

```
1 C:\Users\Crystina\AppData\Local\Programs\Python\
  Python35\python.exe "D:/PycharmProjects/FYP/
  2_feature extraction/
  sougo_word_truncatedsvd_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: 0.516162023651447
7     calinski_harabaz_score: 178.8787630349813
8 Performance with 4 cluster:
9     silhouette_score: 0.5383269999198057
10    calinski_harabaz_score: 147.83355433000114
11 Performance with 5 cluster:
12    silhouette_score: 0.5699780004487793
13    calinski_harabaz_score: 136.75310153676452
14 Performance with 6 cluster:
15    silhouette_score: 0.5614548391616898
16    calinski_harabaz_score: 135.12576157521693
17 Performance with 7 cluster:
18    silhouette_score: 0.5903138505489282
19    calinski_harabaz_score: 145.0143901142449
20 Performance with 8 cluster:
21    silhouette_score: 0.6115275896195084
22    calinski_harabaz_score: 160.41346882955787
23 Performance with 9 cluster:
24    silhouette_score: 0.5589359329029614
25    calinski_harabaz_score: 179.50319330317353
26 Performance with 10 cluster:
27    silhouette_score: 0.5918607370986697
28    calinski_harabaz_score: 200.69120181375234
29 Performance with 11 cluster:
30    silhouette_score: 0.446273907638083
31    calinski_harabaz_score: 223.47490721579902
32 Performance with 12 cluster:
33    silhouette_score: 0.44715151529071906
34    calinski_harabaz_score: 224.79522106358806
35 Performance with 13 cluster:
36    silhouette_score: 0.44266484322720756
37    calinski_harabaz_score: 223.8627188811864
38 Performance with 14 cluster:
```

```
39     silhouette_score: 0.4451636036405687
40     calinski_harabaz_score: 226.73471809471076
41 Performance with 15 cluster:
42     silhouette_score: 0.46601239079294443
43     calinski_harabaz_score: 229.82910046005574
44 Performance with 16 cluster:
45     silhouette_score: 0.44578937547445874
46     calinski_harabaz_score: 217.10082276214433
47 Performance with 17 cluster:
48     silhouette_score: 0.43677130525065005
49     calinski_harabaz_score: 223.1110746070613
50 Performance with 18 cluster:
51     silhouette_score: 0.4612899372812852
52     calinski_harabaz_score: 215.33865654783637
53 Performance with 19 cluster:
54     silhouette_score: 0.4289097237906707
55     calinski_harabaz_score: 217.7539273505959
56 Performance with 20 cluster:
57     silhouette_score: 0.4391248192181621
58     calinski_harabaz_score: 213.6876264015733
59 Performance with 21 cluster:
60     silhouette_score: 0.4262311499499309
61     calinski_harabaz_score: 212.8520884154935
62 Performance with 22 cluster:
63     silhouette_score: 0.4452541679883234
64     calinski_harabaz_score: 217.26602706860854
65 Performance with 23 cluster:
66     silhouette_score: 0.4450819770737564
67     calinski_harabaz_score: 216.55799883399214
68 Performance with 24 cluster:
69     silhouette_score: 0.43325982419221504
70     calinski_harabaz_score: 212.80612083705338
71 Performance with 25 cluster:
72     silhouette_score: 0.4379688795601349
73     calinski_harabaz_score: 215.24945571897587
74 Performance with 26 cluster:
75     silhouette_score: 0.43985417043129155
76     calinski_harabaz_score: 210.66360279906132
77 Performance with 27 cluster:
78     silhouette_score: 0.4432103185880766
79     calinski_harabaz_score: 208.64777538588166
80 Performance with 28 cluster:
```

```
81     silhouette_score: 0.4483272991201555
82     calinski_harabaz_score: 212.53993942903412
83 Performance with 29 cluster:
84     silhouette_score: 0.4519239825516621
85     calinski_harabaz_score: 212.06369041751537
86 Performance with 30 cluster:
87     silhouette_score: 0.4464881246725726
88     calinski_harabaz_score: 210.1439438397575
89 Performance with 31 cluster:
90     silhouette_score: 0.43809249672286216
91     calinski_harabaz_score: 207.32189724777425
92 Performance with 32 cluster:
93     silhouette_score: 0.45032176972212773
94     calinski_harabaz_score: 209.09862597650437
95 Performance with 33 cluster:
96     silhouette_score: 0.44887029424701863
97     calinski_harabaz_score: 210.30665392042485
98 Performance with 34 cluster:
99     silhouette_score: 0.4402271032161949
100    calinski_harabaz_score: 210.3916988492101
101 Performance with 35 cluster:
102    silhouette_score: 0.4305273549871657
103    calinski_harabaz_score: 206.40756741644273
104 Performance with 36 cluster:
105    silhouette_score: 0.43572998039061217
106    calinski_harabaz_score: 213.10590595546367
107 Performance with 37 cluster:
108    silhouette_score: 0.44196748518107254
109    calinski_harabaz_score: 210.8571451049907
110 Performance with 38 cluster:
111    silhouette_score: 0.4432817716954687
112    calinski_harabaz_score: 213.24935321094088
113 Performance with 39 cluster:
114    silhouette_score: 0.4461778294165775
115    calinski_harabaz_score: 212.08147789498994
116 Performance with 40 cluster:
117    silhouette_score: 0.440091999820008
118    calinski_harabaz_score: 212.00958586781613
119 Performance with 41 cluster:
120    silhouette_score: 0.44109732945881946
121    calinski_harabaz_score: 214.34806556729563
122 Performance with 42 cluster:
```

```
123     silhouette_score: 0.4462657092157516
124     calinski_harabaz_score: 221.0586492613413
125 Performance with 43 cluster:
126     silhouette_score: 0.4443565019358745
127     calinski_harabaz_score: 220.42066206978447
128 Performance with 44 cluster:
129     silhouette_score: 0.42769142470966287
130     calinski_harabaz_score: 217.96602606611063
131 Performance with 45 cluster:
132     silhouette_score: 0.44089870308789825
133     calinski_harabaz_score: 219.7404217102048
134 Performance with 46 cluster:
135     silhouette_score: 0.43892847241533284
136     calinski_harabaz_score: 222.58848698205222
137 Performance with 47 cluster:
138     silhouette_score: 0.4412782280210845
139     calinski_harabaz_score: 223.32449267162357
140 Performance with 48 cluster:
141     silhouette_score: 0.43961567724447514
142     calinski_harabaz_score: 228.51663897415818
143 Performance with 49 cluster:
144     silhouette_score: 0.4436851758223676
145     calinski_harabaz_score: 229.6153078565504
146
147 Process finished with exit code 0
148
```