

Summer 2021: CSEE5590 – Special Topics

Python_Lesson_5_Part_2: Word Embedding

Lesson Overview:

In this lesson, we are going to discuss types of ANNs and Recurrent Neural Network.

Use Case Description:

1. Sentiment Analysis on the imdb dataset

Source Code:

Provided in your assignment folder and assignment repo.

In class programming:

1. In the code provided there are three mistake which stops the code from running successfully.
 - a. find those mistakes and explain why they need to be corrected to be able to get the code run.
2. Add embedding layer to the model, did you experience any improvement?
3. Apply the code on 20_newsgroup data set we worked in the previous classes.

```
from sklearn.datasets import fetch_20newsgroups
newsgroups_train = fetch_20newsgroups(subset='train', shuffle=True,
categories=categories,)
```
4. Plot the loss and accuracy using history object.
5. Predict over one sample of data and check what will be the prediction for that.

Bonus Point: (5)

- Plot loss and accuracy in Tensorboard.

Online Submission Guidelines (for Online students):

1. Submit your source code and documentation to GitHub and represent the work in a ReadMe file properly (submit your screenshots as well. The screenshot should have both the code and the output)
2. Comment your code appropriately
3. Video Submission (2 – 5 min video showing the demo of the assignment, with brief voice over on the code explanation)

Note: *Cheating, plagiarism, disruptive behavior and other forms of unacceptable conduct are subject to strong sanctions in accordance with university policy. See detailed description of university policy at the following URL: <https://catalog.umkc.edu/special-notice/academic-honesty/>*