CWordTM Toolkit Usage on the Holy Bible (CUV)

This Jupyter notebook demonstrates how to use the package "CWordTM" on the Holy Bible (Chinese Union Version - Traditional Chinese):

- 1. Utility Features
- 2. Text Visualization Word Cloud
- 3. Text Summarization
- 4. Pivot Table
- 5. OT Quotes
- 6. Topic Modeling LDA, BERTopic and NMF

CWordTM Toolkit's Documentation: https://cwordtm.readthedocs.io

```
In [1]: import warnings
    warnings.filterwarnings('ignore')

In [2]: # Import the Package CWordTM
    import cwordtm
    from cwordtm import *

WARNING:tensorflow:From D:\Dev\Anaconda3\envs\r1\lib\site-packages\tf_keras\src\losses.py:2976: The name tf.losses.sp
    arse_softmax_cross_entropy is deprecated. Please use tf.compat.v1.losses.sparse_softmax_cross_entropy instead.

In [3]: cwordtm.__version__
Out[3]: '0.7.1'
```

1. Utility Features

```
In [4]: # Load the whole Bible (Chinese Union Version)
       bible = "cuv.csv"
       cdf = util.load_word(bible, info=True)
      Loading file 'D:\Dev\Anaconda3\envs\rl\lib\site-packages\cwordtm\data\cuv.csv' ...
      Dataset Information:
      <class 'pandas.core.frame.DataFrame'>
      RangeIndex: 31102 entries, 0 to 31101
      Data columns (total 9 columns):
       # Column Non-Null Count Dtype
       --- -----
                     _____
       0 book 31102 non-null object
       1 book_no 31102 non-null int64
       2 chapter 31102 non-null int64
       3 verse 31102 non-null int64
4 text 31102 non-null object
       5 testament 31102 non-null int64
       6 category 31102 non-null object
       7 cat 31102 non-null object
       8 cat_no
                   31102 non-null int64
      dtypes: int64(5), object(4)
      memory usage: 2.1+ MB
```

Extract Partial Scripture

```
In [5]: # Extract OT Scripture
  otdf = util.extract(cdf, testament=0)
  otdf.head()
```

