

Python 3.8.3 (default, Jul 2 2020, 17:30:36) [MSC v.1916 64 bit (AMD64)]  
Type "copyright", "credits" or "license" for more information.

IPython 7.16.1 -- An enhanced Interactive Python.

```
In [1]: 'E:/Desktop/sem 7/irs/IRS/practical-5/Main.py' = 'E:/Desktop/sem
7/irs/IRS/practical-5'
```

Stop Word Removal

Stemming

Vocabulary

```
['24', '25', 'account', 'actions', 'actors', 'add', 'also', 'amount', 'ask', 'author',
'authoris', 'bank', 'bearer', 'blank', 'borrowing', 'broadcasting', 'cable', 'call',
'camera', 'cameraman', 'cannot', 'carri', 'case', 'cash', 'chequ', 'cinemas',
'commonli', 'communicate', 'computer', 'condit', 'confirm', 'corner', 'cross', 'cry',
'data', 'declar', 'dialogue', 'director', 'directors', 'disk', 'done', 'download',
'draw', 'drawer', 'dvd', 'either', 'empti', 'encash', 'endors', 'entertain',
'equipment', 'every', 'feel', 'field', 'films', 'find', 'frame', 'fun', 'get', 'gets',
'getting', 'hires', 'home', 'identif', 'ie', 'import', 'indian', 'information',
'intact', 'invalid', 'investors', 'kind', 'known', 'later', 'laugh', 'learn', 'left',
'lend', 'liabl', 'like', 'line', 'looks', 'made', 'make', 'mark', 'market', 'may',
'mean', 'medias', 'memorize', 'mention', 'money', 'month', 'motion', 'move', 'movie',
'movies', 'must', 'mutil', 'name', 'need', 'needed', 'older', 'others', 'parti',
'pay', 'paye', 'people', 'period', 'person', 'picture', 'piec', 'process', 'produce',
'producers', 'projector', 'quick', 'rate', 'reach', 'read', 'really', 'record',
'reject', 'relev', 'rented', 'requir', 'say', 'screen', 'screenwriter', 'script',
'scripts', 'second', 'set', 'several', 'shall', 'shown', 'shows', 'signatur', 'sold',
'something', 'sound', 'story', 'stream', 'strike', 'submit', 'takes', 'tapes',
'teach', 'television', 'tell', 'tells', 'thing', 'third', 'time', 'top', 'torn',
'transferred', 'two', 'type', 'use', 'uses', 'usual', 'valid', 'video',
'videocassette', 'view', 'visit', 'visual', 'watch', 'way', 'week', 'whose',
'withdraw', 'words', 'work', 'write']
```

Document Term Matrix

```
Doc1:  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0
0  0  0  0  0  0  1  0  0  0  0  0  1  0  0  0  0  0  0  0  0  0
0  0  0  0  0  1  0  0  1  0  1  0  0  1  0  0  0  0  0  0  0  0
0  0  0  0  0  0  0  1  0  0  0  0  0  0  0  0  1  0  0  0  1  0  0
0  0  0  0  1  0  1  0  0  0  0  0  0  1  0  0  0  1  0  0  1  0  0
0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0
0  0  0  1  1  1  0  0  0  0  0  1  0  1  0  0  0  0  0  0  0  0  1
0  1  0  0  0  0  1  0  1  1  1  0  0  0  0  0  0
Doc2:  0  0  0  0  0  0  1  0  0  0  0  0  0  0  0  0  1  1  0  0  0  0
0  0  0  0  1  0  0  0  0  0  0  0  0  0  0  0  0  0  1  0  1  0  0
1  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  1  0  0  0  0
0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  1  0  0  1  1  0  1  0
0  0  1  0  0  0  1  0  0  0  0  0  1  0  0  1  0  1  1  0  0  0  0
0  0  0  0  0  0  0  0  0  0  0  1  0  0  1  0  0  0  0  1  0  1
```

0	0	1	0	0	0	1	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0
0	0	0	0	0	1	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0
Doc3:	1	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	1	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	1	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0
0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
0	0	1	1	1	0	0	1	1	0	0	0	0	0	0	0	0	0	1	1	0	0
1	0	0	0	1	0	0	0	0	1	0	0	1	0	0	1	0	1	0	0	0	0
0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Doc4:	0	0	0	0	1	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	1	1	1	0	0	0	0
0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
0	1	0	0	0	1	0	0	0	0	0	1	0	0	0	1	0	1	0	0	0	0
1	1	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	0	0	0	0	0
0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	1						
Doc5:	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1
0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	1	1	0	0	0	0
0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0					
Doc6:	0	0	0	0	0	0	0	0	1	1	1	1	1	1	0	0	0	0	0	0	0
1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	1	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0
0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0					
Doc7:	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	1
0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
0	0	0	1	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0
0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0
1	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	1	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1
0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0					
Doc8:	0	0	1	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0
0	1	1	1	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	0	1	1
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	1	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	1	0	1	1
0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0					
Doc9:	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0
0	0	0	1	0	0	0	0	1	1	1	0	0	1	1	0	0	0	0	0	0	1
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	1

```

0 0 0 0 0 1 0 0 0 1 1 0 0 0 0 0 0 0 0 0 1 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 1 0
0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
Doc10: 0 0 0 0 0 0 0 0 0 0 0 0 1 0 1 0 0 0 1 0 0 0
0 0 0 1 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1
0 0 1 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 1
0 0 0 0 0 1 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0
1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1
1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

```

Reduction to lower rank approximation

```

[[ 4.05921467e-16  3.36508785e-16 -1.91941440e-15 ... -7.24561991e-16
   2.68708053e-16  2.68708053e-16]
 [-2.92878142e-16  5.24494651e-16  7.45864733e-16 ...  5.86094173e-16
   1.47041833e-16  1.47041833e-16]
 [ 1.00000000e+00  1.00000000e+00  1.43886679e-15 ... -4.91297355e-16
  -1.78611307e-17 -1.78611307e-17]
 ...
 [-4.86703415e-17  3.22718231e-16  1.00000000e+00 ...  1.09602196e-16
  -2.60485509e-16 -2.60485509e-16]
 [-4.18506194e-17  2.88630584e-16 -1.86584167e-15 ... -1.39458808e-15
   4.33311884e-16  4.33311884e-16]
 [ 3.92527744e-16  3.71925237e-16 -1.42520090e-15 ... -1.29251692e-16
  -2.28227114e-16 -2.28227114e-16]]

```

In [2]: