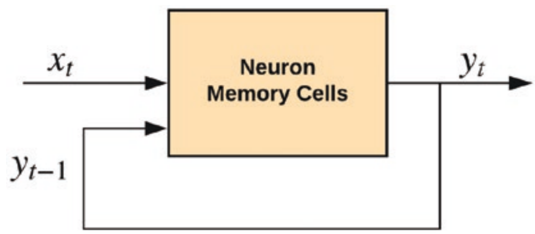

Figure 36-1. *A recurrent neuron*

In Figure 36-1, the recurrent neuron stands in contrast with neurons of the MLP and CNN architectures because instead of transferring a hierarchy of information across the network from one neuron to the other, data is looped back into the same neuron at every new time instant. A time instant can also mean a new sequence.

Hence, the recurrent neuron has two input weights, W_{x_t} and $W_{y_{t-1}}$, for the input at time x_t and for the input at time instant y_{t-1}. See Figure 36-2.

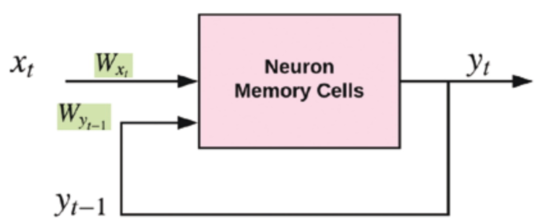

Figure 36-2. *Recurrent neuron with input weights*

Similar to other neurons, the recurrent neuron also injects non-linearity into the network by passing its weighted sums or affine transformations through a non-linear activation function.

Unfolding the Recurrent Computational Graph

A recurrent neural network is formalized as an unfolded computational graph. An unfolded computational graph shows the flow of information through the recurrent layer at every time instant in the sequence. Suppose we have a sequence of five time steps, we will unfold the recurrent neuron five times across the number of instants. The number of sequences constitutes the layers of the recurrent neural network architecture. See Figure 36-3.

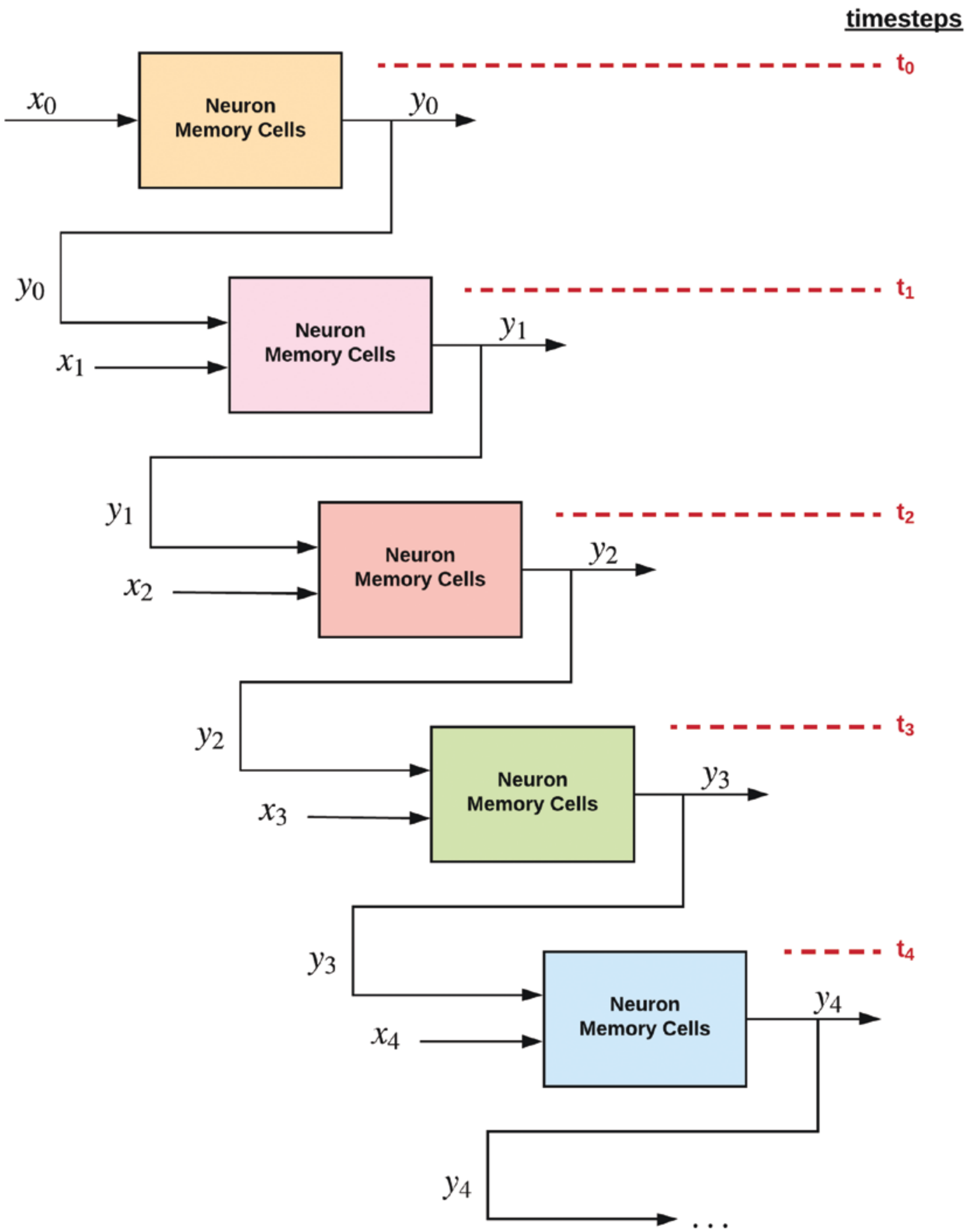

***Figure 36-3.** Unfolding the recurrent neuron into a recurrent neural network*

From the unrolled graph of the recurrent neural network, we can observe how the input into the recurrent layer includes the output of the previous time step $t - 1$ in addition to the current input at time step t. This architecture of the recurrent neuron is central to how the recurrent neural network learns from past events or past sequences.

Up until now, we have seen that the recurrent neuron captures information from the past by storing memory or state in its memory cell. The recurrent neuron can have a much more complicated memory cell (such as the GRU or LSTM cell) than the basic RNN cell as illustrated in the images so far, where the output at time instant $t - 1$ holds the memory.

Basic Recurrent Neural Network

Earlier on, we mentioned that when a recurrent network is unfolded, we can see how information flows from one recurrent layer to the other. Further, we noted that the sequence length of the dataset determines the number of recurrent layers. Let's briefly illustrate this point in Figure 36-4. Suppose we have a time series dataset of ten layers, for each row sequence in the dataset, we will have ten layers in the recurrent network system.

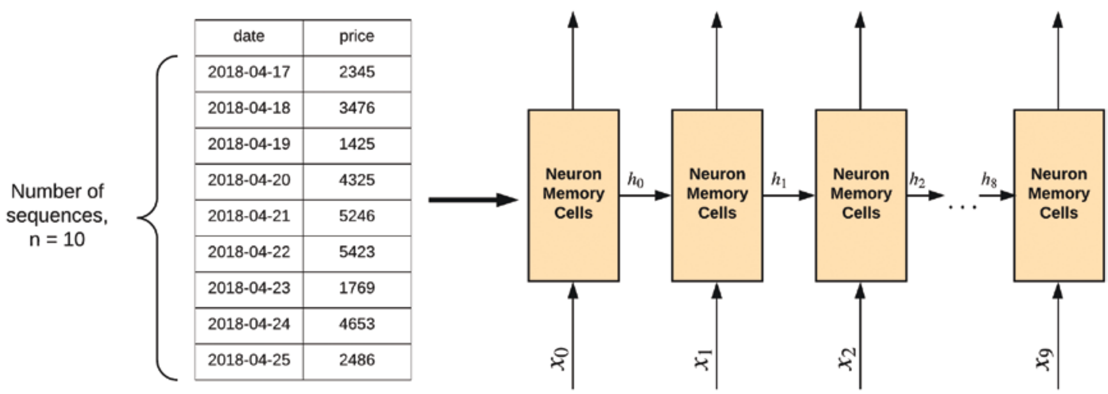

Figure 36-4. *Dataset to layers*

At this point, we must firmly draw attention to the fact that the recurrent layer does not comprise of just one neuron cell, but it is instead a set of neurons or neuron cells as shown in Figure 36-5. The choice of the number of neurons in a recurrent layer is a design decision when composing the network architecture.