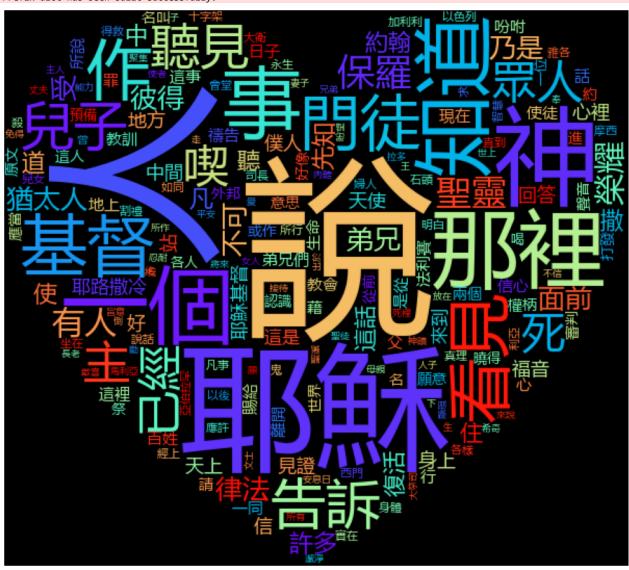
```
Out[5]:
          book book_no chapter verse
                                                                        text testament category cat cat_no
        0
           Gen
                                  1
                                                             起初神創造天地。
                                                                                    0
                                                                                         Torah
                                                                                             tor
                                                                                                      0
        1
                                  2
                                       地是空虛混沌. 淵面黑暗. 神的靈運行在水面上。
                                                                                    0
                                                                                                      0
           Gen
                                                                                         Torah
                                                                                              tor
        2
           Gen
                     1
                            1
                                  3
                                                         神說、要有光、就有了光。
                                                                                   0
                                                                                         Torah tor
                                                                                                      0
            Gen
                                  4
                                                    神看光是好的、就把光暗分開了。
                                                                                         Torah
                                                                                                      0
           Gen
                            1
                                  5 神稱光為畫、稱暗為夜.有晚上、有早晨、這是頭一日。
                                                                                   0
                                                                                        Torah tor
                                                                                                      0
In [6]: # Extract Gospels (The first four book in NT)
        gos = util.extract(cdf, category='gos')
        gos.head()
Out[6]:
              book book_no chapter verse
                                                                          text testament category cat cat_no
                                        亞伯拉罕的後裔、大衛的子孫、耶穌基督的家譜. 〔後
        23145
               Mat
                        40
                                                                                      1
                                                                                                        5
                                                                                          Gospel
                                                                                                gos
                                                           裔子孫原文都作兒子下同〕
                                         亞伯拉罕生以撒. 以撒生雅各. 雅各生猶大和他的弟
        23146
               Mat
                        40
                                1
                                     2
                                                                                                        5
                                                                                          Gospel gos
                                        猶大從他瑪氏生法勒斯和謝拉. 法勒斯生希斯崙. 希斯
        23147
               Mat
                        40
                                1
                                                                                      1
                                                                                                        5
                                                                                          Gospel gos
                                                                      崙生亞蘭.
        23148
               Mat
                        40
                                           亞蘭生亞米拿達. 亞米拿達生拿順. 拿順生撒門.
                                                                                          Gospel
                                                                                                gos
                                                                                                        5
                                        撒門從喇合氏生波阿斯.波阿斯從路得氏生俄備得.俄
        23149
               Mat
                        40
                                                                                                        5
                                                                                          Gospel gos
                                                                    備得生耶西.
In [7]: # Extract the Book of Romans
        rom = util.extract(cdf, book='Rom')
        rom.head()
Out[7]:
              book book_no chapter verse
                                                                        text testament
                                                                                        category
                                                                                                cat cat_no
                                        耶穌基督的僕人保羅、奉召為使徒、特派傳 神的
                                                                                          Pauline
        27930
              Rom
                        45
                                                                                                pau
                                                                                                        7
                                                                                         Epistles
                                                                                          Pauline
        27931
               Rom
                        45
                                        這福音是神從前藉眾先知、在聖經上所應許的、
                                                                                                pau
                                                                                          Epistles
                                        論到他兒子我主耶穌基督. 按肉體說、是從大衛後
                                                                                          Pauline
        27932
              Rom
                        45
                                1
                                                                                                        7
                                                                                                pau
                                                                    裔生的、
                                                                                          Epistles
                                        按聖善的靈說、因從死裡復活、以大能顯明是神
                                                                                          Pauline
        27933
               Rom
                                                                                          Epistles
                                        我們從他受了恩惠、並使徒的職分、在萬國之中叫
                                                                                          Pauline
        27934
              Rom
                        45
                                1
                                                                                                        7
                                                           人為他的名信服真道.
                                                                                          Epistles
In [8]: # Extract Verse 28 of Chapter 8 of Romans
        rom8_28 = util.extract(cdf, book='Rom', chapter=8, verse=28)
        rom8_28.text
               我們曉得萬事都互相效力、叫愛神的人得益處、就是按他旨意被召的人。
Out[8]: 28143
        Name: text, dtype: object
In [9]: # Another way to extract Rom 8:28
        rom8_28 = util.extract2(cdf, filter='Rom 8:28')
        rom8_28.text
               我們曉得萬事都互相效力、叫愛神的人得益處、就是按他旨意被召的人。
Out[9]: 28143
        Name: text, dtype: object
In [10]: # Extract Rom 8:1~10
        rom8_1_10 = util.extract2(cdf, filter='Rom 8:1-10')
        list(rom8_1_10.text)
```

```
Out[10]: ['如今那些在基督耶穌裡的、就不定罪了。'.
         因為賜生命聖靈的律、在基督耶穌裡釋放了我、使我脫離罪和死的律了。',
        '律法既因肉體軟弱、有所不能行的、\u3000神就差遣自己的兒子、成為罪身的形狀、作了贖罪祭、在肉體中定了罪案.',
        '使律法的義、成就在我們這不隨從肉體、只隨從聖靈的人身上。
        '因為隨從肉體的人、體貼肉體的事. 隨從聖靈的人、體貼聖靈的事。',
        '體貼肉體的就是死. 體貼聖靈的乃是生命平安. '
        '原來體貼肉體的、就是與\u3000神為仇. 因為不服\u3000神的律法、也是不能服。',
        '而且屬肉體的人、不能得\u3000神的喜歡。',
        '如果\u3000神的靈住在你們心裡、你們就不屬肉體、乃屬聖靈了. 人若沒有基督的靈、就不是屬基督的。',
        '基督若在你們心裡、身體就因罪而死、心靈卻因義而活.']
       2. Text Visualization - Word Cloud
In [11]: # Extract the NT Scripture for Word Cloud
       text_list = util.get_text_list(cdf[cdf.testament==1]) # Load New Testament Scripture
       text list[:10]
Out[11]: ['亞伯拉罕的後裔、大衛的子孫、耶穌基督的家譜. 〔後裔子孫原文都作兒子下同〕',
        '亞伯拉罕生以撒.以撒生雅各.雅各生猶大和他的弟兄.',
        '猶大從他瑪氏生法勒斯和謝拉. 法勒斯生希斯崙. 希斯崙生亞蘭. ',
        '亞蘭生亞米拿達.亞米拿達生拿順.拿順生撒門.',
        '撒門從喇合氏生波阿斯.波阿斯從路得氏生俄備得.俄備得生耶西.',
        '耶西生大衛王。大衛從烏利亞的妻子生所羅門.',
        '所羅門生羅波安.羅波安生亞比雅.亞比雅生亞撒.',
        '亞撒生約沙法. 約沙法生約蘭. 約蘭生烏西亞. ',
        '烏西亞生約坦. 約坦生亞哈斯. 亞哈斯生希西家. ',
        '希西家生瑪拿西.瑪拿西生亞們.亞們生約西亞.']
In [12]: # White background with no image mask
       viz.chi_wordcloud(text_list)
      Building prefix dict from the default dictionary ...
      Loading model from cache C:\Users\johnnyc\AppData\Local\Temp\jieba.cache
      Loading Chinese vocabulary 'D:\Dev\Anaconda3\envs\rl\lib\site-packages\cwordtm\data\bible_vocab.txt' ...
      Loading model cost 0.517 seconds.
      Prefix dict has been built successfully.
      Building prefix dict from D:\Dev\Anaconda3\envs\rl\lib\site-packages\cwordtm\dictionary\dict.txt.big.txt ...
      Dumping model to file cache C:\Users\johnnyc\AppData\Local\Temp\jieba.u056162d159e6cf3cec64c34333130f67.cache
      Loading model cost 1.183 seconds.
      Prefix dict has been built successfully
```



```
In [13]: # Use internal image mask
         viz.chi_wordcloud(text_list, bg='black', image=1)
```

