D213 Course Chatter

Task 2 – Sentiment analysis

Robert G.

COMMENT: 09/24/21 at 04:14 PM

Like • 0 Likes

Helpful hint for those of you running Anaconda for task 2. The latest version of Anaconda comes with python 3.8.8. There seems to be a bug with Python 3.7+ and TensorFlow. So, when you try to put certain functions after your embedded layer, it will crash. Bug discussion can be found here https://github.com/tensorflow/models/issues/9706. The solution that I found helpful and discussed in this thread is to downgrade your python to 3.6. Note, if you import sys and print(sys.version) on the data camp courses, you will indeed see they are running python version 3.6.7. It is not recommended that you downgrade your current anaconda environment. It can become unstable. The better solution can be found here: https://conda.io/projects/conda/en/latest/user-guide/getting-started.html#managing-python. You want to open up Anaconda Prompt and call ‘conda create --name d213 python=3.6’in the command prompt (remove ‘’). Then it will ask you if you are sure, and you type ‘y’. Then you open anaconda in this environment and reinstall your tensorflow, pandas, etc. Then magically, it works. If you have your code set up correctly. I spent over a week trying to figure out what I was missing. I hope this helps someone out there!

Robert G.

COMMENT: 09/09/21 at 09:54 AM

Like • 1 Likes

I chatted with Dr. Sewell this evening, it sounds like he recommends we use the file located here: https://archive.ics.uci.edu/ml/machine-learning-databases/00331/ and that we combine all 3 files in the zip folder. It was under the UCI Sentiment Labeled Sentences Data Set folder. Dr. Gagner said he does not teach this course, so I am not sure why he is on the instructor list. Dr. Elleh is on vacation. So, right now the only instructor in the office seems to be Dr. Sewell. I know he teaches a lot of other courses, but you best bet is to contact him. When I tried emailing d213@wgu.edu I had issues of my email actually getting to anyone.

Aziza B.

QUESTION: 11/04/21 at 11:21 PM

Like • 0 Likes

Greetings, I am not sure how do we identify categories of sentiments in B4 Task.2. Are the categories: "positive" and "negative"?

A person wearing glasses

Description automatically generated with low confidence

Festus E.

COMMENT: 11/10/21 at 11:12 PM

Like • 0 Likes

The answer is 'Yes' Aziza. You are looking for a binary classification of positive or negative in your model



Aziza B.

COMMENT: 11/13/21 at 08:03 PM

Like • 0 Likes

Thank you so much!

Sean B.

QUESTION: 09/21/21 at 10:23 AM

Like • 0 Likes

For task 2, I'm having trouble finding literature or video describing the appropriate method of choosing the word embedding length and a statistical justification for the chosen max sequence length. Any help would be appreciated. - Thanks.

A person wearing glasses

Description automatically generated with low confidence

Festus E.

COMMENT: 09/21/21 at 02:50 PM

Like • 0 Likes

Hello Sean, you can try reviewing the following resources. Hopefully, it provides some insights. Resources: Brownlee, J. (2021, Feb 02). How to use word embedding layers for deep learning with keras. https://machinelearningmastery.com/use-word-embedding-layers-deep-learning-keras/ https://developers.googleblog.com/2017/11/introducing-tensorflow-feature-columns.html Medium. (2020, Apr 03). padding for nlp. https://medium.com/@canerkilinc/padding-for-nlp-7dd8598c916a TensorFlow. (n.d). tf.keras.preprocessing.text.tokenizer. https://www.tensorflow.org/api\_docs/python/tf