Course > Unit 3 Neural networks (2.5 weeks) > Homework 4 > 4. Word Embeddings

4. Word Embeddings

Extension Note: Because EdX site was down for a while yesterday, Homework 4 due date has been extended by 1 day to **April 1 23:59UTC** for those affected. **Future deadlines will remain unchanged.**

Word Embeddings

1/1 point (graded)

Training a neural network using back-propagation and SGD moves the network weights in a direction that minimizes the loss function. If the network contains a *bottleneck*, a layer in which many inputs are reduced to only a few outputs, training will adjust the weights to maximize the useful information contained in the layer's output. In this way, a sparse input representation can be embedded in a lower-dimensional space to become a dense, *distributed* representation. Embeddings often have interesting properties like transforming semantic or visual similarity into geometric proximity.

For example, imagine you have the words "cat", "lion", "car", "bridge". You could have a possible dense representation like: "cat" : [0.9, 0.2], "lion" : [0.9, 0.5], "car" : [0.01, 0.8], "bridge" : [0.01, 1.0]. In this representation, the first component give (not exactly) a measure of "animalness" and the second, size. In addition, the vectors for similar or related words may be close together in space. In this question, we will examine the utility of (the highly popular) word embeddings.

Consider two neural networks for classifying sequences of words that differ only in their input representation:

- The first of which uses a sparse *one-hot* encoding of each word in which word i is identified by a vector that contains a 1 in position i and 0s elsewhere. For instance, a dictionary containing the words *word* and *embedding* might be represented as $[0\ 1]$ and $[1\ 0]$, respectively. You may assume that the dictionary used contains all words in both the training and testing sets.

- The second neural network, instead, uses a pre-trained embedding of the dictionary that you may assume represents every word in the dictionary.

Assuming that both networks use $\tanh$ activations and have randomly initialized weights, and they are trained using the same training set, but different input representation.

Now, at test time, each network is presented with a sequence of words not seen during training. Which of the following statement(s) is/are true about the output of the network for this sequence?

- [x] The first network cannot classify the sequence correctly
- [ ] The second network cannot classify the sequence correctly
- [x] The second network has a fighting chance at classifying the sequence
- [ ] The first network has a fighting chance at classifying the sequence

✔

Solution:

The first network cannot classify the sequence correctly. As a result of the words not being seen during training, the corresponding input units always had an output of zero, which means that no gradient was backpropagated into the weights and the randomly-initialized weights remain unchanged.

The second network, however, has a fighting chance at classifying the sequence.

Since the embedding process causes words that appear in similar contexts to have similar locations in vector space, the unseen words will likely fall into a region that contains words that the network **has** seen and knows how to correctly classify.

Submit You have used 2 of 2 attempts

Answers are displayed within the problem

Discussion

Hide Discussion

Topic: Unit 3 Neural networks (2.5 weeks):Homework 4 / 4. Word Embeddings

Add a Post

Show all posts by recent activity

- How I got it right — 5
  Here we are comparing "sparse word representation" (one-hot enconded) vs "dense distribut...
- Great Homework — 3
  I think this homework is great because 1) It relates nicely to all the concepts we have covered...
- Regarding "a fighting chance" — 3
  Isn't describing something with "a fighting chance" a bit subjective? I already got the answer ...
- Sorry, but that is a truly ridiculous problem — 1
  Everything is in the title. (If I wanted to develop a little, I think I would just have to put side to ...
- Cannot is a very strong statement — 3
  Does **cannot** mean that is is NEVER possible for it to happen? Or just highly unlikely. Tha...
- Fighting chance — 2
  I know there is another discussion post about what fighting chance means, but is it emphasiz...
- ? word embeding — 3
  what does fighting chance means here in the statement ?
- Stuck here; any hint — 5
  I am confused on how to approach this question. To me, both one-hot encoding and embedd...

what is the meaning of 'classify sequence'?
I am confused.
1

Hint
I saw other hints about the complexity issues, which is very fair but I thought it about differe...
1

Interesting problem
The example was very helpful (for me) in providing intuition.
1

Meaning of "sequence of words not seen during training"
To my understanding this means that all the words had been seen during training but the se...
2

[Staff] Homework 4 due date
It is said in the text above that: "**Extension Note: Homework 4 due date has been extended
3