

Word embedding Techniques.ipynb ☆

File Edit View Insert Runtime Tools Help Changes will not be saved

+ Code + Text Copy to Drive Connect ▾

Word Embedding Techniques using Embedding Layer in Keras

```
[ ] ### Libraries USed Tensorflow> 2.0 and keras
[ ] !pip install tensorflow-gpu

Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-wheels/public/simple/
Collecting tensorflow-gpu
  Downloading tensorflow_gpu-2.9.1-cp37-cp37m-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (511.7 MB)
    |████████| 511.7 MB 6.2 kB/s
Requirement already satisfied: astunparse>=1.6.0 in /usr/local/lib/python3.7/dist-packages (from tensorflow-gpu) (1.6.3)
Requirement already satisfied: numpy>=1.20 in /usr/local/lib/python3.7/dist-packages (from tensorflow-gpu) (1.21.6)
Requirement already satisfied: wrapt>=1.11.0 in /usr/local/lib/python3.7/dist-packages (from tensorflow-gpu) (1.14.1)
Collecting keras<2.10.0,>=2.9.0rc0
  Downloading keras-2.9.0-py3-none-any.whl (1.6 MB)
    |████████| 1.6 MB 51.8 MB/s
Requirement already satisfied: h5py>=2.9.0 in /usr/local/lib/python3.7/dist-packages (from tensorflow-gpu) (3.1.0)
Requirement already satisfied: setuptools in /usr/local/lib/python3.7/dist-packages (from tensorflow-gpu) (57.4.0)
Requirement already satisfied: packaging in /usr/local/lib/python3.7/dist-packages (from tensorflow-gpu) (21.3)
Requirement already satisfied: grpcio<2.0,>=1.24.3 in /usr/local/lib/python3.7/dist-packages (from tensorflow-gpu) (1.46.3)
Requirement already satisfied: six>=1.12.0 in /usr/local/lib/python3.7/dist-packages (from tensorflow-gpu) (1.15.0)
Requirement already satisfied: tensorflow-io-gcs-filesystem>=0.23.1 in /usr/local/lib/python3.7/dist-packages (from tensorflow-gpu) (0.26.0)
Requirement already satisfied: keras-preprocessing>=1.1.1 in /usr/local/lib/python3.7/dist-packages (from tensorflow-gpu) (1.1.2)
Requirement already satisfied: absl-py>=1.0.0 in /usr/local/lib/python3.7/dist-packages (from tensorflow-gpu) (1.1.0)
Requirement already satisfied: libclang>=13.0.0 in /usr/local/lib/python3.7/dist-packages (from tensorflow-gpu) (14.0.1)
Collecting tensorboard<2.10,>=2.9
  Downloading tensorboard-2.9.1-py3-none-any.whl (5.8 MB)
    |████████| 5.8 MB 18.4 MB/s
Requirement already satisfied: protobuf<3.20,>=3.9.2 in /usr/local/lib/python3.7/dist-packages (from tensorflow-gpu) (3.17.3)
Requirement already satisfied: typing-extensions>=3.6.6 in /usr/local/lib/python3.7/dist-packages (from tensorflow-gpu) (4.1.1)
Requirement already satisfied: opt-einsum>=2.3.2 in /usr/local/lib/python3.7/dist-packages (from tensorflow-gpu) (3.3.0)
Collecting tensorflow-estimator<2.10.0,>=2.9.0rc0
  Downloading tensorflow_estimator-2.9.0-py3-none-any.whl (438 kB)
    |████████| 438 kB 58.8 MB/s
Requirement already satisfied: google-pasta>=0.1.1 in /usr/local/lib/python3.7/dist-packages (from tensorflow-gpu) (0.2.0)
Requirement already satisfied: termcolor>=1.1.0 in /usr/local/lib/python3.7/dist-packages (from tensorflow-gpu) (1.1.0)
Collecting gast<=0.4.0,>=0.2.1
  Downloading gast-0.4.0-py3-none-any.whl (9.8 kB)