Building prefix dict from D:\Dev\Anaconda3\envs\rl\lib\site-packages\cwordtm\dictionary\dict.txt.big.txt ... Loading model from cache C:\Users\johnnyc\AppData\Local\Temp\jieba.u056162d159e6cf3cec64c34333130f67.cache



3. Text Summarization

```
In [14]: # Extract and summarize Chapter 8 of Romans
rom8 = util.extract2(cdf, 'Rom 8')
ta.summary_chi(rom8)

Building prefix dict from D:\Dev\Anaconda3\envs\rl\lib\site-packages\cwordtm\dictionary\dict.txt.big.txt ...
Loading model from cache C:\Users\johnnyc\AppData\Local\Temp\jieba.u056162d159e6cf3cec64c34333130f67.cache
Loading Chinese vocabulary 'D:\Dev\Anaconda3\envs\rl\lib\site-packages\cwordtm\data\bible_vocab.txt' ...
Loading model cost 0.880 seconds.
Prefix dict has been built successfully.

Out[14]: ['因為隨從肉體的人、體貼肉體的事. 隨從聖靈的人、體貼聖靈的事',
'而且屬肉體的人、不能得神的喜歡',
```

4. Pivot Table

Show Bible Scripture Statistics through a Pivot Table

'如果神的靈住在你們心裡、你們就不屬肉體、乃屬聖靈了. 人若沒有基督的靈、就不是屬基督的']

```
In [15]: util.set_rows()
    pivot.stat(cdf, chi=True)
```

			chapter	verse	text
category	book_no	book			
Torah	1	Gen	50	1533	51460
	2	Exo	40	1213	40057
	3	Lev	27	859	29228
	4	Num	36	1288	41654
	5	Deu	34	959	35904
Sub-Total			187	5852	198303
OT History	6	Jos	24	658	25794
	7	Jug	21	618	24375
	8	Rut	4	85	3362
	9	1Sa	31	810	32218
	10	2Sa	24	695	26697
	11	1Ki	22	816	30530
	12	2Ki	25	719	29626
	13	1Ch	29	942	30231
	14	2Ch	36	822	33445
	15	Ezr	10	280	10094
	16	Neh	13	406	14739
	17	Est	10	167	6571
Sub-Total			249	7018	267682
Ketuvim	18	Job	42	1070	24294
	19	Psm	150	2461	64303
	20	Pro	31	915	19986
	21	Ecc	12	222	7228
	22	Son	8	117	3979
Sub-Total			243	4785	119790
Major Prophets	23	Isa	66	1292	50630
	24	Jer	52	1364	58005
	25	Lam	5	154	4721
	26	Eze	48	1273	49218
	27	Dan	12	357	14646
Sub-Total			183	4440	177220
Minor Prophets	28	Hos	14	197	7455
	29	Joe	3	73	2642
	30	Amo	9	146	5487
	31	Oba	1	21	838
	32	Jon	4	48	1675
	33	Mic	7	105	4307
	34	Nah	3	47	1716
	35	Hab	3	56	1974
	36	Zep	3	53	2216

