Introduction to Deep Learning (CS474)

Lecture 25

Rupam

Outline

Module 3

- **Long Short Term Memory networks (LSTMs)**
- **Bidirectional RNN**

Rupam

Motivation

- As you read this essay, you understand each word based on your understanding of previous words.
- You don’t throw everything away and start thinking from scratch again.
- Your thoughts have ***persistence***.
- RNNs are networks with loops in them, allowing information to persist.
- We are going to start with a very special kind of recurrent neural network!

Slide credit: Christopher Olah

Rupam

Introduction to LSTMs

- Long Short Term Memory networks – usually just called "LSTMs" – are a special kind of RNN, capable of learning long-term dependencies.
- They work tremendously well on a large variety of problems, and are now widely used.
- Remembering information for long periods of time is practically their default behavior, not something they struggle to learn!
- All recurrent neural networks have the form of a chain of repeating modules of neural network.

Slide credit: Christopher Olah

Rupam

Introduction to LSTMs

$$\tilde{c}^{<t>} = \tanh(W_c[a^{<t-1>}, x^{<t>}] + b_c)$$

$$\Gamma_u = \sigma(W_u[\, a^{<t-1>}, x^{<t>}] + b_u)$$

$$\Gamma_f = \sigma(W_f[\, a^{<t-1>}, x^{<t>}] + b_f)$$

$$\Gamma_o = \sigma(W_o[\, a^{<t-1>}, x^{<t>}] + b_o)$$

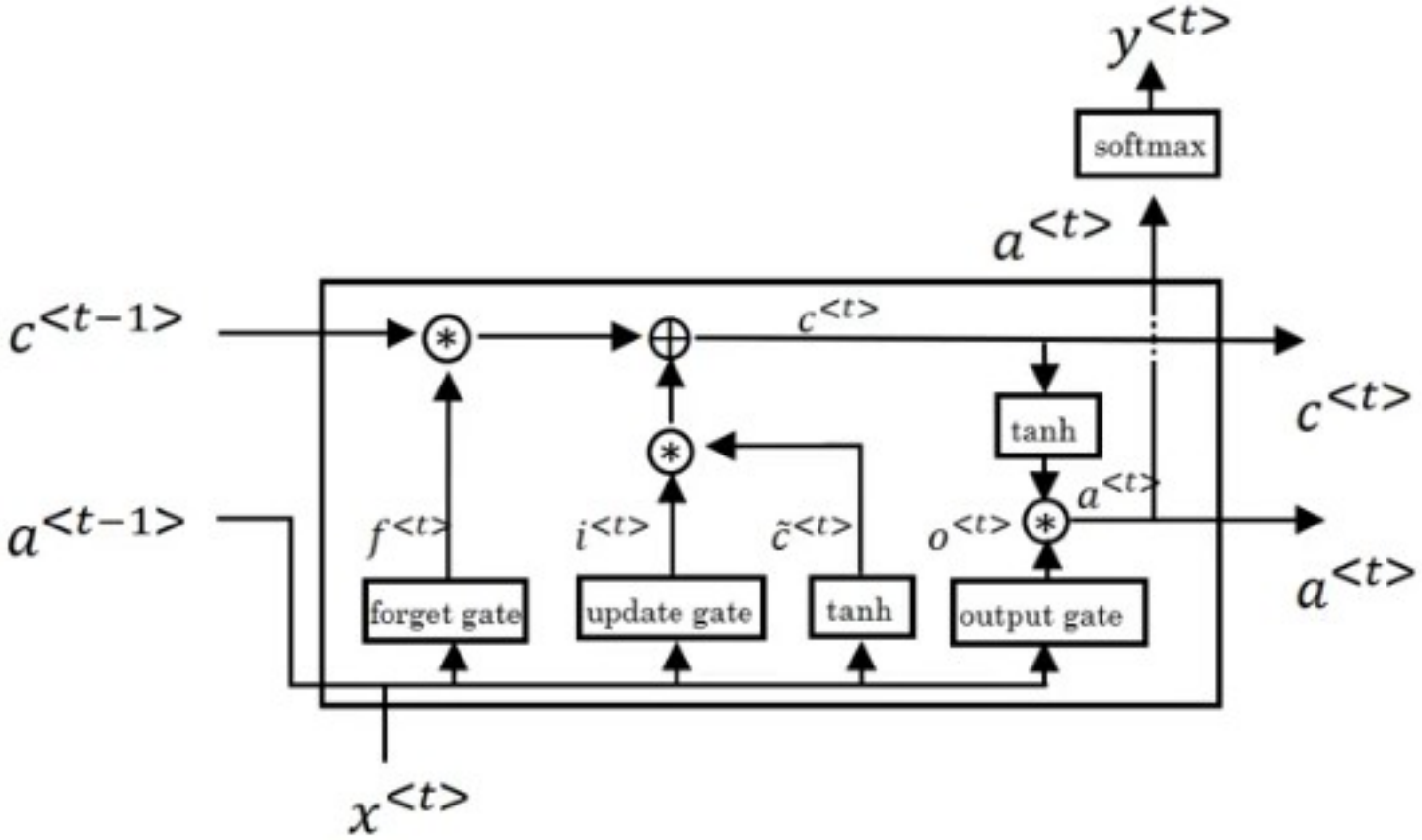
$$c^{<t>} = \Gamma_u * \tilde{c}^{<t>} + \Gamma_f * c^{<t-1>}$$