Collecting flatbuffers<2,>=1.2
  Downloading flatbuffers-1.12-py2.py3-none-any.whl (15 kB)
Requirement already satisfied: wheel<1.0,>=0.23.0 in /usr/local/lib/python3.7/dist-packages (from astunparse>=1.6.0>tensorflow-gpu) (0.37.1)
Requirement already satisfied: cached-property in /usr/local/lib/python3.7/dist-packages (from h5py>=2.9.0>tensorflow-gpu) (1.5.2)
Requirement already satisfied: tensorboard-data-server<0.7.0,>=0.6.0 in /usr/local/lib/python3.7/dist-packages (from tensorboard<2.10,>=2.9>tensorflow-gpu) (0.6.1)
Requirement already satisfied: markdown>=2.6.8 in /usr/local/lib/python3.7/dist-packages (from tensorboard<2.10,>=2.9>tensorflow-gpu) (3.3.7)
Requirement already satisfied: google-auth<3,>=1.6.3 in /usr/local/lib/python3.7/dist-packages (from tensorboard<2.10,>=2.9>tensorflow-gpu) (1.35.0)
Requirement already satisfied: requests<3,>=2.21.0 in /usr/local/lib/python3.7/dist-packages (from tensorboard<2.10,>=2.9>tensorflow-gpu) (2.23.0)
Requirement already satisfied: tensorboard-plugin-wit>=1.6.0 in /usr/local/lib/python3.7/dist-packages (from tensorboard<2.10,>=2.9>tensorflow-gpu) (1.8.1)
Requirement already satisfied: google-auth-oauthlib<0.5,>=0.4.1 in /usr/local/lib/python3.7/dist-packages (from tensorboard<2.10,>=2.9>tensorflow-gpu) (0.4.6)
Requirement already satisfied: werkzeug>=1.0.1 in /usr/local/lib/python3.7/dist-packages (from tensorboard<2.10,>=2.9>tensorflow-gpu) (1.0.1)
Requirement already satisfied: rsa<5,>=3.1.4 in /usr/local/lib/python3.7/dist-packages (from google-auth<3,>=1.6.3>tensorboard<2.10,>=2.9>tensorflow-gpu) (4.8)
Requirement already satisfied: pyasn1-modules>=0.2.1 in /usr/local/lib/python3.7/dist-packages (from google-auth<3,>=1.6.3>tensorboard<2.10,>=2.9>tensorflow-gpu) (0.4.7)
Requirement already satisfied: cachetools<5.0,>=2.0.0 in /usr/local/lib/python3.7/dist-packages (from google-auth<3,>=1.6.3>tensorboard<2.10,>=2.9>tensorflow-gpu) (2.4.0)
Requirement already satisfied: requests-oauthlib>=0.7.0 in /usr/local/lib/python3.7/dist-packages (from google-auth-oauthlib<0.5,>=0.4.1>tensorboard<2.10,>=2.9>ten
Requirement already satisfied: importlib-metadata>=4.4 in /usr/local/lib/python3.7/dist-packages (from markdown>=2.6.8>tensorboard<2.10,>=2.9>tensorflow-gpu) (4.11)
Requirement already satisfied: zipp>=0.5 in /usr/local/lib/python3.7/dist-packages (from importlib-metadata>=4.4>markdown>=2.6.8>tensorboard<2.10,>=2.9>tensorflow
Requirement already satisfied: pyasn1<0.5.0,>=0.4.6 in /usr/local/lib/python3.7/dist-packages (from pyasn1-modules>=0.2.1>google-auth<3,>=1.6.3>tensorboard<2.10,>=
Requirement already satisfied: chardet<4,>=3.0.2 in /usr/local/lib/python3.7/dist-packages (from requests<3,>=2.21.0>tensorboard<2.10,>=2.9>tensorflow-gpu) (3.0.4)
Requirement already satisfied: urllib3!=1.25.0,!=1.25.1,<1.26,>=1.21.1 in /usr/local/lib/python3.7/dist-packages (from requests<3,>=2.21.0>tensorboard<2.10,>=2.9>te
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.7/dist-packages (from requests<3,>=2.21.0>tensorboard<2.10,>=2.9>tensorflow-gpu) (2022.
