Geoff Hinton of the University of Toronto, a pioneer and giant in the field, was able to devise a method for training deep nets. His work led to the creation of the Restricted Boltzmann Machine, or RBM. Structurally, an RBM is a shallow neural net with just two layers – the visible layer and the hidden layer. In this net, each node connects to every node in the adjacent layer. The “restriction” refers to the fact that no two nodes from the same layer share a connection. The goal of an RBM is to recreate the inputs as accurately as possible. During a forward pass, the inputs are modified by weights and biases and are used to activate the hidden layer. In the next pass, the activations from the hidden layer are modified by weights and biases and sent back to the input layer for activation. At the input layer, the modified activations are viewed as an input reconstruction and compared to the original input. A measure called KL Divergence is used to analyze the accuracy of the net. The training process involves continuously tweaking the weights and biases during both passes until the input is as close as possible to the reconstruction. If you’ve ever worked with an RBM in one of your own projects, please comment and tell me about your experiences. Because RBMs try to reconstruct the input, the data does not have to be labelled. This is important for many real-world applications because most data sets – photos, videos, and sensor signals for example – are unlabelled. By reconstructing the input, the RBM must also decipher the building blocks and patterns that are inherent in the data. Hence the RBM belongs to a family of feature extractors known as auto-encoders.