			chapter	verse	text
category	book_no	book			
	37	Hag	2	38	1484
	38	Zec	14	211	8235
	39	Mal	4	55	2658
Sub-Total			67	1050	40687
Gospel	40	Mat	28	1071	32774
	41	Mak	16	678	20482
	42	Luk	24	1151	35526
	43	Jhn	21	878	27378
Sub-Total			89	3778	116160
NT History	44	Act	28	1007	33762
Sub-Total			28	1007	33762
Pauline Epistles	45	Rom	16	433	14365
	46	1Co	16	437	14154
	47	2Co	13	257	9260
	48	Gal	6	149	4817
	49	Eph	6	155	4591
	50	Phl	4	104	3408
	51	Col	4	95	3312
	52	1Ts	5	89	2983
	53	2Ts	3	47	1624
	54	1Ti	6	113	3579
	55	2Ti	4	83	2639
	56	Tit	3	46	1544
	57	Phm	1	25	717
Sub-Total			87	2033	66993
General Epistles	58	Heb	13	303	10428
	59	Jas	5	108	3534
	60	1Pe	5	105	3905
	61	2Pe	3	61	2343
	62	1Jn	5	105	3779
	63	2Jn	1	13	463
	64	3Jn	1	15	478
	65	Jud	1	25	1030
Sub-Total			34	735	25960
Apocalypse	66	Rev	22	404	15606
Sub-Total			22	404	15606
	Total		1189	31102	1062163

In [16]: util.bible_cat_info()

	category	cat	book_list	nbooks
0	Torah	tor	Gen Exo Lev Num Deu	5
1	OT History	oth	Jos Jug Rut 1Sa 2Sa 1Ki 2Ki 1Ch 2Ch Ezr Neh Est	12
2	Ketuvim	ket	Job Psm Pro Ecc Son	5
3	Major Prophets	map	Isa Jer Lam Eze Dan	5
4	Minor Prophets	mip	Hos Joe Amo Oba Jon Mic Nah Hab Zep Hag Zec Mal	12
5	Gospel	gos	Mat Mak Luk Jhn	4
6	NT History	nth	Act	1
7	Pauline Epistles	pau	Rom 1Co 2Co Gal Eph Phl Col 1Ts 2Ts 1Ti 2Ti Ti	13
8	General Epistles	ері	Heb Jas 1Pe 2Pe 1Jn 2Jn 3Jn Jud	8
9	Apocalypse	аро	Rev	1

5. OT Quotes

Out[16]:

Identify Cited Sources in OT Scripture for some NT verses

```
In [17]: rom10 = util.extract2(cdf, 'Rom 10')
       quot.show_quot(rom10, lang='chi')
      Building prefix dict from D:\Dev\Anaconda3\envs\rl\lib\site-packages\cwordtm\dictionary\dict.txt.big.txt ...
      Loading model from cache C:\Users\johnnyc\AppData\Local\Temp\jieba.u056162d159e6cf3cec64c34333130f67.cache
      Loading Chinese vocabulary 'D:\Dev\Anaconda3\envs\rl\lib\site-packages\cwordtm\data\bible_vocab.txt' ...
      Loading file 'D:\Dev\Anaconda3\envs\rl\lib\site-packages\cwordtm\data\cuv.csv' ...
      Loading file 'D:\Dev\Anaconda3\envs\rl\lib\site-packages\cwordtm\data\book_categories.csv' ...
      (1)羅 10:5 摩西寫著說、『人若行那出於律法的義、就必因此活著。』
      Loading model cost 0.882 seconds.
      Prefix dict has been built successfully.
      (2)羅10:6惟有出於信心的義如此說、『你不要心裡說、誰要升到天上去呢.就是要領下基督來
      (3)羅 10:8 他到底怎麼說呢. 他說、『這道離你不遠、正在你口裡、在你心裡。』就是我們所傳信主的道。
      (4)羅10:11經上說、『凡信他的人、必不至於羞愧。』
      (5)羅10:13因為『凡求告主名的、就必得救。』
      (6)羅10:15若沒有奉差遣、怎能傳道呢.如經上所記、『報福音傳喜信的人、他們的腳蹤何等佳美。』
      (7)羅10:16 只是人沒有都聽從福音. 因為以賽亞說、『主阿、我們所傳的有誰信呢。』
                         我們所傳的、〔或作所傳與我們的〕有誰信呢. 耶和華的膀臂向誰顯露呢。
          -> 0.6723 賽 53:1
      (8)羅10:18 但我說、人沒有聽見麼.誠然聽見了. 『他們的聲音傳遍天下、他們的言語傳到地極。』
      (9)羅10:19 我再說、以色列人不知道麼.先有摩西說、『我要用那不成子民的、惹動你們的憤恨.我要用那無知的民、觸動你們的怒氣。』
                         他們以那不算為神的、觸動我的憤恨、以虛無的神、惹了我的怒氣、我也要以那不成子民的、觸動他們的憤恨、以
          -> 0.5403 申 32:21
      愚昧的國民、惹了他們的怒氣。
      (10) 羅 10:20 又有以賽亞放膽說、『沒有尋找我的、我叫他們遇見.沒有訪問我的、我向他們顯現。』
                         素來沒有訪問我的、現在求問我.沒有尋找我的、我叫他們遇見.沒有稱為我名下的、我對他們說、我在這裡、我
         -> 0.6651 賽 65:1
      在這裡。
      (11) 羅 10:21 至於以色列人、他說、『我整天伸手招呼那悖逆頂嘴的百姓。』
                         我整天伸手招呼那悖逆的百姓、他們隨自己的意念行不善之道.
          -> 0.6086 賽 65:2
```

6. Topic Modeling

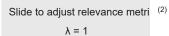
LDA Modeling

```
In [18]: # Build an LDA Model on the OT Scripture
lda = tm.lda_process(bible, cat=1, chi=True, eval=True, timing=True)
```

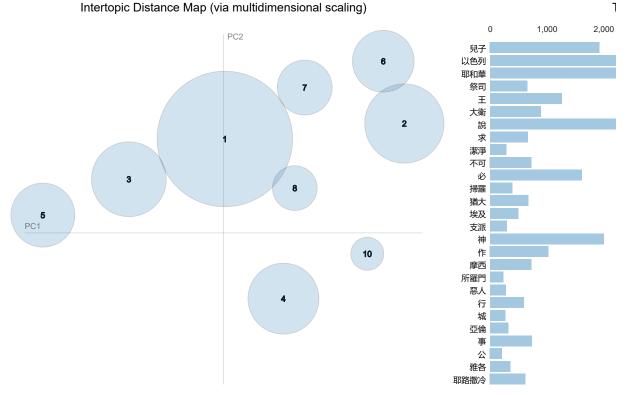
```
Loading Bible 'D:\Dev\Anaconda3\envs\rl\lib\site-packages\cwordtm\data\cuv.csv' ...
       Corpus loaded!
       Text preprocessed!
       Text trained!
       You may execute the following commands to show the visualization:
           import pyLDAvis
           pyLDAvis.display(lda.vis_data)
       Visualization prepared!
       Topics from LDA Model:
       [(0,
         '0.016*"耶和華" + 0.015*"求" + 0.009*"必" + 0.008*"智慧" + 0.008*"神" + 0.008*"惡人" + '
         '0.007*"願" + 0.006*"不要" + 0.006*"慈愛" + 0.006*"永遠"'),
        (1,
         '0.011*"支派" + 0.009*"說" + 0.009*"屬城" + 0.007*"城" + 0.007*"猶大" + 0.006*"行" + '
         '0.006*"安營" + 0.005*"生" + 0.005*"地業" + 0.005*"宗族"'),
         '0.024*"兒子" + 0.021*"說" + 0.016*"大衛" + 0.009*"耶和華" + 0.007*"掃羅" + '
         '0.007*"以色列" + 0.006*"干" + 0.006*"子孫" + 0.006*"一個" + 0.006*"作"'),
         '0.011*"王" + 0.010*"說" + 0.009*"埃及" + 0.008*"耶和華" + 0.007*"以色列" + 0.007*"神" '
         '+ 0.006*"猶大" + 0.006*"法老" + 0.005*"肘" + 0.005*"兒子"'),
        (4,
         '0.023*"耶和華" + 0.018*"以色列" + 0.013*"王" + 0.011*"神" + 0.010*"事" + 0.010*"兒子" '
         '+ 0.009*"所羅門" + 0.009*"說" + 0.008*"耶路撒冷" + 0.007*"殿"'),
         '0.043*"耶和華" + 0.035*"說" + 0.014*"以色列" + 0.012*"神" + 0.009*"必" + 0.006*"百姓" '
         '+ 0.006*"王" + 0.006*"摩西" + 0.005*"沒有" + 0.005*"中"'),
         '0.016*"潔淨" + 0.014*"祭司" + 0.009*"耶和華" + 0.008*"作" + 0.007*"屬" + 0.006*"被數" '
         '+ 0.006*"災病" + 0.005*"兒子" + 0.005*"共有" + 0.005*"以色列"'),
         '0.025*"說" + 0.023*"兒子" + 0.017*"以色列" + 0.015*"耶和華" + 0.009*"掃羅" + '
         '0.009*"大衛" + 0.009*"約瑟" + 0.007*"父親" + 0.007*"雅各" + 0.005*"王"'),
        (8,
         '0.036*"耶和華" + 0.016*"不可" + 0.009*"喫" + 0.009*"神" + 0.008*"公" + 0.008*"作" + '
         '0.007*"祭司" + 0.007*"說" + 0.007*"燔祭" + 0.007*"以色列"'),
        (9.
         '0.017*"說" + 0.016*"耶和華" + 0.015*"必" + 0.007*"神" + 0.005*"中" + 0.004*"沒有" + '
         '0.004*"喫" + 0.003*"好像" + 0.003*"知道" + 0.003*"作"')]
       Model Evaluation Scores:
         Coherence: 0.3733050435396904
         Perplexity: -8.225260396967666
         Topic diversity: 0.001289792973924288
         Topic size distribution: 0.0024937655860349127
       Finished 'lda_process' in 36.0294 secs
In [19]: # Show LDA Model Visualization
         import pyLDAvis
```