Requirement already satisfied: idna>=2.5 in /usr/local/lib/python3.7/dist-packages (from requests<3,>=2.21.0>tensorboard<2.10,>=2.9>tensorflow-gpu) (2.10)
Requirement already satisfied: oauthlib>=3.0.0 in /usr/local/lib/python3.7/dist-packages (from requests-oauthlib>=0.7.0>google-auth-oauthlib<0.5,>=0.4.1>tensorboar
Requirement already satisfied: pyparsing!=3.0.5,>=2.0.2 in /usr/local/lib/python3.7/dist-packages (from packaging>tensorflow-gpu) (3.0.9)
Installing collected packages: tensorflow-estimator, tensorboard, keras, gast, flatbuffers, tensorflow-gpu
  Created wheel for tensorflow-estimator: tensorflow-estimator-2.9.0-py3-none-any.whl
  Created wheel for tensorboard: tensorboard-2.9.1-py3-none-any.whl
  Created wheel for keras: keras-2.9.0-py3-none-any.whl
  Created wheel for gast: gast-0.4.0-py3-none-any.whl
  Created wheel for flatbuffers: flatbuffers-1.12-py2.py3-none-any.whl
  Created wheel for tensorflow-gpu: tensorflow-gpu-2.9.1-cp37-cp37m-manylinux_2_17_x86_64.manylinux2014_x86_64.whl
Successfully installed tensorflow-estimator-2.9.0 tensorboard-2.9.1 keras-2.9.0 gast-0.4.0 flatbuffers-1.12 tensorflow-gpu-2.9.1
```

```
[ ] import tensorflow as tf
print(tf.__version__)

2.9.1

[ ] ##tensorflow >2.0
from tensorflow.keras.preprocessing.text import one_hot

[ ]

[ ] ### sentences
sent=[ 'the glass of milk',
      'the glass of juice',
      'the cup of tea',
      'I am a good boy',
      'I am a good developer',
      'understand the meaning of words',
      'your videos are good']

[ ] sent

['the glass of milk',
 'the glass of juice',
 'the cup of tea',
 'I am a good boy']
```

```
'I am a good developer',
'understand the meaning of words',
'your videos are good']
```

```
[ ] ### Vocabulary size
voc_size=500
```

▼ One Hot Representation

```
[ ] onehot_repr=[one_hot(words,voc_size)for words in sent]
print(onehot_repr)

[[180, 405, 264, 53], [180, 405, 264, 8], [180, 92, 264, 33], [291, 43, 307, 242, 275], [291, 43, 307, 242, 98], [362, 180, 144, 264, 188], [354, 52, 496, 242]]
```

▼ Word Embedding Representation

```
[ ] from tensorflow.keras.layers import Embedding
from tensorflow.keras.preprocessing.sequence import pad_sequences
from tensorflow.keras.models import Sequential

[ ] import numpy as np

[ ] ## pre padding
sent_length=8
embedded_docs=pad_sequences(onehot_repr,padding='pre', maxlen=sent_length)
print(embedded_docs)

[[ 0   0   0   0 180 405 264  53]
 [ 0   0   0   0 180 405 264   8]
 [ 0   0   0   0 180  92 264  33]
 [ 0   0   0 291  43 307 242 275]
 [ 0   0   0 291  43 307 242  98]
 [ 0   0   0 362 180 144 264 188]
 [ 0   0   0 354  52 496 242]]
```

```
[ ] ## 10 feature dimesnions
dim=10
```

```
[ ] model=Sequential()
model.add(Embedding(voc_size,10,input_length=sent_length))
model.compile('adam','mse')
```

```
[ ] model.summary()
```

```
Model: "sequential"
-----  
Layer (type)          Output Shape         Param #
embedding (Embedding) (None, 8, 10)        5000
-----  
Total params: 5,000
Trainable params: 5,000
Non-trainable params: 0
```

```
[ ] ##'the glass of milk',
embedded_docs[0]

array([ 0,  0,  0,  0, 180, 405, 264,  53], dtype=int32)

[ ] model.predict(embedded_docs[0])

WARNING:tensorflow:Model was constructed with shape (None, 8) for input KerasTensor(type_spec=TensorSpec(shape=(None, 8), dtype=tf.float32, name='embedding_input'), na
