Long Short-term Memory TensorFlow (software library) +2

What is the meaning of “The number of units in the LSTM cell?

This question previously had question details. You can find them in the question comments.

Answer Request Follow 23 Comment 1 Share Downvote

Have this question too? Request Answers:

Request From Quora
We will distribute this question to writers, and notify you about new answers.

Alex James Jordie, studied Natural Language Processing
11 Answers in TensorFlow (software library)

Juan Leon, studied Biomedical Engineering at Universidad de los Andes, Colombia
3 Answers in Artificial Neural Networks

Julien Guillaumin, MSc Image Processing & Deep Learning, Télécom Bretagne (2018)
3 Answers in Machine Learning

View More or Search

4 Answers

Hieu Pham, Has done some machine learning
Answered Jul 28 2016 · Upvoted by Matthew Li, PhD Natural Language Processing & Machine Learning (2017)

LSTM is by definition, a function F: $\mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}^d \times \mathbb{R}^d$. It takes in $x, p_c, p_h \in \mathbb{R}^d$ and performs the following computations:

$$\begin{bmatrix} i \\ f \\ o \\ g \end{bmatrix} = \begin{bmatrix} \sigma \\ \sigma \\ \sigma \\ \tanh \end{bmatrix} W \cdot \begin{bmatrix} x \\ p_h \end{bmatrix}$$

$$c = i \otimes g + f \otimes p_c$$

$$h = o \otimes \tanh(c)$$

where $\otimes$ denotes the elementwise product of vectors. The LSTM returns $(c, h) = F(x, p_c, p_h)$.

The number d is loosely referred to as “the number of units” in the LSTM cell.

6.9k Views · 30 Upvotes · Answer requested by Pavel Brans

Add a comment... Recommended All

Mostafa Samir, Independently resareshing AI, while using TensorFlow
Answered Jul 24 2016

Upvote Downvote

Related Questions

What is in the cell state in an LSTM? What information does it contain?

What is the difference between the cell state and cell state in LSTM?

What is the meaning of the number of LSTM as 256? Is it the number of cells or the vector dimension from one LSTM?

What the difference between an LSTM memory cell and an LSTM layer?

How is the hidden state (h) different from the memory (c) in an LSTM cell?

To what extent does the dimension of LSTM hidden states affect performance?

Where exactly is information stored in an LSTM cell?

Is one LSTM block better then more LSTM blocks?

What does it mean for LSTM to be multilayer?

What is the meaning of 'projection layer' in lstm?

More Related Questions

Question Stats

23 Public Followers

12,888 Views

Last Asked Nov 10 2016

Edits

- The **previous time-step output**.

Tensorflow’s `num_units` is the size of the LSTM’s hidden state (which is also the size of the output if no projection is used). To make the name *num_**units*** more intuitive, you can think of it as the number of hidden *units* in the LSTM cell, or the number of memory *units* in the cell.

6.1k Views · 16 Upvotes


Add a comment... Recommended All


Dan Elton, studied Physics
Answered 11d ago

The previous answerer (Hieu Pham) is mostly (but not entirely) correct, but I felt his explanation was hard to follow.

It took me a little while to figure out that I was thinking of LSTMs wrong. If you’re like I was (or people who’ve answered this incorrectly on other forums), you might be visualizing “a LSTM” as something with a scalar (1D) hidden cell state (denoted c) and a scalar output (denoted h). So it takes vector input (x) and gives a scalar output. If you think of LSTMs this way, then it is tempting to think of the “number of units = d” as taking d serial LSTMs and running them parallel, with a total of d hidden states and d output states.

This is not a good way to think about it, though. The number of units is a parameter in the LSTM, referring to the dimensionality of the hidden state and dimensionality of the output state (they must be equal). a LSTM comprises an entire layer. There is crosstalk between the hidden states via the weight matrix, so its not correct to think of it as d serial LSTMs running in parallel. Some day I might make a diagram to visualize this, as most of the diagrams of LSTMs do not show this (they represent the hidden cell by a single line, so the vectoral nature is not shown).

If the input x is of size $n \times 1$, and we have d hidden states, then the size of the total weight matrix is $4\,d \times (n + d)$

(the total weight matrix includes parameters for all 4 operations and the sub matrices W and U)

Hieu Pham’s answer contains an error because the dimensionality of the input vector (n) does not need to be equal to the dimensionality of the outputs (d) . Furthermore, he seems to be assuming a scalar hidden state, while it actually needs to be dimension d.

So an LSTM is a function that maps $\mathbb{R}^n \times \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}^d \times \mathbb{R}^d$

484 Views · 6 Upvotes


Add a comment... Recommended All

architecture that was designed to model temporal sequences and their long-range dependencies more accurately than conventional RNNs. The LSTM does have the ability to remove or add information to the cell state, carefully regulated by structures called gates.

Gates are a way to optionally let information through. They are composed out of a sigmoid neural net layer and a pointwise multiplication operation. An LSTM has three of these gates, to protect and control the cell state. If I’m not wrong, the number of units in the LSTM cell should be referring to the number of gates i.e. 3. Please correct me incase I’m wrong.

4.2k Views

Add a comment...

Recommended All

Top Stories from Your Feed

Answer written · Secrets · Topic you might like · Jul 9

What is the dark secret in your life?

Anonymous
Answered Jul 9

This is regarding the most stressful moments of my life. I never thought I would share it with anyone. But I think this is the platform where I can vent it out and feel light. It happened long back ...

Read In Feed

Answer written · Experiences in Life · Topic you might like · Jul 21

What is the most badass thing your parent has ever done?

Elena Ledoux, Founder (2017-present)
Answered Jul 21

My dad’s photo should be in a dictionary under the definition of *badass*:

1. Our neighbors’ son - a promising young engineer - suddenly

Read In Feed

Answer written · Family Relationships a... · Topic you might like · Jul 28

What is the most messed up thing that your family member has ever done?

Anonymous
Updated 40m ago

Going anonymous for this one. My brother, let's call him Josh, was spoiled. Much more spoiled than I. He was the youngest and my mom always wanted a boy. Spoiled boy turned into spoiled adult and h...

Read In Feed