

```
1 C:\Users\Crystina\AppData\Local\Programs\Python\
  Python35\python.exe "D:/PycharmProjects/FYP/
  2_feature_extraction/sougo_word_nmf_kmeans.py"
2 extracting tf-idf feature...
3 Fitting the NMF model (Frobenius norm) with TF-
  IDF features
4 Performance with 3 cluster:
5     silhouette_score:  0.4266968962858341
6     calinski_harabaz_score:  47.474354294697925
7 Performance with 4 cluster:
8     silhouette_score:  0.46321303065125163
9     calinski_harabaz_score:  48.372100625075184
10 Performance with 5 cluster:
11     silhouette_score:  0.5002597814272685
12     calinski_harabaz_score:  50.06083739428638
13 Performance with 6 cluster:
14     silhouette_score:  0.5345237404172782
15     calinski_harabaz_score:  57.26249843623332
16 Performance with 7 cluster:
17     silhouette_score:  0.5392397619808799
18     calinski_harabaz_score:  71.34669140610356
19 Performance with 8 cluster:
20     silhouette_score:  0.5840971021552549
21     calinski_harabaz_score:  90.370158049469
22 Performance with 9 cluster:
23     silhouette_score:  0.5801481349858659
24     calinski_harabaz_score:  118.13652162609269
25 Performance with 10 cluster:
26     silhouette_score:  0.6465152023576902
27     calinski_harabaz_score:  170.2822504117251
28 Performance with 11 cluster:
29     silhouette_score:  0.7048102710717995
30     calinski_harabaz_score:  247.21012664150095
31 Performance with 12 cluster:
32     silhouette_score:  0.7111616823423401
33     calinski_harabaz_score:  257.8898545894717
34 Performance with 13 cluster:
35     silhouette_score:  0.7151091847299585
36     calinski_harabaz_score:  267.52082174586286
37 Performance with 14 cluster:
38     silhouette_score:  0.6828406459942639
39     calinski_harabaz_score:  271.31299201146044
```

```
40 Performance with 15 cluster:
41     silhouette_score: 0.713573044348318
42     calinski_harabaz_score: 284.0422238662047
43 Performance with 16 cluster:
44     silhouette_score: 0.7062729574944083
45     calinski_harabaz_score: 293.317279091164
46 Performance with 17 cluster:
47     silhouette_score: 0.6723365604513053
48     calinski_harabaz_score: 306.5072897707933
49 Performance with 18 cluster:
50     silhouette_score: 0.6709792855967328
51     calinski_harabaz_score: 316.547274813899
52 Performance with 19 cluster:
53     silhouette_score: 0.6689166383047348
54     calinski_harabaz_score: 338.4673399764017
55 Performance with 20 cluster:
56     silhouette_score: 0.6654097029051783
57     calinski_harabaz_score: 370.49157774760573
58 Performance with 21 cluster:
59     silhouette_score: 0.6582541585948897
60     calinski_harabaz_score: 391.62685781909454
61 Performance with 22 cluster:
62     silhouette_score: 0.6555906855754841
63     calinski_harabaz_score: 401.57227199737025
64 Performance with 23 cluster:
65     silhouette_score: 0.6502230024447481
66     calinski_harabaz_score: 400.72726834517965
67 Performance with 24 cluster:
68     silhouette_score: 0.6273258217724623
69     calinski_harabaz_score: 399.7229348100768
70 Performance with 25 cluster:
71     silhouette_score: 0.6168236175138242
72     calinski_harabaz_score: 402.10158186515275
73 Performance with 26 cluster:
74     silhouette_score: 0.6097536537164991
75     calinski_harabaz_score: 400.8170475801369
76 Performance with 27 cluster:
77     silhouette_score: 0.629854422470853
78     calinski_harabaz_score: 399.58003215394416
79 Performance with 28 cluster:
80     silhouette_score: 0.617597439661694
81     calinski_harabaz_score: 404.20680834493055
```

```
82 Performance with 29 cluster:
83     silhouette_score: 0.6338885774325613
84     calinski_harabaz_score: 417.02871760773706
85 Performance with 30 cluster:
86     silhouette_score: 0.6172272108813389
87     calinski_harabaz_score: 403.7428428969682
88 Performance with 31 cluster:
89     silhouette_score: 0.602826566314675
90     calinski_harabaz_score: 416.15079743319853
91 Performance with 32 cluster:
92     silhouette_score: 0.6015378345562735
93     calinski_harabaz_score: 413.36936007328615
94 Performance with 33 cluster:
95     silhouette_score: 0.6033706580019889
96     calinski_harabaz_score: 413.4072635361434
97 Performance with 34 cluster:
98     silhouette_score: 0.6025416787506642
99     calinski_harabaz_score: 432.0327437585838
100 Performance with 35 cluster:
101     silhouette_score: 0.5852698254265468
102     calinski_harabaz_score: 447.2773244895394
103 Performance with 36 cluster:
104     silhouette_score: 0.5851369935308631
105     calinski_harabaz_score: 429.496282996276
106 Performance with 37 cluster:
107     silhouette_score: 0.5835382823026436
108     calinski_harabaz_score: 466.96158188900586
109 Performance with 38 cluster:
110     silhouette_score: 0.5894970661455787
111     calinski_harabaz_score: 467.58790066785014
112 Performance with 39 cluster:
113     silhouette_score: 0.5663622548511232
114     calinski_harabaz_score: 457.6331057775336
115 Performance with 40 cluster:
116     silhouette_score: 0.5744102695512933
117     calinski_harabaz_score: 470.4535038599692
118 Performance with 41 cluster:
119     silhouette_score: 0.576701595972323
120     calinski_harabaz_score: 474.47558910304366
121 Performance with 42 cluster:
122     silhouette_score: 0.584923298236542
123     calinski_harabaz_score: 475.9185173386504
```

```
124 Performance with 43 cluster:
125     silhouette_score: 0.5657143162198287
126     calinski_harabaz_score: 489.3365527887897
127 Performance with 44 cluster:
128     silhouette_score: 0.5720271629211563
129     calinski_harabaz_score: 489.78806631448276
130 Performance with 45 cluster:
131     silhouette_score: 0.5748093894079882
132     calinski_harabaz_score: 505.3140278365244
133 Performance with 46 cluster:
134     silhouette_score: 0.5615920844526328
135     calinski_harabaz_score: 523.4759126065427
136 Performance with 47 cluster:
137     silhouette_score: 0.5612381473844562
138     calinski_harabaz_score: 520.4392328534617
139 Performance with 48 cluster:
140     silhouette_score: 0.5518898678175651
141     calinski_harabaz_score: 520.2670137238731
142 Performance with 49 cluster:
143     silhouette_score: 0.5608335698830443
144     calinski_harabaz_score: 555.3568831628236
145
146 Process finished with exit code 0
147
```