

```
1 C:\Users\Crystina\AppData\Local\Programs\Python\
  Python35\python.exe "D:/PycharmProjects/FYP/
  2_feature_extraction/sougo_word_pca_kmeans.py"
2 extracting tf-idf feature...
3 Fitting the NMF model (Frobenius norm) with TF-
  IDF features
4 (200, 10)
5 Performance with 3 cluster:
6     silhouette_score:  0.5137292519146941
7     calinski_harabaz_score:  170.798149020914
8 Performance with 4 cluster:
9     silhouette_score:  0.5356992244052237
10    calinski_harabaz_score:  140.27915421207462
11 Performance with 5 cluster:
12    silhouette_score:  0.5880425141984996
13    calinski_harabaz_score:  130.1903365298701
14 Performance with 6 cluster:
15    silhouette_score:  0.5695756185196429
16    calinski_harabaz_score:  126.80879428361744
17 Performance with 7 cluster:
18    silhouette_score:  0.6107517301305907
19    calinski_harabaz_score:  136.14837532730223
20 Performance with 8 cluster:
21    silhouette_score:  0.6157326363368925
22    calinski_harabaz_score:  149.23332330013397
23 Performance with 9 cluster:
24    silhouette_score:  0.5923226792543576
25    calinski_harabaz_score:  168.0966113753411
26 Performance with 10 cluster:
27    silhouette_score:  0.6072603378891119
28    calinski_harabaz_score:  190.79967729291948
29 Performance with 11 cluster:
30    silhouette_score:  0.47604622975599964
31    calinski_harabaz_score:  207.56941788346887
32 Performance with 12 cluster:
33    silhouette_score:  0.4768713143599711
34    calinski_harabaz_score:  209.92954505193123
35 Performance with 13 cluster:
36    silhouette_score:  0.425645641902886
37    calinski_harabaz_score:  211.70258363715044
38 Performance with 14 cluster:
39    silhouette_score:  0.4411668009941704
```

```
40     calinski_harabaz_score: 213.6583538772524
41 Performance with 15 cluster:
42     silhouette_score: 0.4436743286874013
43     calinski_harabaz_score: 218.79499062568445
44 Performance with 16 cluster:
45     silhouette_score: 0.4452422790009794
46     calinski_harabaz_score: 217.2151233864357
47 Performance with 17 cluster:
48     silhouette_score: 0.45386199097640373
49     calinski_harabaz_score: 217.25800843873907
50 Performance with 18 cluster:
51     silhouette_score: 0.4229665832830948
52     calinski_harabaz_score: 213.35231583375742
53 Performance with 19 cluster:
54     silhouette_score: 0.41230510561026706
55     calinski_harabaz_score: 215.66140315154306
56 Performance with 20 cluster:
57     silhouette_score: 0.4459702477350749
58     calinski_harabaz_score: 220.7803672432602
59 Performance with 21 cluster:
60     silhouette_score: 0.461007811004116
61     calinski_harabaz_score: 217.35601185922047
62 Performance with 22 cluster:
63     silhouette_score: 0.4300536170428615
64     calinski_harabaz_score: 212.929006574054
65 Performance with 23 cluster:
66     silhouette_score: 0.43943313612591395
67     calinski_harabaz_score: 216.95423273314807
68 Performance with 24 cluster:
69     silhouette_score: 0.4286566040383548
70     calinski_harabaz_score: 208.45222950189682
71 Performance with 25 cluster:
72     silhouette_score: 0.43141919830571374
73     calinski_harabaz_score: 211.91416796471313
74 Performance with 26 cluster:
75     silhouette_score: 0.44065923230992354
76     calinski_harabaz_score: 213.45216370642436
77 Performance with 27 cluster:
78     silhouette_score: 0.42091786462113473
79     calinski_harabaz_score: 211.0074385868482
80 Performance with 28 cluster:
81     silhouette_score: 0.4291556712680499
```

```
82     calinski_harabaz_score: 209.1807888372702
83 Performance with 29 cluster:
84     silhouette_score: 0.4317276777940273
85     calinski_harabaz_score: 214.06452491619586
86 Performance with 30 cluster:
87     silhouette_score: 0.4313818251283372
88     calinski_harabaz_score: 216.90556423390422
89 Performance with 31 cluster:
90     silhouette_score: 0.43478917177399745
91     calinski_harabaz_score: 217.51355967024517
92 Performance with 32 cluster:
93     silhouette_score: 0.4293444451781883
94     calinski_harabaz_score: 223.41257233401274
95 Performance with 33 cluster:
96     silhouette_score: 0.4226386763659971
97     calinski_harabaz_score: 217.3350860735972
98 Performance with 34 cluster:
99     silhouette_score: 0.4353450509205954
100    calinski_harabaz_score: 223.96675091317096
101 Performance with 35 cluster:
102    silhouette_score: 0.43849579201938765
103    calinski_harabaz_score: 221.29463638106802
104 Performance with 36 cluster:
105    silhouette_score: 0.4301860012653458
106    calinski_harabaz_score: 224.43060640725744
107 Performance with 37 cluster:
108    silhouette_score: 0.4297663042241409
109    calinski_harabaz_score: 223.11754348118916
110 Performance with 38 cluster:
111    silhouette_score: 0.43457179015489417
112    calinski_harabaz_score: 233.73822572422847
113 Performance with 39 cluster:
114    silhouette_score: 0.43330427345984246
115    calinski_harabaz_score: 230.5258468009648
116 Performance with 40 cluster:
117    silhouette_score: 0.4453814152632931
118    calinski_harabaz_score: 243.23122913639634
119 Performance with 41 cluster:
120    silhouette_score: 0.4365707990826479
121    calinski_harabaz_score: 241.395184103699
122 Performance with 42 cluster:
123    silhouette_score: 0.43985471146970556
```

```
124     calinski_harabaz_score: 241.51019184021644
125 Performance with 43 cluster:
126     silhouette_score: 0.44042518842974315
127     calinski_harabaz_score: 248.02425746421414
128 Performance with 44 cluster:
129     silhouette_score: 0.44554845473183485
130     calinski_harabaz_score: 254.0054092766995
131 Performance with 45 cluster:
132     silhouette_score: 0.4395480706773632
133     calinski_harabaz_score: 249.7873282538188
134 Performance with 46 cluster:
135     silhouette_score: 0.4526174347787692
136     calinski_harabaz_score: 264.4949969727844
137 Performance with 47 cluster:
138     silhouette_score: 0.45255109146717953
139     calinski_harabaz_score: 261.5622220367517
140 Performance with 48 cluster:
141     silhouette_score: 0.4354782192151081
142     calinski_harabaz_score: 254.776218258966
143 Performance with 49 cluster:
144     silhouette_score: 0.4532900289587933
145     calinski_harabaz_score: 264.2920946768401
146
147 Process finished with exit code 0
148
```