$$a^{<t>} = \Gamma_o * \tanh c^{<t>}$$

Slide credit: Prof. A. Ng

Rupam

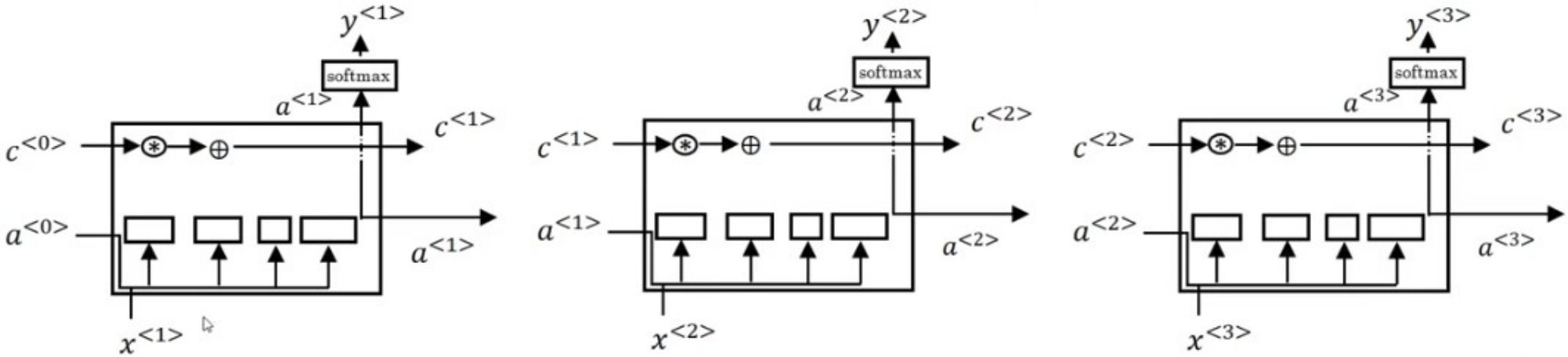
Introduction to LSTMs



Slide credit: Prof. A. Ng

Rupam

Introduction to LSTMs



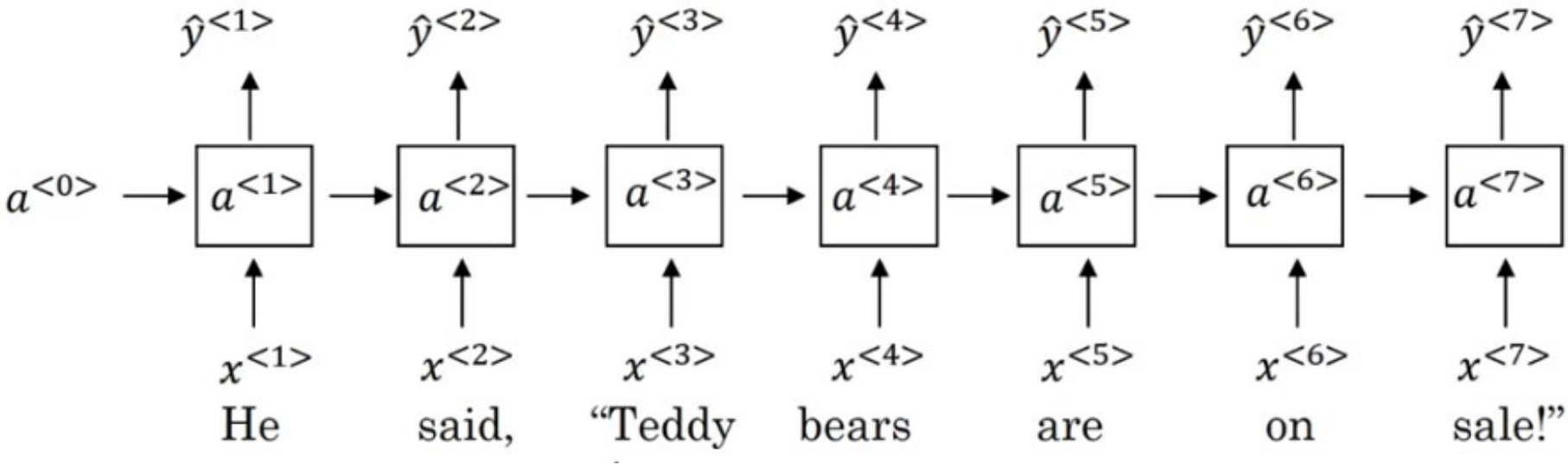
Slide credit: Prof. A. Ng

Rupam

Bidirectional RNN

He said, “Teddy bears are on sale!”

