

Recurrent Neural Network



- 
1. RNN Architecture
 2. LSTM



RNN Architecture



Sequential Data



Sentences

1. Fay and Jane play on the same day.

2. Gail made a fake cake.



3. It may rain cats and dogs today.

4. Jed cut the cake and ate the cake.

5. Jay and Ray hate the name game.

6. Dave made a cake for Jane's sake.

7. Kate's mate hates the rain.



8. Dave can rake hay and play in the rain.

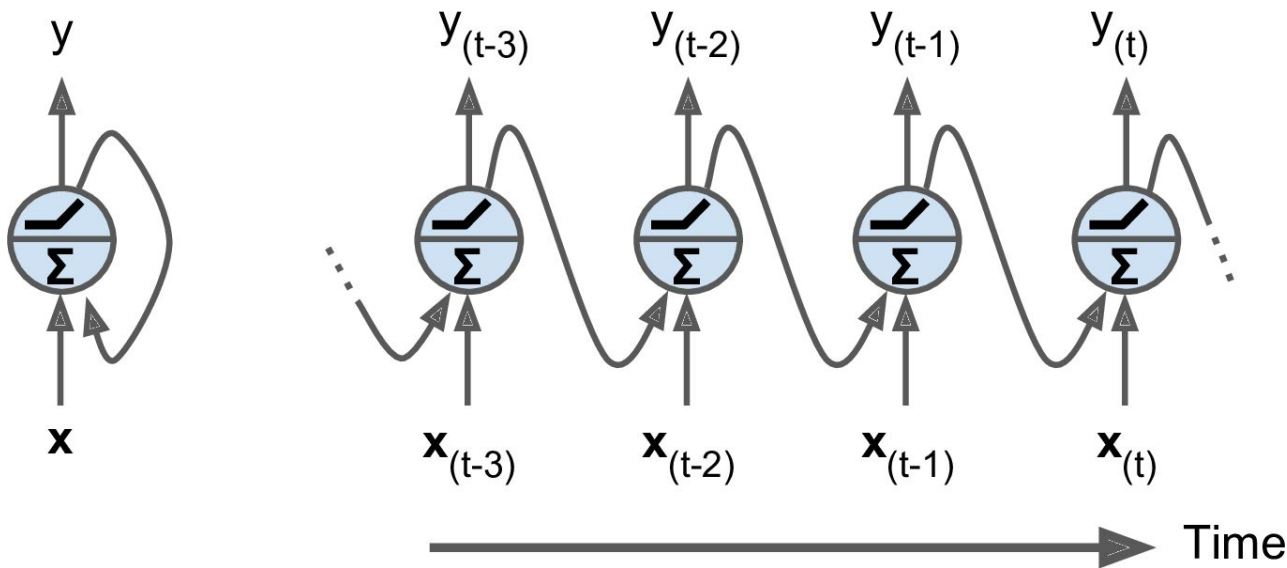
9. May Jake play the maze game?

10. Fay ate cake at the bay with Gabe.

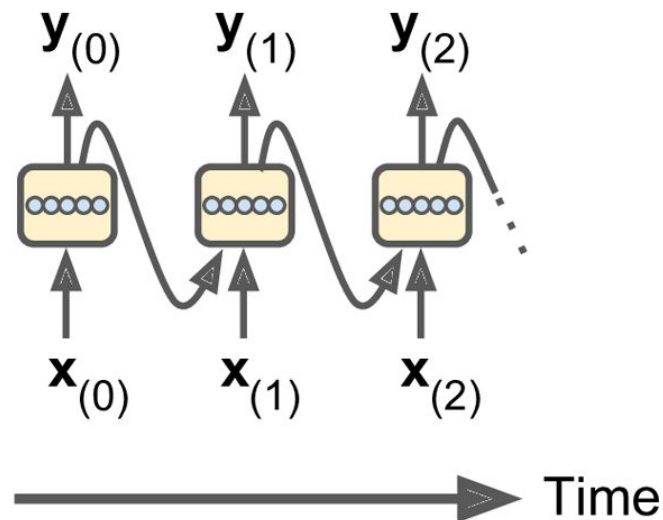
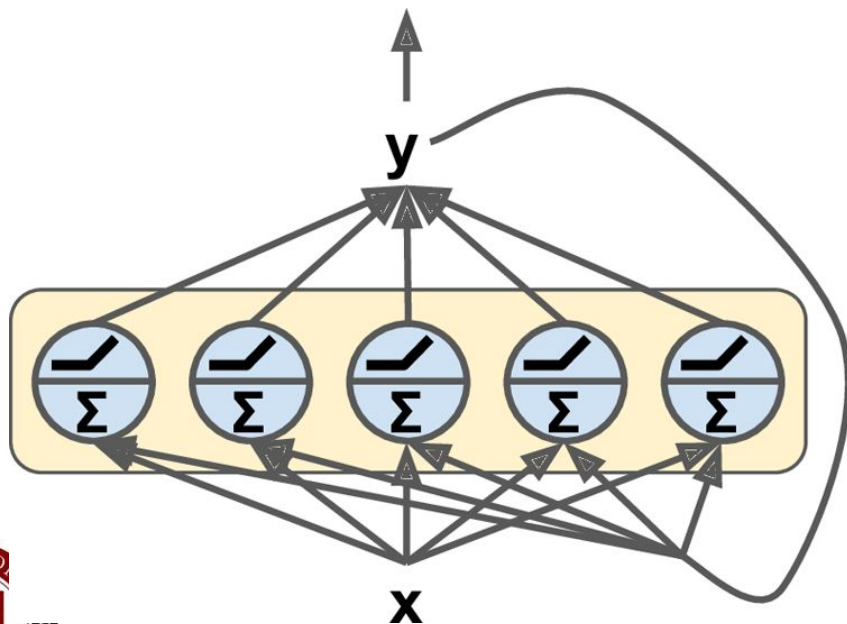


