Introduction to Artificial Intelligence
Recurrent Neural Networks

Janyl Jumadinova
December 5, 2016

Alex Graves, “Supervised Sequence Labelling with Recurrent Neural Networks”
http://colah.github.io/posts/2015-08-Understanding-LSTMs/
Recurrent Neural Networks

The output depends ONLY on the current input.

Output Layer

Hidden Layers

Input Layer
Recurrent Neural Networks

The hidden layers and the output depend from previous states of the hidden layers.
Recurrent Neural Networks

An unrolled recurrent neural network.
Recurrent Neural Networks

Each node represents a layer of neurons at a single timestep.
Recurrent Neural Networks

The input is a \textbf{SEQUENCE} $x(t)$ of any length.
Recurrent Neural Networks

Common visual sequences:

Still image + Spatial scan (zigzag, snake)

The input is a **SEQUENCE** $x(t)$ of any length.
Recurrent Neural Networks

Must learn temporally shared weights $w_2$; in addition to $w_1$ & $w_3$. 

$t-1$  $t$  $t+1$
Long Short-Term Memory (LSTM)

LSTMs are really mainstream now … just referenced in the @Apple #WWDC2016 keynote for iOS QuickType auto-completion
Long Short-Term Memory (LSTM)

Based on a standard RNN whose neuron activates with tanh

Cristopher Olah, “Understanding LSTM Networks” (2015)
Long Short-Term Memory (LSTM)

The repeating module in an LSTM contains four interacting layers.
$P(\text{next words} \mid \text{a history of previous words})$

https://www.tensorflow.org/versions/r0.12/tutorials/recurrent/

python ptb_word_lm.py --data_path=data/ --model small