pyLDAvis.display(lda.vis_data)





Intertopic Distance Map (via multidimensional scaling)



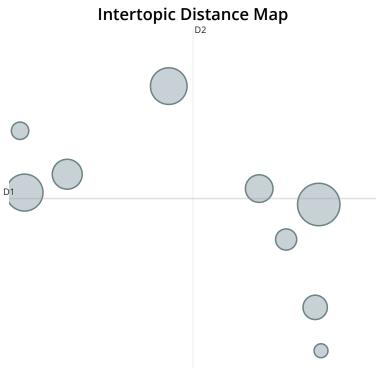
BERTopic Modeling

```
In [20]: # Build a BERTopic Model on the OT Scripture
        btm = tm.btm_process(bible, cat=1, chi=True, eval=True, timing=True)
      Loading Bible 'D:\Dev\Anaconda3\envs\rl\lib\site-packages\cwordtm\data\cuv.csv' ...
      Corpus loaded!
      Chinese text preprocessed!
                                   | 0.00/412M [00:00<?, ?B/s]
      model.safetensors: 0%
      Text trained!
      Topics from BERTopic Model:
      Topic 0: 耶和華 | 不可 | 以色列 | 雅各 | 兒子 | 沒有 | 面前 | 埃及 | 地上 | 知道
      Topic 1: 耶和華 | 以色列 | 摩西 | 支派 | 沒有 | 兒子 | 百姓 | 吩咐 | 一個 | 埃及
      Topic 2: 兒子 | 耶和華 | 以色列 | 大衛 | 巴比倫 | 子孫 | 猶大 | 耶路撒冷 | 萬軍 | 非利士人
      Topic 3: 祭司 | 耶和華 | 兒子 | 潔淨 | 以色列 | 利未人 | 不可 | 子孫 | 燔祭 | 亞倫
      Topic 4: 兒子 | 大衛 | 雅各 | 以掃 | 父親 | 亞伯拉罕 | 僕人 | 一個 | 女子 | 女兒
      Topic 5: 耶和華 | 讚美 | 永遠 | 的詩 | 伶長 | 公義 | 之詩 | 交與 | 上行 | 大衛
      Topic 6: 智慧 | 惡人 | 愚昧 | 義人 | 知識 | 正直 | 日光 | 之下 | 虚空 | 言語
      Topic 7: 耶和華 | 獻給 | 摩西 | 火祭 | 不可 | 以色列 | 素祭 | 亞倫 | 基路伯 | 羊羔
      Topic 8: 歌頌 | 我要 | 耶和華 | 歡呼 | 稱謝 | 詩歌 | 稱讚 | 萬民 | 救恩 | 細拉
```

