**RNN Notation**

: the *t*-th element of input sequence in the *i*-th training example

: the input sequence length of the *i*-th training example

: the *t*-th element of output sequence in the *i*-th training example

: the output sequence length of the *i*-th training example

**RNN Forward Propagation**

: the weight parameter from *x<t>* to the hidden layer, shared for every time step

: the weight parameter of the horizontal connection, shared for every time step

: the weight parameter from the hidden layer to *y<t>*, shared for every time step

Hidden Layer Activation:

Simplified Hidden Layer Activation:

Output Layer Activation:

**RNN Backpropagation through Time**

Loss Function:

**Different Types of RNN**

* Many-to-many architecture
  + Name entity recognition
  + Machine translation (*TX* != *TY*)
* Many-to-one architecture
  + Sentiment classification
* One-to-one architecture
* One-to-many architecture
  + music generation

**Gated Recurrent Unit (GRU)**

Candidate Cell:

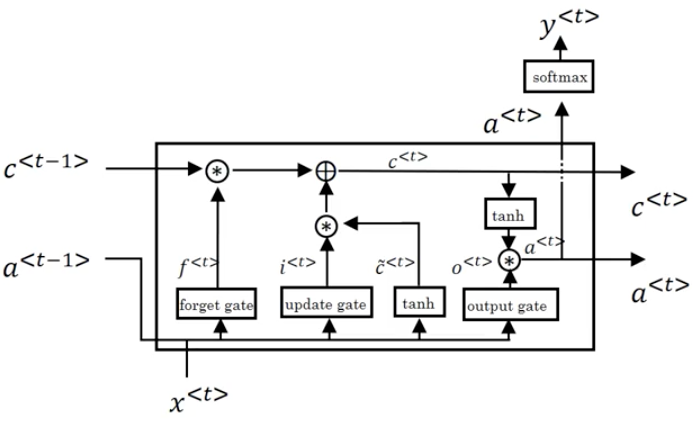
Update Gate:

Relevance Gate

Memory Cell:

Unit Activation:

**Long Short-Term Memory (LSTM)**



Candidate Cell:

Update Gate:

Forget Gate:

Output Gate:

Memory Cell:

Unit Activation:

**Deep RNN**

: the activation value of the *l*-th hidden layer for the *t*-th element

Deep RNN Activation: