

# 2022-01-24: Questions and answers

Made with a warm hug

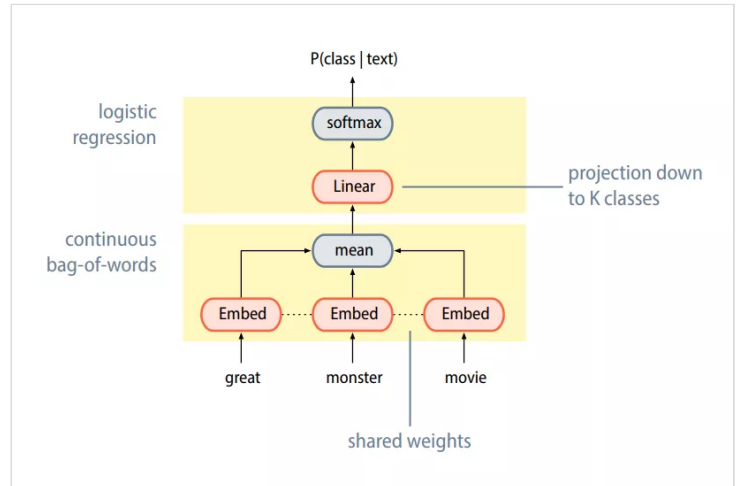
CHEERYTEXT3944 JAN 20, 2022 09:24PM

## (Max) pooling layers

With CNNs you often see max pooling in 2x2 blocks, which I interpret as retaining some locality of features but at a coarser level. When you presented max pooling you took the max over the whole feature matrices, but do you ever apply pooling (max or otherwise) over smaller intervals, similar to CNNs? Or is the concept of locality already captured sufficiently by how the text is processed piece-wise?

*You are right, you can pool both over filters (as in the Kim model that I presented in the lecture) or time-steps (= positions in the sequence, i.e., words or characters). One example of the latter is Lee et al.: <https://aclanthology.org/Q17-1026.pdf>*

– CHEERYTEXT3944



*I answered that question in class; see the video recording.*

– CHEERYTEXT3944

## CBOW

When printing the parameter shapes of the CBOW model, I get  
`torch.Size([13690, 30])`  
`torch.Size([5, 30])`  
`torch.Size([5])`

Could you explain how these parameter shapes fit into the picture of the CBOW-model, considering that in the notebook our input is a batch of 24 sentences, which afaik the embedding layer translates to a 24x20x30 tensor of word embeddings?

**Yes, I can publish here.**

HOORAY! – CHEERYTEXT3944

## Questions and answers

Post your questions about the course content on this padlet.  
 Btw, Marco is cheerytext3944.

\*\*\*\*\*