2. Introduction

Bookmark this page

Exercises due Oct 28, 2020 05:29 IST *Completed*

Introduction to Recurrent Neural Networks

Start of transcript. Skip to the end.

Hello.

Welcome back.

Today we're going to be talking about how to model sequences.

In particular, we're going to be using neural networks

to model sequences.

And as part of that, we will introduce

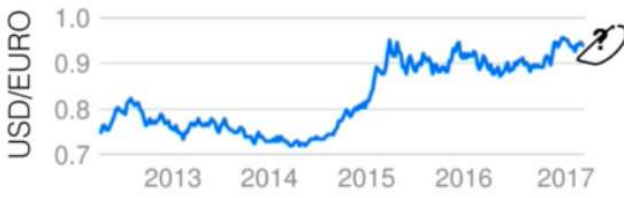
Video
Download video file

Transcripts
Download SubRip (.srt) file
Download Text (.txt) file

Encoding Sequences with Feed-Forward Neural Networks

1/1 point (graded)

We have a temporal dataset of USD/EURO conversion rate from late 2012 to early 2017. Our goal is to predict the value of USD/EURO at the next timestep of early 2017.



If we are trying to encode the data into feature vectors for a feed-forward neural network, which of the following is the most viable strategy?

- (●) slide a window of size 10 and use the most recent 10 points as a feature vector
- ( ) calculate the mean and the standard deviation of the entire sequence, and use them as a feature vector
- ( ) Use the length of the sequence and the standard deviation as a feature vector

✔

Solution:

As discussed in the lecture, a common scheme to encode sequences is to use sliding windows and use data inside the most recent sliding window.

Submit You have used 1 of 2 attempts

Answers are displayed within the problem

Context for Predicting Next Word

1/1 point (graded)
What is the issue with predicting the next word in the sentence using the previous **three** words as context? (Choose all that apply.)

- [x] Some words might need more context to predict
- [x] Some words might need less context to predict, and additional words could be inefficient
- [x] Some words might be closely related to words far away in the sentence
- [ ] Longer words are harder to predict because they have more letters

✔

Solution:

The amount of context we need to predict a word depends on the word. Therefore, some words could need more or less than 3 previous words as context. In addition, words could be related to other words far away in the sentence. Lastly, since we one-hot encode each of our words, the length of the word is irrelevant. All words, regardless of word length, have the same length one-hot encoding.

Submit You have used 1 of 2 attempts

Answers are displayed within the problem

Discussion

Hide Discussion

Topic: Unit 3 Neural networks (2.5 weeks):Lecture 10. Recurrent Neural Networks 1 / 2. Introduction

Add a Post

Show all posts by recent activity

? How to one-hot encode a word? 3
Hi, can you elaborate on how you would one-hot encode, let's say, the previous 3 words in a sentence? If I understand correctly, each w...

? Question 2 and its relation to the video 7
The video mentions 2 Problems with predicting the next word based on previous n words, 1. the length of the history. -> how far shall o...

Previous Next

© All Rights Reserved

© All Rights Reserved

edX

About
Affiliates
edX for Business
Open edX
Careers
News

Legal

Terms of Service & Honor Code
Privacy Policy
Accessibility Policy
Trademark Policy
Sitemap

Connect

Blog
Contact Us
Help Center
Media Kit
Donate

© 2020 edX Inc. All rights reserved.
深圳市恒宇博科技有限公司 粤ICP备17044299号-2