Model Evaluation Scores:

Coherence: 0.33409250158597886

BERTopic Model Visualization:

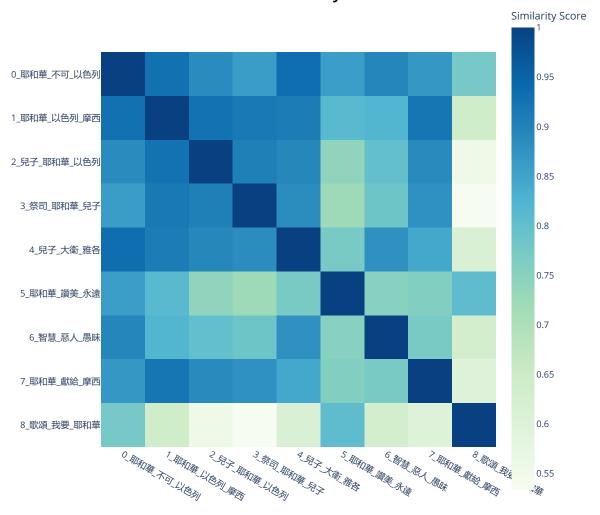




Topic Word Scores



Similarity Matrix



Finished 'btm_process' in 77.9931 secs

NMF Modeling

In [21]: nmf = tm.nmf_process(bible, num_topics=8, cat=1, chi=True, eval=True, timing=True)

 $\label{libsite-packages} Loading Bible 'D:\Dev\Anaconda3\envs\rl\lib\site-packages\cwordtm\data\cuv.csv' \dots Corpus loaded!$

Text preprocessed!

 $\label{libsite-packages} D: \Dev\Anaconda \envs\rl\lib\site-packages\gensim\mbox{models} nmf.py: 578: Deprecation \Warning: \Deprecation \Warning: \Packages \Packag$

Please import `csc_matrix` from the `scipy.sparse` namespace; the `scipy.sparse.csc` namespace is deprecated and will be removed in SciPy 2.0.0.

 $\label{libsite-packages lens} D: \Dev\Anaconda \envs\rl\lib\site-packages \gensim\mbox{models nmf.py:} 607: Deprecation \Warning: \Deprecation \Warning: \Berry \Libsite-packages \gensim\mbox{models nmf.py:} 607: Deprecation \Warning: \Berry \gensim\mbox{models nmf.py:} 607: Depre$

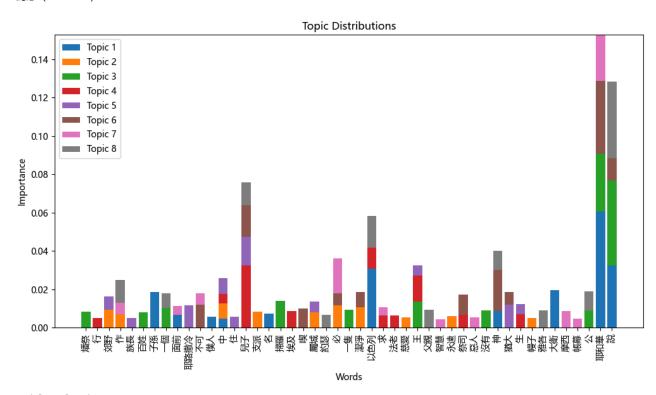
Please import `csc_matrix` from the `scipy.sparse` namespace; the `scipy.sparse.csc` namespace is deprecated and will be removed in SciPy 2.0.0.

D:\Dev\Anaconda3\envs\rl\lib\site-packages\gensim\models\nmf.py:620: DeprecationWarning:

Please import `csc_matrix` from the `scipy.sparse` namespace; the `scipy.sparse.csc` namespace is deprecated and will be removed in SciPy 2.0.0.

