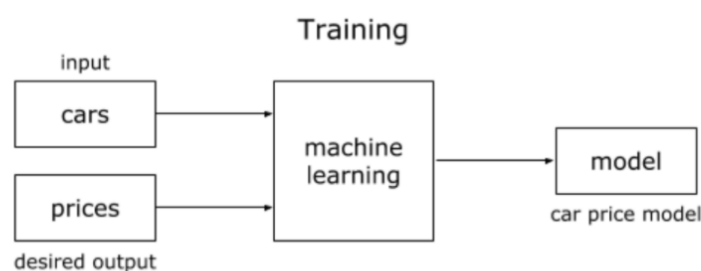

Figure 1.1 A machine learning algorithm takes in input data (descriptions of cars) and desired output (the cars' prices). Based on that data, it produces a model.

When training is done, we can use the model by asking it to predict car prices that we don't know yet.

Figure 1.3 In traditional software, patterns are discovered manually and then encoded with a programming language. A human does all the work.

As we see, all we need for machine learning is a dataset in which for each input item (a car), we have the desired output (the price).

This process is quite different from traditional software engineering. Without machine learning, analysts and developers look at the data they have and try to find patterns there manually. After that, they come up with some logic: a set of rules for converting the input data to the desired output. Then they explicitly encode these rules with a programming language such as Java or Python, and the result is called software. So, in contrast with machine learning, a human does all the difficult work in traditional programming.


Figure 1.3 In traditional software, patterns are discovered manually and then encoded with a programming language. A human does all the work.

We can summarize the difference between a traditional software system and a system based on machine learning with the images below.

In machine learning, we give the system the input and output data, and the result is a model (code) that can transform the input into the output. The difficult work is done by the machine; we need only supervise the training process to make sure that the model is good (A). By contrast, in traditional systems, we first find the patterns in the data ourselves and then write code that converts the data to the desired outcome, using the manually discovered patterns (B).

(A) In traditional software we discover patterns manually and encode them with a programming language.
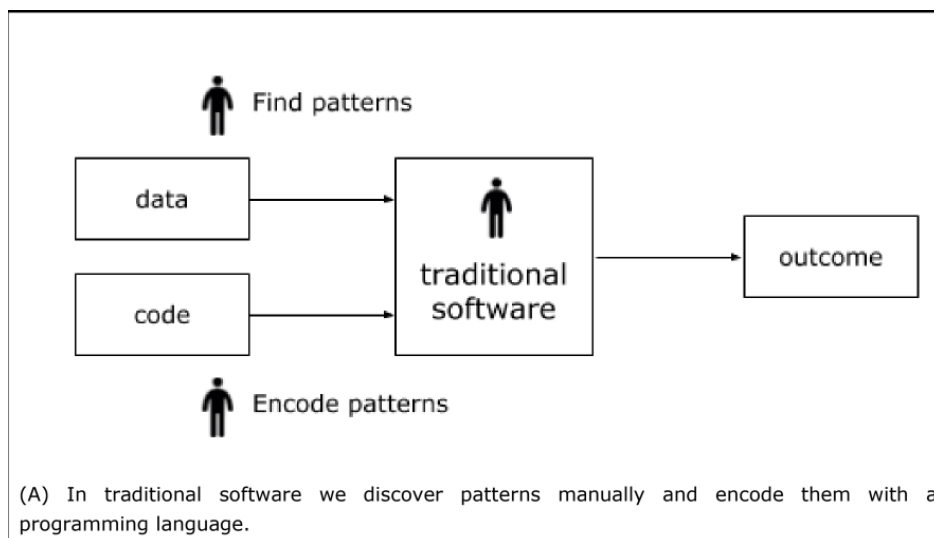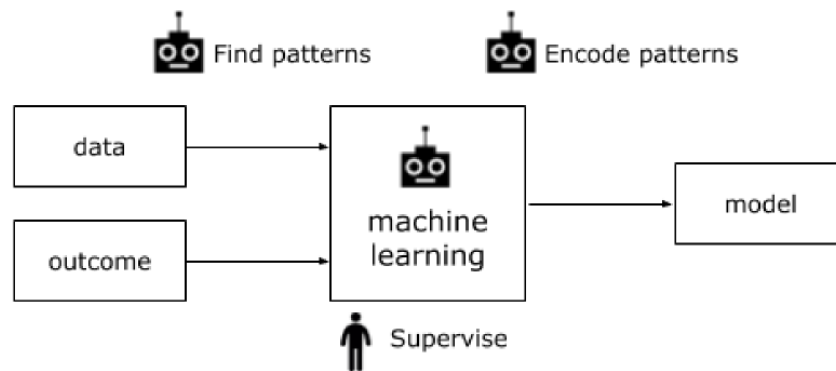
(B) A machine learning system discovers patterns automatically by learning from examples. After training it produces a model that "knows" these patterns, but we still need to supervise it to make sure the model is correct.

Now it's time to jump into the hands-on exercises. Happy Learning!