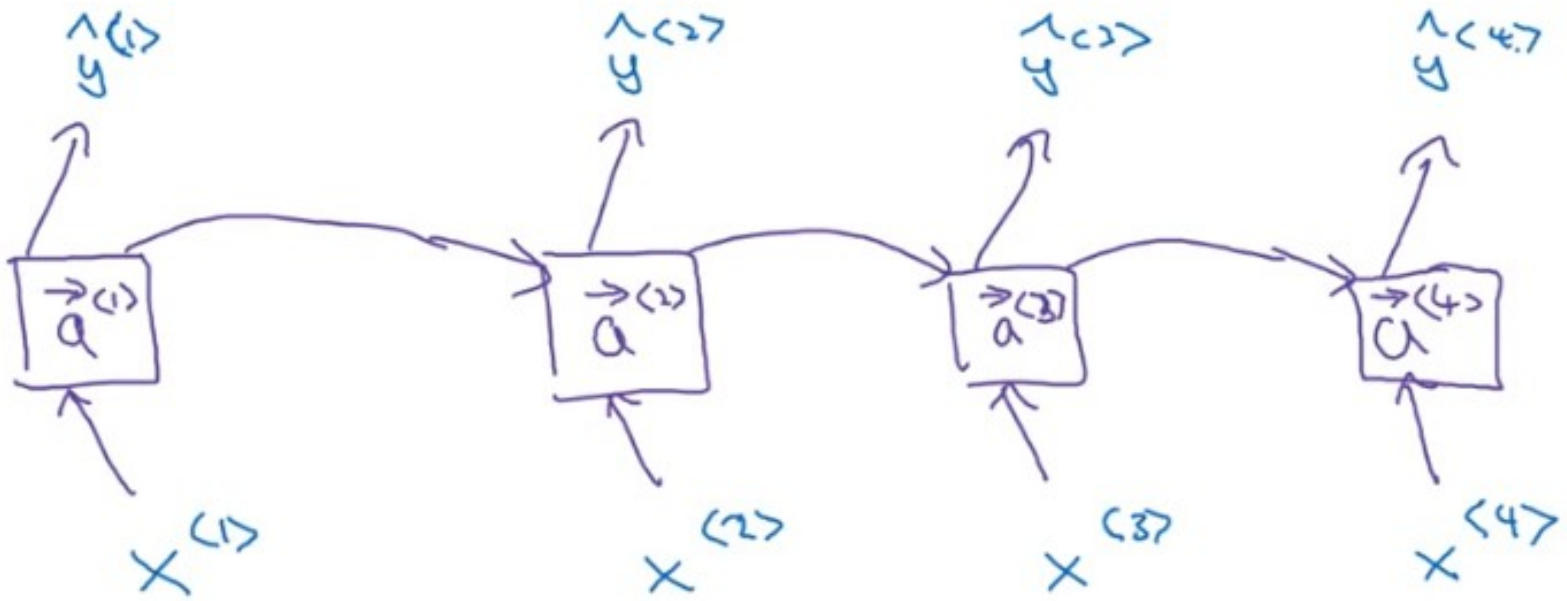
He said, “Teddy Roosevelt was a great President!”



