CHAPTER 4

Fully Connected Deep Networks

This chapter will introduce you to fully connected deep networks. Fully connected networks are the workhorses of deep learning, used for thousands of applications. The major advantage of fully connected networks is that they are "structure agnostic." That is, no special assumptions need to be made about the input (for example, that the input consists of images or videos). We will make use of this generality to use fully connected deep networks to address a problem in chemical modeling later in this chapter.

We delve briefly into the mathematical theory underpinning fully connected networks. In particular, we explore the concept that fully connected architectures are "universal approximators" capable of learning any function. This concept provides an explanation of the generality of fully connected architectures, but comes with many caveats that we discuss at some depth.

While being structure agnostic makes fully connected networks very broadly applicable, such networks do tend to have weaker performance than special-purpose networks tuned to the structure of a problem space. We will discuss some of the limitations of fully connected architectures later in this chapter.

What Is a Fully Connected Deep Network?

A fully connected neural network consists of a series of fully connected layers. A fully connected layer is a function from $\mathbb{R}^m$ to $\mathbb{R}^n$. Each output dimension depends on each input dimension. Pictorially, a fully connected layer is represented as follows in Figure 4-1.