Recurrent Neuron

- A recurrent neuron takes in some input at time step t and the output from time step $t - 1$



Multiple Recurrent Neurons





LSTM



Training Over Many Time Steps

- When the length of the sequence is long
 - Training can be very slow
 - The “memory” of the initial inputs are slowly forgotten

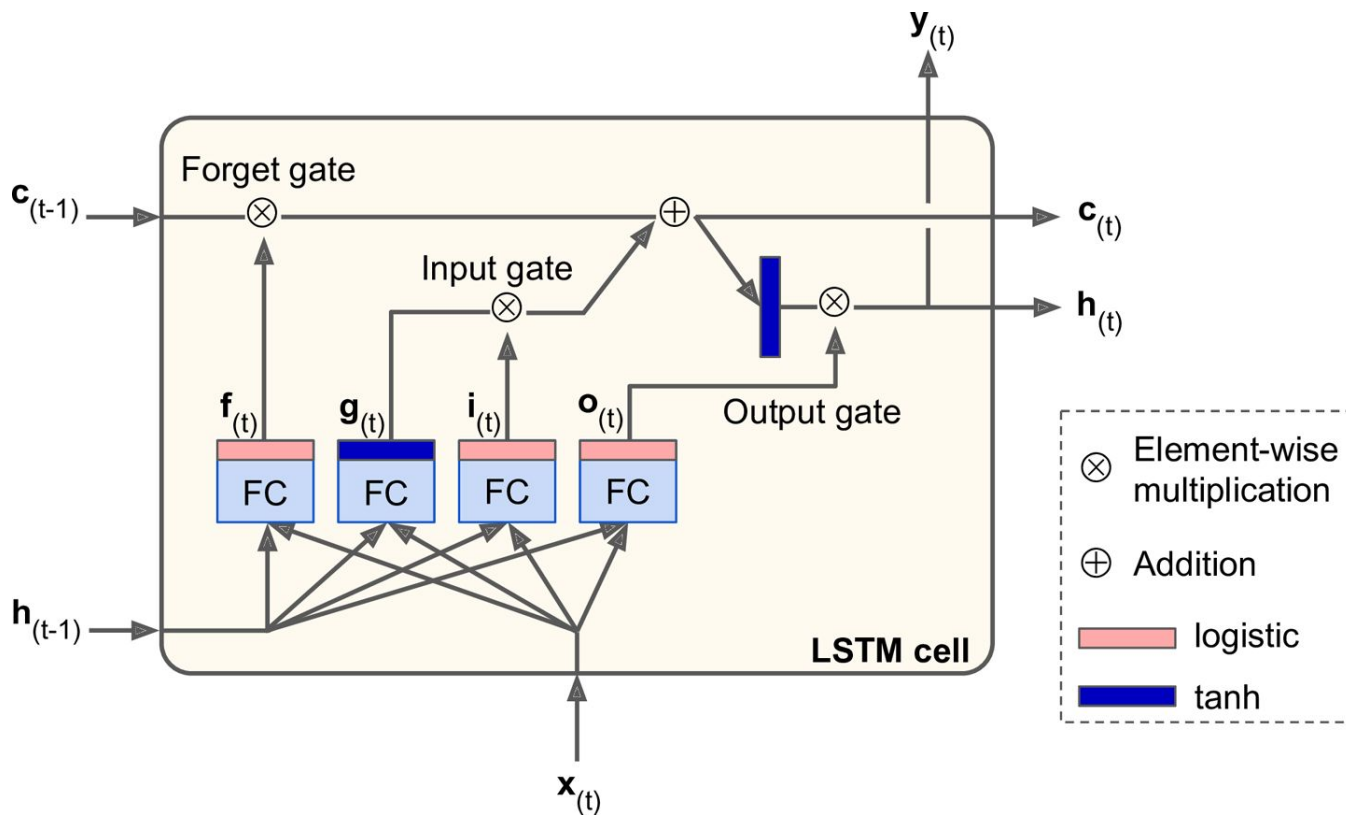


Long Short-Term Memory (LSTM)

- LSTMs learn to recognize important inputs, store it in the long-term state, learn to preserve it for as long as it is needed, and learns to extract it when it is needed
- Performs better, training converges faster, and it detects long-term dependencies



LSTM Architecture



Questions to Answer

1. What can you do to sequences to speed up training for RNNs?
2. What are some downsides of RNNs and how can you resolve them?
3. Are you likely to find exploding / vanishing gradient problem in RNNs / LSTMs?