Slide credit: Prof. A. Ng

Rupam

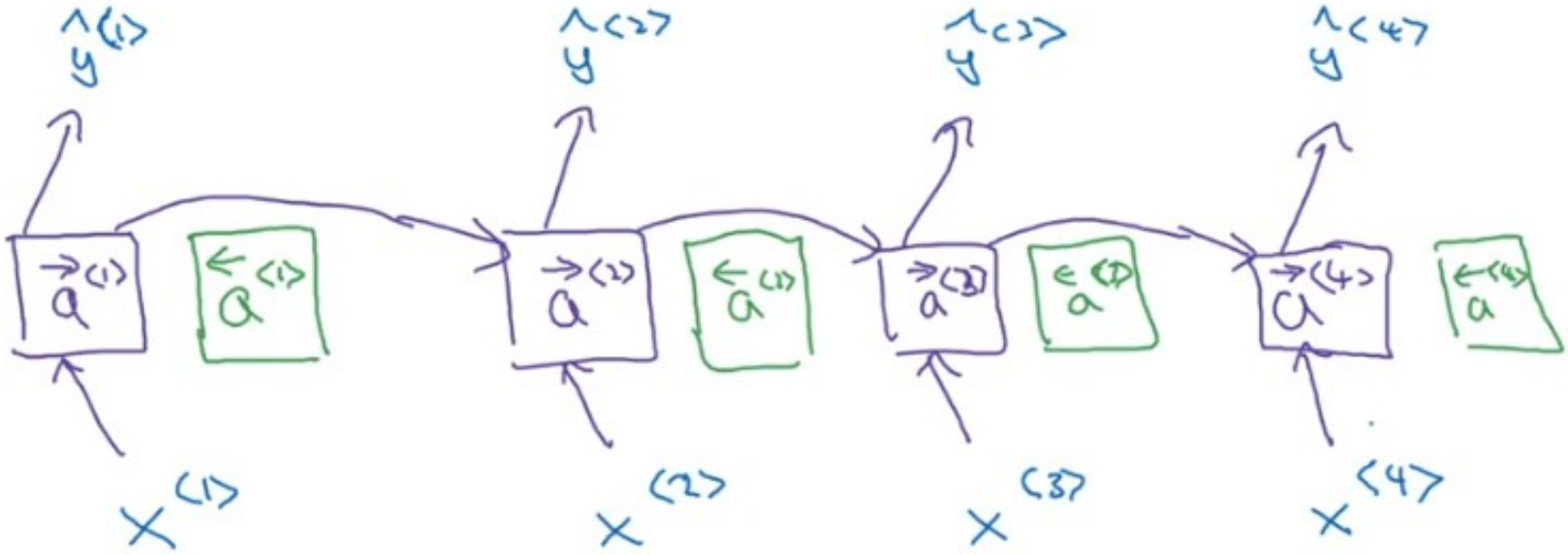
Bidirectional RNN



Slide credit: Prof. A. Ng

Rupam

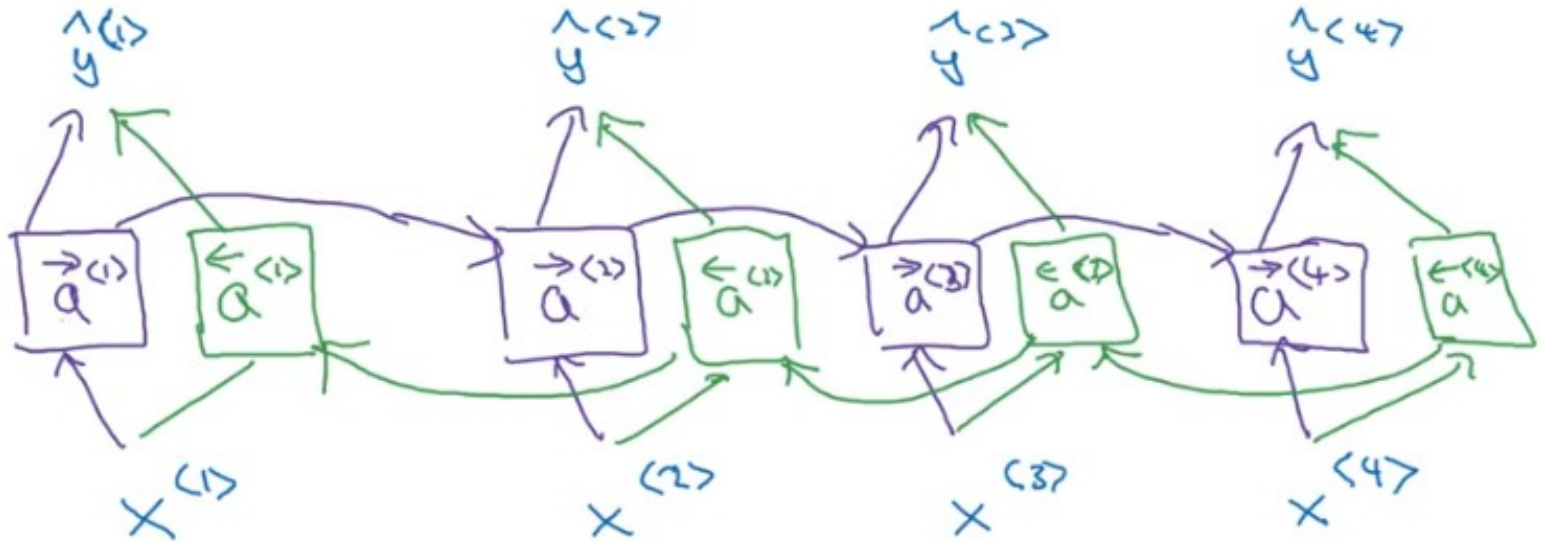
Bidirectional RNN



Slide credit: Prof. A. Ng

Rupam

Bidirectional RNN



Slide credit: Prof. A. Ng

References

- All the contents present in the slides are taken from various online resources. Due credit is given in the respective slides.These slides are used for *academic* purposes only.