1/1 [=====] - 0s 72ms/step
array([[ 0.03938437, -0.02009605, -0.03878935, -0.04955565,  0.00419912,
-0.01431773,  0.02523251,  0.01653036,  0.04291571, -0.00864979],
[ 0.03938437, -0.02009605, -0.03878935, -0.04955565,  0.00419912,
-0.01431773,  0.02523251,  0.01653036,  0.04291571, -0.00864979],
[ 0.03938437, -0.02009605, -0.03878935, -0.04955565,  0.00419912,
-0.01431773,  0.02523251,  0.01653036,  0.04291571, -0.00864979],
[ 0.03938437, -0.02009605, -0.03878935, -0.04955565,  0.00419912,
-0.01431773,  0.02523251,  0.01653036,  0.04291571, -0.00864979],
[ 0.03938437, -0.02009605, -0.03878935, -0.04955565,  0.00419912,
-0.01431773,  0.02523251,  0.01653036,  0.04291571, -0.00864979],
[ 0.03059326, -0.04286614,  0.00899569,  0.00743791, -0.000781 ,
0.04186494,  0.03977301,  0.00326709,  0.00619651, -0.01993654],
[ 0.02512412, -0.0087087 ,  0.03144198,  0.00704668, -0.00177735,
-0.03415867, -0.00100178,  0.01562483,  0.03178963,  0.02784893],
[-0.00653008,  0.02340979, -0.01967902, -0.00494973, -0.02693756,
-0.03746525,  0.01460877, -0.00449115, -0.00130982, -0.0039017 ],
[-0.03150218,  0.01950303, -0.01415605, -0.00183152,  0.01207731,
0.02444079,  0.0140041 ,  0.0070256 ,  0.04950741, -0.03602346]],  
dtype=float32)
```

```
[ ] print(model.predict(embedded_docs))
```

```
1/1 [=====] - 0s 31ms/step
[[ 0.03938437 -0.02009605 -0.03878935 -0.04955565  0.00419912
-0.01431773  0.02523251  0.01653036  0.04291571 -0.00864979]]
```

```

[ 0.03938437 -0.02009605 -0.03878935 -0.04955565  0.00419912
-0.01431773  0.02523251  0.01653036  0.04291571 -0.00864979]
[ 0.03938437 -0.02009605 -0.03878935 -0.04955565  0.00419912
-0.01431773  0.02523251  0.01653036  0.04291571 -0.00864979]
[ 0.03938437 -0.02009605 -0.03878935 -0.04955565  0.00419912
-0.01431773  0.02523251  0.01653036  0.04291571 -0.00864979]
[ 0.03059326 -0.04286614  0.00899569  0.00743791 -0.000781
 0.04186494  0.03977301  0.00326709  0.00619651 -0.01993654]
[ 0.02512412 -0.0087087  0.03144198  0.00704668 -0.00177735
-0.03415867 -0.00100178  0.01562483  0.03178963  0.02784893]
[ 0.00653008  0.02340979 -0.01967902 -0.00494973 -0.02693756
-0.03746525  0.01460877 -0.00449115 -0.00130982 -0.0039017 ]
[ 0.03150218  0.01950303 -0.01415605 -0.00183152  0.01207731
 0.02444079  0.0140041  0.0070256  0.04950741 -0.03602346]]
```



```

[[ 0.03938437 -0.02009605 -0.03878935 -0.04955565  0.00419912
-0.01431773  0.02523251  0.01653036  0.04291571 -0.00864979]
[ 0.03938437 -0.02009605 -0.03878935 -0.04955565  0.00419912
-0.01431773  0.02523251  0.01653036  0.04291571 -0.00864979]
[ 0.03938437 -0.02009605 -0.03878935 -0.04955565  0.00419912
-0.01431773  0.02523251  0.01653036  0.04291571 -0.00864979]
[ 0.03938437 -0.02009605 -0.03878935 -0.04955565  0.00419912
-0.01431773  0.02523251  0.01653036  0.04291571 -0.00864979]
[ 0.03938437 -0.02009605 -0.03878935 -0.04955565  0.00419912
-0.01431773  0.02523251  0.01653036  0.04291571 -0.00864979]
[ 0.03059326 -0.04286614  0.00899569  0.00743791 -0.000781
 0.04186494  0.03977301  0.00326709  0.00619651 -0.01993654]
[ 0.02512412 -0.0087087  0.03144198  0.00704668 -0.00177735
-0.03415867 -0.00100178  0.01562483  0.03178963  0.02784893]
[ 0.00653008  0.02340979 -0.01967902 -0.00494973 -0.02693756
-0.03746525  0.01460877 -0.00449115 -0.00130982 -0.0039017 ]
