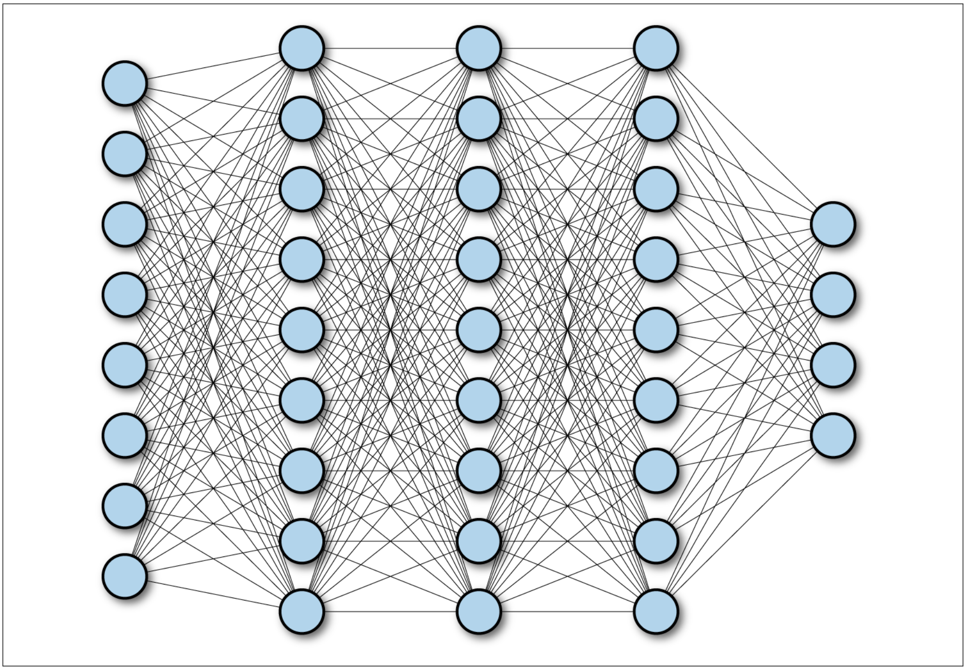

Figure 4-2. A multilayer deep fully connected network.

As a quick implementation note, note that the equation for a single neuron looks very similar to a dot-product of two vectors (recall the discussion of tensor basics). For a layer of neurons, it is often convenient for efficiency purposes to compute y as a matrix multiply:

$$y = \sigma(wx)$$

where sigma is a matrix in $\mathbb{R}^{n \times m}$ and the nonlinearity σ is applied componentwise.

“Neurons” in Fully Connected Networks

The nodes in fully connected networks are commonly referred to as “neurons.” Consequently, elsewhere in the literature, fully connected networks will commonly be referred to as “neural networks.” This nomenclature is largely a historical accident.

In the 1940s, Warren S. McCulloch and Walter Pitts published a first mathematical model of the brain that argued that neurons were capable of computing arbitrary functions on Boolean quantities. Successors to this work slightly refined this logical model by making mathematical “neurons” continuous functions that varied between zero and one. If the inputs of these functions grew large enough, the neuron “fired” (took on the value one), else was quiescent. With the addition of adjustable weights, this description matches the previous equations.

Is this how a real neuron behaves? Of course not! A real neuron (Figure 4-3) is an exceedingly complex engine, with over 100 trillion atoms, and tens of thousands of different signaling proteins capable of responding to varying signals. A microprocessor is a better analogy for a neuron than a one-line equation.

