Spatial dimension:

We can give every point in our three-dimensional space an *(x, y, z)* coordinate. Since these points are represented as three real numbers, we say that it belongs to the set $\mathbb{R}^3$. These coordinates are usually not very meaningful, but they encode useful notions of distance and magnitude.

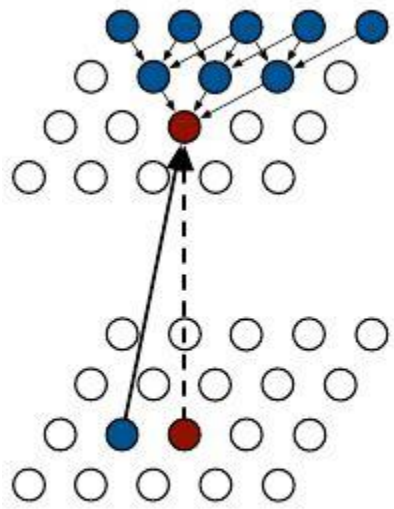
The important thing above is that when we had *three* real-valued numbers, we could represent it in our three-dimensional space. Let's now think about an image. An image is simply a big collection of pixels, with each pixel representing an intensity in some range.

LSTM:

LSTM recurrent unit tries to "remember" all the past knowledge that the network is seen so far and to "forget" irrelevant data.

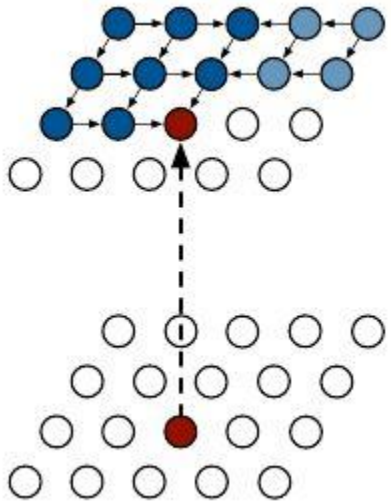
4 gates control how much information have to memorize.

Row LSTM:



Row LSTM

Bi-Diagonal LSTM:



The CIFAR-10 dataset

The CIFAR-10 dataset consists of 60000 32x32 color images in 10 classes, with 6000 images per class. There are 50000 training images and 10000 test images