Topics-Words from NMF Model: Topic 1: 耶和華 (0.060434) 說 (0.032314) 以色列 (0.030670) 大衛 (0.019521) 子孫 (0.018642) 神 (0.008524) 名 (0.007180) 面前 (0.006592) 僕人 (0.005790) 中 (0.004794) Topic 2: 必 (0.011584) 潔淨 (0.010808) 郊野 (0.009289) 支派 (0.008285) 屬城 (0.008021) 中 (0.007965) 作 (0.006976) 永遠 (0.005992) 慈愛 (0.005299) 幔子 (0.004960) Topic 3: 說 (0.044386) 耶和華 (0.030060) 掃羅 (0.013935) 王 (0.013464) 一個 (0.010244) 一隻 (0.009441) 公 (0.008918) 沒有 (0.008851) 燔祭 (0.008414) 百姓 (0.007840) Topic 4: 兒子 (0.032498) 王 (0.013861) 以色列 (0.010919) 埃及 (0.008594) 生 (0.006986) 祭司 (0.006756) 求 (0.006480) 法老 (0.006437) 行 (0.004900) 中 (0.004820) Topic 5: 兒子 (0.014829) 猶大 (0.011927) 耶路撒冷 (0.011567) 中 (0.008391) 郊野 (0.006903) 住 (0.005849) 屬城 (0.005585) 生 (0.005265) 王 (0.005068) 族長 (0.005021) Topic 6: 耶和華 (0.038127) 神 (0.021477) 兒子 (0.016476) 不可 (0.011861) 說 (0.011471) 祭司 (0.010620) 喫 (0.010078) 潔淨 (0.007774) 猶大 (0.006703)

```
Topic 7:
耶和華 (0.024071)
必 (0.018089)
摩西 (0.008711)
不可 (0.006204)
作 (0.005945)
惡人 (0.005274)
帳幕 (0.004705)
面前 (0.004632)
智慧 (0.004322)
求 (0.004062)
Topic 8:
說 (0.040113)
以色列 (0.016668)
兒子 (0.011958)
作 (0.011920)
神 (0.009935)
公 (0.009887)
父親 (0.009470)
雅各 (0.008914)
一個 (0.007758)
約瑟 (0.006721)
```



Model Evaluation Scores:
Coherence: 0.28930507302481484
Topic diversity: 0.001151257977212733
Topic size distribution: 0.0018497109826589595

Finished 'nmf_process' in 21.8099 secs

Save NMF Model

```
In [22]: nmf.save("models/nmf_cuv")

NMF model has been saved: 'models/nmf_cuv_model.gensim' and 'models/nmf_cuv_dictionary.gensim'
```

Load NMF Model

```
In [23]: nmf2 = tm.NMF("", nmf.num_topics)
nmf2.model, nmf2.dictionary = nmf2.load("models/nmf_cuv")
nmf2.show_topics_words()
```

Topics-Words from NMF Model: Topic 1: 耶和華 (0.060434) 說 (0.032314) 以色列 (0.030670) 大衛 (0.019521) 子孫 (0.018642) 神 (0.008524) 名 (0.007180) 面前 (0.006592) 僕人 (0.005790) 中 (0.004794) Topic 2: 必 (0.011584) 潔淨 (0.010808) 郊野 (0.009289) 支派 (0.008285) 屬城 (0.008021) 中 (0.007965) 作 (0.006976) 永遠 (0.005992) 慈愛 (0.005299) 幔子 (0.004960) Topic 3: 說 (0.044386) 耶和華 (0.030060) 掃羅 (0.013935) 王 (0.013464) 一個 (0.010244) 一隻 (0.009441) 公 (0.008918) 沒有 (0.008851) 燔祭 (0.008414) 百姓 (0.007840) Topic 4: 兒子 (0.032498) 王 (0.013861) 以色列 (0.010919) 埃及 (0.008594) 生 (0.006986) 祭司 (0.006756) 求 (0.006480) 法老 (0.006437) 行 (0.004900) 中 (0.004820) Topic 5: 兒子 (0.014829) 猶大 (0.011927) 耶路撒冷 (0.011567) 中 (0.008391) 郊野 (0.006903) 住 (0.005849) 屬城 (0.005585) 生 (0.005265) 王 (0.005068) 族長 (0.005021) Topic 6: 耶和華 (0.038127) 神 (0.021477) 兒子 (0.016476) 不可 (0.011861) 說 (0.011471) 祭司 (0.010620) 喫 (0.010078) 潔淨 (0.007774) 猶大 (0.006703)

Topic 7:

必 (0.006357)

必 (0.018089)
摩西 (0.008711)
不可 (0.006204)
作 (0.005945)
惡人 (0.005274)
帳幕 (0.004705)
面前 (0.004632)
智慧 (0.004322)
求 (0.004062)

Topic 8: 說 (0.040113) 以色列 (0.016668) 兒子 (0.011958) 作 (0.011920) 神 (0.009935) 公 (0.009887) 父親 (0.009470) 雅各 (0.008914) 一個 (0.007758) 約瑟 (0.006721)