```
                                                    oov_char=2,
                                                    index_from=3)
```

The meanings of all of these arguments are here. But in a nutshell, the most important ones are:

- **num_words**: Top most frequent words to consider. This is useful if you don't want to consider very obscure words such as "Ultracrepidarian."
- **skip_top**: Top words to ignore. This is useful if you don't want to consider the most common words. For example, the word "the" would add no information to the review, so we can skip it by setting `skip_top` to 2 or higher.

Pre-processing the data

We first prepare the data by one-hot encoding it into (0,1)-vectors as follows: If, for example, we have 10 words in our vocabulary, and the vector is (4,1,8), we'll turn it into the vector (1,0,0,1,0,0,0,1,0,0).

Building the model

Now it's your turn to use all you've learned! You can build a neural network using Keras, train it, and evaluate it! Make sure you also use methods such as dropout or regularization, and good Keras optimizers to do this. A good accuracy to aim for is 85%. Can your model achieve this?

Help

This is a self-assessed lab. If you need any help or want to check your answers, feel free to check out the solutions notebook in the same folder, or click here.

Search or ask questions in Knowledge.

Ask peers or mentors for help in Student Hub.

NEXT