EdX and its Members use cookies and other tracking technologies for performance, analytics, and marketing purposes. By using this website, you accept this use. Learn more about these technologies in the Privacy Policy.





MITx: 6.86x

Machine Learning with Python-From Linear Models to Deep Learning

<u>Help</u>



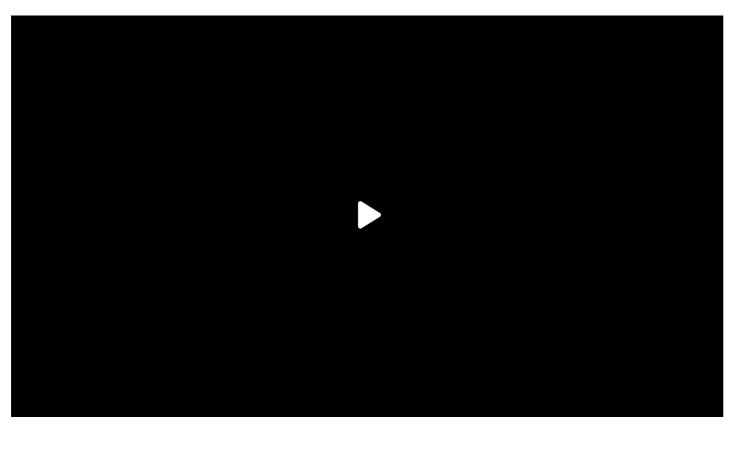
📆 <u>sandipan_dey</u>

Lecture 10. Recurrent Neural

Course > Unit 3 Neural networks (2.5 weeks) > Networks 1

> 5. Gating and LSTM

5. Gating and LSTM **Gating and LSTM**



If feed forward architectures suffer from any exploding

gradient issues, so do recurrent neural networks.

Specific architectures such as the LSTM recurrent neural

network maintain a better control over the information that's retained,

updated along the sequence, and they are therefore

easier to train to do what we want them to do.

9:21 / 9:21

▶ Speed **1.50**x



CC

66

End of transcript. Skip to the start.

Video Download video file **Transcripts** Download SubRip (,srt) file Download Text (.txt) file

Gating

1/1 point (graded)

Recall that the most simple, single-layered RNN can be written in equation as:

$$s_t = \tanh(W^{s,s} s_{t-1} + W^{s,x} x_t).$$

Recognize that, in the above formulation, s_t is always overwritten with the calculated result $anh(W^{s,s}s_{t-1}+W^{s,x}x_t)$.

Now, we introduce a gate vector g_t of the same dimension as s_t , which determines "how much information to overwrite in the next state." In equation, a single-layered gated RNN can be written as:

$$g_t \ = ext{sigmoid} \left(W^{g,s} s_{t-1} + W^{g,x} x_t
ight) s_t \ = (1-g_t) igodots s_{t-1} + g_t igodots anh \left(W^{s,s} s_{t-1} + W^{s,x} x_t
ight).$$

where the sign \bigcirc denotes element-wise multiplication. Now, which of the following is true about the gate g_t ? Choose all those apply.

- \square If the *i*th element of g_t is 1, the *i*th element of s_t and that of s_{t-1} are equal
- lacklose I If the ith element of g_t is 0, the ith element of s_t and that of s_{t-1} are equal \checkmark
- lacksquare If g_t is a vector whose elements are all 1, s_t and s_{t-1} are equal
- lacksquare If g_t is a vector whose elements are all 0, s_t and s_{t-1} are equal lacksquare



Solution:

Let the ith element of s_t , g_t , s_{t-1} be s_t^i , g_t^i , s_{t-1}^i .

If the ith element of g_t is 0, $(1-g_t^i)=1$, so

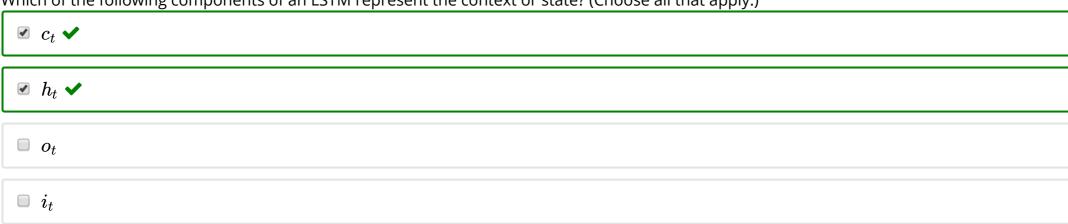
$$s_t^i = s_{t-1}^i.$$

Thus, if the ith element of g_t is 0, the ith element of s_t and that of s_{t-1} are equal. Also, if g_t is a vector whose elements are all 0, s_t and s_{t-1} are equal.

You have used 1 of 2 attempts Submit **1** Answers are displayed within the problem LSTM

1/1 point (graded)

Which of the following components of an LSTM represent the context or state? (Choose all that apply.)



Solution:

 c_t represents the memory cell, and h_t represents the visible state. Together they make up the context or state. The other two choices are the output and input gate, respectively. They simply accommodate new inputs and output predictions, and are not part of the context/state

Submit

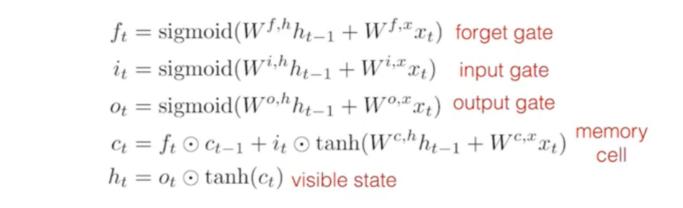
You have used 1 of 2 attempts

• Answers are displayed within the problem

LSTM Calculations

1/1 point (graded)

Let all the neural network's weight matrices, the hidden state, and the memory cell be a scalar 1. Let the new x-value be 5. Calculate the value of the new hidden state. Round sigmoid to 1 or 0, and round tanh to -1 or 1.



O -1



0

O 5

Solution:

The forget gate is equal to sigmoid(6), or 1. The same applies for the input and output gate. The memory cell is equal to 1+ tanh(1+6), which is 2. The new hidden state is therefore tanh(2), or 1.

Submit

You have used 1 of 2 attempts

1 Answers are displayed within the problem

Discussion

Show all posts ▼

Hide Discussion

by recent activity ▼

Topic: Unit 3 Neural networks (2.5 weeks):Lecture 10. Recurrent Neural Networks 1 / 5. Gating and LSTM

Add a Post

[STAFF] Please, fix first question equation

▲ Community TA

Learn About Verified Certificates

© All Rights Reserved