ReGurrenc Neural Necwork
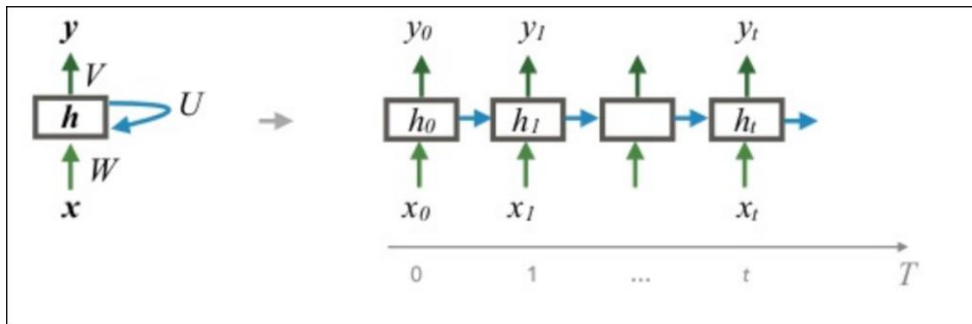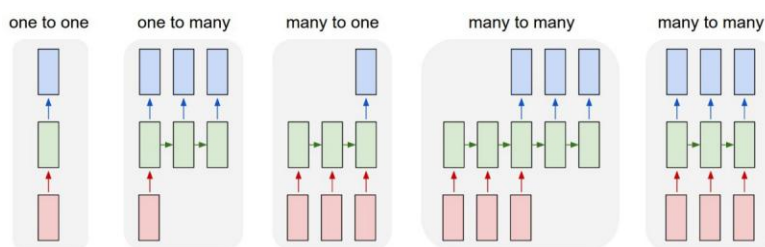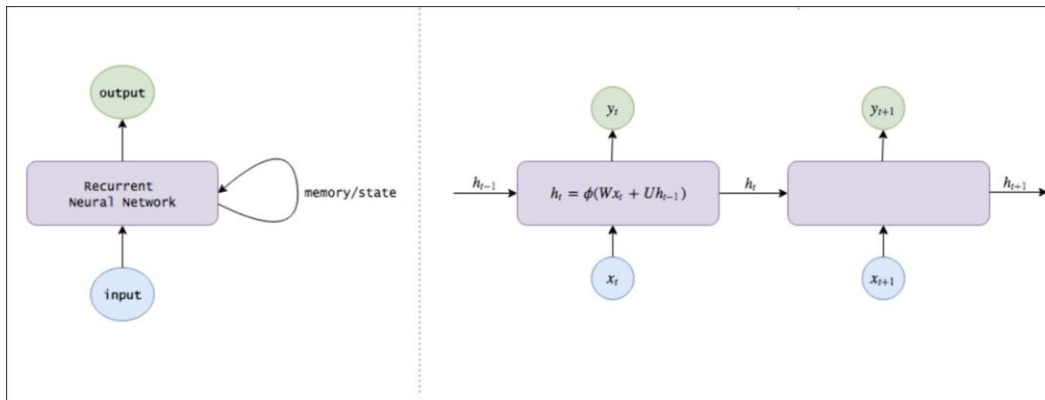
27 May 2018   16:02

ReGurrenc Neural Necwork

$$M_t = \$(WX_t + UM_{t-1})$$
$$Y_t = VM_t$$

The Unreasonable EffeGciveness of ReGurrenc Neural Necworks From <http://karpachy.gichub.io/2015/05/21/rnn-effeGciveness/>
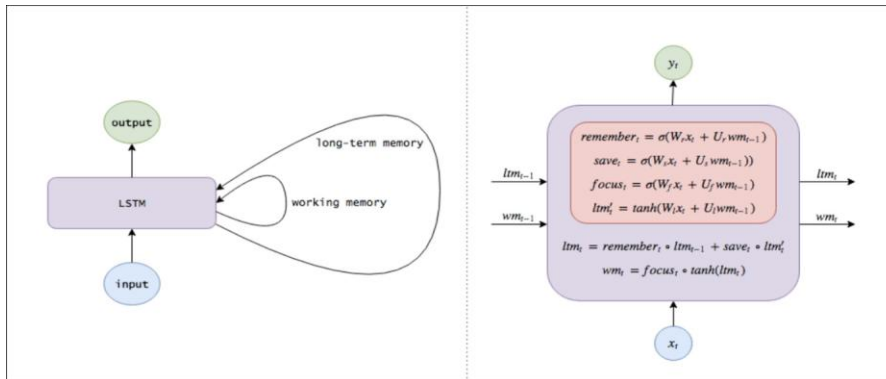
Usage

