Gated Recurrent Units (GRU)

The complexity, both conceptual and computational, for LSTM cells has motivated a number of researchers to attempt to simplify the LSTM equations while retaining the performance gains and modeling capabilities of the original equations.

There are a number of contenders for LSTM replacement, but one of the frontrunners is the gated recurrent unit (GRU), shown in Figure 7-5. The GRU removes one of the subcomponents of the LSTM but empirically seems to achieve performance similar to that of the LSTM. The GRU might be a suitable replacement for LSTM cells on sequence modeling projects.

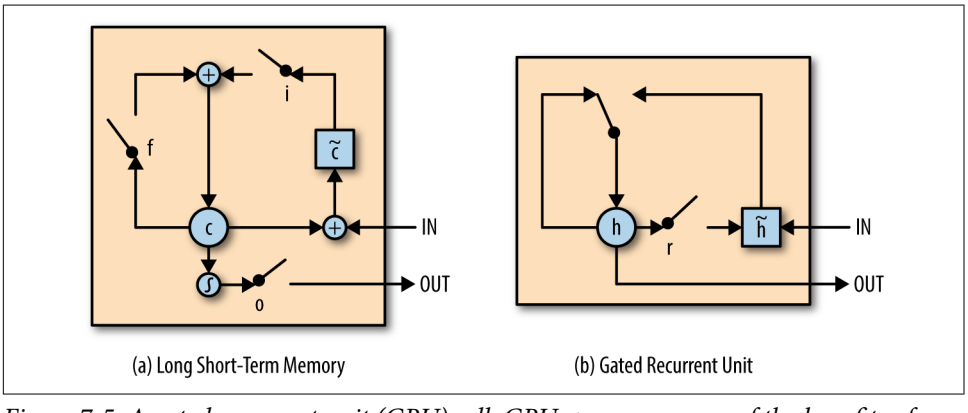

Figure 7-5. A gated recurrent unit (GRU) cell. GRUs preserve many of the benefits of LSTMs at lower computational cost.

Applications of Recurrent Models

While recurrent neural networks are useful tools for modeling time-series datasets, there are a host of other applications of recurrent networks. These include applications such as natural language modeling, machine translation, chemical retrosynthesis, and arbitrary computation with Neural Turing machines. In this section, we provide a brief tour of some of these exciting applications.

Sampling from Recurrent Networks

So far, we've taught you how recurrent networks can learn to model the time evolution of sequences of data. It stands to reason that if you understand the evolution rule for a set of sequences, you ought to be able to sample new sequences from the distribution of training sequences. And indeed, it turns out that that good sequences can be sampled from trained models. The most useful application thus far is in language modeling. Being able to generate realistic sentences is a very useful tool that underpins systems such as autocomplete and chatbots.