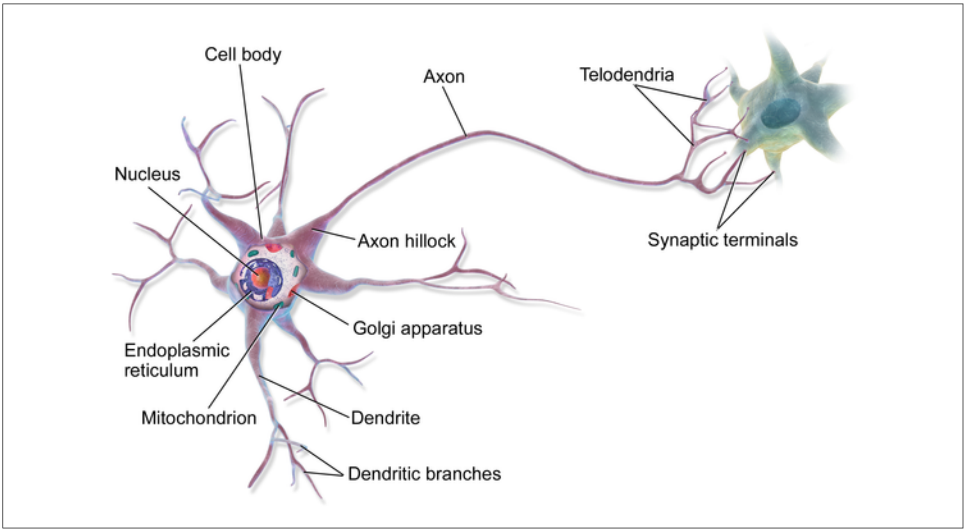

Figure 4-3. A more biologically accurate representation of a neuron.

In many ways, this disconnect between biological neurons and artificial neurons is quite unfortunate. Uninitiated experts read breathless press releases claiming artificial neural networks with billions of "neurons" have been created (while the brain has only 100 billion biological neurons) and reasonably come away believing scientists are close to creating human-level intelligences. Needless to say, state of the art in deep learning is decades (or centuries) away from such an achievement.

As you read further about deep learning, you may come across overhyped claims about artificial intelligence. Don't be afraid to call out these statements. Deep learning in its current form is a set of techniques for solving calculus problems on fast hardware. It is not a precursor to *Terminator* (Figure 4-4).

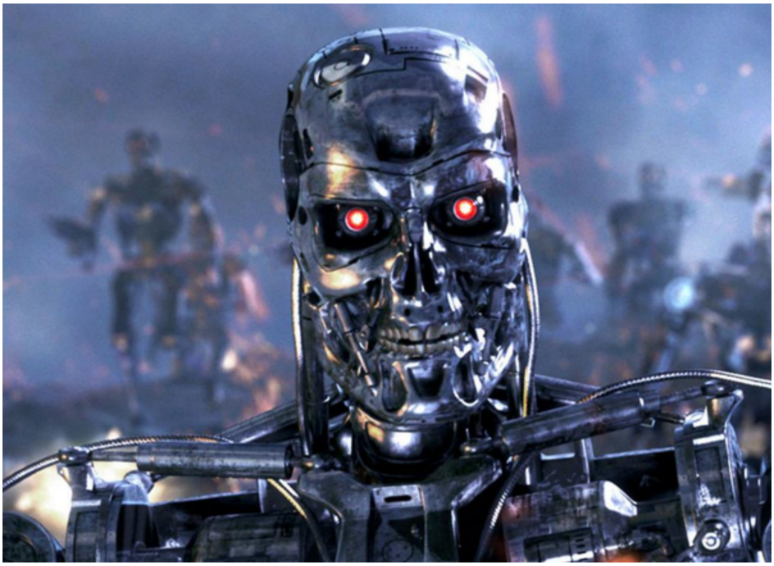

Figure 4-4. Unfortunately (or perhaps fortunately), this book won't teach you to build a Terminator!

AI Winters

Artificial intelligence has gone through multiple rounds of boom-and-bust development. This cyclical development is characteristic of the field. Each new advance in learning spawns a wave of optimism in which prophets claim that human-level (or superhuman) intelligences are incipient. After a few years, no such intelligences manifest, and disappointed funders pull out. The resulting period is called an AI winter.

There have been multiple AI winters so far. As a thought exercise, we encourage you to consider when the next AI winter will happen. The current wave of deep learning progress has solved many more practical problems than any previous wave of advances. Is it possible AI has finally taken off and exited the boom-and-bust cycle or do you think we're in for the "Great Depression" of AI soon?

Learning Fully Connected Networks with Backpropagation

The first version of a fully connected neural network was the Perceptron, (Figure 4-5), created by Frank Rosenblatt in the 1950s. These perceptrons are identical to the "neurons" we introduced in the previous equations.