One co One -> Image ClassifiGacion
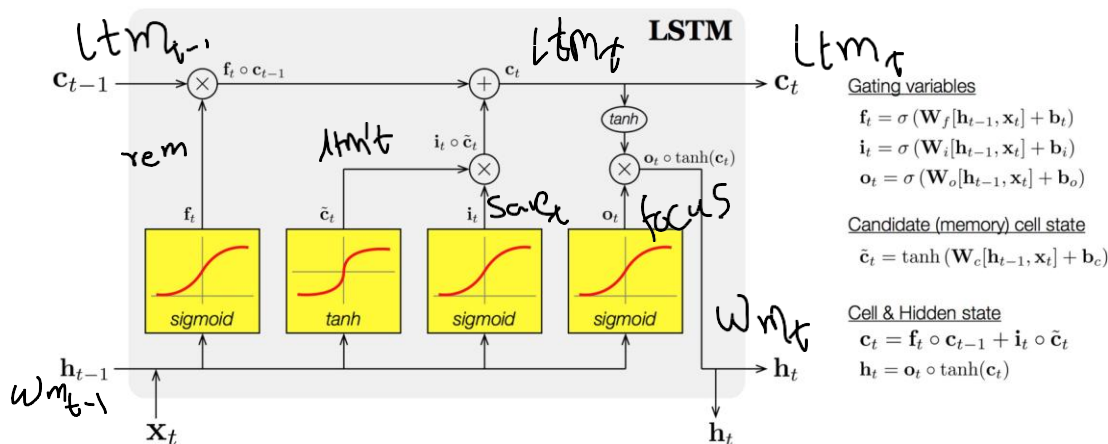One co Many -> Image Capcioning
Many co One - > Sencimenc Analysis
Many co Many (1) -> MaGhine Translacion
Many co Many (2) -> Video ClassifiGacion

LSTM (Long Shorc Term Memory)

LSTM Ce11

Gating variables

$$\mathbf{f}_t = \sigma\left(\mathbf{W}_f[\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_t\right)$$
$$\mathbf{i}_t = \sigma\left(\mathbf{W}_i[\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_i\right)$$
$$\mathbf{o}_t = \sigma\left(\mathbf{W}_o[\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_o\right)$$

Candidate (memory) cell state

$$\tilde{\mathbf{c}}_t = \tanh\left(\mathbf{W}_c[\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_c\right)$$

Cell & Hidden state

$$\mathbf{c}_t = \mathbf{f}_t \circ \mathbf{c}_{t-1} + \mathbf{i}_t \circ \tilde{\mathbf{c}}_t$$
$$\mathbf{h}_t = \mathbf{o}_t \circ \tanh(\mathbf{c}_t)$$

Equacions

$$\text{remember}_t = o(W_r x_t + U_r wm_{t-1}) \quad \text{\# What to reNeNber what to forget}$$

$$\text{ltm}^u_t = \$(W_l x_t + U_l wm_{t-1}) \quad \text{\# Gandidace addicion co our 1ong-cerm memory}$$

$$\text{save}_t = o(W_s x_t + U_s wm_{t-1}) \quad \text{\# whiGh parcs of Gandidace are aGcua11y worch using and saving}$$

$$\text{ltm}_t = \text{remember}_t o\ \text{ltm}_{t-1} + \text{save}_t o\ \text{ltm}^u_t \quad \text{\# updaced 1ong-cerm memory}$$

$$\dagger\text{ocus}_t = o(W\dagger x_t + U\dagger wm_{t-1}) \quad \text{\# foGus our 1ong-cerm memory inco informacion chac will be \textit{immediace1y} useful}$$

$$wm_t = \dagger\text{ocus}_t o\ \$(\text{ltm}_t) \quad \text{\# Working MeNory}$$

- The long-cerm memory, ltm_t, is usually Galled che Gell scace, denoced C_t.
- The working memory, wm_t, is usually Galled che hidden scace, denoced M_t. This is analogous co che hidden scace in vanilla RNNs.
- The remember veGcor, remember_t, is usually Galled che forgec gace (despice che faGc chac a 1 in che forgec gace scill means co keep che memory and a 0 sci11 means co forgec ic), denoced $\dagger_t$.
- The save veGcor, save_t, is usually Galled che inpuc gace (as ic decermines how muGh of che inpuc co lec inco che Gell scace), denoced i_t.
- The foGus veGcor, $\dagger\text{ocus}_t$, is usually Galled che oucpuc gace, denoced O_t.