[ -0.0434371  0.01733501 -0.0254814  0.03025435 -0.0460149
 0.00874413  0.04856688  0.03454936 -0.02982813 -0.00472248]]
```



```

[[ 0.03938437 -0.02009605 -0.03878935 -0.04955565  0.00419912
-0.01431773  0.02523251  0.01653036  0.04291571 -0.00864979]
[ 0.03938437 -0.02009605 -0.03878935 -0.04955565  0.00419912
-0.01431773  0.02523251  0.01653036  0.04291571 -0.00864979]
[ 0.03938437 -0.02009605 -0.03878935 -0.04955565  0.00419912
-0.01431773  0.02523251  0.01653036  0.04291571 -0.00864979]
[ 0.03938437 -0.02009605 -0.03878935 -0.04955565  0.00419912
-0.01431773  0.02523251  0.01653036  0.04291571 -0.00864979]
[ 0.03059326 -0.04286614  0.00899569  0.00743791 -0.000781
 0.04186494  0.03977301  0.00326709  0.00619651 -0.01993654]
[ -0.0165449 -0.0245487 -0.04049651 -0.03379797  0.03827978
-0.01891239  0.03935272  0.03601534  0.04047254  0.02626738]
[ 0.00653008  0.02340979 -0.01967902 -0.00494973 -0.02693756
-0.03746525  0.01460877 -0.00449115 -0.00130982 -0.0039017 ]
[ -0.01796211  0.03012553  0.02499587 -0.03019696 -0.02619814
-0.04370998  0.02499839  0.0259905 -0.02895923 -0.03277919]]
```



```

[[ 0.03938437 -0.02009605 -0.03878935 -0.04955565  0.00419912
-0.01431773  0.02523251  0.01653036  0.04291571 -0.00864979]
[ 0.03938437 -0.02009605 -0.03878935 -0.04955565  0.00419912
-0.01431773  0.02523251  0.01653036  0.04291571 -0.00864979]
[ 0.03938437 -0.02009605 -0.03878935 -0.04955565  0.00419912
-0.01431773  0.02523251  0.01653036  0.04291571 -0.00864979]
[ 0.03938437 -0.02009605 -0.03878935 -0.04955565  0.00419912
-0.01431773  0.02523251  0.01653036  0.04291571 -0.00864979]]
```

```

[ ] embedded_docs[0]

array([ 0, 0, 0, 6654, 998, 8966, 1609])

[ ] print(model.predict(embedded_docs)[0])

[[ -0.00425554 -0.00159295 -0.04714153  0.04425247 -0.00973954 -0.04325813
  0.04007108 -0.0143286 -0.03659749 -0.02379028]
[ -0.00425554 -0.00159295 -0.04714153  0.04425247 -0.00973954 -0.04325813
  0.04007108 -0.0143286 -0.03659749 -0.02379028]
[ -0.00425554 -0.00159295 -0.04714153  0.04425247 -0.00973954 -0.04325813
  0.04007108 -0.0143286 -0.03659749 -0.02379028]
[ -0.00425554 -0.00159295 -0.04714153  0.04425247 -0.00973954 -0.04325813
  0.04007108 -0.0143286 -0.03659749 -0.02379028]
[ -0.00425554 -0.00159295 -0.04714153  0.04425247 -0.00973954 -0.04325813
  0.04007108 -0.0143286 -0.03659749 -0.02379028]
[ -0.00425554 -0.00159295 -0.04714153  0.04425247 -0.00973954 -0.04325813
  0.04007108 -0.0143286 -0.03659749 -0.02379028]
[ -0.03786233 -0.02628061  0.02974111 -0.03307171  0.0271405  0.00945134
  0.03787127  0.04176904  0.00514941  0.0152082 ]
[ 0.04834186  0.04388311 -0.02802253 -0.01475487 -0.01212303  0.03762435
-0.01166249 -0.02141088  0.04654533  0.01537322]
[ 0.03276015 -0.00637691  0.03907344 -0.019612468  0.02177186 -0.04630325
 0.00800942 -0.03115667 -0.00486455 -0.04843524]
[ -0.04173617  0.03438064  0.02880521 -0.01896455  0.0323303 -0.00109453
-0.01675171 -0.00941917 -0.03309294 -0.04779492]]
```

```

[ ] ### Assignment

sent=[ "The world is a better place",
      "Marvel series is my favourite movie",
      "I like DC movies",
      "the cat is eating the food",
      "Tom and Jerry is my favourite movie",
      "Python is my favourite programming language"
    ]
```

[Colab paid products](#) - [Cancel contracts here](#)

